

Uncertainty reduction in reservoir parameters prediction from multiscale data using machine learning in deep offshore reservoirs.

ARIGBE, O.D.

2020

The author of this thesis retains the right to be identified as such on any occasion in which content from this thesis is referenced or re-used. The licence under which this thesis is distributed applies to the text and any original images only – re-use of any third-party content must still be cleared with the original copyright holder.

UNCERTAINTY REDUCTION IN RESERVOIR
PARAMETERS PREDICTION FROM MULTISCALE DATA
USING MACHINE LEARNING IN DEEP OFFSHORE
RESERVOIRS

OVOKE DANIEL ARIGBE

Dedication

I dedicate this work to the glory of God Almighty for supplying all the grace I needed even in very trying times. This work is also dedicated to the memory of my late father, Mr Richard Arigbe who always told me that education is the key; and also to my sweet mother, Rebecca Arigbe who even though did not have a formal education herself prayed and worked tirelessly to ensure her children get one.

Acknowledgement

I offer all the praise to the Almighty God whose name is YAHWEH in Hebrew and OGHENE in Urhobo my native dialect, the Father of our Lord and Saviour Jesus Christ. His word in Philippians Chapter 1 verse 6 says "Being confident of this very thing, that he who hath begun a good work in you will finish it until the day of Jesus Christ" has indeed kept His word and seen me through this journey of my PhD. Even when it looked like the results were not forthcoming, this blessed assurance gave me hope in these trying time. To Him alone be all the glory.

My profound gratitude goes to my supervisory team in persons of Prof. Babs Oyeneyin, Dr Ines Arana and Dr. Ghazi Droubi for their very important technical and moral supports through the period of the PhD programme. Their believe in me and understanding of the subject making them to give informed suggestions while also allowing me to express myself, were a major source of inspiration for me. Special thanks also to the Niger Delta Development Commission for all the support before and during my PhD. I also want to thank Dr. Gbenga Oluyemi, Dr. Sani Sadiq and all the staff of both the Schools of Engineering and Computing for their technical support and encouragement; Pst. and Pst Mrs. Dapo Olarenwaju as well as the members especially Evangelism and Choir department in Jesus House Torry for their prayers and encouragement; and to my friends and colleagues, Pst. and Pst. Mrs Ben Fakoya, Pst. Chris, Kenneth Afebu, Temienor Nikasekpe, Andrew Oni, Voke Salubi, Ovonically Otarigho, Dammy, Lolu, Afolabi Morakinyo, Bernard Oghenekevwe, Maureen Ani, Chimere Obodo, James Olugbade, Iton Whiteney, Vanessa Cliff-Ekubo, Pst Felix, Joyce Adeyele, Dorothy McDonald, Bola, Coleman and Danielle, Martin Simpson, Dr Rosslyn Shanks, Folashade, Dr. Arindam, Russein, Balogun, Dr Elizabeth Wuyep, Ofasa, Mariam, Larry Oyibo, Seyi, Ogochukwu Unachukwu and many others whom were not mention, for their support and prayers in the course of my research.

This acknowledgement will not be complete without thanking my sweet mother, Rebecca Arigbe; my brothers and sister, Israel and Rita Arigbe, Edith Ideke, Ufuoma Arigbe, Abel Iyamu, Believe Iyamu, Jacob Iyamu, Monday Iyamu, Efe Arigbe, Joshua Arigbe, Akpos Arigbe, Tive for their love, prayers and support. Many thanks to all of you!

Abstract

Developing a complete characterization of reservoir properties involved in subsurface multiphase flow is a very challenging task. In most cases, these properties such as porosity, water saturation, permeability (and their variants), pressure, wettability, bulk modulus, Young Modulus, shear modulus, fracture gradient cannot be directly measured and if measured are only available at small number of well locations. The limited data are then combined with geological interpretation to generate a model. Also increasing the degree of this uncertainty is that the reservoir properties from different data sources like well logs, cores and well test most times produce different results thus making predictions less accurate.

The present study focussed on three reservoir parameters which are porosity, fluid saturation and permeability selected based on literature and sensitivity analysis using Monte Carlo simulations on net present value, reserve estimates and pressure transients. Sandstone assets from the North Sea were used to establish the technique for uncertainty reduction using machine learning as well as empirical models after data digitization and cleaning. These models were built (trained) with observed data using other variables as inputs after which they were tested by then using the input variables (not used for the training) to predict their corresponding observed data. Root Mean Squared Error (RMSE) of the predicted and the actual observed data was calculated. Model tuning was done in order to optimize its key parameters to reduce RMSE. Appropriate log, core and test depth matching was also ensured including upscaling combined with Lorenz plot to identify the dominant flow interval. Nomographic approach involving a numerical simulation run iteratively on multiple non-linear regression model obtained from the dataset was also run. Sandstone reservoirs from the North Sea not used for developing the models were then used to validate the different techniques earlier developed.

Based on the aforementioned, the degree of uncertainty associated with porosity, permeability and fluid saturation usage was demonstrated and reduced. For example, improved accuracies of 1-74%, 4-77% and 40% were achieved for Raymer, Wyllie and Modified Schlumberger respectively. Raymer and Wyllie were also not suitable for unconsolidated sandstones while machine learning models were the most accurate. Evaluation of logs, core and test from several wells showed permeability across board to be different which also highlights the uncertainty in their interpretation. The gap between log, core and test was also closed using machine learning and nomographic methods. The machine learning model was then coded into a dashboard containing the inputs for its training. Their relationship provides the bench mark to calibrate one against the other and also create the platform for real time reservoir properties prediction. The technology was applied to an independent dataset from the Central North Sea deep offshore

sandstone reservoir for the validation of these models with minimum tuning and thus effective for real time reservoir and production management. While uncertainties in measurements are crucial, the focus of this work was on the intermediate models to get better final geological models since the measured data used were from the industry.

Table of Contents

Dedication	ii
Acknowledgement.....	iii
Abstract	iv
Table of Contents	v
List of Appendices	xi
List of Figures	xii
List of Tables	xviii
Nomenclature	xix

Chapter One Introduction

1.1 Introduction	1
1.2 Evaluation Method	6
1.3 Literature Review	6
1.4 Project Driver	8
1.5 Aim of Project	10
1.6 Objectives	10
1.7 Research Questions	11
1.8 Conceptual Plan	12
1.9 Contributions to Knowledge	13
1.10 Thesis Layout/Arrangement	14

Chapter Two Review of Sandstone Reservoir Parameters

2.1 Introduction	15
2.2 Sandstone Reservoirs	16
2.2.1 North Sea Basin	18
2.3 Well Log Practice	21
2.3.1 Gamma Ray	22

2.3.2	Spontaneous Potential	22
2.3.3	Resistivity	22
2.3.4	Density	22
2.3.5	Neutron	22
2.3.6	Sonic	23
2.4	Cores Analysis.....	30
2.5	Well Testing Analysis.....	31
2.5.1	Pressure Build-Up	32
2.5.2	Pressure Drawdown	33
2.5.3	Drill Stem Test	34
2.5.4	Interference Testing	34
2.5.5	Pulse Test	35
2.5.6	Injection Tests	36
2.5.7	Falloff Tests	37
2.5.8	Modular Dynamic Tester	38
2.5.9	Reservoir Saturation Tool	38
2.6	Sandstone Reservoir Rock and Fluid Properties	39
2.6.1	Porosity	39
2.6.2	Fluid Saturation	41
2.6.2.1	Total Water Saturation for Sandstone Reservoirs	41
2.6.2.1.1	Well Logs Determination of Water Saturation	41
2.6.2.1.2	Nuclear Magnetic Resonance	43
2.6.2.1.3	Capillary/Saturation Measurements	44
2.6.2.1.4	Oil-Based Mud	44
2.6.2.2	Influencing Parameters on Irreducible Water Saturation	44
2.6.2.2.1	Surface Area to Pore Volume	45
2.6.2.2.2	Flow Zone Indicator	47
2.6.2.2.3	Volume of Shale	48

2.6.2.3	Residual Oil Saturation Determination	49
2.6.3	Permeability	53
2.7	Choice of Variables	53
2.7.1	Based on Net Present Value	53
2.7.2	Based on Reserve Estimation	54
2.7.3	Based on Pressure Transient Analysis	57
2.8	Summary	59

Chapter Three Methods, Concepts and Models

3.1	Introduction	60
3.2	Data Gathering and Well Distribution	61
3.3	Data Preparation	62
3.3.1	Kataharan sand/shale cutoff	62
3.4	Fluid Identification	63
3.5	Uncertainty Recognition and Reduction in Log, Core and Well Test Analysis	65
3.6	State-of-the-Art of Uncertainty Analysis Methods	66
3.6.1	Deterministic Approach	66
3.6.2	History Matching Technique	67
3.6.3	Probabilistic/Stochastic Technique	67
3.6.4	Geostatistical Approach	68
3.6.5	Machine Learning Algorithms	68
3.6.5.1	Reinforced Learning	68
3.6.5.2	Unsupervised Learning	69
3.6.5.3	Supervised Learning	69
3.6.5.3.1	Neural Networks	70
3.6.5.3.2	Deep Learning	72
3.6.5.3.3	Support Vector Machines	73
3.6.5.3.4	Ensemble Model	74

Chapter Four Uncertainty Reduction in Porosity, Fluid Saturation and Permeability Usage

4.1	Uncertainty Reduction in Porosity Usage.....	76
4.1.1	Introduction.....	76
4.1.2	Methodology.....	78
4.1.2.1	Uncertainty in Porosity Determination.....	80
4.1.2.2	Support Vector Machines Regression Modeling.....	81
4.1.3	Result and Discussion.....	86
4.1.3.1	North Sea Porosity Modeling.....	86
4.1.3.1.1	Unconsolidated Sandstones.....	89
4.1.3.1.2	Semi-Consolidated Sandstones.....	92
4.1.3.1.3	Consolidated Sandstones.....	95
4.1.3.2	Effect on Reserve and Production Forecast.....	97
4.1.4	Summary and Conclusion.....	99
4.2	Solving the Fluid Saturation Problem Using Stacked Ensembles and Empirical Methods	
4.2.1	Introduction.....	100
4.2.2	Methodology.....	103
4.2.2.1	Machine Learning Approach (Stacked Ensemble) Approach.....	103
4.2.2.2	Empirical Model Approach.....	103
4.2.3	Results and Discussion.....	105
4.2.3.1	Ensemble Total Water Saturation Model.....	105
4.2.3.2	Ensemble and Empirical Irreducible Water Saturation.....	111
4.2.3.3	Ensemble Residual Oil Saturation Model.....	118
4.3	Real-Time Permeability from Logs	
4.3.1	Introduction.....	123
4.3.2	Methodology.....	129
4.3.2.1	Deep Neural Networks.....	131
4.3.2.2	Cross Validation.....	133
4.3.3	Results and Discussion.....	134

4.3.3.1	Absolute Permeability Modelling of the North Sea	134
4.3.3.2	Oil- Water Relative Permeability Prediction	138
4.3.4	Summary and Conclusion	152

Chapter Five Well Test Uncertainty Analysis

5.1	Introduction.....	153
5.2	Methodology	156
5.2.1	Identification of the Interpretation Model	156
5.2.2	Verification of the Interpretation Model	157
5.2.3	Calculation of the Interpretation Model Parameters (Direct Problem).....	157
5.2.4	Pressure Build-Up Analysis	157
5.2.5	Drawdown Analysis	159
5.2.6	Superposition, Type Curves and Pressure Derivatives	160
5.3	Results and Discussion	162
5.3.1	Effective Flow Interval and Heterogeneity	162
5.3.2	Pressure Transient Analysis	166
5.4	Summary and Conclusion	174

Chapter Six Log, Core, Test Across Scale Uncertainty Reduction (LCTUn)

6.1	Introduction.....	175
6.2	Methodology	176
6.2.1	Scale Analysis	176
6.2.2	Reconciling Permeability	178
6.2.3	Depth Matching MultiScale Data	183
6.3	Results and Discussion	184
6.3.1	Scenarios of More Accurate Scale	184
6.3.2	Log, Core, Test Permeability Relationship Using Machine Learning	196
6.3.3	Log, Core, Test Permeability Relationship Using Chart Method	201
6.3.4	Machine Learning Model Validation	203
6.4	Summary and Conclusion	205

Chapter Seven Conclusion, Review and Recommendation for Further Work

7.1	Conclusion.....	206
7.2	Review of Work	208
7.3	Recommendation for Further Work	209
	References.....	211
	Appendices.....	222

List of Appendix

Appendix A- 1 Input parameters for well test analysis of well	222
Appendix A- 2 Input parameters for well test analysis of the infinite conductive reservoir.....	222
Appendix A- 3 History plot indicating pressure buildup analysis	223
Appendix A- 4 Log-Log diagnostic plot showing pressure and derivative	223
Appendix A- 5 Superposition plot of actual and simulated plot for well	224
Appendix A- 6 Numerical simulation of the tested well.....	224
Appendix A- 7 Log-Log diagnostic plot showing model sensitivity to porosity	225
Appendix A- 8 Log-Log diagnostic plot showing model sensitivity to pay thickness	225
Appendix A- 9 Log-Log diagnostic plot showing model sensitivity to permeability.....	226
Appendix A- 10 Log-Log diagnostic plot showing model sensitivity to compressibility.....	226
Appendix A- 11 History plot showing model sensitivity to porosity.....	227
Appendix A- 12 History plot showing model sensitivity to pay thickness	227
Appendix A- 13 History plot showing model sensitivity to permeability.....	228
Appendix A- 14 History plot showing model sensitivity to compressibility	228
Appendix B- 1 Codes for Support Vector Porosity Model	229
Appendix B- 2 Codes for Ensemble Total Water Saturation Model.....	230
Appendix B- 3 Codes for Relative permeability Deep Learning Model.....	234
Appendix B- 4 Log, Core and Test Support Vector Regression Model Alongside other tested Models	238
Appendix C- 1 Log, Core and Test permeability data.....	246
Appendix D- 1 Journal paper on real time relative permeability prediction using deep learning.....	250

List of Figures

Figure 1. 1 Discrepancy widened with increasing complexity or heterogeneity of the reservoir..	9
Figure 1. 2 Conceptual plan for the work.....	12
Figure 2. 1 Schematic Geology of Petroleum Resource	18
Figure 2. 2 Schematic of the North Sea Graben.....	20
Figure 2. 3 Example footprints of the different types of well logs from a North Sea well.....	21
Figure 2. 4 Schematic for Pressure buildup technique	32
Figure 2. 5 Schematic for Pressure drawdown technique.....	33
Figure 2. 6 Drill Stem Test procedure (Chaudhry, 2004).....	34
Figure 2. 7 Pressure Data of Interference Test (Heinemann and Zoltan, 2003)	35
Figure 2. 8 Schematic pulse test rate and pressure history showing definition of time and pulse response amplitude.....	36
Figure 2. 9 Rate schedule and pressure response in injectivity test (Chaudhry, 2004)	37
Figure 2. 10 Rate schedule and pressure response for falloff testing (Chaudhry, 2004)	38
Figure 2. 11 Schematic showing cross plot of core porosity versus density, neutron and sonic porosities explaining the uncertainty inherent in its determination.	40
Figure 2. 12 Different shaly sand water saturation models.....	43
Figure 2. 13 Relationship between irreducible water saturation and.	46
Figure 2. 14 Simplified Influence Diagram Showing Relationship between Reservoir Parameters and Net Present Value (Adapted from Bouchard & Fox, 1999)	54
Figure 2. 15 Distribution function showing the P100, P90, P50 and P10 using 1000 simulation runs for the Ness reservoir, Northern North Sea.....	55
Figure 2. 16 Showing porosity and water saturation as the most sensitive for reserve estimation with both P10 and P90 cases for Ness Reservoir, Don Field, North Sea.....	56
Figure 2. 17 Distribution function showing the P100, P90, P50 and P10 using 1000 simulation runs for the Sherwood reservoir, North Sea.	57
Figure 2. 18 Showing permeability as the most sensitive for pressure transient estimation with both P10 and P90 cases for Sherwood Sandstone Reservoir, North Sea.	58
Figure 3. 1 Distribution of wells used for the different aspects of the study.....	61
Figure 3. 2 Kataharan plot discriminating shale from sand rich facies.....	63
Figure 3. 3 Gamma ray, neutron-density and pressure data with depths in feet.	64
Figure 3. 4 Different groups of machine learning algorithms.....	69
Figure 3. 5 Neural network model elements.....	71
Figure 3.6 Schematic for the final neural network used for prediction.....	72

Figure 4. 1 Log and Core porosity measurements.....	77
Figure 4. 2 Gamma ray showing lithology with photomicrograph of core.	79
Figure 4. 3 Fluid type uncertainty in porosity determination from density logs.....	80
Figure 4. 4 Influence of lithology on Raymer-Hunt-Gardner porosity	81
Figure 4. 5 Feature engineering showing the relationship between the different data types considered and to the independent variable (core) at a glance.	83
Figure 4. 6 Porosity SVM model performance	84
Figure 4. 7 Errors associated with estimation of porosities from sonic using derived best fit for Raymer, Wyllie and Density.....	88
Figure 4. 8 Improvement accuracy of estimated based on new coefficients over the default Raymer, Wyllie and Schlumberger Density model coefficients.	88
Figure 4. 9 Porosity prediction using default Schlumberger, modified Schlumberger, and Support Vector Machine Regression models for the unconsolidated Galley sandstone turbidites.....	90
Figure 4. 10 Comparison between predicted and actual core porosity for the unconsolidated Galley Sandstone formation.	91
Figure 4. 11 Showing the different porosity models alongside gamma ray for semi consolidated Captain Sandstone Member., Carrack Formation, Blake Field, 13/24a-6.....	93
Figure 4. 12 Comparison between predicted and actual core porosity for the Semi-consolidated Captain Sandstone formation, Carrack Formation, Blake Field, 13/24a-6.....	94
Figure 4. 13 Plot of Gamma Ray, Modified Schlumberger, Default Schlumberger for consolidated Dogger Reservoir, Brent Formation, Northern North Sea (3/9a-1).....	96
Figure 4. 14 Comparison between predicted and actual core porosity for the Consolidated Captain Sandstone formation, Carrack Formation, Blake Field, 3/09a-1.....	97
Figure 4. 15 Showing the influence of irreducible water saturation on permeability.....	104
Figure 4. 16 Correlation between the different base algorithms	107
Figure 4. 17 Stacked Ensemble, Archie, Schlumberger, Simandoux, Indonesian and Core alongside Gamma Ray for the Brae field, Northern North Sea	108
Figure 4. 18 Cross plot of actual against Simandoux, Schlumberger, Indonesia, Hosin, Archie and Stacked Ensemble.....	109
Figure 4. 19 Stacked Ensemble, Archie, Schlumberger, Simandoux, Indonesia and Core alongside Gamma Ray for the Brae field, Central North Sea.	110
Figure 4. 20 Cross plot of actual against Simandoux, Schlumberger, Indonesia, Hosin, Archie and Stacked Ensemble.....	111

Figure 4. 21 Comparison between GBM, Neural Network, Random Forest, Stacked Ensemble and Empirical predicted with actual for the Brent validation set, Hutton Field, Northern North Sea.	112
Figure 4. 22 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM, Random Forest, and Empirical Models for the Brent Reservoir, Hutton Field, Northern North Sea.....	114
Figure 4. 23 Comparison between GBM, Neural Network, Random Forest, Stacked Ensemble and Empirical predicted with actual for the Test set, Hutton Field, Northern North Sea.....	115
Figure 4. 24 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM, Random Forest, and Empirical Models for the Brent Reservoir, Hutton Field, Northern North Sea.....	116
Figure 4. 25 Parameter sensitivity of the machine learning model and training scoring history.	117
Figure 4. 26 Sensitivity of different parameters used in the empirical model.....	118
Figure 4. 27 Comparison between GBM, Neural Network, Random Forest and Stacked Ensemble with actual for the validation set, Hutton Field, Northern North Sea.....	119
Figure 4. 28 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM and Random Forest models for the Brent Reservoir, Hutton Field, Northern North Sea.....	120
Figure 4. 29 Comparison between GBM, Neural Network, Random Forest and Stacked Ensemble with actual for the validation set, Hutton Field, Northern North Sea.....	121
Figure 4. 30 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM and Random Forest models for the Brent Reservoir, Hutton Field, Northern North Sea.....	122
Figure 4. 31 Schematic of oil-water relative permeability curve.....	125
Figure 4. 32 Water fractional flow curve with its derivative for the field considered.....	131
Figure 4. 33 Deep Neural Network model architecture showing input, hidden and output layers	132
Figure 4. 34 How data science techniques scale with amount of data (Andrew, 2013)	133
Figure 4. 35 Rate of permeability change in Northern North Sea sandstone reservoirs.....	134
Figure 4. 36 Rate of permeability distribution in Southern North Sea sandstone reservoirs	135
Figure 4. 37 Rate of permeability distribution in Central North Sea sandstone reservoir.....	135
Figure 4. 38 Distribution of permeabilities across the Northern, Central and Southern North Sea indicating their different depositional environments	137
Figure 4. 39 Tixier, Morris, Coates, Timur modelled over North Sea data. Broken line is obtained from the data and passes through the centre.....	137
Figure 4. 40 Actual vs predicted value for different models of validation set.....	139
Figure 4. 41 Actual vs predicted value for different models of the test set.	139
Figure 4. 42 Actual vs predicted value for deep neural networks.....	140
Figure 4. 43 Experimental and predicted relative permeability models.....	142

Figure 4. 44 Experimental (actual) and predicted relative permeability models for out of sample datasets.....	142
Figure 4. 45 Comparison of Wyllie, Corey, Parker, Stone, Baker, Honarpour, Deep Neural Networks for the Brent reservoir, North Sea.....	143
Figure 4. 46 Comparison of Wyllie, Corey, Baker, Honarpour, Deep Neural Networks models for the Stratjford reservoir, NorthSea.....	144
Figure 4. 47 Comparison of Deep Neural Networks and Baker with the measured wetting and non-wetting relative permeability models for the validation set (Brent reservoir).....	145
Figure 4. 48 Comparison of Deep Neural Networks and Baker with the wetting and non-wetting phase relative permeability models with for the test sets (Stratjford reservoir).....	145
Figure 4. 49 Water saturation against distance along pay	146
Figure 4. 50 Dimensionless distance against dimensionless time describing saturation history	147
Figure 4. 51 Wetting and non-wetting phase saturation history of case field.....	148
Figure 4. 52 Real time wetting and non-wetting phase relative permeability of a case field.	148
Figure 4. 53 Sensitivity analysis of individual variables used for building the wetting phase Deep Learning relative permeability model.	149
Figure 4. 54 Sensitivity analysis of individual variable used for building the non-wetting phase Deep Learning relative permeability model.....	150
Figure 5. 1 Signal theory schematic for the reservoir (Gringarten 1985a).....	154
Figure 5. 2 Interpretation model identification process (Gringarten 2012)	155
Figure 5. 3 Signal theory schematic for the chosen model (Gringarten 1985a).....	156
Figure 5. 4 Lorenzo plot and modified Lorenzo plot for well 1 sandstone formation.....	163
Figure 5. 5 Lorenzo plot and Modified Lorenzo plot for well 4 sandstone reservoir. The plot shows that about 60% of flow is from approximately 45% of storage.....	164
Figure 5. 6 Lorenzo plot and Modified Lorenzo plot for well 3 sandstone reservoir. Here 85% of flow is from 25% of storage.	164
Figure 5. 7 Lorenzo plot and Modified Lorenzo plot for well 2 sandstone reservoir. Here 90% of flow is from less than 20% of storage.....	165
Figure 5. 8 Lorenzo plot and Modified Lorenzo plot for well 6 sandstone reservoir. Here 70% of flow is from about 25% of storage.	165
Figure 5. 9 History plot indicating pressure buildup analysis.....	166
Figure 5. 10 Log-log diagnostic plot showing pressure change and derivative	167
Figure 5. 11 Superposition plot of actual and simulated pressure	168
Figure 5. 12 Numerical modelling of the tested well.....	169

Figure 5. 13 Log-log plot showing model sensitivity to porosity	170
Figure 5. 14 Log-log plot showing model sensitivity to permeability	171
Figure 5. 15 Log-log plot showing model sensitivity to pay thickness.....	171
Figure 5. 16 Porosity sensitivity on history plot	173
Figure 5. 17 Permeability sensitivity on history plot.....	173
Figure 5. 18 Reservoir thickness sensitivity on history plot.....	174
Figure 6. 1 Relative scales of permeability sources.....	178
Figure 6. 2 Parameter tuning of the SVM model.....	182
Figure 6. 3 Core, log, test permeabilities to model observed production rate for well 6..	185
Figure 6. 4 Core, log, test permeabilities to model observed production rate for well 5	185
Figure 6. 5 Core, log, test permeabilities to model observed production rate for well 4	186
Figure 6. 6 Core, log, test permeabilities to model observed production rate in well 3	186
Figure 6. 7 Core, log, test permeabilities to model observed production rate for well 7	187
Figure 6. 8 Core, log and test permeability for estimating bottom hole pressure compared with observed bottom hole pressure for well 6.....	188
Figure 6. 9 Core, Log and test permeabilities for estimating bottom hole pressure compared with observed bottom hole pressure for well 4.....	188
Figure 6. 10 Core, Log and test permeabilities for estimating bottomhole pressure compared with observed bottom hole pressure for well 3.....	189
Figure 6. 11 Log Permeability distribution in the i-direction	190
Figure 6. 12 Log permeability distribution in the j-direction	190
Figure 6. 13 Log permeability distribution in the k-direction.....	191
Figure 6. 14 Core permeability distribution in the i-direction.....	191
Figure 6. 15 Core permeability distribution in the j-direction.....	192
Figure 6. 16 Core permeability distribution in the k-direction.....	192
Figure 6. 17 Well test permeability distribution in the i-direction.....	193
Figure 6. 18 Well test permeability distribution in the j-direction.....	193
Figure 6. 19 Well test permeability distribution in the k-direction	194
Figure 6. 20 Correlation of the pay zone across the different wells.....	195
Figure 6. 21 Flow diagram showing Uncertainty Reduction with SVM Technique	197
Figure 6. 22 Actual well test permeability versus model on a randomly selected test set.	198
Figure 6. 23 Actual test permeability versus model on a second randomly selected test set.	198
Figure 6. 24 Screen shot of well test permeability prediction dashboard result.....	200
Figure 6. 25 Nomographic technique showing the relationship between core, log and well test permeability.....	202

Figure 6. 26 Schematic showing the Upper and Lower sand penetrated by some of the wells used for validation	203
Figure 6. 27 Actual well test permeability versus model predicted well test permeability from core and logs permeability.	204

List of Tables

Table 1. 1 Showing key challenges in the E&P industry (modified after Aminzadeh, 2005).....	3
Table 1.2 Indicating the variables predicted with data summary.....	5
Table 1. 3 Research questions, key concepts and contents.....	11
Table 2. 1 Showing the classification of sandstones in terms of grain size (Wentworth, 1922).....	17
Table 2. 2 Prospective Recoverable Oil and Gas in the North Sea (OGA, 2016)	19
Table 2. 3 Log and example equations of how their properties were obtained as well as their limitations.....	24
Table 2. 4 Pros and Cons of the different residual oil saturation measurements.....	50
Table 4. 1 Typical ranges in porosity of common rock types.....	78
Table 4. 2 Different combinations of training data sets.....	86
Table 4. 3 Distribution of best fit (BF) coefficients and their associated errors for Raymer, Wyllie and Modified Schlumberger models.....	84
Table 4. 4 Porosity uncertainty in reserve estimation assuming other parameters have no uncertainty	98
Table 4. 5 Error in reserve estimation arising from porosity uncertainty assuming other parameters have no uncertainty	98
Table 4. 6 Some water saturation models for shaly sands	101
Table 4. 7 Assumptions and application of the commonly used two and three phase relative permeability correlations.....	126
Table 4. 8 Comparison of relative permeability models (vertical) with different datasets (horizontal) using correlation coefficient (Modified after Baker 1988)	130
Table 4. 9 Accuracy of the deep learning model for the wetting phase	140
Table 4. 10 Accuracy of the deep learning model for the non-wetting phase	141
Table 4. 11 Sensitivity analysis showing the importance of the different to both water and oil relative permeabilities.	151
Table 6. 1 Factors to consider when reconciling permeability data.....	180
Table 6. 2 Accuracy in terms of RMSE of the dofferent features.....	183
Table 6. 2 Averages for upscaling core and log data for comparism with well test data.....	184

Nomenclature

Φ_t = Total Porosity

Φ_e = Effective Porosity

S_w = Water saturation

S_o = Oil saturation

S_{wi} = Irreducible water saturation

S_{or} = Residual oil saturation

K_{rw} = Water relative permeability

K_{ro} = Oil relative permeability

K_{eff} = Effective permeability

K_a = Absolute permeability

P_i = Initial pressure

C = Wellbore storage

I_{GR} = Gamma ray index

GR = Gamma ray

V_{sh} = Volume of shale

k = Permeability

R_t = True resistivity

F = Formation factor

Δt_{log} = Interval transit time from sonic log

V_p = Primary velocity

ρ = Density

T_2 = Cutoff time

p_{wf} = Flowing bottomhole pressure

B = Formation volume factor

c_t = Total compressibility

μ = Viscosity

S = Skin

r_w = Wellbore radius

h = Pay thickness

q = Flowrate

τ = Tortuosity

S_{gv} = Specific grain volume

Φ_z = Normalized porosity

A = Area

f_{zi} = Flow zone indicator

Δp = Change in pressure

Δt = Change in time

p_{ws} = Well shutin pressure

t_D = Dimensionless time

p_D = Dimensionless pressure

Chapter One

1.1 Introduction

Despite the huge returns on investment in the oil and gas industry, several researchers have asserted that the level of performance is usually below expectation in terms of production efficiency (McVay and Dossary, 2014). Aminzadeh (2005) also pointed out challenging areas in the oil and gas industry influencing the required expenditure on facilities and field development (Table 1.1). Exploration and Production companies delivered only about half of its predicted reserves in the last twenty years of the twentieth century (Rose, 2004). In fact, Merrow (2011) stated that success rate of petroleum projects worth more than 1 billion USD i.e. megaprojects, declined from 50 to 22%. Although Leach (2006) asserted that the crucial source of value in business is uncertainty, the underperformance of projects was ascribed to uncertainty and geotechnical parameters estimate (Rose, 2004; Begg et al., 2003) together with calculation of hydrocarbon reserves, which represents the main asset of exploration and production companies.

Huge amount of world oil and gas have been bypassed on both microscopic and macroscopic scale in already discovered and developed reservoirs (Bassiouni and Velic, 1996). These bypassed hydrocarbon on a microscopic scale are immobile oil and gas trapped in the pores of reservoirs by capillary and viscous forces and cannot be displaced by water hence recovery is enhanced using chemical flooding techniques and miscible gas. On a macroscopic scale, mobile oil remains in the reservoir due to inaccurate delineation of reservoir boundaries, reservoir heterogeneity leaving considerable volumes of oil bypassed in isolated pockets, water flooding with oil left behind the water front in these water-swept zones, attic oil trapped updip from the highest well in a reservoir, water coning causing the oil cut to prematurely get to its economic limit leaving a significant amount of mobile oil unrecovered, misinterpretation of well logs due to complex rock-fluid system, relatively more focus on structural traps than stratigraphic and combination traps especially in sandstone reservoirs. The magnitude of the bypassed oil is substantial in that for example in offshore Gulf of Mexico nearly two-thirds of the known resources will remain unrecovered after using conventional techniques with about 5.53 billion barrels i.e. 878 million cubic metre (19% of the original oil in place) in the class of bypassed oil. Bear in mind that this huge amount is just for the Gulf of Mexico and also known reserve cases let alone considering a global scale and unknown cases even in matured fields. In all of these reasons for the bypass and abandonment summarised into reservoir development and management, permeability plays a key role in determining whether a well should be completed and brought online, choosing optimal

drainage point and production rate, designing EOR configurations and injection settings as well as optimizing perforation and completion design.

Although cores are direct measurements in some way, they are not sufficient to show reservoir heterogeneity since it is only possible to drill limited number of wells due to time and cost. Again they are taken on samples that have been moved to surface and also cleaned thus are different from insitu conditions such as clay alteration, relief cracks, saturation although others like reservoir pressure, effective stress (overburden minus pore pressure) and temperature might be created in the laboratory. Well logs just like core give a layer by layer estimate of permeability at a scale greater than cores. As pointed out by Haldorsen (1986), the volume of rock investigated by a sonic log is 162 times the volume of a one-inch diameter core plug and their porosity relationship is shown in Equation 4.2. Permeability is obtained indirectly using models like Timur, Tixier, Coates and Dumanoir which have all also been modelled on North Sea data with Timur as shown in Chapter Four.

Even with their objectives, it is incredibly important to admit the limit to the level of detail that well testing provides whether in terms of reservoir evaluation or description since pressure transmission is diffusive hence it is controlled mainly by average conditions rather than local heterogeneities. This underscores the point that well test is insensitive to most local scale heterogeneities but gives average or bulk reservoir properties. Coupled with the fact that tests can be compromised by say a leaking packer or a leak at the wellhead due to wellbore or interference effects from other injectors or producers although surface recorders could be used to nip this in the bud. Also mini-frac conducted in very low permeability reservoirs make it difficult to guess when reservoir dominated radial flow has been attained. Tests need to be designed to be long enough so as to get reliable and useful data which also means more money and productive time. Kuchuk (2015) stated that the near well bore region dominates the drawdown and buildup pressure transient behaviour and it's true for all types of reservoirs. There have been attempts to determine average reservoir pressure without shutting in wells as is the case with buildup tests which could cause companies money. In 1991, it was estimated that just BP Exploration in Alaska alone lost \$1.3 million in lost-production due to shutting in wells for test purposes (Ahmed et al., 1994). Thus the conventional methods for estimating average reservoir pressure do not only require a prior knowledge of shape factor but also result in lost production.

Table 1. 1 Showing key challenges in the E&P industry with challenges 2, 7, 11, & 12 part of the motivating factors for this work (modified after Aminzadeh, 2005).

S/N	Important Exploration & Production Challenges
1	Accurate positioning of complex structures (salt and overthrust)
2	Characterization of thinly laminated sand/shale sequences
3	Deep exploration and accurate depth imaging
4	Distinguishing commercial gas from non-commercial gas
5	Fault detection and their types (e.g sealing vs non-sealing)
6	Environmental issues (gas clouds, mud volcano, basalt)
7	Fluid/permeability prediction, quantifying uncertainty and changes
8	Exploration in difficult areas
9	Frequency, orientation, connectivity and type of Fractures
10	Evaluation and prediction of over-pressured reservoir
11	Carbon sequestration and storage
12	Accurate calculation of hydrocarbon reserves

There are several reservoir rock properties such as total porosity, effective porosity, absolute permeability, effective permeability, relative permeability, reservoir pressure, grain size, sorting, grain shape, fluid saturation, irreducible water saturation, critical gas saturation, critical oil saturation, movable oil saturation, capillary pressure, surface and interfacial tension, wettability, temperature, net pay thickness, reservoir heterogeneity, anisotropy, compressive strength, poisson ratio, shear modulus, young modulus, depletion profile, FZI while the fluid properties include viscosity, compressibility, solubility, molecular weight, formation volume factor, API specific gravity, liquid specific gravity, bubble point pressure, solution gas-oil ratio, solution oil-gas ratio, retrograde dew point pressure, saturation pressure, critical point, interfacial tension, diffusion coefficient. Most of these properties have several models which give different results as well (Table 1.1). Assumptions made in the necessary simplification of these models are quite different from reality. Box (1979) stated that all models are wrong but some are useful which is no less true considering the fact that the earth's subsurface is too complex than we can ever accurately model and thus requires an interdisciplinary approach for better results (Catuneanu, 2006).

Reservoir characterization which involves describing the various reservoir properties to provide a reservoir model for accurate prediction is an essential aspect of reservoir engineering (Jong-Se, 2005) although developing a complete characterization of multiphase flow parameters is a very challenging task. Reservoir properties like porosity, reservoir pressure, fluid saturation, permeability amongst others are indirectly measured and even if directly measured are only available at relative small number of well locations compared to the size of the field. These limited data are then combined with geological and geophysical interpretations to generate a geostatistical models. The degree of uncertainty is further increased since the different data sources like well logs, cores and well testing all give different results.

Based on the aforementioned, there is need to close the gap between especially well logs, core and well test (Figure 1.1) by updating through machine learning to:

- a) provide the bench mark to calibrate one against the other especially when any is absent and also create a platform for real time reservoir properties prediction.
- b) provide continuous quality improvement in such large and complex process environment
- c) allow time cycle reduction and efficient utilization of resources.

Table 1.2 summarises the different variables predicted alongside the algorithm and input data used. North Sea siliciclastic reservoirs were first used to build the different models for uncertainty reduction while a part of the dataset not used for the development of the models was used for validation (due to unavailability of Gulf of Guinea Deepwater sandstone reservoir data which was supposed to be the focus of this research). The technology should be applicable to global provinces with minimum tuning and effective for reservoir and production management.

Table 1.2: Indicating which variables have been predicted, algorithms and input variables used for the prediction

S/N	Variable Predicted	Data Summary Statistics		Algorithm Used	Input Data
		Amount of Data	Mean		
1	Porosity	446	16.4	Support Vector Regression	ϕ_n, ϕ_s
2	Water Saturation	471	32.03	Stacked Ensemble	$gr, ild, \phi_n, rhob$
3	Irreducible Water Saturation	290	42.21	Stacked Ensemble	$gr, ild, \phi_n, rhob$
4	Residual Oil Saturation	243	22.37	Stacked Ensemble	$gr, ild, \phi_n, rhob$
5	Relative Permeability (water and oil)	132	27.13; 17.35	Deep Learning	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w, (S_w - S_{wc}), (S_o - S_{or}), (1 - S_{wc}), (\mu_o/\mu_w)$
6	Well Test Permeability	114	48.34	Support Vector Regression	$klog, kcore, (klog)^{1/2}, (kcore)^{1/2}, (klog)^2, (kcore)^2$

1.2 Evaluation Methods

The datasets used in this work were described in Table 1.2 comprising those used for porosity, water saturation, irreducible water saturation, residual oil saturation, relative permeability and well test permeability. As supervised learning methods (Section 3.6.5.3) were used for the work, the datasets were divided into training, validation and test sets. The dataset split ratio depends on the total number of samples in the data and the actual model been trained. Apart from the relative permeability dataset, all others were randomly split. Training datasets are those used to fit the model, validation dataset provide the unbiased evaluation of the model fit of the training dataset while tuning model hyper parameters while the test dataset provides an unbiased evaluation of the final model fit on the training dataset.

1.3 Literature Review

Modeling of petrophysical properties from cores, well logs and well test has an important role in oil and gas exploration and production. These datasets pose two main challenges (a) diverse and nonlinear nature of reservoir variables and (b) from a theoretical perspective, absence of a direct relationship between the datasets. Permeability, fluid saturation, porosity which are some of the important reservoir variables, have a non-uniform distribution spatially. Nonlinear machine learning models like artificial neural networks, support vector regression, ensemble models, fuzzy logic have gotten recognition as the potential tool for solving nonlinear and complex problems (Nikraves and Aminzadeh, 2001; Hamada and Elshafei, 2010; Borsh et al., 2010; Majdi et al., 2010). Nikraves and Aminzadeh (2001), used an ensemble of neural and fuzzy logic to optimally develop a set of rules for nonlinear mapping between grain size, porosity, permeability. Good prediction performance and clear variable relationship was discovered.

Helle and Bhatt (2002) developed a committee machine (CM) networks for the prediction of porosity and permeability is input data from logs like density, sonic, neutron, gamma ray from NorthSea wells. Two separate networks using different inputs were built and then combines to form a committed machine. They opined that the accuracy of their model is restricted only by the accuracy of the real data used for its development claiming that their approach is more accurate than multiple linear regression. Al-Anazi and Gates, (2012) evaluated the capability of support vector regression (SVR) to make predictions of porosity and permeability with small sample size heterogeneous sandstone reservoir data. Compared results with multilayer perceptron neural network showed that the SVR model consistently yielded better predictions of the porosity and

permeability even with the small sample size than the MLP method using less computational time as well.

Ahmadi et al., (2018) predicted reservoir permeability using Neural Networks, Genetic Algorithms and Particle Swarm Optimization, fuzzy decision tree as well as their combination. Their results showed that the deviation the different model was found to be less than 1% for the hybridized models thereby indicating the hybridized techniques for permeability and porosity can result in more reliable models. Huang et al., (1998a&b) also used genetic algorithms (GA) to optimize the connection weights of a neural network for the prediction of permeability from well logs. They found that the GA-trained networks gave consistently smaller errors compared to the conventional backpropagated networks although comparatively, the convergence of the former were slower.

Guler et al. (1999) developed several neural network models for relative permeability considering different parameters influencing the parameter and selected the best model to make predictions for the test set. Important to state that ordinary neural neural networks require far more neurons to perform like deep learning models. Arigbe et al., (2018), used deep learning to predict 2- and 3-phase relative permeability. The common factors influencing relative permeability as well as parameter combinations from Wyllie and Baker, were used as input for the model. The deep neural network performed better than the existing empirical models especially for test data due to their ability to generalize and regularize.

Zhang et al., (2019) used Long Short-Term Memory (LSTM) to predict reservoir water saturation using data derived from monitoring and simulating an actual reservoir. They got prediction with overall AARD of less than 14.82% outperforming other models such as Gated Recurrent Unit (GRU) and Recurrent Neural Network (RNN) presenting another way for fast and robust water saturation prediction. Al bulushi et al., 2009 used neural network models to make prediction of water saturation from well logs. They developed two case studies and presented that they performed better than multiple linear regression. The neural network model proved its robustness from water saturation prediction.

Training multiple linear regression and multilayer perceptron neural networks with core and well logs often suffer from poor generalization to test data especially when the training sample is small probably due to the empirical risk minimization (ERM) principle which converges asymptotically when sample size increases. On the contrary, the structural risk minimization (SRM) principle which is used by support vector regression and other models allows the training samples to cover the complexity of the parameter space so that the model is not only able to match the training sample but also able to generalize to new data. For this work SVR models were

developed for porosity and well test permeability, deep learning model was used for relative permeability and stacked ensemble models were used for fluid saturation.

1.4 Project Driver

Reservoir properties from different sources such as core, log and well test as well as models (for porosity, saturation and permeability) give different results hence the work is driven by uncertainty reduction in order to close the gap between the sources, determine the most accurate and establish a relationship between them. If well test is the most accurate for permeability prediction for example, it should be deduced from log and core data if not available. Recent technological advances and the increasing drive on maximizing the recovery from especially marginal fields consequent on the energy demand currently experienced in the world, has led to the upsurge of interest and quantitative modelling of rock properties (Bryant and Flint., 1993).

Characterization of heterogenous geological systems need not just a multi-disciplinary approach involving data from all relevant measurements but also understanding the scale and limits of these measurements. These geological, petrophysical and engineering characterization combines both their static and dynamic behaviour to ensure great improvements in their description and predictions. Fethi et al., 2010 asserted that permeability is probably the most challenging issue for both geologists, petrophysicists and reservoir engineers especially in reservoirs where facies, petrophysical properties and petroleum systems are completely different which in turn leads to more disparity between the different sources for its determination. Figure 1.1 shows the increasing disparity between core permeability and well test permeability as the reservoir becomes more and more complex from layered to channel-levee complexes. This discrepancy is as a result of the collective effect of:

- a) scale of measurement of the static porosity component of core and log permeability model which is smaller than the radius of investigation of well test
- b) tensorial nature of permeability not being reflected in the log permeability owing to the fact that its models are from parameters that are non-directional,
- c) averaging technique used for log permeability which gives different values for the different types of averaging as $\bar{k}_{harm} \leq \bar{k}_{geom} \leq \bar{k}_{arith}$ as well as
- d) the heterogeneity of the flow system whether pores and/or fractures.

As all these factors are magnified with increasing volume of investigation, the discrepancy in permeability of the different sources is also increased. The uncertainty lies in both the measured

data and derived variables which are then propagated to the static models. Getting it right (related to accurate knowledge of porosity and saturation for log models) will enhance reservoir description and thus reduce the vast amount of oil and gas bypassed and left behind pipe in many reservoirs around the world. A true estimate of permeability and its distribution in a large number of wells is one of the main objectives of reservoir characterization (Cozzi et al., 2006).

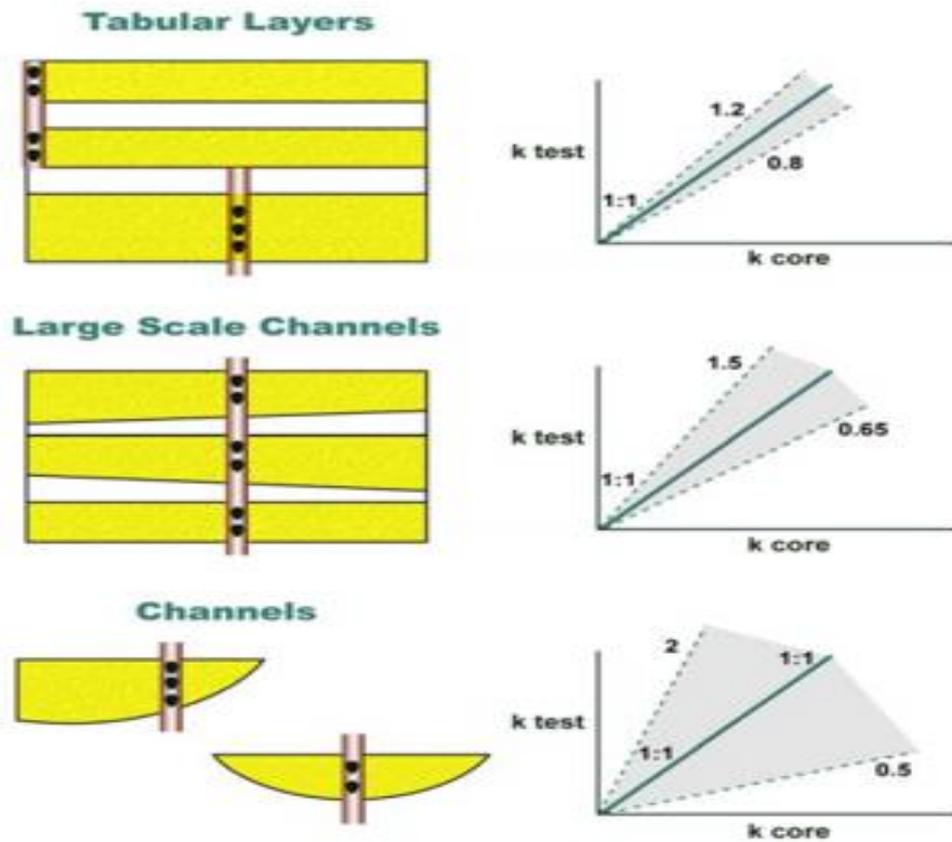


Figure 1. 1 Discrepancy between the different sources widened with increasing complexity or heterogeneity of the reservoir.

1.5 Aim of Project

To reduce the uncertainty associated with reservoir rock properties

1.6 Objectives

The objectives of this work are:

1. Demonstrate why the parameters focussed on in this work were chosen
2. Demonstrate the level of uncertainty inherent in the determination of these reservoir properties.
3. Develop porosity, permeability and water saturation models from well logs to save time and cost associated with the acquisition of cores and well test.
4. Reduce uncertainty in well test analysis.
5. Ascertain which permeability source is more accurate and develop a methodology to reduce the uncertainty in the integration of well logs, cores and well test establishing a relationship between them K_{WT} , K_{CORE} and K_{LOG} to determine the most accurate source when not available due to time and/or cost from the other two.
6. Validation with other North Sea reservoirs not used for the development to test the repeatability and reproducibility of the models.

1.7 Research Questions

Table 1. 3 Research questions, key concepts and contents

RESEARCH QUESTIONS	OBJECTIVES	CONCEPT	RESEARCH CONTENT
How do we use porosity values from neutron, density and sonic to get best results	Sources of porosity and the implications for using porosity data from different sources	Guidelines for porosity usage from different to obtain best results	<p>GAP</p> <p>No systematic guideline for usage of porosity considering its sensitivity in several models</p> <p>RESEARCH</p> <p>Map porosity sources with their usage</p>
Are there lapses in the models currently used for porosity, fluid saturation and permeability	Demonstrating the extent of the uncertainty between the different models for porosity, permeability, relative permeability, water and oil saturation, irreducible water and residual oil saturation	More generic models that are theoretically correct	<p>GAP</p> <p>Uncertainty in choosing the right model from the huge numbers available each with several assumptions and hence not generic.</p> <p>RESEARCH</p> <p>The need for generic models that honour the physical model for these properties.</p>
Which of the permeability data from core, log and well test is more accurate	To achieve better modelling from more accurate data Reduce uncertainty in well test analysis	Demonstrate the impact on model when discrepancy in results from the various sources is large	<p>GAP</p> <p>Technique for determining which data source is more accurate</p> <p>RESEARCH</p> <p>Systematic approach to determine what data sources to use for a particular case study</p>
How to close the gap between well log, core and well test and to establish a relationship between them	Reduce the uncertainty in the integration of core, log and well test	Establish a relationship k_{core} , k_{log} and k_{test} to enable the prediction of any when not available due to time constraint and/or expense	<p>GAP</p> <p>Expense and time for cores and buildup test. Real time parameter prediction</p> <p>RESEARCH</p> <p>Model the relationship between the different sources for permeability.</p>

1.8 Conceptual Plan

The plan involves getting models of the listed parameters from the most readily available source(s) using North Sea data. The uncertainty associated with these parameters is thereby reduced and confidence in their interpretation increased. These analysis results in these parameters being available across board (where such data is present). For this work permeability from the different sources was then modelled using machine learning and nomographic approach thus making scale wise precise prediction possible from relatively smaller scale. Same could have been done for water saturation if Reservoir Saturation Tool (RST) well test data were available.

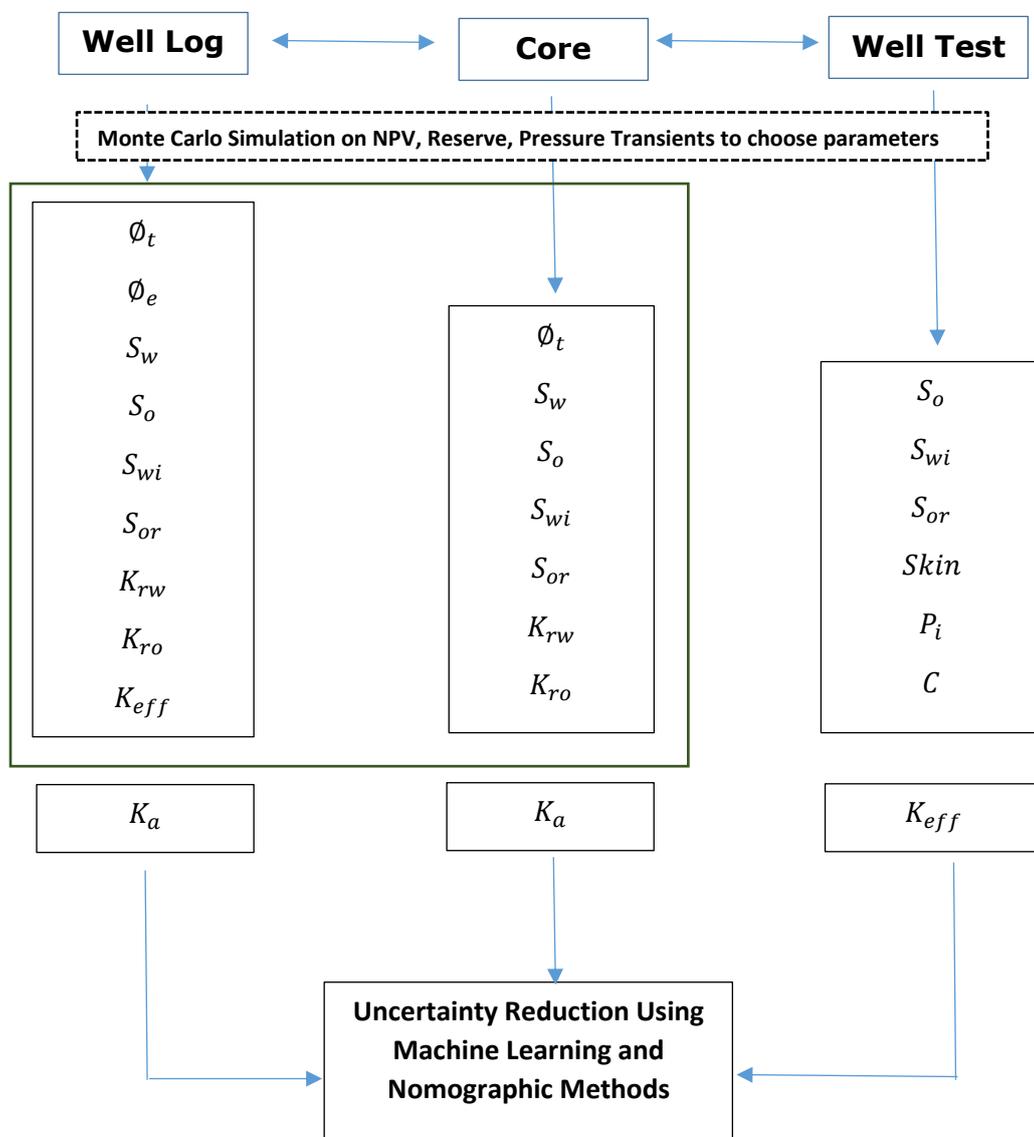


Figure 1. 2 Conceptual plan for the work

1.9 Contributions to Knowledge

This sections entails the contributions this work has made to existing knowledge as listed below:

- The Developed Support Vector Machine Regression porosity model consistently gave errors of about 1% in all sandstones which could be less if more tuning was done. The commonly used default Schlumberger, Raymer and Wyllie Time Average could lead to errors of up to 17% in unconsolidated sandstone formation, about 5% in semi-consolidated sandstones and about 3% in consolidated sandstone reservoirs for reserve estimation separate from those caused by other parameters. This can be found in Chapter Four.
- Data driven fluid saturation models to reduce uncertainty associated with its prediction both in terms of which one to choose from the myriads of empirical models available and the cost associated with their determination from cores. Generally the machine learning models performed better than the empirical models. Chapter Four contains details of this work.
- Deep Learning for real-time relative permeability prediction from common well log parameters especially in an industry where big data is now available providing the platform to systematically forecast reservoir fluid and rock properties in order to drastically optimize the cost and time needed for laboratory experiments. Details can also be found in Chapter Four.
- A 50% decrease in the pay thickness appears to cause more destabilization on constant flowrate pressure drawdown, buildup and its derivative than a 100% increase in pay thickness. Overall a high decrease in porosity, permeability and pay thickness was more detrimental than a high increase. With this been said, it is very important to accurately define Effective Flow Interval. This was discussed in Chapter Five but its calibration with other data sets such as core and logs is discussed in Chapter Six.
- It has been shown from using the techniques presented in this work that making good use of cores and well logs (macro and meso-scale respectively) can evaluate well and reservoir performance (well test) with accuracy which is particularly important in deep water offshore environment where the cost of DST's in every well in an entire field as well as in the different pay zones which the well penetrates, will not be justified. A relationship between the different dataset was learnt using appropriate machine learning algorithms. Chapter Six contains this.
- Developed a relationship between core, log and well test permeability using nomographic technique as well to provide a quick look for permeability determination. This can be found in Chapter Six.

1.10 Thesis Layout/Arrangement

A general overview of the thesis layout with the underlying theories and concepts with the aim of giving a first hand grasp of the concepts shown in each chapter.

Chapter One is general introduction of the issues the research is to address, what drives this work, the project's aim and objectives, research questions and the contributions made to existing body of knowledge. Chapter Two details the different properties together with the different data sources such as log, core and well test. It also describes why porosity, fluid saturation and permeability (and their variants) were selected amongst the different rock and fluid properties using Monte Carlo simulations to generate frequency and tornado plots with both reserve and pressure transient data. Chapter Three entails the methodology of the research highlighting data digitization and wrangling, describing the different machine learning algorithms used for the research in detail.

Chapter Four introduces porosity concept in siliciclastic reservoirs discussing the uncertainty, guidelines for its usage in unconsolidated, semi-consolidated and consolidated sandstones also highlighting the implication of using default porosity models on reserve estimation. Total and irreducible water saturations as well as residual oil saturation were modelled from conventional well logs. The Chapter also shows absolute permeability modelling for North Sea reservoirs noting how the existing empirical models fair in the North Sea knowing which one to take forward in its relation with core and well test permeability later in Chapter Six. It continues with a deep learning modelling of relative permeability for real time application comparing it with other networks and common software empirical models.

Chapter Five offers well test analysis to determine reservoir properties especially permeability as well as consider well test uncertainty reduction. Chapter Six analyses the derived log, core and well test permeability to ascertain the disparity between them as well as develop a relationship between them using both machine learning and nomographic models including validation with an independent North Sea data not used for model development. Chapter Seven presents the conclusions of the research study, gives a review of the entire work and recommendations for further work.

Chapter Two

Review of Sandstone Reservoir Variables

2.1 Introduction

For oil and gas to form, there has to be some play elements such as a source rock, migratory path, reservoir rock, trap to prevent it from getting to the earth surface and ofcourse the formation of the hydrocarbon and the above elements take time (millions of years). The source rock is mainly shales (clays with fissility) e.g. the Agbada Formation in the Niger Delta and the Kimmeridge clays of the North Sea. Migration from source to reservoir due to factors like buoyancy, compression, maturation, thermal expansion, topography, gravitational separation of the fluids. Again traps which are not just cul de sac i.e. passive containers or receivers but an active force draft system that can discharge water at depth and can also be structural caused by tectonics (anticlines, salt domes, faults), stratigraphic due to sediment deposition (pinch outs, unconformity, reef, bedding plane), hydrodynamic due to differences in pressure initiated by water flow and lastly combination traps due to a combination of the above types with all these traps needing to be insitu at the right time to prevent oil and gas migration to the surface. The reservoir element of the play fairway can be sandstones or carbonates. Over 60% of the world's oil and gas reserves are found in siliclastic rocks (sandstones) having interconnected pore spaces which are more common than carbonate reservoirs (Bjørlykke and Jahren, 2010). Sandstones are the focus of this study although the models can also be applied to carbonates especially if the database is improved.

In this Chapter, a brief description of North Sea petroleum system is presented after which a review of well logs, core and well test data sources was undertaken together with the tools and methods used for evaluating reservoir properties. This is necessary since the study focusses on sandstone reservoirs hence carbonates were not exclusively dealt with here. In addition previous works in the literature using Net Present Value to justify the chosen parameters focussed on in this work, a Monte Carlo simulation was also run on static and dynamic data to systematically justify them. The stochastic approach randomly selects combinations of these parameter values (and their assigned distribution) using a computer algorithm and a selected number of simulation runs to build a reserve distribution curve. A pressure transient parameter sensitivity analysis was also carried out using the same technique. These helped to know the important properties on both the static and dynamic aspects of reservoir characterization. The working principles of the

different logs including gamma ray, resistivity, density, neutron, sonic alongside the different properties that they measure was discussed.

2.2 Sandstone Reservoirs

Around the world, reservoir rocks are dominated by sedimentary rocks since they contain spaces available when the petroleum was formed. These sedimentary reservoir rock can be siliclastic (allochthonous) meaning they were weathered from other rocks, transported and deposited. This depositional environment is a geographical location where physical, chemical and biological processes are taking place to permit the accumulation of sediments e.g. lacustrine (lake), aeolian (wind), fluvial (river), deltaic distributory channel, fluvial point bar, marine shelf, marine deep water reservoirs (Ganapathy et al. 2000). Each of these sandstones has its distinct internal geometry and features controlling the nature of petroleum production. They contain particles sizes of between 1/16 to 2 mm in diameter with the lower limit tending towards silt and clay while the upper limit towards coarser grains like pebbles (conglomeratic sandstones) as seen in Table 2.1. These rocks are referred to as sandstone reservoirs if they contain hydrocarbon (Figure 2.1). The clays acts as lubricants for the mechanical compaction of the sands and also reduces porosity. Quartz cementation increases the rock strength at burial depth of 2-3km (i.e. 80-100°C) although due to grain fracturing, coarse grained sands could additionally show substantial compaction. Reservoirs can also be carbonates (autochthonous) with carbonate fractions of more than 50% e.g. reef, limestone, dolomite. This research focusses on sandstone assets of the Northern, Central and Southern North Sea.

Table 2. 1 Showing the classification of sandstones in terms of grain size (Wentworth, 1922).

$\phi = -\log_2 D$	Size range (metres)	Wentworth size scale	Lithology
<-8	>256	Boulder	Gravel
-6 to -8	64-256	Cobble	
-2 to -6	4-64	Pebble	
-1 to -2	2-4	Granule	
0 to -1	1-2	Very coarse sand	Sand
1 to 0	0.5-1	Coarse sand	
2 to 1	0.25-0.5	Medium sand	
3 to 2	0.125-0.25	Fine sand	
4 to 3	0.0625-0.125	Very fine sand	
8 to 4	0.0039-0.0625	Silt	Mud
<8	<0.0039	Clay	

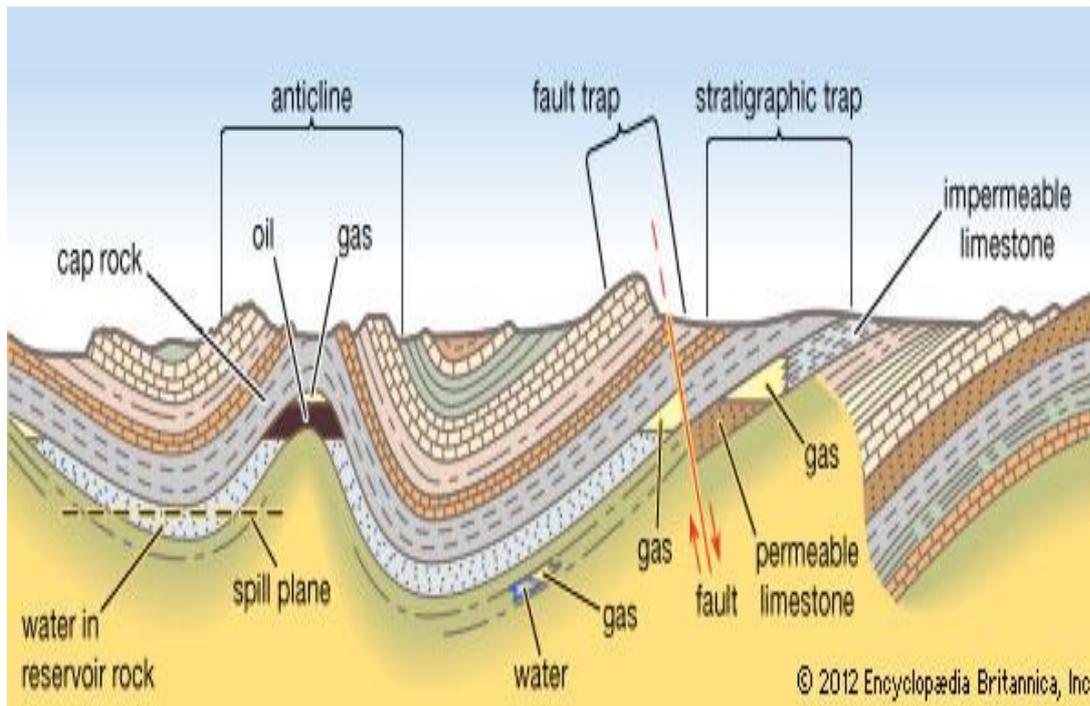


Figure 2. 1 Schematic Geology of Petroleum Resource (Encyclopædia Britannica, 2012)

2.2.1 North Sea Basin

The North Sea which covers an area of about 750,000 squared kilometres, is located at the margin of the Atlantic Ocean and is divided into three geographic areas (Figure 2.2) namely: The Viking Graben in the Northern part (includes both United Kingdom and Norway); The Moray Firth/Witch Ground in the west-central part (United Kingdom only); The Central Graben in the Southern parts (Denmark, Germany, Netherlands, Norway and the United Kingdom). The Caledonian Orogeny encompassing the Ordovician to Early Devonian (490-390 ma) assembled the configuration of the igneous and metamorphic basement rocks which lie beneath the North Sea sedimentary basins. This was done through the closure of the Lapetus Ocean and the Tornquist Sea (Andrews et al. 1990; Johnson et al. 1993; Gatliff et al. 1994; Glennie and Underhill 1998). The Southern North Sea was also formed during this time as a result of a long history of basinal subsidence.

The main source rock is the Kimmeridge Clay Formation which became mature with the help of Cretaceous and Cenozoic post-rift thermal subsidence and burial (Johnson and Fisher 1998). Sandstone reservoirs caused by mass flows contain about 20% of the proven reserves of this oil province (Pengrum and Spencer 1990). Due to an evolution from the emplacement of lateral sheet sands on the basin floor, apparently all UK sector Palaeogene sand systems became gradually distal to the east and SE.

The Central North Sea fields mainly produce from syn-rift Upper Jurassic and/or post-rift Lower Cretaceous or Palaeogene reservoir sandstones. The area has fault traps (Piper and Brae fields), stratigraphic traps (Britannia field) which are either in shallow or deep depths. In the Southern North Sea, Aeolian dune sandstone of the pre-Zechstein Permian Rotliegend Group, produces about 85% of all gas while 13% is from Triassic fluvial sandstone. This huge percentage of gas on the Permian sandstone shows efficiency of the seal provided by Zechstein Group (Permian evaporate-rich succession).

Table 2. 2 Prospective Recoverable Oil and Gas in the North Sea (OGA, 2016)

North Sea Graben	Recoverable Oil (BBO)			Recoverable Gas (TCF)		
	Lower	Mid	Upper	Lower	Mid	Upper
Northern North Sea	0.53	0.84	1.27	0.67	1.17	2.19
Central North Sea	2.29	3.05	4.05	5.97	8.65	12.7
Southern North Sea	0	0	0.23	3.21	4.17	5.47
West of Shetland	0.47	1.07	2.03	2.47	4.66	7.7
West of Scotland	0	0.50	2.00	0.49	1.48	6.00
Land	0.05	0.16	0.32	71	0.25	0.85
Other areas of the UKCS	0	0.02	0.18	0	0.035	0.42

Table 2.2 provides a broad indication of the remaining potential with greater than 5% geological chance of success cut off. Some of these leads with their high level of risk may not be produced in the immediate future, a data driven uncertainty reduction method discussed in this work should be part of the industry and OGA stewardship process since commercially recoverable reserves will strongly rely on whether the level of uncertainty in these and many other leads are adequate. The total undiscovered oil and gas in the North Sea is estimated at between 3.3 to 10.1 billion barrels and 12.9 to 35.3 trillion cubic feet respectively (OGA, 2016). Data from Northern, Central and Southern North Sea were used to develop the models and as stated earlier, the emphasis is on sandstone assets.

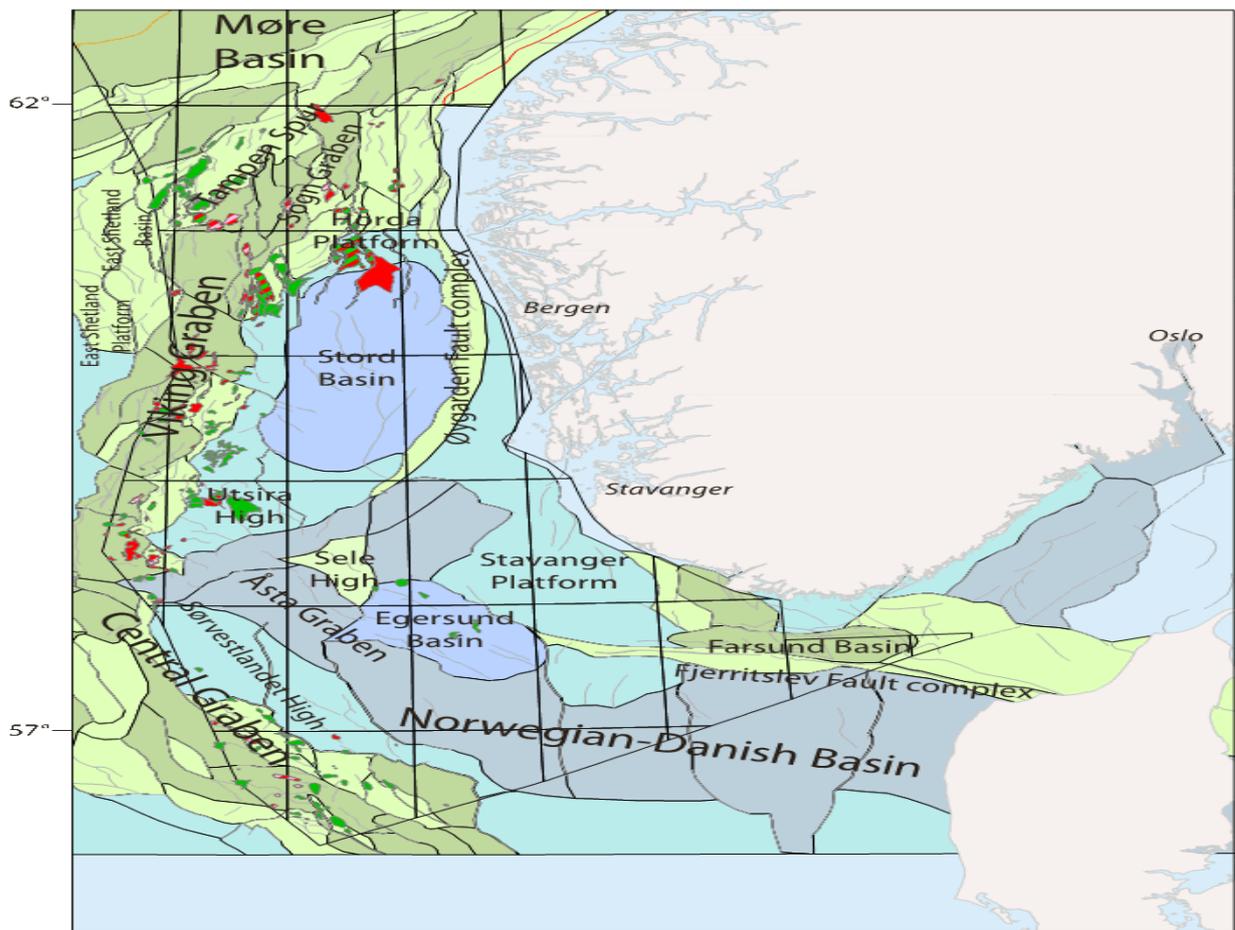


Figure 2. 2 Schematic of the North Sea Graben

2.3 Well Log Practice

Well logs give a concise and detailed plot of reservoir parameters against depth (Figure 2.3). The various properties obtained from well logs and their governing equations are summarily shown in Table 2.3. Models obtained from the different logs like gamma ray, resistivity, neutron, density, sonic as well as nuclear magnetic resonance logs are highlighted and some of their limitations also pointed out. Different results are also obtained for the same variable from the same log, adding to the uncertainty. The ability to interpret a log depends on knowing the significance of each measurement as summarised in the following sections.

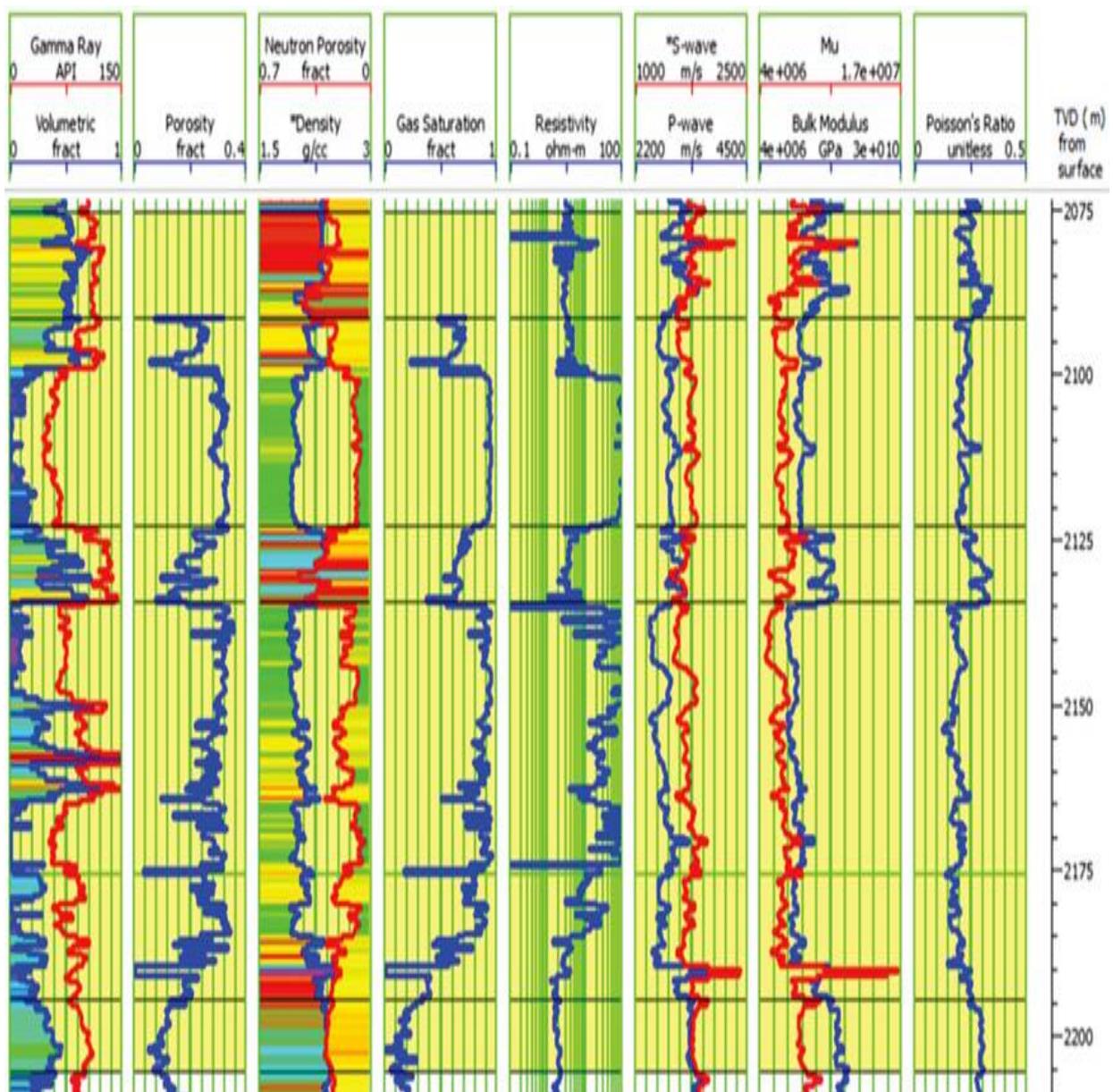


Figure 2. 3 Example footprints of the different types of well logs (CSEG, 2017).

2.3.1 Gamma Ray (GR)

GR measures radioactivity that occurs naturally in the formation thus being able to delineate reservoir rocks such as sandstone and carbonates from non-reservoir rocks like shales and clays. The latter contain naturally occurring radioactive elements like uranium, potassium, thorium and gives a high GR deflection while sands give a low GR value. This is found in track one of Figure 2.3.

2.3.2 Spontaneous Potential (SP)

This log gives a measurement of the difference in voltage between a fixed electrode at the surface and a movable electrode in the well bore. The potential difference is caused by exchange of fluids with different salinities i.e formation fluid and mud filtrate which invades the permeable formation during drilling. If filtrate salinity is less than formation fluids, chlorine ions from formation water causes the SP log to deflect to the left and vice versa. This was not used in this study as they were not available.

2.3.3 Resistivity

Resistivity logs measure the resistivity of the flushed zone around the borehole (shallow), the mud filtrate invaded zone (medium) and the uncontaminated zone which is the true resistivity of the pristine formation (deep). The diameter of mud filtrate invasion and delineation of permeable zone can be evaluated using the separation of the curves. Resistivity log is shown if track 2 of Figure 2.3.

2.3.4 Density

This log measures the electron density of a formation. The tool emits gamma rays into the formation which collides with formation electrons thus giving off energy and scattering in a process referred to as Compton scattering. The number of collision is related to the number of electrons in the formation hence in low density formations, more scattered GR reach the detector than in high density formations. In Figure 2.3, they are shown in track 3 alongside the neutron log.

2.3.5 Neutron

Rock pores are filled with water and hydrocarbon which has hydrogen as the main constituent. This concentration can be used to determine porosity. Hydrogen atoms have similar mas as neutron which emitted by the tool using a chemical source. The neutrons loose maximum energy, slow down and reach very low energy (thermal state) when they collide with hydrogen. The rate at which this happens is proportional to hydrogen index (hydrogen concentration) when is then

converted to porosity. Neutron porosity assumes clean, liquid filled formation and hence needs accurate calibration to be useful. It captures clay bound water thus used in the determination of irreducible water saturation which is dependent on lithology as shown in Chapter Four. They are shown in track 3 alongside the density log.

2.3.6 Sonic

Sonic logs measure the interval transit time Δt as a compressional wave in microseconds per foot. It is affected by lithology and porosity. Fluid filled porosity slows down waves while they are faster in dense and consolidated formations. Interval transit time is the inverse of velocity. Sonic log response are adversely affected by non-uniform hole size, sonde tilt, and gas cut bore hole fluid leading to erroneously high or low reading depending on the size of the non-uniform hole and geometry of the tilted sonde. They are affected by fractures as well depending on the hole orientation. Compressional wave amplitudes is drastically reduced across fracture planes with dip angles of 15-85° but much less affected by near horizontal, horizontal or vertical fractures. In Figure 2.3, the sonic log is shown in track 4.

Table 2. 3 Log and example equations of how their properties were obtained as well as their limitations.

LOGS	PROPERTIES MEASURED	MODELS	REMARKS
Gamma Ray (GR)	Lithology identification Gamma Ray Index (I_{GR}) Volume of Shale (V_{sh}) Correlating cores with logged depth Estimate bed boundaries Clay volume Permeability calculation (k) Wave velocity calculation Perforation depth control	$I_{GR} = \frac{GR_{log} - GR_{min}}{GR_{max} - GR_{min}}$ <p>Larionov (1969)</p> $V_{sh} = 0.083(2^{3.7I_{GR}} - 1)$ <p>Xue et al., 1997</p> $\log(k) = 0.151\Delta t - 0.019\phi - 0.0392I_{GR} + 0.0222RR - 7.7$	Need to be sure of the age of the rock
Resistivity (LLS, LLD, ILD) i.e. shallow and deep laterolog and deep induction log.	Formation water, true and flushed zone resistivities estimation (R_w, R_t, R_{xo}) Porosity (ϕ) Permeability (k) Water saturation (S_w) Movable hydrocarbon Lithology delineation Qualitative estimate of shaliness	<p>Archie, 1950</p> $S_w^n = \frac{aR_w}{\phi^m R_t}$ <p>Waxman-Smit (1968)</p> $\frac{1}{R_t} = \frac{S_w^2}{F^* \cdot R_w} + \frac{BQ_v S_w}{F^*}$ <p>Schlumberger (1972)</p>	Archie fails for shaly sands since it was formulated with brine as the only conductive medium. W-S requires an iterative process since water saturation is on both sides

		$\frac{1}{R_t} = \frac{S_w^2}{F(1 - V_{sh})R_w} + \frac{V_{sh}S_w}{R_{sh}}$ <p>Timur (1968)</p> $k = 0.136 \frac{\phi^{4.4}}{S_{wi}^2}$ <p>Tixier, 1949</p> $K^{1/2} = 250 \frac{\phi^3}{S_{wi}}$ <p>Kozeny-Carman (1927)</p> $k = A_1 \frac{\phi^3}{S_o^2(1 - \phi)^2}$ <p>Saner et al., 1997</p> $\log(k) = 7.04 - 4.19(F)$	<p>Applicable where condition of residual water saturation exists.</p> <p>Physically limited due to the paucity of the logs to show a valid OWC.</p> <p>K-C are suitable for packs of uniformly sized spheres and also Surface area can only be determined using special equipment for core analysis.</p> <p>Determined by multiple correlation from relatively few data. Not used for high gravity crudes (API>40°) and for depths greater than 6500ft.</p>
Sonic (DT)	Porosity evaluation Fluid typing Interval transit time	<p>Wyllie et al., (1958)</p> $\phi = \frac{\Delta t_{log} - \Delta t_{ma}}{\Delta t_f - \Delta t_{ma}}$	<p>Need to be sure of the fluid type</p>

	<p>Identification of gas bearing intervals</p> <p>Cement evaluation</p> <p>Wellbore and Perforation stability</p> <p>Anisotropy for determination of orientation of natural fractures</p> <p>Abnormal pressure</p> <p>Gas detection i.e compressive fluid detection (better seismic int)</p> <p>Assists in Lithology estimate and identification</p> <p>Permeability index</p> <p>Saturation evaluation</p> <p>Understanding rock mechanical properties from shear velocity</p> <p>Completion design (Hydraulic fracture evaluation)</p> <p>Estimate rock permeability (k)</p>	<p>Sen et al., 1990</p> $k = 10^{6.59}(\phi^m V_p/S)^{2.08}$	
<p>Density i.e. bulk(RHOB) and formation compensated (FDC)</p>	<p>Porosity and permeability estimation</p> <p>Gas zone identification (with other logs)</p> <p>Estimates mechanical properties (with sonic)</p> <p>Mineral identification esp. in evaporates</p> <p>Evaluation of shaly sands and complex lithologies</p>	<p>Schlumberger (1975)</p> $\phi = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f}$ <p>Tixier, 1949</p> $k = z \left(a \frac{2.3}{\rho_w - \rho_o} \right)^2$	<p>Need to know the fluid type with certainty</p> <p>Necessity to estimate exact hydrocarbon density.</p>

	Reservoir type Reservoir thickness Fracture, LDT Inter-well correlation Over-pressured zone detection Determination of hydrocarbon density Determination of oil shale yield		
Neutron i.e. compensated log (CNL) and neutron porosity (NPHI)	Porosity determination Porous formation delineation Gas zone identification Complex lithology determination Evaluation of shale content Depth correlation		
Nuclear Magnetic Resonance (NMR)	Petrophysical Analyses Porosity estimation and distribution <ul style="list-style-type: none"> - Effective porosity, ϕ_e - Total porosity, ϕ_t, vugs, fracturing, pores Clay bound water and bulk volume irreducible water in shaly sands evaluation Permeability evaluation (k) Lithologic Characterization	Timur-Coates $k_{Coates} = \left[\left(\frac{\phi}{C} \right)^2 \frac{MFFI}{MBVI} \right]^2$ Quintero et al., 1999 $k = Cpf4.6\phi^4T_{2tm}^2$	T-C represent matrix k hence not applicable to estimation of fracture k

	<p>Grain size and sorting</p> <p>Rock composition</p> <p>Mineralogy</p> <p>Clay presence</p> <p>Grain size</p> <p>Hydrocarbon identification, characterization and typing</p> <p>Hydrocarbon vs water</p> <p>Gas vs oil</p> <p>Heavy vs light oil</p> <p>Viscosity estimation</p> <p>Type of fluid</p> <p>Quality of fluid hydrocarbons</p> <p>Reserve estimates</p> <p>Hydrocarbon producibility by means of free fluid index (FFI)</p> <p>Low-resistivity/low-contrast pay evaluation</p> <p>Low porosity/low permeability, tight formations</p> <p>Residual oil saturation determination</p> <p>Other applications</p>	<p>Swanson (1981)</p> $k = 4.6\phi^4 T_{2lm}^2$ <p>SDR Model</p> $k_{SDR} = C \times T_{2gm}^2 \times \phi^4$ <p>Sen et al., 1990</p> $k = 10^{-0.1(\phi^m T_1)}^{2.15}$ <p>Quintero et al., 1999</p> $\phi_T = FFI + BFV$	<p>SDR works only for water saturated formations but fails for oil and gas bearing unlike T-C model</p>
--	--	--	---

	Enhancing stimulation design by selection of best sites for well paths and perforations.		
--	--	--	--

2.4 Core Analysis

Core analysis information is considered as one of the main corner stones of formation evaluation. The improvements in core technology have been driven by the need to acquire core to support reservoir exploration and development in new environments and in new play types, including deep- and ultra-deep water, high-temperature/high-pressure (HT/HP) reservoirs, and unconventional reservoirs, including shale-gas and tight-oil (Owens and Evans, 2013). Whether the coring objective is focused on routine core analysis, special core analysis (SCAL) or other geologic objectives, significant advancements in core handling have been made to ensure that acquired core reaches the laboratory in the best possible state for analysis like being undisturbed, contains its fluid saturation, prevention of excessive pressure differential when tripping out of hole e.t.c (Skopec, 1994).

Once the cores are brought to the surface, in most situations they cannot be immediately used to calibrate logs. Core depths are different from log depths and thus a basic problem is common depth scale (Worthington, 1991). A very time consuming pre-processing job has to be done to prepare core data for further quantitative analysis. Core-log matching is a difficult task because it involves the comparison of measurements obtained at different scales (Figure 1.1). Moreover, factors such as the type of sampling carried out on cores, the laboratory measurement conditions, the volume of investigation and the precision of each measurement may affect the comparison of wireline log interpretations with core measurements. Furthermore, core data are usually not available in all lithological zones due to cost or technical feasibility.

Over the years there appears to have been a shift from laboratory core measurements to the use of logging measurements to derive similar information. Examples include Laboratory determination of mineralogy (XRD) vs. “geochemical” logging (elemental analysis) mineralogy; core-plug wettability vs. NMR logging measurements; laboratory relative-permeability measurements vs. those predicted/estimated from combined NMR and resistivity log analysis. Core analysis are still required as “ground truth” for calibrating these new logging-derived measurements.

Today, the time interval separating exploration/exploitation and project management needs a faster analysis of core data. One way of achieving this is to estimate reservoir rock properties from well log data. Based on papers presented over a decade ago, the focus was on developing new experimental protocols that could make core data analysis faster but today, there appears to be less emphasis along these lines. Digital rock physics is currently the solution to obtaining core-

analysis data more quickly but at present still needs experimental data for validation. The experimental data used for calibration e.g., temperature, fluid saturation, pore pressure and stresses which ideally should be obtained at reservoir pressure, are often not existing.

As already stated, core analysis could be special or routine. Special core analysis gives a more detailed measurement than routine analysis and must take into account the effects of oil-based mud filtrate on especially wettability, variation of relative permeability across different rock products and the effect of extraction, drying and test procedures on laboratory water flood performance amongst others. Their measurements include capillary pressure, relative permeability, NMR analysis, wettability determination, Archie exponents determination (a, n, m), pore volume compressibility, core mechanical properties, rock fluid sensitivity, residual oil saturation. The common technique for relative permeability determination from core is steady-state method that involves saturating the core with the wetting phase and the non-wetting phase to sustain the needed water saturation. Their relative permeabilities is then determined from the resulting pressure and flow rates. Another technique is to first saturate with oil and then displace with water while measuring the resulting flow rate and pressure to determine their permeabilities. The routine core analysis involves the determination of porosity, permeability (vertical and horizontal), rock and grain density, salinity analysis, CT scanning and slab orientation.

2.5 Well Test Analysis

A well test records changes in downhole pressure that follows a change (starting, stopping or abridging) in flow or injecting fluid involving a single or multistep rate. Wells are tested to obtain reservoir parameters which cannot be measured adequately (in context of scope and quality) with other methods like mud and electrical logging, seismic surveys and coring. Well testing comes in different modes, sizes and can be done at any stage in the life of the reservoir. Openhole testing with Modular Formation Dynamics Tester (MDT), Drill Stem Test (DST) to cased hole testing mode using Cased Hole Dynamics Tester tool (CHDT), Reservoir Saturation Tester (RST). Production well testing is possible using slickline bottomhole pressure surveys while others monitor pressure during shut-in. Some objectives are achieved in hours but some run for days or even weeks (pressure build-up test) which could be complicated by many factors such as wellbore/boundary conditions and production time. Some of the different types of well testing techniques are explained below:

2.5.1 Pressure Build-Up (PBU)

These are the measurement and analysis of the bottomhole pressure data obtained following the shutting-in (zero rate) of a producing well. The well is first produced at a constant stabilized rate and then shutin at time t_p (Figure 2.4). Reservoir properties like permeability thickness, skin effects. It is said to be the most preferred well testing technique as precise control of the rate and initial pressure can be ensured. However it leads to loss of production time costing companies huge sums of money annually.

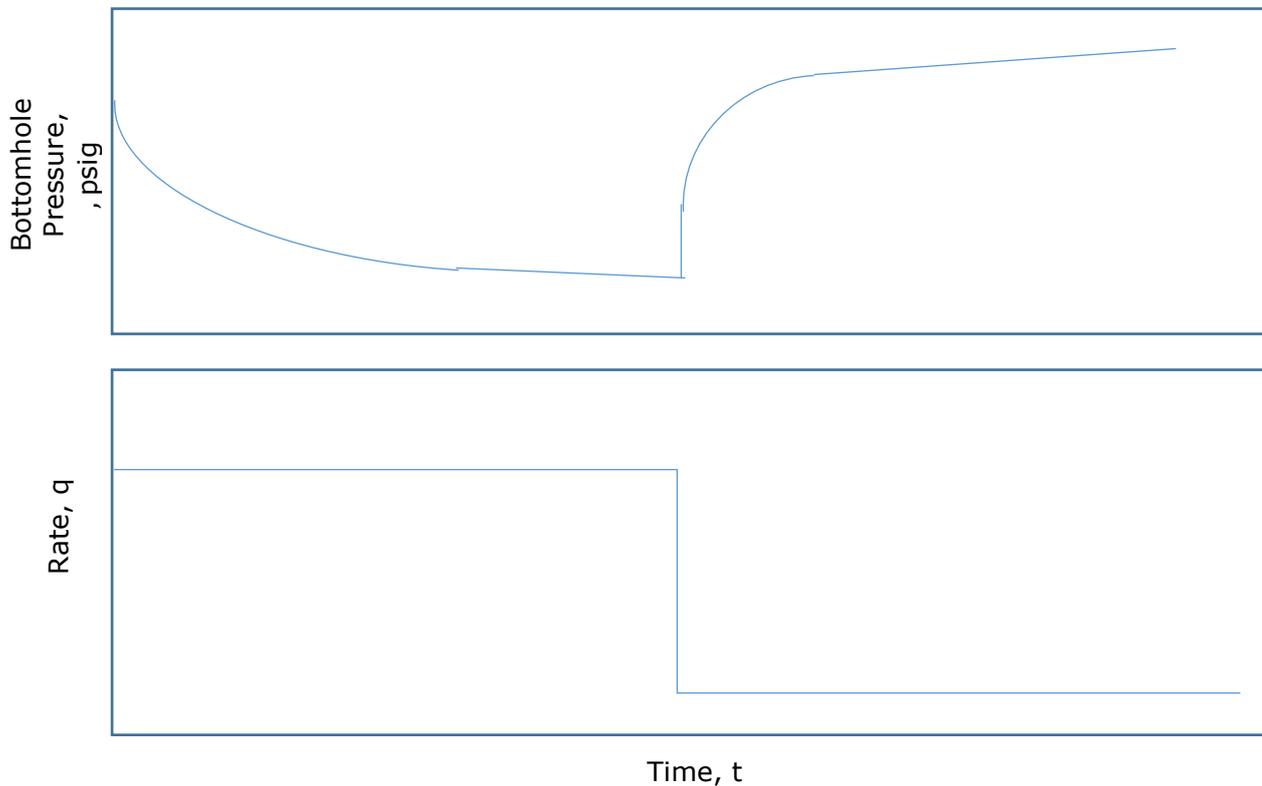


Figure 2. 4 Schematic for Pressure buildup technique

2.5.2 Pressure Drawdown (PDD)

This is a test in which the rate is held approximately constant as the bottomhole pressure is measured (Figure 2.5). It involves shutting-in the well till pressure reaches a static level after which it is flowed at a constant rate. They are used to determine permeability thickness as well just like build up but they are quite noisy and are quite difficult to interpret, difficult to maintain constant production rate and initial pressure determination requires long shut-in period.

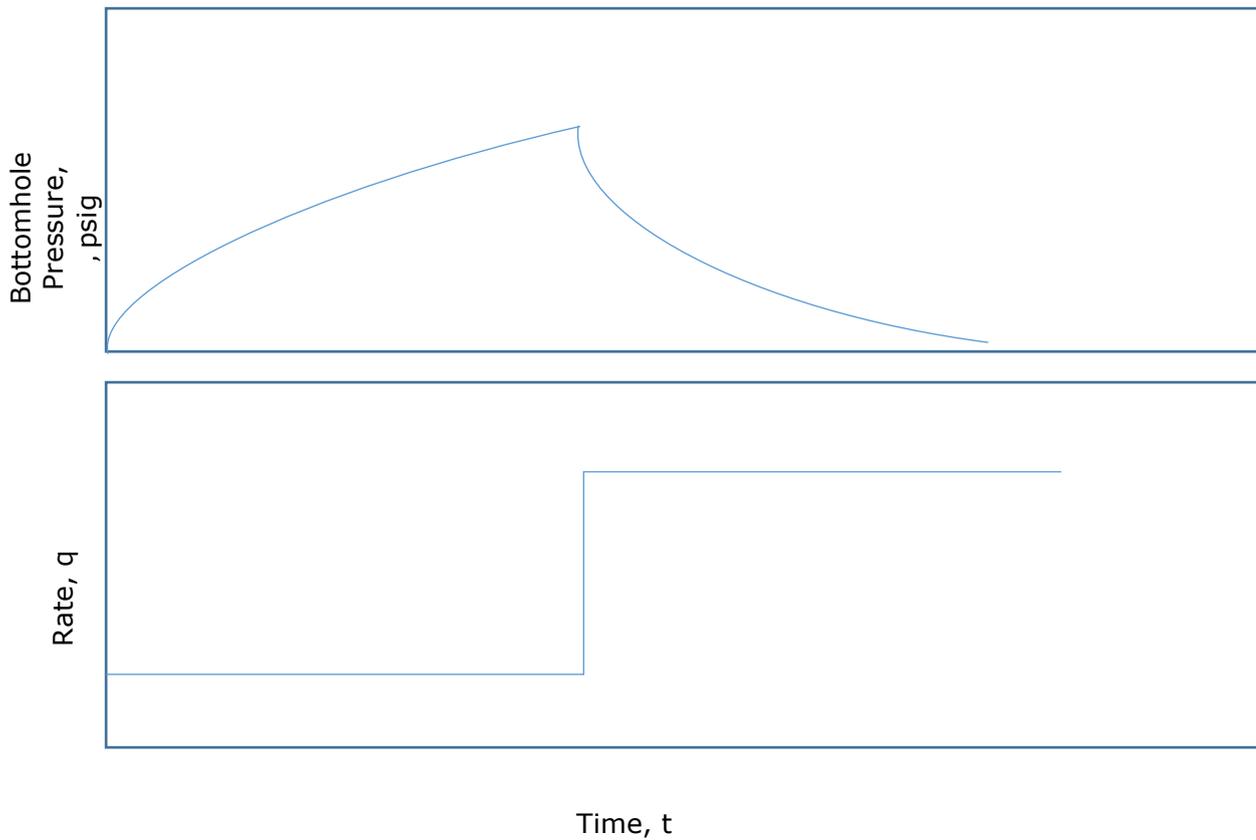


Figure 2. 5 Schematic for Pressure drawdown technique

2.5.3 Drill Stem Test (DST)

DST is a technique to temporarily complete a well in order to determine productive characteristics of a zone of interest. They provide an indication of the fluid content of that zone, the extent of formation damage and the need for stimulation to also increase effective wellbore size. The sequence of pressure recording during DST are shown in Figure 2.6 below:

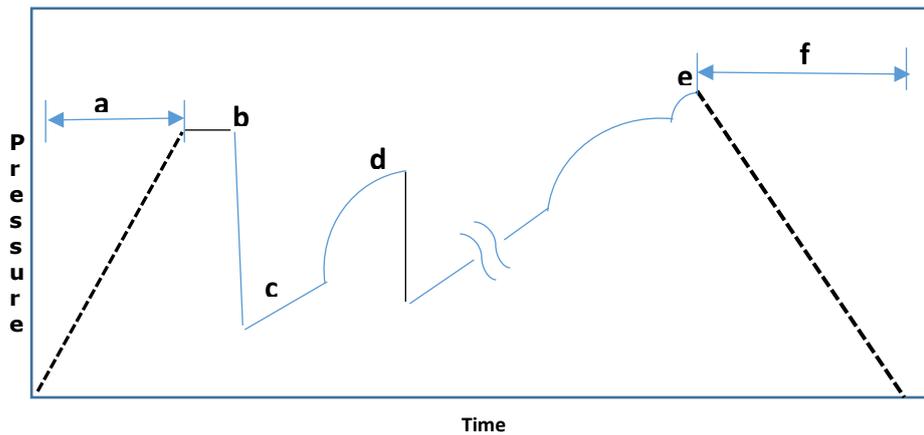


Figure 2. 6 Drill Stem Test procedure (Chaudhry, 2004)

- a) Lowered the tool into the hole causes hydrostatic mud pressure to increase
- b) Setting of the packers causes compression of the mud in the annulus in the test interval leading to a corresponding increase in pressure
- c) The tool is opened causing an inflow from the zone of interest
- d) Period of pressure build up resulting from closure of the test tool
- e) The test has ended and the packers are pulled loose leading to a return to hydrostatic mud pressure
- f) Tool is pulled.

2.5.4 Interference Testing

In this test, a disturbance is caused in one layer or well while the response is measured from another layer or well (Figure 2.7). From the above definition, it can be vertical interference testing or multiwell interference testing. The former involves pressure transient applied to one perforated horizon and the response is measured at another perforated horizon in the same well. This is used to determine vertical permeability and assessment of presence and degree of vertical

communication. In multiwell interference testing, the transient is applied at one horizon in one well and the pressure is measured at the same genetic unit in another well. This gives average permeability as well as an assessment of horizontal communication.

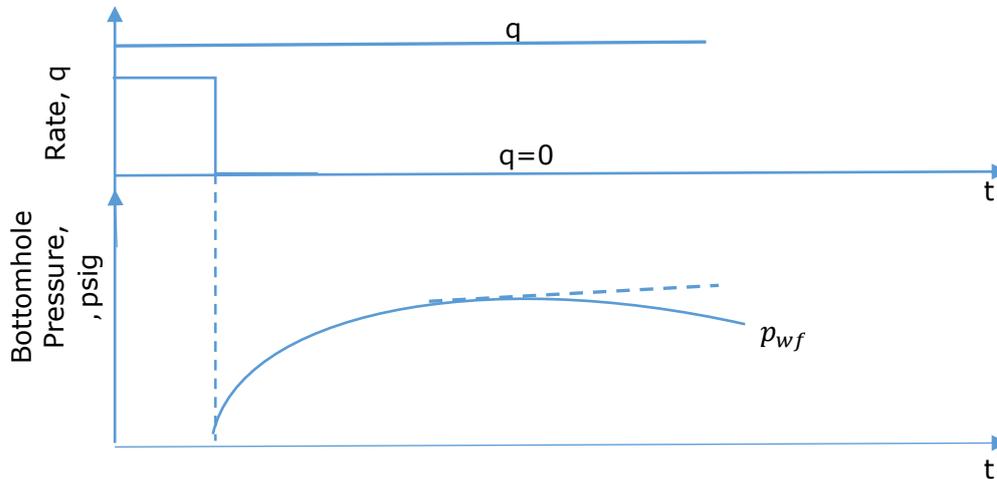


Figure 2. 7 Pressure Data of Interference Test (Heinemann and Zoltan, 2003)

2.5.5 Pulse test

This is similar to interference test but usually within a considerably shorter amount of time with smaller pressure changes thus requiring special differential pressure gauges (Figure 2.8). The production and shutin periods are different but the cycles must be the same. The tangent is drawn to the maximum and the minimum of the pressure waves and the time lags t_{L1} , t_{L2} , t_{L3} and the corresponding pressure differences ΔP_1 , ΔP_2 , ΔP_3 are measured.

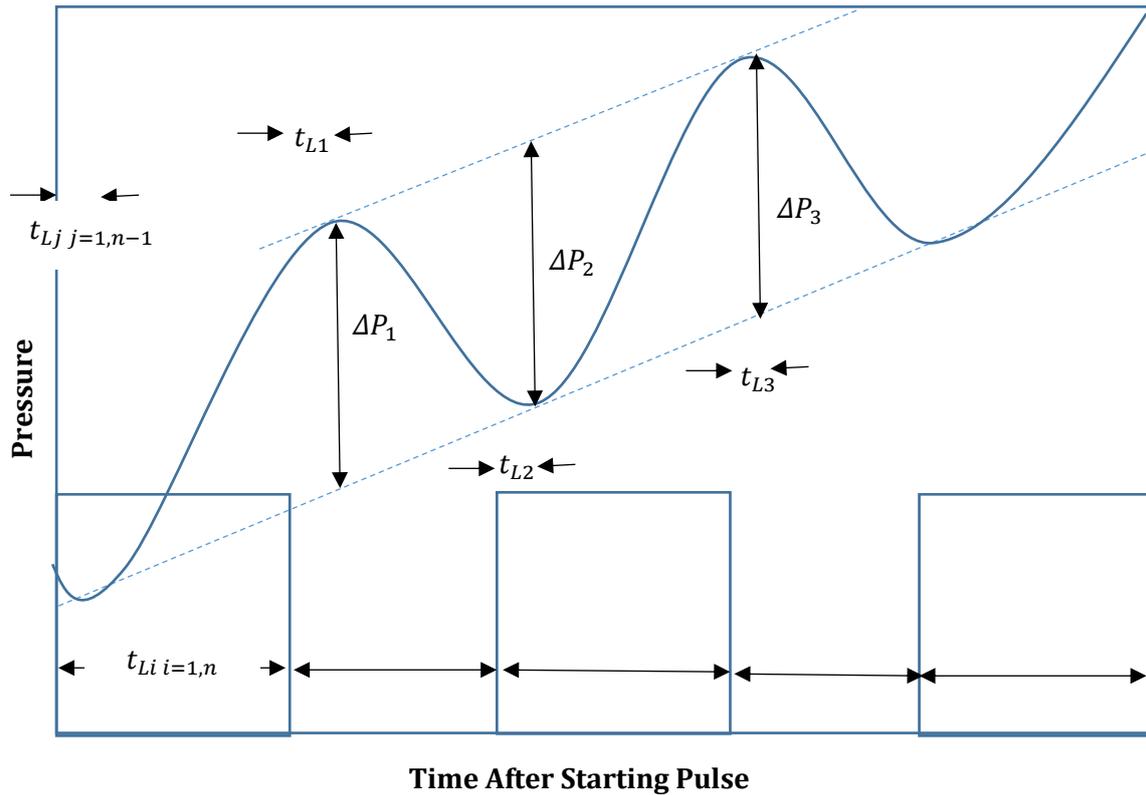


Figure 2. 8 Schematic pulse test rate and pressure history showing definition of time and pulse response amplitude.

2.5.6 Injection Tests

This is the testing of wells by fluid injection into the reservoir. The injection and falloff tests are used to determine the reservoir properties of injection wells during secondary and tertiary recovery. Injection well testing and analysis are simple as long as the mobility ratio of the displacing and displaced insitu fluids is one and the radius of investigation is within the injected water bank. This test is similar to pressure drawdown test except that constant rate is negative with a certain value of q_{inj} . Its procedure involves shutting-in the well until pressure is stabilised at initial reservoir pressure followed by injecting at a constant rate q_{inj} and recording the bottom hole flowing pressure p_{wf} (Figure 2.9).

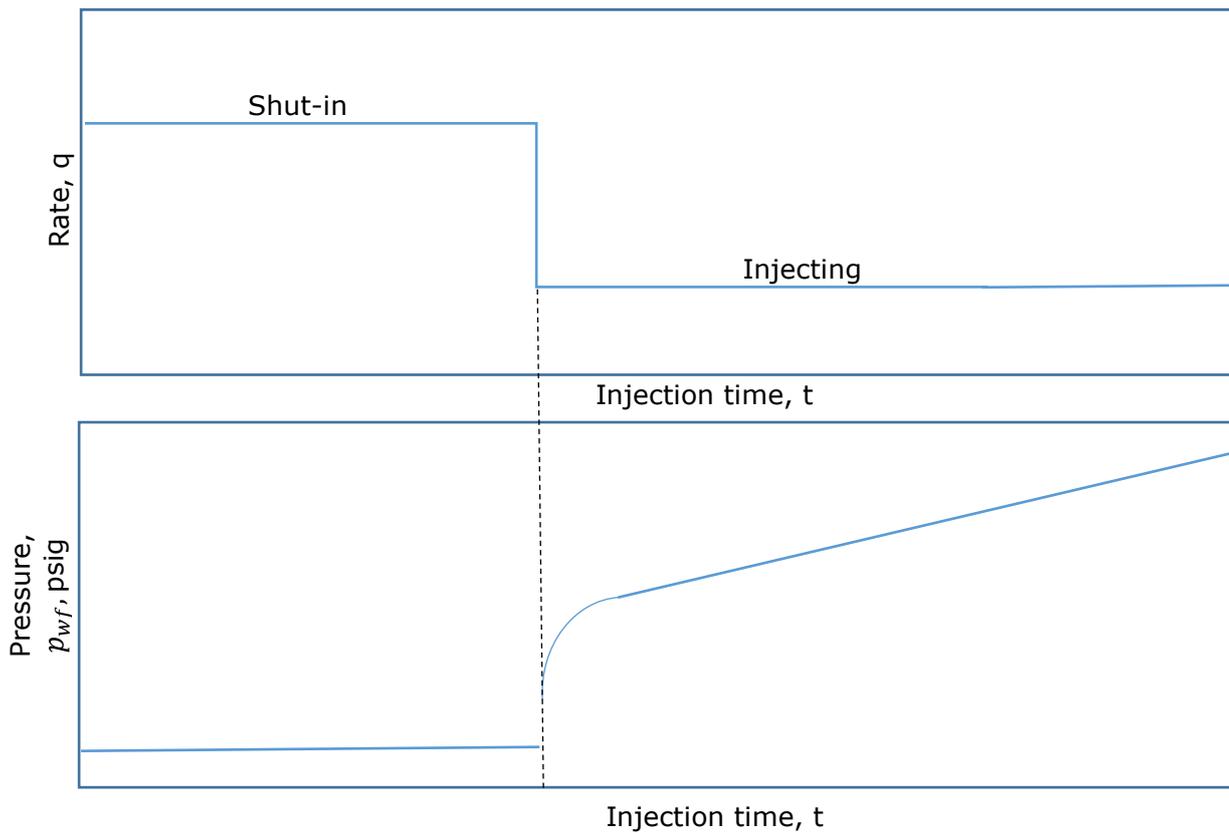


Figure 2. 9 Rate schedule and pressure response in injectivity test (Chaudhry, 2004)

2.5.7 Falloff Tests

A pressure falloff test is preceded by a long duration injectivity test. Shutting in an injection well results in a pressure fall-off that is analogous to a pressure buildup. Thus the procedure includes carrying out an injectivity test that lasted for a total injection time t_p at a constant injection rate as well as well shut in followed by pressure analysis by Horner's method (Figure 2.10).

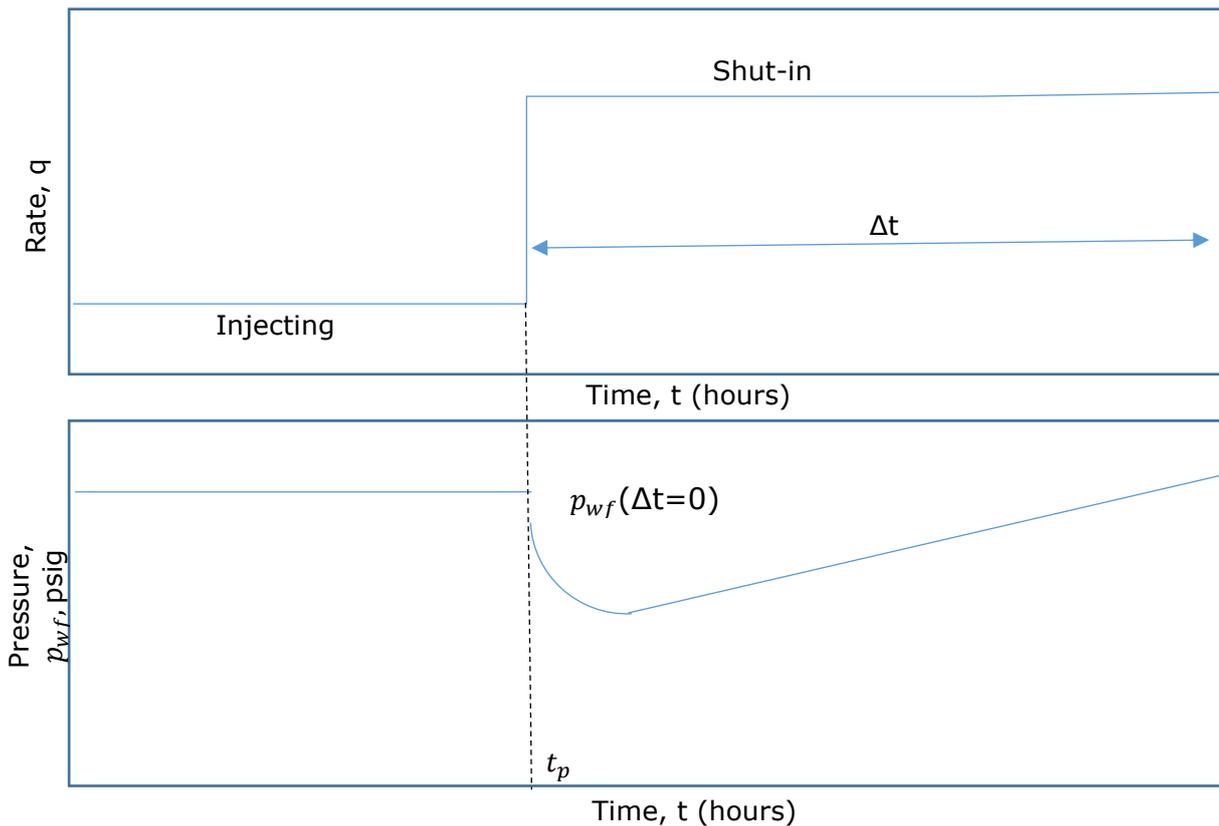


Figure 2. 10 Rate schedule and pressure response for falloff testing (Chaudhry, 2004)

2.5.8 Modular Dynamic Tester (MDT)

Modular Dynamic Tester consists of multiple probes and dual inflatable packers capable of multiple sampling in one trip for timely and accurate decision making. To obtain the sample, the unwanted fluid is discarded before collecting the reservoir fluid sample. Discrimination between formation fluid and filtrate from water and oil based mud is done using the flowline resistivity measurements taken by the probe module. A pump-out module makes sampling possible while monitoring the flowline resistivity by pumping filtrate contaminated fluid into the mud column. They give real time fluid gradients, permeability and contamination assessment.

2.5.9 Reservoir Saturation Tool (RST)

The RST is used to determine water and hydrocarbon saturations while avoiding the lapses of the conventional techniques for doing this like thermal decay time (TDT) logging and carbon/oxygen (C/O) logging. The TDT tool uses the rate of thermal neutron absorption explained by the capture cross section Σ of the formation to estimate water saturation. Saline water has high absorption rate as it contains lots of chlorine which abundantly absorbs thermal neutron efficiently. Fresh

water or hydrocarbon has low absorption rate. When formation water salinity is high, constant and known, this tool provides good saturation estimates. With water injection to enhance recovery comes a reduction in salinity (less than 35,000 parts per million) causing problems for the tool as it cannot accurately distinguish between oil and water whose neutron capture cross sections are similar. C/O tools on the otherhand, are used when salinity is too low or unknown. It measures gamma ray emitted from the inelastic neutron scattering to ascertain the relative concentrations of carbon and oxygen in the formation. Oil bearing formations have high C/O ratio while water and gas bearing zones show low C/O ratio. They have large diameters meaning wells have to be killed and production tubing removed to accommodate the C/O tools. They also have slow logging speeds and are more sensitive to borehole fluid than the actual reservoir fluids thus measurements are less precise.

The RST combines both methods in a tool slim enough to fit through tubing hence no need to kill the well, avoids reinvasion of perforated intervals, allows well to be observed under operating conditions, provides a log of borehole oil holdup, and saves money and time. However its radius of investigation, response to gas and use in uncharacterized environment are still not well known. RST data was not used since they were not available.

In this work, pressure build-up and drawdown were the focus and are discussed in more detail in Chapter Five. Permeability was obtained from these analysis and then compared with their corresponding log and core values to generate a model relationship for these different sources.

2.6 Sandstone Reservoir Rock Properties

2.6.1 Porosity

Porosity can be the number of total or interconnected pore spaces present. Bhatt (2002) recalls that laboratory measurements of porosity most often than not, are higher than in insitu conditions owing to the fact that it is not corrected for overburden pressure. In this regard, porosity from density logs has the potential of being a better estimate since it considers the insitu compressibility of the rock. It is well known that there are lots of uncertainty associated with its measurement and it's detrimental to the accurate determination of other parameters such as permeability, fluid saturation, hydrocarbon in place etc. Figure 2.11 is an arbitrary plot (not drawn to scale) between core porosity and porosity inverse from density, neutron and sonic sources. Figure 2.11a is an ideal case of no uncertainty with each one of density, neutron or sonic porosity at a time on the x-axis having the same value as core porosity which is on the y-axis (A-A¹) while Figure 2.11b is a case with uncertainty thus leading to overestimation (B-B¹) or underestimation (C-C¹) of the property. Although core have their own uncertainty, the focus here

is on the uncertainty arising from the x-axis parameters which are density, neutron and sonic porosities.

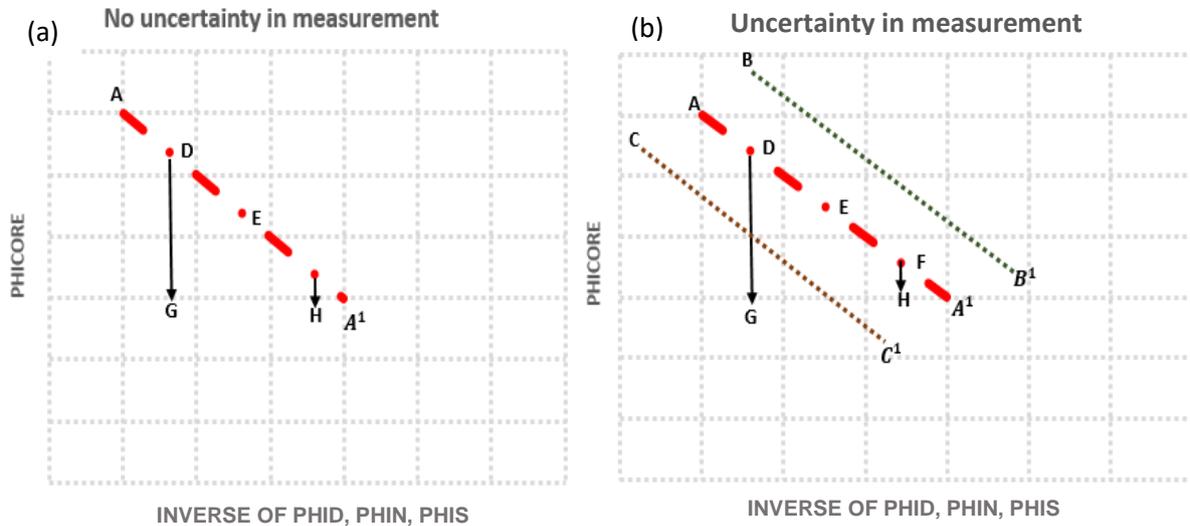


Figure 2. 11 Schematic showing cross plot of core porosity versus density, neutron and sonic porosity inverse explaining the uncertainty inherent in its determination with (a) showing a case where density, neutron and sonic porosities matched core i.e no uncertainty in measurement (A-A¹) while (b) density, neutron, sonic porosities all give different values from core porosity due to uncertainty leading to over-estimation (B-B¹) or underestimation (C-C¹).

Wyllie Time Average

$$\phi_s = \left(\frac{\Delta t_{log} - \Delta t_{mat}}{\Delta t_f - \Delta t_{mat}} \right) \quad (2.1)$$

Raymer-Hunt-Gardner

$$\phi_s = 0.625 \left(\frac{\Delta t_{log} - \Delta t_{mat}}{\Delta t_{log}} \right) \quad (2.2)$$

For unconsolidated formations

$$\phi_s = \left(\frac{\Delta t_{log} - \Delta t_{mat}}{\Delta t_f - \Delta t_{mat}} \right) \cdot \frac{1}{C_p} \quad (2.3)$$

Where, $C_p = \frac{\Delta t_{log} \times C}{100}$

Schlumberger, 1972

$$\phi_D = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f} \quad (2.4)$$

Δt_{mat} , Δt_f , C_p , ρ_{ma} , are very uncertain parameters and thus affects the accuracy of porosity predictions.

2.6.2 Fluid Saturation

Gas, oil and water can be all present in a reservoir and their proportion is defined as saturation. The lowest that this saturation can be is termed as irreducible or residual. This parameter can be estimated from resistivity logs, direct measurement on core plugs. Widarsono (2008) pointed out that assuming an error of 20% in water saturation, large errors as high as 50% may occur. This could potentially be higher considering the difference in the results from the different models. Both total and irreducible water saturation models were developed in this work as further described in Chapter Four.

2.6.2.1 Total Water Saturation Determination for Sandstone Reservoirs

Water saturation is determined using the following methods listed below

- Using resistivity well logs
- Nuclear Magnetic Resonance (NMR)
- Laboratory capillary/saturation measurements
- Dean-Stark water volume determination using oil based mud (OBM)

A hybrid of these techniques can also be used as well to include all wells and the entire hydrocarbon column.

2.6.2.1.1 Well logs determination of Water Saturation

As have been shown in Figure 2.12, there are different models used for the determination of water saturation from resistivity logs. Starting with Archie in 1942 model for clean sands and since several models for shaly sands have been developed with some considering volume of shale while others consider cation exchange capacity of the clays. To determine water saturation accurately from resistivity logs, connate brine resistivity (R_w) or conductivity (C_w) needs to be accurately determined as well. This is achieved by using the Spontaneous potential log in wells drilled with water based mud, back calculate using invasion corrected logs, a and m with Archie models in the

aquifer since S_w is 100% in this region, samples obtained from flow tests of the aquifer interval (which must be checked for contamination by mud filtrate invasion). Since R_w varies with temperature at constant connate water composition, S_w is affected by temperature which is measured with most log and DST. At any given temperature level, there appears to be no significant effect of oil viscosity on irreducible water saturation. An increase in temperature decreases residual oil saturation and increases wettability and therefore irreducible water saturation and the relationship is said to be partially reversible (Poston et al. 1970). Electrical concepts application require distinct salinity contrasts between formation waters and the conductive muds as well as well-defined invasion profiles in order to quantify and separate the conventional volume conductivity term from the surface conductivity term (Pape et al. 1987).

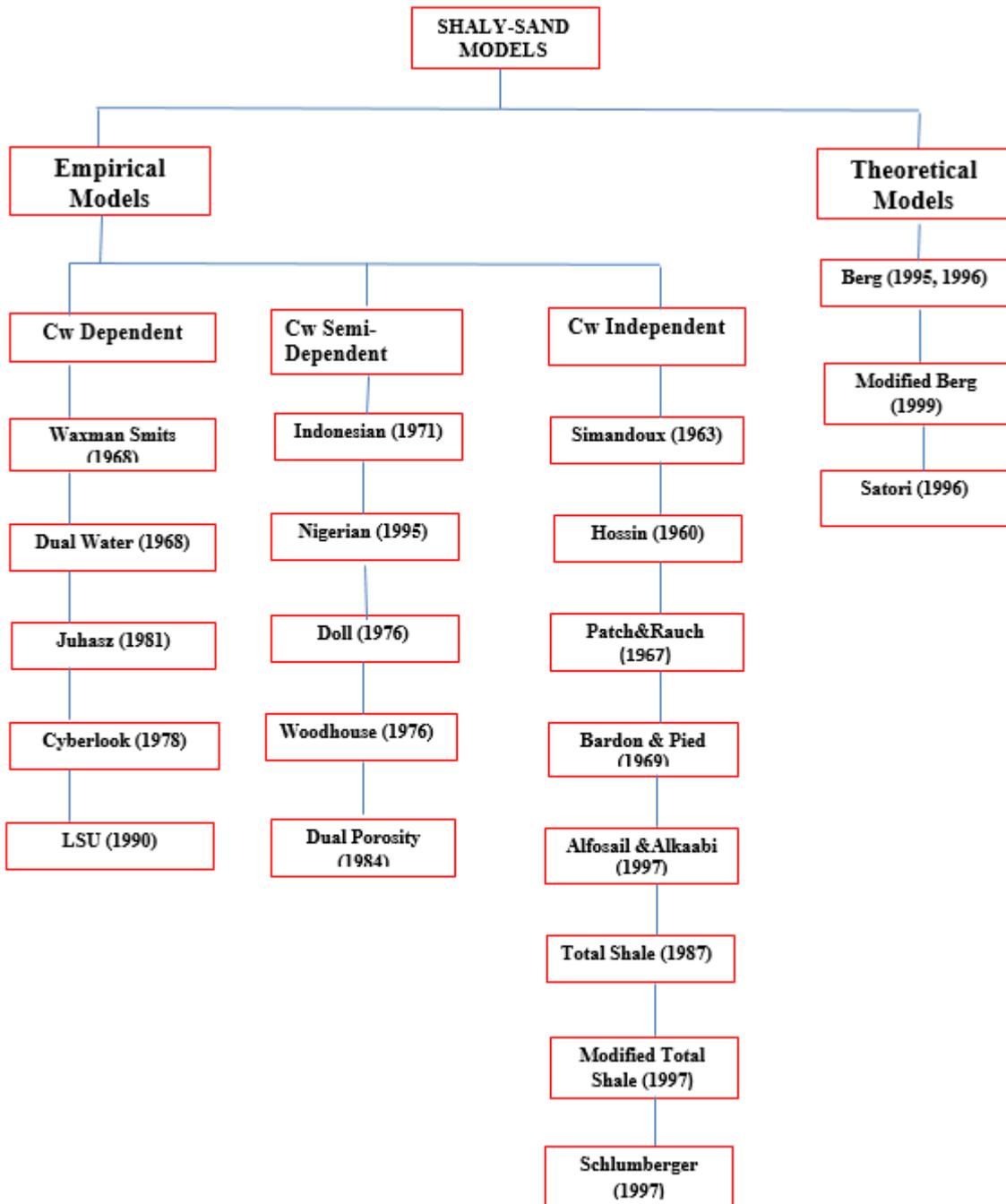


Figure 2. 12 Different shaly sand water saturation models with Cw as water conductivity which is the inverse of resistivity (Dawood et al., 2002)

2.6.2.1.2 Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) logs has shown they can solve petrophysical problems. The pulse-echo technique and the capability to measure T2 signals below 1ms, can give insitu recording of T2 spectra and its BVI component. This is used for differentiating between movable

and immovable fluids and they are lithology independent. Despite its importance, these logs are expensive, relies on T_2 cutoff which is difficult to determine in the laboratory, affected by tool run speed, presence of hydrocarbon which is inevitably present in the reservoir and needs environmental corrections.

2.6.2.1.3 Capillary/Saturation Measurements

The concept underlying this is that the reservoir has attained a state of equilibrium over geologic time and it is this equilibrium that is reproduced in these laboratory experiments using centrifuge, mercury injection capillary pressure (MICP) and porous plate techniques. This is an indirect method of water saturation determination and requires several assumptions and inferences.

2.6.2.1.4 Oil-Based Mud Water Saturation Determination

This method uses distillation extraction where water in the sample is vaporized using a boiling solvent which is later condensed and flowed back into the core sample to extract the oil as well. They provide one of the most direct methods of water saturation determination above the transition zone and requires less assumptions. The pitfall of this technique been that it does apply to the lowest parts of the transition zone where the water phase has mobility. Also due to expense, the number of OBM core data is limited to fewer wells. This technique of water saturation determination requires proper planning and monitoring from mud formulation to actual measurements hence lots of time and effort is required to achieve success of the project.

2.6.2.2 Influencing Parameters on Irreducible Water Saturation

Historically, the water saturation at the top of the pay zone is approximated as the irreducible water saturation and then assumed to be constant throughout the reservoir thus indicating that the reservoir is homogeneous which means that permeability is underestimated, implying to be dependent only on porosity and not on variations of pore throat distributions, irreducible water saturation, grain size, texture, capillary pressure. This underestimation is valid if there is no considerable separation between the free water level (FWL) and the reservoir top hence the water saturation at this position may also be way more than the irreducible water saturation. It has also been shown by some authors that water rather than clay is the primary contributing factor to low resistivity pays which is currently an industry problem.

Irreducible water saturations may vary between 27 and 55% of the total core porosity (Ringen et al., 2001). The industry standard for determining irreducible water saturation has been to drill cores with oil based muds (since water based muds will increase contamination). The connate water can itself be mobile and maybe moved by mud invasion (oil or water) or gas expansion in the transition zone. This is key since at some point into the transition zone, the S_w from the core may not be trusted and thus it may then be critical to determine the residual oil saturation using retort test. Even with the irreducible water above the transition zone determined and residual oil in the aquifer known, there is still going to be uncertainty in the lower section of the transition zone.

Due to the foregoing, there is need for a systematic approach for the estimation of irreducible water saturation for accurate application in other models to be valid for each product of rock that forms the reservoir. Both a machine learning stacked ensemble model using well logs and an empirical model computed from influencing parameters such as specific surface area to pore volume ratio, volume of shale and flow zone indicator hence obtaining the property as continuous curve across reservoir intervals from well logs is discussed. Their relationship with irreducible water saturation is shown in Figure 2.13. Some of these parameters are discussed below:

2.6.2.2.1 Surface Area to Pore Volume

Specific surface area is a key parameter to understand the physics of porous media. Laboratory measurements on a large number of different sedimentary reservoir rocks indicates a close relationship between specific surface and several petrophysical parameters (Ripe 1979, 1984). Results of several experiments have shown that surface area is not just a function of grainsize but more importantly of mineralogy as similar sizes of quartz and clay give remarkably different surface area values. Furthermore, for the common reservoir clays, studies have shown that smectite gives a very high surface area of an order of magnitude higher than illite which is a factor of 2 or 3 times more than kaolinite (Zamanek, 1989). Hence it is logical to say that irreducible water saturation is strongly dependent on lithology and mineralogy. Hence, where grains are small, the formation has high surface to volume ratios leading to high irreducible water saturation resulting in reduced bulk resistivities. The use of the resistivity log as pay indicators could result to bypass of pay zones which is a challenge.

This powerful geometrical parameter is also affected by variations in permeability, porosity, tortuosity and formation resistivity factor which encapsulates the size, shape, distribution, sorting and packing of the grains. In addition to its role in reservoir engineering for understanding

hydrocarbon recovery (Chilingarian, 1992; Barlai, 1976), it finds extensive use in environmental studies (Salem, 1992; Sen et al., 1990); chemical engineering (Carman, 1956; Bear, 1972; Chalkley et al., 1949); sedimentology (Pettijohn et al., 1973); geophysics, petrophysics, hydrogeology (Carman, 1937, 1938; Bear, 1972; Salem, 1992).

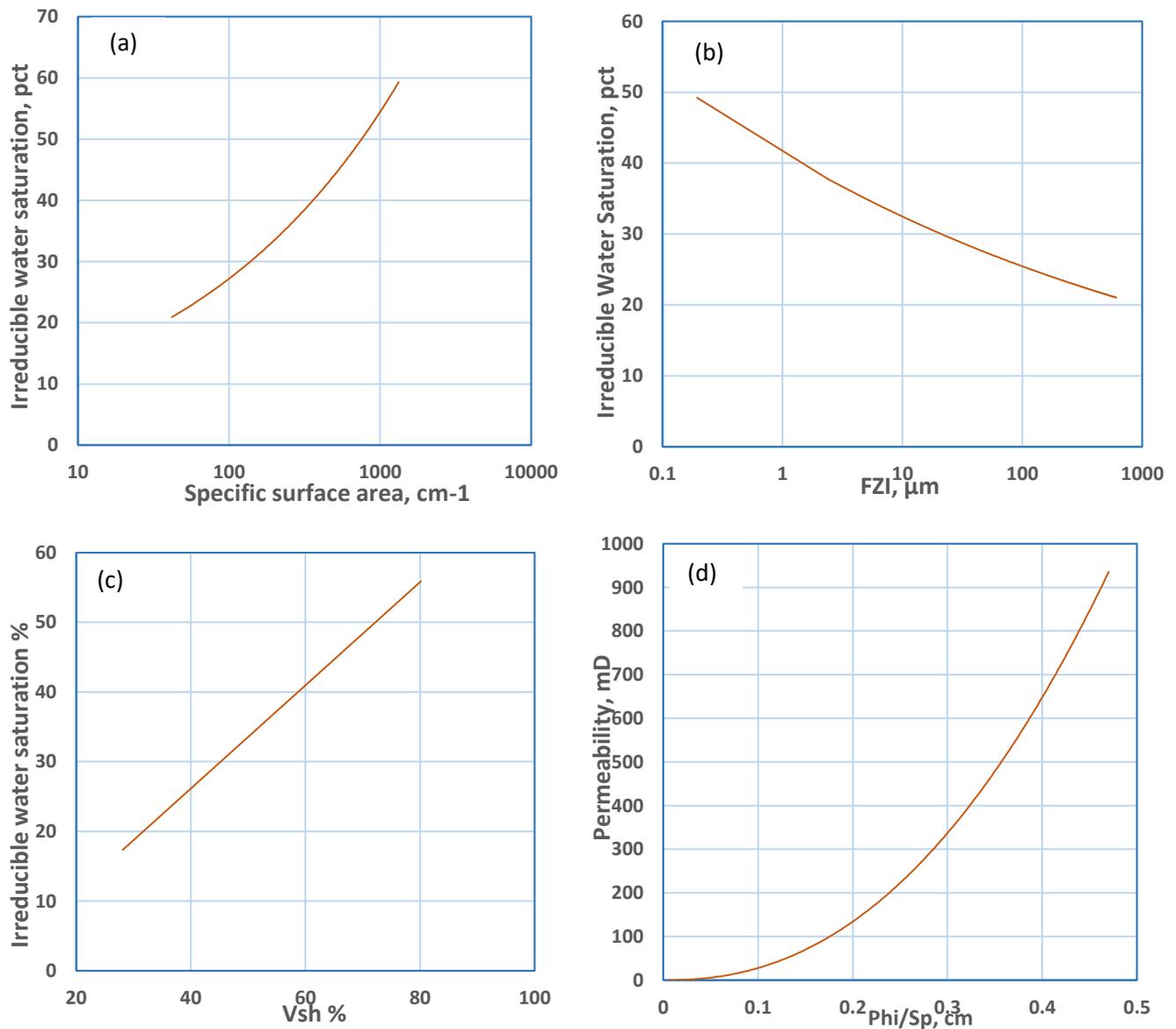


Figure 2. 13 Shows relationship between irreducible water saturation and (a) specific surface area to be directly proportional (b) flow zone indicator as inversely proportional (c) volume of shale to be directly proportional while (d) shows the role of surface area in reducing the permeability-porosity scatter.

2.6.2.2.2 Flow Zone Indicator (FZI)

FZI is a unique and useful value for quantifying the flow character of a reservoir. Fluid flow quality of a rock is a function of pore geometry, grain shape, size, packing, sorting, mineralogy and they help to zone reservoir into several units. Reservoirs with large pore throats and high permeability have short transition zones, and the transition zone at a gas-oil contact will be shorter than that at an oil-water contact simply because of the inter-phase density differences involved. Amaefule et al. (1993) described how porosity and permeability can be related to the wetting surface of a porous medium by the concept of mean hydraulic radius (r_{mh}).

That is,

$$r_{mh} = \frac{\text{Cross sectional area}}{\text{Wetted perimeter}} \quad (2.6)$$

Assuming cylindrical capillary tube from

$$r_{mh} = \frac{\text{Pore radius, } r}{2} \quad (2.7)$$

Using 82 corrected and 240 uncorrected samples, Winland developed a correlation for pore radius, r , using the 35th percentile mercury injection R_{35} ,

$$\log(R_{35}) = 0.732 + 0.588\log(k_{win}) - 0.864\log(\phi) \quad (2.8)$$

For a bundle of capillary tube modelled reservoir (Kozeny-Carman), a relationship between porosity and permeability can be derived by applying Darcy and Poisseuille's equation to obtain:

$$k = \frac{\phi_e r^2}{8\tau^2} = \frac{\phi_e}{2\tau^2} \left(\frac{r}{2}\right)^2 = \frac{\phi_e r_{mh}^2}{2\tau^2} \quad (2.9)$$

And hence derived surface area per grain volume to be:

$$S_{gv} = \frac{2}{r} \left(\frac{\phi_e}{1 - \phi_e}\right) = \frac{1}{r_{mh}} \left(\frac{\phi_e}{1 - \phi_e}\right) \quad (2.10)$$

$$r_{mh} = RQI = 0.0314 \sqrt{\frac{k}{\phi_e}} \quad (2.11)$$

Thus

$$FZI = \frac{rqi}{\phi_z} \quad (2.12)$$

Where

$$\phi_z = \frac{\phi_e}{1 - \phi_e} \quad (2.13)$$

2.6.2.2.3 Volume of Shale

The water-wet microporosity can cause high immobile water saturations resulting in low electric resistivities (1-5ohmm) of the pay zones. As stated earlier, several reservoir clays have different affinity to water but generally the greater the volume of shale the more the irreducible water held back from flow (Figure 2.13). Several models like Larionov, Clavier, Stieber could be used for determining this property although Larionov's model for older rocks (Equation 2.14) was used in this work since most of the sandstone formation are of Jurassic age.

Larionov for older rocks (1969):

$$Vsh = 0.33 (2^{(2*IGR)} - 1) \quad (2.14)$$

Larionov for Tertiary rocks (1969):

$$Vsh = 0.083 (2^{(3.7*IGR)} - 1) \quad (2.15)$$

Clavier (1971):

$$Vsh = 1.7 - [3.38 - (IGR + 0.7)^2]^{\frac{1}{2}} \quad (2.16)$$

Stieber (1970):

$$Vsh = \frac{IGR}{3 - 2 \times IGR} \quad (2.17)$$

2.6.2.3 Residual Oil Saturation Determination

Several methods are available for the determination of residual oil saturation and these are summarized in Table 2.4. None of the methods for determination of the property can be regarded as the best method (Teklu et al. 2013). Although there is a difference between remaining oil saturation and residual oil saturation, they are used interchangeably in this work since innovative methods can reduce even the residual value just like the remaining oil saturation. An example of this was described by Chun and Gary (2008) that while a tertiary polymer flood (after water flood) could not mobilize residual oil in homogenous and water wet core, a secondary polymer flood i.e without waterflood, could displace the residual oil as noticed in the same core.

Table 2. 4 Pros and Cons of the different residual oil saturation measurements (Chang et al., 1988)

Method	Models	Pros	Cons
Core	<p>Rathmell et al.1973</p> $(\bar{S}_{or})_{res} = (\bar{S}_{or})_{core} B_o E$ <p>Craig (1971)</p> $(\bar{S}_{or})_{res} = (\bar{S}_{or})_{core} B_o E \frac{M}{1 - V^2}$	<p>Excellent accuracy</p> <p>Widely available</p>	<p>Difficult to get gas saturation</p> <p>Poor to fair core recovery</p> <p>Difficult to obtain insitu S_{or}</p> <p>New well required</p>
Well log	<p><u>Resistivity log</u></p> <p>Archie, (1942)</p> $S_{or} = 1 - \left(\frac{R_{ox}}{R_t} \right)^{1/n}$	<p>Good accuracy</p> <p>Good radius of investigation</p> <p>Widely available</p>	<p>Uncertainty with det. of n</p>
	<p><u>NMR log</u></p> <p>Neuman and Brown, (1982)</p> $S_{or} = \frac{FFI}{\phi}$	<p>Direct measurement</p> <p>Lithology independent</p>	<p>Expensive</p>

	<p><u>PNC log</u></p> <p>Randall et al., (1983)</p> $S_{or} = 1 - \frac{\Sigma_{b2} - \Sigma_{b1}}{\phi(\Sigma_{w2} - \Sigma_{w1})}$ <p>Where</p> $\Sigma_b = \Sigma_{ma}(1 - \phi) + \Sigma_w S_w \phi + \Sigma_{hc}(1 - S_w)\phi$	<p>Porosity not required</p> <p>Excellent accuracy</p> <p>Can measure movable oil saturation</p> <p>Different btw gas, oil and water bearing fm</p> <p>Excellent accuracy of ROS det. behind casing</p>	<p>Cross section determination uncertainty</p> <p>Can be sensitive to chlorine content of formation water</p> <p>Indirect measurement</p>
	<p><u>Carbon/Oxygen log</u></p> <p>Horner and Sanyal (1984)</p> $S_{or} = \frac{C/O_{log} - C/O_{100\%water}}{C/O_{100\%oil} - C/O_{100\%water}}$	<p>Not sensitive to chlorine content of water</p> <p>Independent of shaliness</p> <p>Good accuracy</p>	<p>Uncertainty of some of its parameters</p>
<p>Well Tracer</p>	<p>Bragg et al., (1976)</p> $S_{or} = \beta / (\beta + k_1)$ <p>Where</p> $\beta = k_i S_{or} / (1 - S_{or})$	<p>Large radius of investigation</p> <p>Applicable in a wide range of conditions</p>	<p>Weighted average single value per pay zone</p> <p>Expensive</p> <p>Sometimes difficult to achieve</p> <p>Tracee recovery an issue</p>

	and $k_1 = c_1(oil)/c_1(water)$	Insitu measurement	
Well Test	Cordiner et al., (1972) $S_{or} = \frac{c_t - c_w - c_\phi}{c_o - c_w}$ Where $c_t = \frac{0.0002637(k/\mu)t}{\phi r^2} \left(\frac{t_D}{r_D^2}\right) M$	Large radius of investigation Excellent accuracy	Less accurate in saturated reservoirs Expensive Requires RP Curve from cores Weighted average single value per pay zone
Material Balance	Terry and Rogers (2014) $S_{or} = \frac{(N_{oi} - N_p)B_{OWF}}{7758Ah\phi}$	Large radius of investigation	Weighted average single value per pay zone Uncertainty in reserve estimate Accuracy in doubt
Production Data	Dietrich (1975) $S_{or} = \frac{V_o}{V_p}$ $= \frac{(N_{oi} - N_p)B_o}{N_{oi}B_{oi}(1 - c_\phi\Delta p)/(1 - S_{wi})}$	Insitu measurement Large radius of investigation Simple calculations	Reliable input parameters difficult to obtain
Reservoir simulation	Numerical	Large radius of investigation	Geologic model uncertainty Poor accuracy

It is important to note that these residual oil are found in the pore hence as can be deduced from Table 2.4, the larger the radius of investigation the lower the resolution and accuracy of measurement as well as obtaining the property as a profile. Machine learning model was developed from logs and core capable of giving accurate predictions of the parameter.

2.6.3 Permeability

Permeability refers to the ease with which fluids flow in a reservoir. While porosity and water saturation are scalars, permeability is a tensor (Matheron 1967). Even though permeability variation is more pronounced than that of porosity, its determination usually requires extensive petrophysical analysis, pressure and rate transient analysis. Reservoir simulation models require permeability values to be as accurate as possible in order to forecast performance with less uncertainty. Permeability can be absolute, effective or relative. Permeability measurements from core are absolute influenced by the microscopic nature of the measurement, absence of insitu temperature, pressure and saturation conditions (environment). As stated before special core analysis can determine relative permeability. Most logs on the otherhand, measure absolute permeability. The complexity of sandstone reservoir structures and the scanty parameterization make their models nonuniversal although their parameters are measured at insitu conditions. In this work, deep learning models with their special ability to generalise were used to determine relative permeability from logs. Lastly unlike most core and well log measurements, well test measures effective permeability when the effective flow interval (EFI) is known but lacks the foot by foot resolution of the other two. In this work, Lorenzo plots were used to determine the effective flow interval.

2.7 Choice of Variables

2.7.1 Based on Net Present Value

Although Helle et al., (2001) stated that porosity and permeability are the most important reservoir parameters, it was Bouchard and Fox, (1999) who presented a methodology using value of information based decision making by means of Monte Carlo integrating uncertainty to determine the most important parameters controlling the Net Present Value of an asset (Figure 2.14) and concluded that absolute permeability, relative permeability, heterogeneity, residual oil saturation, porosity, irreducible water saturation and wettability are the most important parameters influencing the value for different reservoir assets like complex carbonates, fractured oil reservoir and over pressured mini basin at almost every stage of their life.

$$NPV = \sum_{t=1}^n \frac{R_t}{(1+i)^t}$$

Where R_t is the net cash inflow-outflows during a single period of time t , i is the discount rate or return that could be earned in alternative investments, t is the number of timer periods.

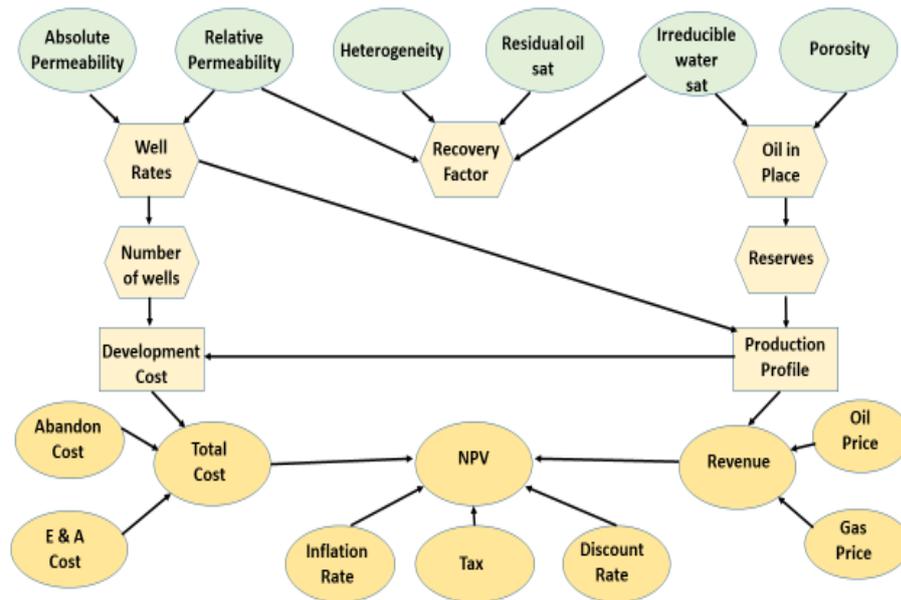


Figure 2. 14 Simplified Influence Diagram Showing Relationship between Reservoir Parameters and Net Present Value (Adapted from Bouchard & Fox, 1999)

Bjørlykke and Jahren (2010) also mentioned that the most important reservoir parameter are porosity and permeability critical for petroleum and even groundwater recovery.

2.7.2 Based on Reserve Estimation

Monte Carlo simulation was used to calculate the uncertainty range for a reserve from independent parameters (Equation 2.18). The stochastic approach randomly selects combinations of these parameter values (and their assigned distribution) using a computer algorithm and 1000 simulation runs was selected to build a reserve distribution curve (Figure 2.15). The vertical axis depends on the number of times each reserve estimate occur. The P10 (upside), P50, P90 (downside) and the base case were developed.

$$N = 7758Ah\phi[1 - S_w]/B_{oi} * RF \tag{2.18}$$

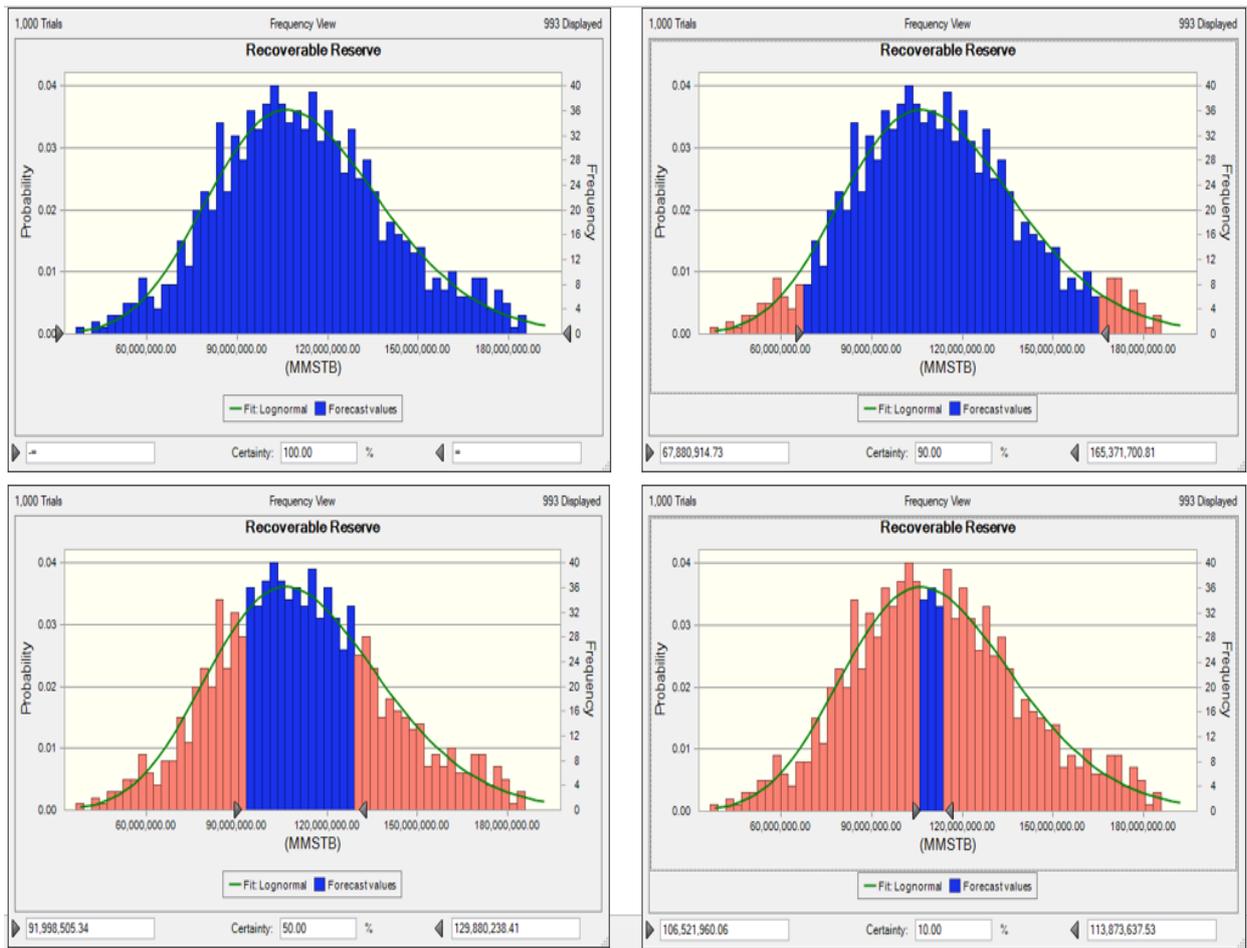


Figure 2. 15 Distribution function showing the P100, P90, P50 and P10 using 1000 simulation runs for the Ness reservoir, Northern North Sea.



Figure 2. 16 Showing porosity and water saturation as the most sensitive for reserve estimation with both P10 and P90 cases for Ness Reservoir, Don Field, North Sea.

From Figure 2.16 it is clear that amongst the parameters used for reserve estimation such as porosity, reservoir area, pay thickness, formation volume factor (Bo), water saturation and Recovery Factor, it was porosity and water saturation that were the most sensitive. Error in the calculation of these parameters could greatly increase the uncertainty of base case, downside and upside reservoir estimation (Garb, 1988).

2.7.3 Based on Pressure Transients

The stochastic Monte Carlo simulation was also run on pressure transient parameters involved in radial flow regime in infinitely acting reservoirs (Equation 2.19) as shown in Figure 2.17 and 2.18. The parameters used were permeability, reservoir thickness, flow rate, formation volume factor, viscosity, skin, well radius, compressibility, time, porosity. It is clear to see that permeability is the most sensitive and an uncertainty in it can result in greater error in pressure transient calculations compared to equal uncertainty in other parameters involved.

$$\Delta p = 162.6 \frac{qB\mu}{kh} \left[\log \Delta t + \log \frac{k}{\phi \mu c_t r_w^2} - 3.23 + 0.87S \right] \quad (2.19)$$

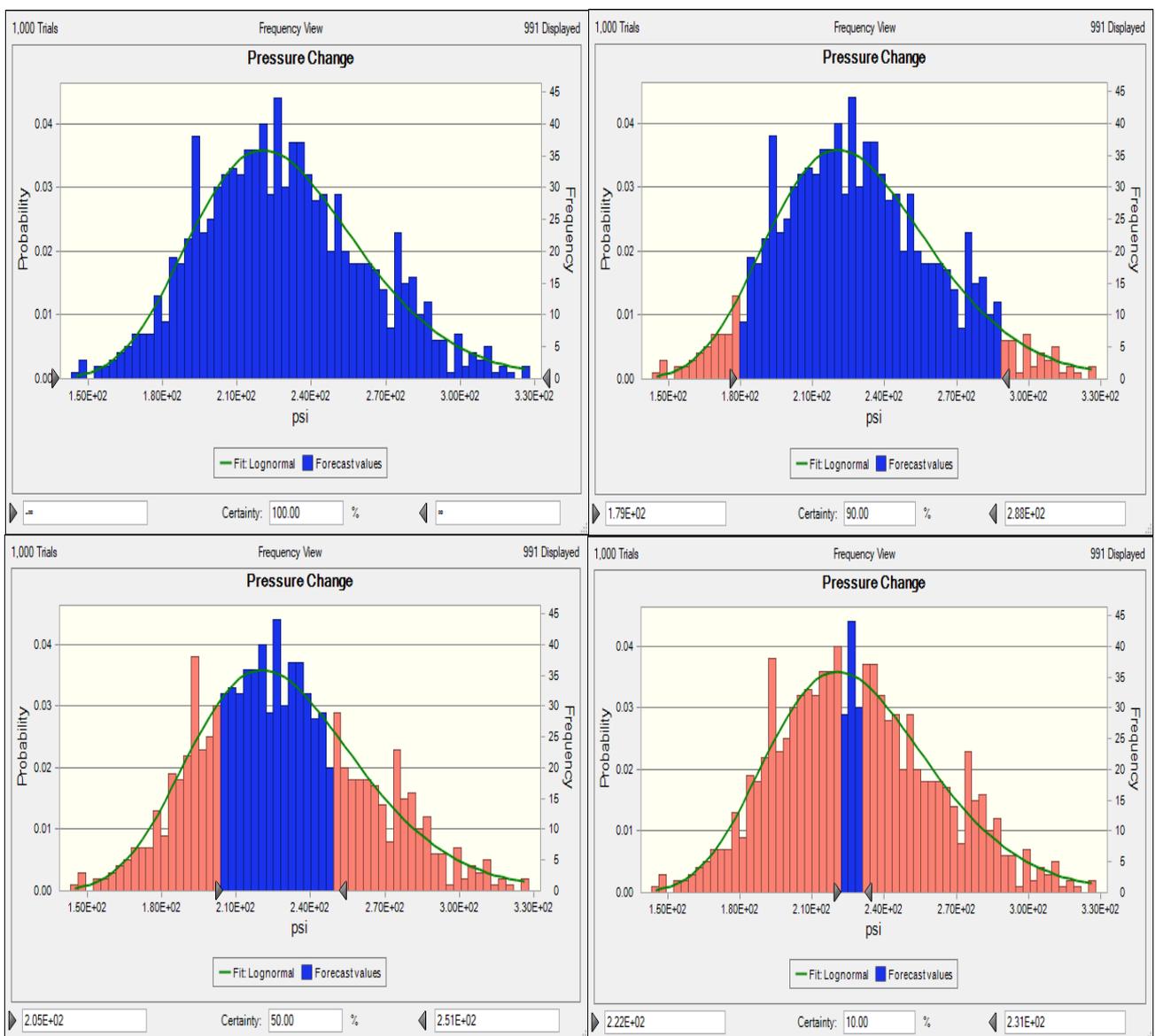


Figure 2. 17 Distribution function showing the P100, P90, P50 and P10 using 1000 simulation runs for the Sherwood reservoir, North Sea.

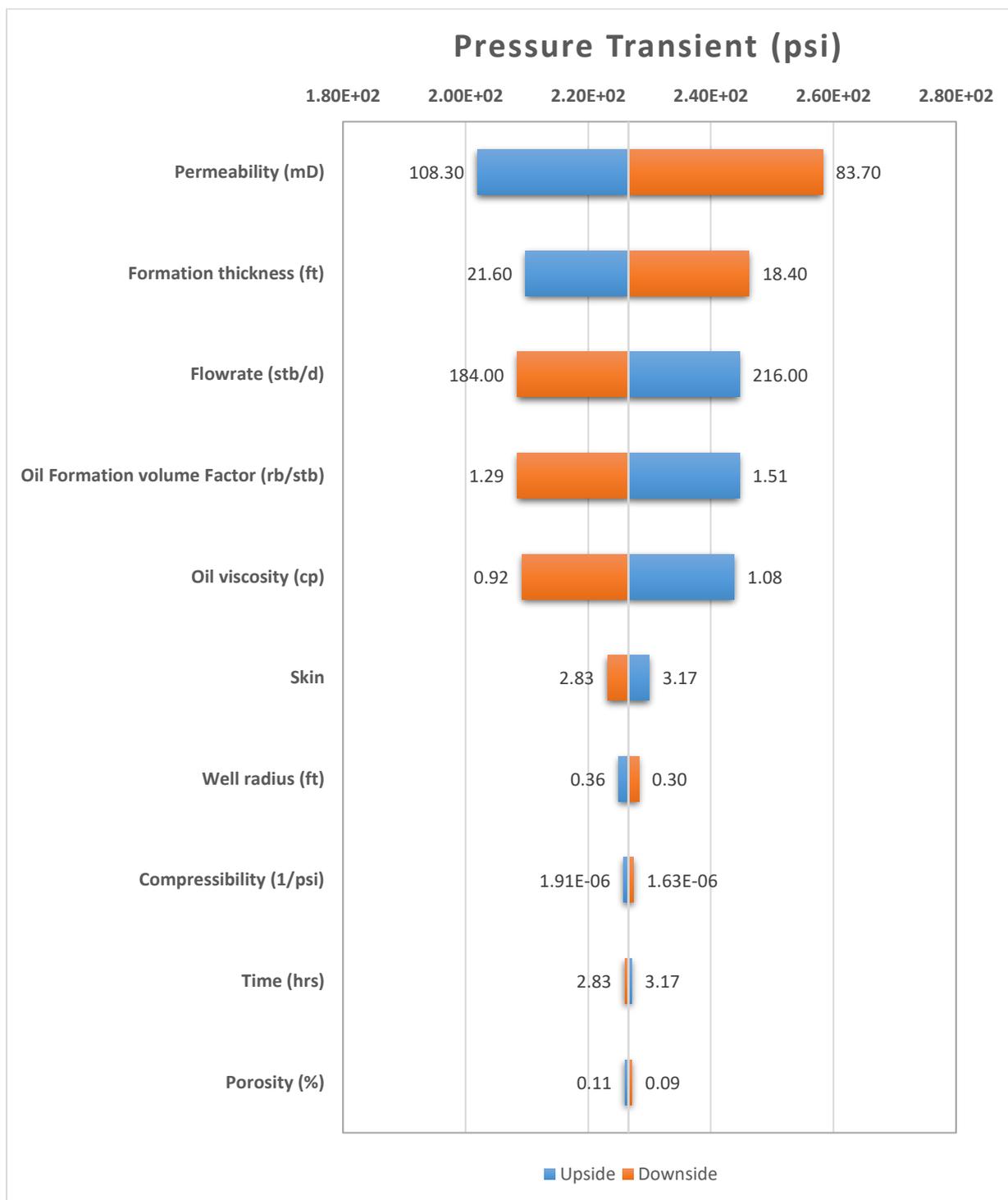


Figure 2. 18 Showing permeability as the most sensitive for pressure transient estimation with both P10 and P90 cases for Sherwood Sandstone Reservoir, North Sea.

2.8 Summary

Porosity, fluid saturation and permeability have been demonstrated both from the literature and analysis of reserve and pressure transients, to be the most important parameters of porous media. Monte Carlo simulations were run to carry out sensitivity analysis for all the parameters using Oracle Crystal Ball software. Optimistic, most likely, pessimistic and base case scenarios were also modelled. The simulation runs show porosity as the most sensitive parameter followed by water saturation in terms of reserve estimation while permeability was the most sensitive with respect to pressure transients (among several other parameters considered) hence we are considering them in this work. Oil and gas companies are interested in accurate reserve estimate which are rarely so thereby making projects less economically profitable.

Chapter Three

Methods, Concepts and Models

3.1 Introduction

An overview of the approach, theories and techniques used in this study to reduce the uncertainty in the determination of the chosen reservoir properties, are shown in this section. The research methodology involves data gathering of sandstone assets from about fifty (50) different wells in Northern, Central and Southern North Sea. The datasets include well logs like gamma ray, deep and shallow resistivity, neutron, density, caliper, sonic, photoelectric factor where available; core and well test data. They were digitized both manually and with the use of digitization software packages. The digitized data was then wrangled and outliers removed. Logs cross checked with caliper for diameter consistency to identify borehole rugosity while log headers were checked for depth datum. Lithology identification was done with gamma ray as well the photoelectric factor log (PEF) while fluids were identified using pressure profile from modular dynamic tester (MDT), PEF, neutron-density cross plot as well as from the resistivity logs.

Porosity, water saturation and permeability (and their variants) were determined from well logs using both machine learning and empirical models. Several machine learning models were tried using R software package and literature searched for the appropriate one to use. Training of the models was done with log data as the independent variable and cores as the dependent variable. In line with these, support vector regression models were used to determine porosity, an ensemble model was used for the saturation models, deep learning oil and water relative permeability models were also developed. Modeling with existing permeability data (using porosity and water saturation) with North Sea data found Equation 4.22 (Timur, 1968) to be the closest to the data although a modified Equation 4.26 (Coates, 1973) was formulated due to its flexibility. Well test buildup and drawdown analysis were done using Kappa software package for permeability determination.

Appropriate log, core and test depth matching was also ensured including upscaling combined with Lorenzo plot to identify the dominant flow interval and give an idea of the heterogeneity of the formation. The database of the log, core and well test permeabilities was then formed from these analysis as well as from the literature. A support vector regression model using well test permeability as the dependent variable and core together with log as independent variables. This is useful to predict well test permeability when not available due to time and cost especially in

marginal fields. A dashboard was developed for real time prediction using the machine learning model. A nomogram forming a relationship between the three sources of this important flow parameter was also formed. Important to say that more still has to be done in terms of data gathering to improve these models.

3.2 Data Gathering and Well Distribution

Well log suites which included Gamma Ray (GR), Sonic Log (Vp), Bulk Density (Rho_b), Density Porosity (DPHI), Neutron Porosity (NPHI), Caliper Log and Photoelectric Factor (PEF) were obtained from several wells in the North Sea chosen based on several criteria like (a) if their reservoir is siliclastic (b) have good quality data (c) wide coverage of the basin i.e. drawn from Northern, Central and Southern North Sea (d) they contain both core and well test information as well

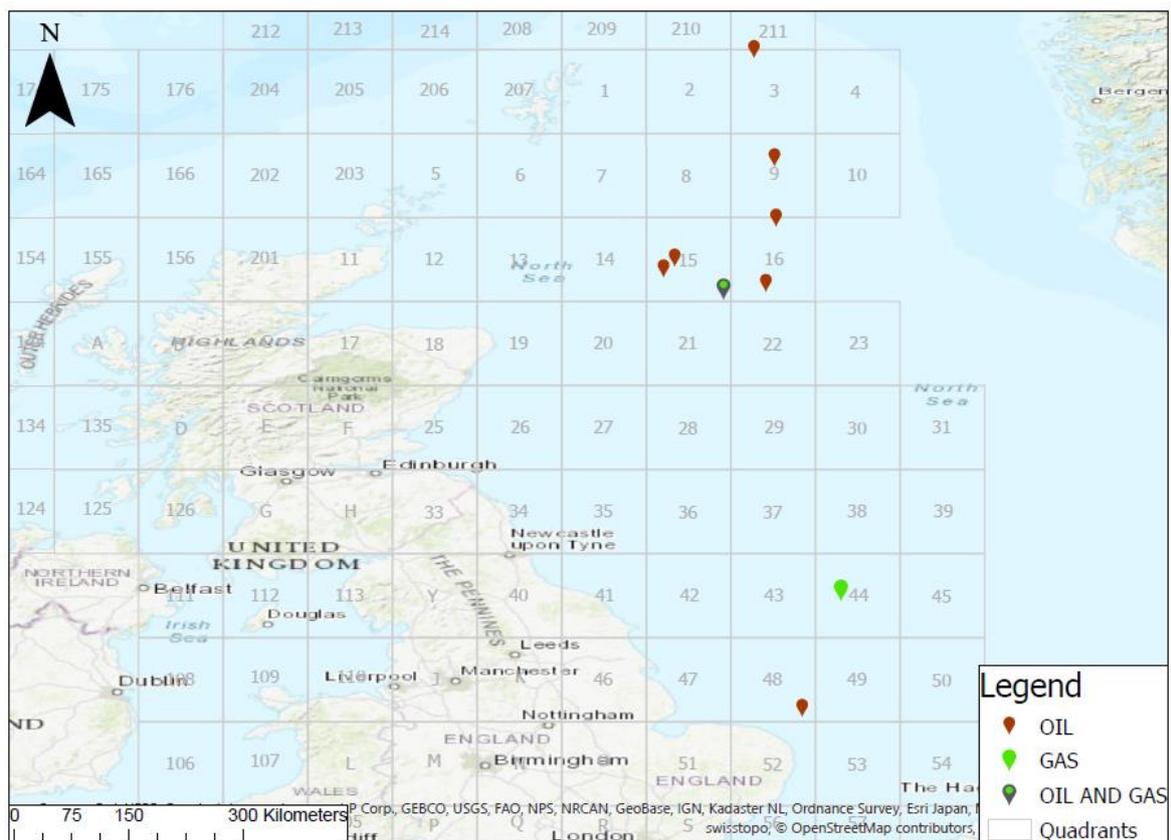


Figure 3. 1 Distribution of wells used for the different aspects of the study

3.3 Data Preparation

Logs were quality checked using caliper logs for wellbore diameter consistency hence rugose and compromised depths were identified. Caliper readings are from a scale of 5-15 with most normal areas lying between 6 and 7. Values greater than this draws attention to check more for rugosity. Of the fifty (50) wells considered, this outlying areas were encountered in reservoir sections of 5 wells. All log readings collected from that rugose section and were therefore removed. Log headers were checked to ensure quoted depths were relative to subsea. More verifications were made using information from their well files and composite logs. Since the work focusses on sand stone assets, top and bases of reservoir sands were delineated from well reports and composite logs as well as visual examination for consistency of general trend of velocity and density behaviours with respect to depth. Routine identification of different lithologies most often depends on gamma ray logs with readings typically higher in shales due to high concentration of potassium of clay minerals and other radioactive elements. These absorbed radioactive minerals can occur during condensation which is the preservation of relatively long geologic time by a relatively thin layer of sediments. Condensation can also form hot sands (highly radioactive sands) thus giving similar gamma ray readings as shales causing misinterpretation. To ensure that hot sands are not mistaken for shales, the photoelectric factor, the Kataharan-type plot, neutron-density logs in addition to the resistivity log was used.

3.3.1 Katahara sand/shale cutoff

Katahara (2008) makes use of the differential between neutron porosity and density porosity for shale discrimination from other lithologies such as sand since bulk density calculates rock porosity assuming water filled pore and grain densities (Figure 3.2). When plotted, these porosities overlap across sands filled with fluid but are separated across shales.

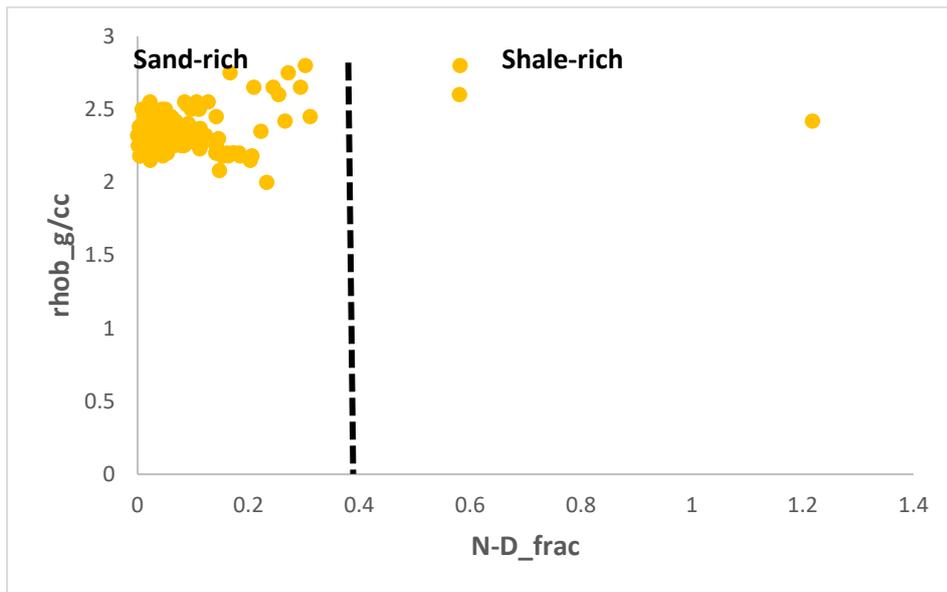


Figure 3. 2 Kataharan plot discriminating shale from sand rich facies.

3.4 Fluid Identification

The modular formation tester was used to ascertain the fluid type and contact in the reservoir where available. The hydrostatic pressure differential produced with depth. Fluid levels are obtained from open-hole pressure data and fluid contacts are gotten from core and logs to cross check the fluid level reading. In general, there is good agreement between the different methods to pick the gas/oil contact with a maximum difference of 2 ft between the log and pressure data possible due to depth control.

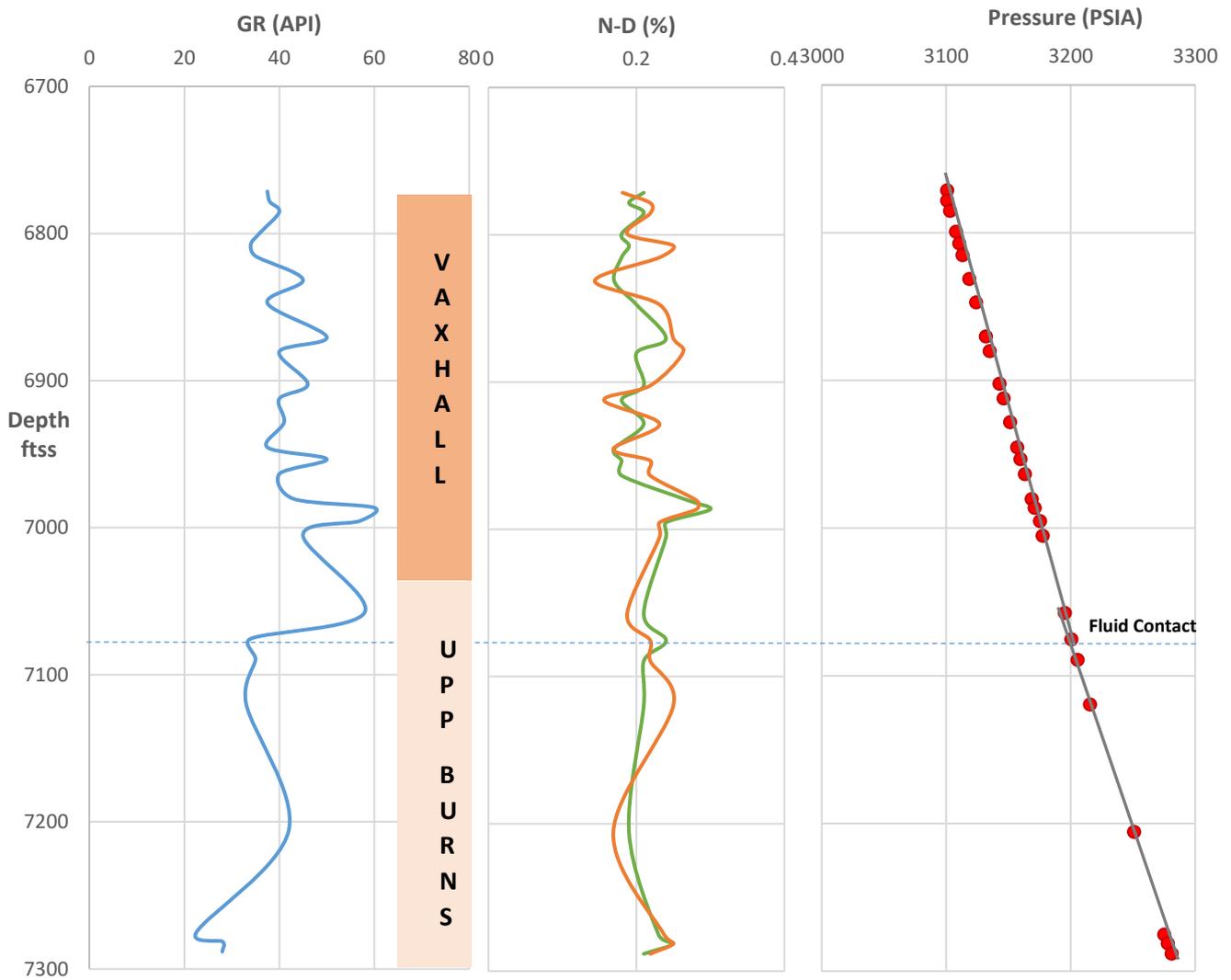


Figure 3. 3 Gamma ray, neutron-density and pressure data with depths in feet. The Intersection gives the fluid contact as shown by the broken line coinciding with the base of the coarsening upward sequence and the intersection of the neutron and density logs.

Fluid identification was also done using the shallow and deep resistivity logs in relation to the type of drilling mud used.

3.5 Uncertainty Recognition and Reduction in Log, Core and Well Test Analysis.

Schlumberger (2016) described uncertainty as the extent to which predicted values deviates from measured data. With respect to reservoir characterisation, it is reservoir modelers best estimate of how a modelled reservoir quantity might deviate from the true value of that quantity. Thus the challenge of the Oil and Gas industry is to accurately characterise reservoir parameters made difficult due to these uncertainties. Each analysis begins with demonstrating the degree of uncertainty inherent in the determination of the reservoir rock properties. An appreciation of the degree of uncertainty in porosity from density and sonic logs by just getting the reservoir fluid type right (hence MDT if available was used to verify this) or whether the formation is consolidated, semi-consolidated or unconsolidated. Well test data provides one of the few effective methods of direct reservoir analysis, and therefore it is important for petroleum engineers to overcome as much of the uncertainty as possible, and to understand in more detail the part which cannot be overcome. There are several sources of uncertainty in cores, log and well testing. The principal ones include:

- Physical error in the core, log and pressure data e.g. noise, drift, temperature effects and time shift, poor calibration, hysteresis and creep disturbs pressure sensors.
- Uncertainty due to the fact that processed data can be interpreted in many ways with both processing and interpretation requiring models themselves.
- Uncertainty from type of geological setting used which are inturn interpreted from models and data which are uncertain themselves.
- Errors in the flow rate information. Since flow rate measurements are commonly made at the surface and pressure measurements are made downhole, there is a significant disconnection between the knowledge on the impulse which caused it. These measurements, both surface and downhole, are very less accurate than pressure measurements and are treated with less care. Flowrate can be subject to liquid carryover, gas entrainment, poor metre calibration, foamy crudes as well as other operational problems that disrupts flowrate such as uncertainty from low rates been read below the flow meter reading.
- Ambiguity in the response. With a sufficient duration of test, most pressure transient information can provide an unambiguous indicator of the character of the reservoir. However many well tests cannot or are not run for sufficient duration to reach a point which excludes ambiguity. It is not uncommon therefore to encounter well test data that

can be matched to several different interpretation models with apparently equal resemblance.

- Parameter estimation ill-posedness. The use of nonlinear regression analysis to solving problems of well tests has revealed that the estimation of common reservoir parameters is difficult or even impossible in cases where the reservoir response is a weak function of the parameter values while strong correlations can imply that the estimated values are meaningless.
- Uncertainties in properties. The estimates of reservoir parameters such as permeability k depends on the knowledge of fluid and rock properties such as viscosity and compressibility C_t . In practice, these parameters are not as well-known as they ought to and therefore the estimates of reservoir parameters are affected.
- Test are carried above fluid saturation conditions to maintain single phase.
- Determination of the Middle Time Region (MTR) representing the infinite acting radial flow period based on which permeability is determined especially when there are several plateaus on the pressure derivative. The derived permeability from each plateau can be different by an order of magnitude.
- Trade-off between spatial and resolution uncertainty as well test gives larger coverage but lesser detail compared to logs and cores.

3.6 State-of-the-Art of Uncertainty Analysis Methods

3.6.1 Deterministic Approach

This works by using a single value of each input parameter into a suitable mathematical model to give exact solutions of the expected outcomes. It is often used at the start of a field development when understanding of the field is low. Washburne (1916) stated that the earliest efforts at analysing and managing uncertainties in oil reserve estimation were done by contouring the geological structure and basal water of the map. Afterwards came the Mapping and Estimation Method with experimentally obtained porosity. This was then followed by more advanced computing (Sheldon et al., 1960). These techniques had outcomes that are specifically determined through known relationships and expressed in three levels such as “Best Case, Most Likely and Worst Case”. A single model (Keijzer and Kortekas, 1990) or a few more models (Ringrose and Bentley, 2015) are built to describe some scenarios of the reservoir.

3.6.2 History Matching Technique

History matching could be referred to as a hybrid between deterministic and non-deterministic approaches. Conventional history matching could be considered as a separate approach but modern history is very much dependent on probability methods e.g. Ensemble Kalman Filter (EnKF) and stochastic methods. It is the procedure for comparing the reservoir simulation model with the observed data and adjust the uncertain parameters in the simulation model to reduce its mismatch with the observed data, is known as history matching. This procedure can be deterministic or probabilistic. The deterministic method entails obtaining one simulation model between many probable matches to the field data by running the initial simulation model with several input values. The probabilistic procedure considers different reservoir model scenarios thereby using the uncertainty analysis technique. Since the input space to be searched can be high dimensional, outputs collection to be matched may be very huge and each single valuation may take a long time thus identifying the input parameters within which the simulation match the observed data can be a difficult assignment.

3.6.3 Probabilistic /Stochastic Technique

This involves a stochastic (random) description of the specific parameters of the reservoir model through predefined probabilistic algorithms and thus able to model the likelihood scenarios. Monte Carlo simulation is an example of this method as shown in Figure 2.12-2.14 conveyed as percentiles of the cumulative probability distribution as P90, P50 and P10 indicating pessimistic, median and optimistic case respectively (Sanyal and Sarmiento, 2005). It was run on both STOIP and pressure transient models where their inputs were randomly drawn from their distribution to understand how the output changes (O'Hagan, 2006; Uusitalo et al., 2015, Ani et al., 2016). The observed change could be high or low depending on how sensitive the model output is to a parameter. Some workers favour changing a few parameter values (Nakayama, 2000), while others favour randomly picking likely values for all parameters (Kumar and Varghese, 2005; Baroni and Tarantola, 2014). Apart from MC, Experimental Design (ED) is another example of this method (Lee et al., 2006; Lawal, 2009).

Caers (2011) stated that deterministic and statistical/probabilistic numerical models can only not fully capture true variability since they do not consider the spatial function. This prompted the application of geostatistics to allow the modelling of reservoir property distribution to capture spatial variability and uncertainty.

3.6.4 Geostatistical Approach

These examine the data distribution, trends, directional components as well as outliers of geological parameters across the reservoir (Holden et al., 1988; Caers, 2011; Ringross and Bentley, 2015). They use variograms which describes the degree of spatial dependence between sample values on separation distance (lag). The semivariogram depicts the spatial autocorrelation of measured sample points. Although semivariograms are sometimes used interchangeably as variograms, they should be avoided (Bachmaier and Backes, 2011). The difference between the two is that the latter uses each pair of data elements just once while all possible data pair are used by the former. Thirdly, they could use a technique known as kriging which interpolates between sample values considering their separation distance as well as determine from known values in a nearby grid, the value of a point in heterogenous grid. As stated by Yarus and Chambers (2006), inappropriate use of variables and poor technique implementation will never be compensated for by geostatistical analysis which are also not automatic. Machine learning and artificial intelligence is therefore needed to fill this gap.

3.6.5 Machine Learning Algorithms

Due to the ever increasing need to accurately model reservoir parameters and constant development of better techniques for reservoir characterization and uncertainty analysis, artificial intelligence (AI) has become the tool of choice evident in self driving cars and several automation systems. Machine learning is an application of Artificial Intelligence based upon giving machines data to learn from and then make predictions or decisions. The trend of soft computing in reservoir and uncertainty modelling will continue to rise (Wong, Aminzadeh and Nikraves, 2013). Machine learning can be divided into supervised, unsupervised and reinforcement learning (Figure 3.2).

3.6.5.1 Reinforced Learning

Reinforcement algorithms is a type of machine learning algorithm where intelligent systems receives a delayed reward in the next time step in other to evaluate its previous action i.e. it learns from past experience and then makes future decisions. It does not need correct input/output pairs and sub-optimal actions do not need to be corrected explicitly. Its focus is on finding a balance between exploration of uncharted territory and exploitation of current knowledge (Kaelbling et al., 1996). Examples include Markov decision process, Q-Learning, State-Action-Reward-State-Action (SARSA), Deep Q Network (DQN), Deep Deterministic Policy Gradient (DDPG), Monte Carlo Algorithms.

3.6.5.2 Unsupervised Learning

Unsupervised learning does not have a target variable and can be grouped into clustering and association problems. The goal is to learn more about the data by modelling the structure or distribution underlying the data (Hinton et al., 1999). Unsupervised learning infers an apriori probability distribution (Hastie et al., 2009). Examples include apriori algorithm for association rule learning problems and k-means clustering etc.

3.6.5.3 Supervised Learning

Supervised learning algorithms maps inputs to desired outputs in order to efficiently approximate the mapping function when there is new data e.g. regression, decision tree, random forest, nearest neighbour, logistic regression etc. A major issue with supervised learning being that a set of solved examples are given to the algorithm (including all possible results) so that it devises a way of determining the outcome for problems with unknown outcomes. Other concerns are the dimensionality of the input space which can confuse the algorithm causing it to have high variance, bias vs variance tradeoff, function complexity and amount of training data, effect of noise in the output values (Geman et al., 1992; James, 2003; Brodely and Friedl, 1999; Smith and Martinez, 2011). Examples include Support Vector Machines (SVM), Neural Networks, Random Forest, Gradient Boosting Model (GBM).

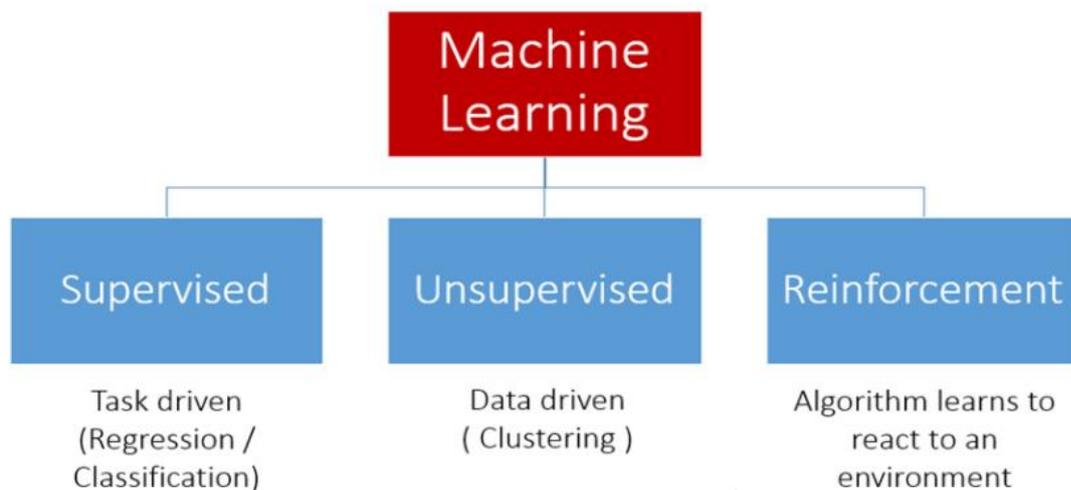


Figure 3. 4 Different groups of machine learning algorithms

This work utilized Supervised Learning algorithms such as Deep Learning for modelling complex multiphase processes like relative permeability due to its ability to perform better with more data where other algorithms flatten out in performance (Figure 4.34). Support Vector Machine for both porosity and uncertainty reduction across log, core and well test permeability due to its ability to perform well with irregular and smaller number of data mainly due to its kernel function and other hyperparameters. Other algorithms like Random Forest, Neural Networks were also tried. Stacked Ensemble of Neural Networks and GBM models for fluid saturation modelling since they gave best predictions. Other models tried include Random Forest, Generalized Linear Model (GLM) but were not used in the Stacked Ensemble. Their results were compared with both existing and developed empirical models.

3.6.5.3.1 Neural Networks

Artificial neural networks are an artificial intelligence technique that tries to imitate the behaviour of the human brain and nervous system. They are well suited for modelling systems with complex relationships between input and output which is what is obtainable in natural earth systems. In such cases with no prior knowledge of the nature of non-linearity, traditional regression analysis is not adequate (Gardner and Dowlings, 1998). Its typical structure consist of a number of processing elements (nodes) arranged in layers (input, hidden and output). The input signal from each node (x_i) at the input of synapse i connected to neuron j is multiplied by an adjustable connection weight (w_{ij}). The weighted input signals at each node are summed and a threshold value is added and the combined input is passed through a non-linear transfer function (Equation 1) to produce the outputs (y_i) which then provides the input for the next neuron. An externally applied b_k which increases or decreases the net input (depending on whether it is positive or negative) of the activation function is also part of the architecture. A neural work with four (4) hidden layers and sixteen (16) neurons were used for training the relative permeability models used for comparison with the deep learning model in Chapter Four.

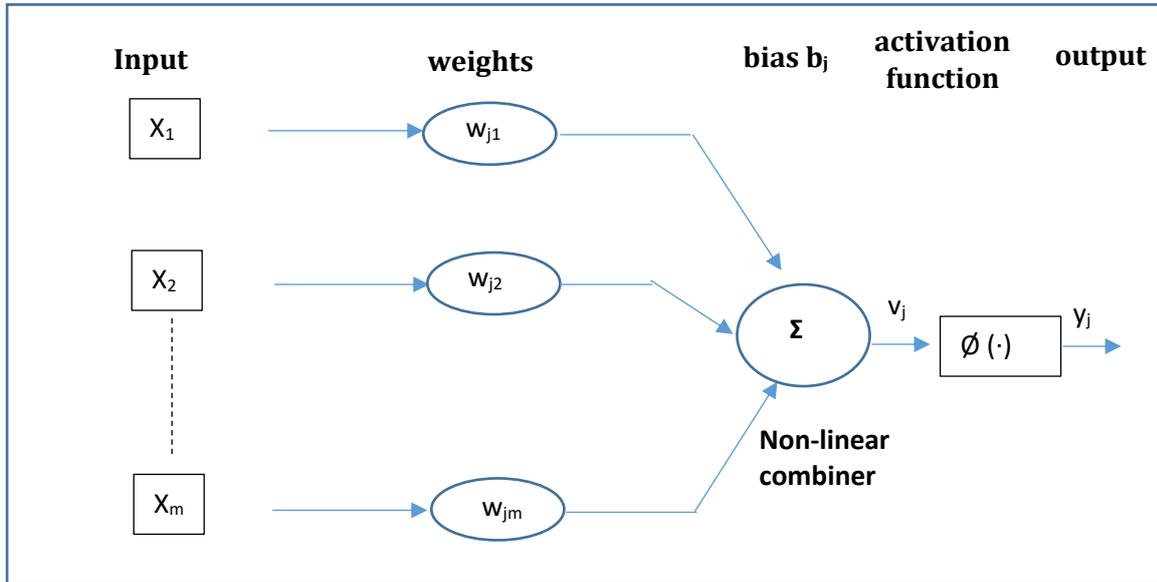


Figure 3. 5 Neural network model elements

Mathematically, the neuron (j) function is given by:

$$y_j = \emptyset(u_j + b_j) \quad (3.1)$$

Where

$$u_j = \sum_{i=1}^m w_{ji}x_i \quad (3.2)$$

x_i is the input signal from an m dimensional input, w_{ji} is the synaptic weights of the neuron j , u_j is the linear combiner output as a result of the input signal, b_j is the bias, $\emptyset(\cdot)$ is the activation function and y_j is the output signal of the neuron.

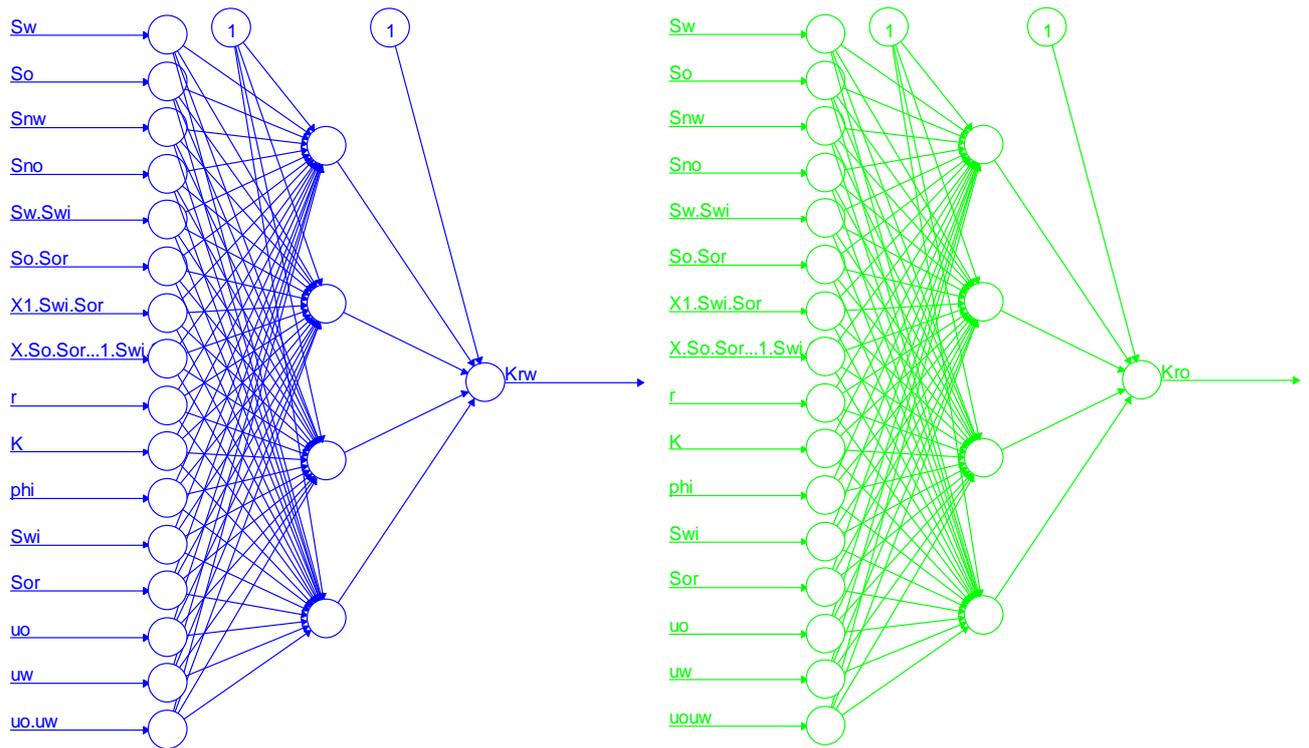


Figure 3.6 Schematic for the final neural network used for the prediction of water (blue) and oil (green) relative permeabilities showing sixteen (16) neurons and four (4) hidden layers. Same input variables were used for both.

3.6.5.3.2 Deep Learning

Deep learning is neural networks with a large number of hidden layers. It seeks to exploit the unknown structure in the input distribution so as to discover good representations at different levels of abstraction with higher level learnt features defined in terms of lower level features thus allowing the system to learn complex functions mapping of the input directly to the output from data without depending totally on human-crafted features. It provides an effective way to initialize the weights which allows the network to learn these low-dimensional codes which works much better than principal component analysis for reduce data dimensionality. Deep Learning was applied for real time prediction of relative permeability since it has the scalability advantage of being able to continuously perform better as more data is obtained compared to other algorithms that flatten out in performance at some point. Ten (10) sets of water-oil relative permeability data with 132 data points from a NorthSea field with four-fifths used as training set and one-fifth as validation set. Another set of water-oil relative permeability data from a separate field were used as the testing set after data wrangling and normalization. There were also 12 input parameters fed into the network, having one hundred hidden layers with twelve neurons.

3.6.5.3.3 Support Vector Machines

This is a supervised machine learning technique used for both regression and classification. Due to their ability to generalize for unseen data, they are increasingly replacing other algorithms as the tool of choice for pattern recognition and prediction. The main idea is to minimize error, individualizing the hyperplane which maximizes the margin of tolerance, keeping in mind that part of the error is tolerated. The kernel functions it uses transforms the data into a higher dimensional feature space to make it possible to perform the linear separation that it does.

$$H_{IJ} = y_i y_j k(x_i, x_j) = x_i \cdot x_j = x_i^T x_j \quad (3.3)$$

$k(x_i, x_j) = x_i^T x_j$ which is an example of a kernel function and are based on calculating inner products of two vectors. That is, if the functions can be recast into a higher dimensionality space by some potentially non-linearly feature mapping function $x \mapsto \phi(x)$ only inner products of the mapped inputs in the feature space need be determined without us needing to explicitly calculate ϕ .

For regression the method includes:

Selection of how meaningfully misclassifications should be treated as well as how large the insensitive loss region should be by selecting suitable value for the parameter ϵ (epsilon). The linear insensitive loss function ϵ , ignores errors that are within its distance by treating them as equal to zero. The loss is calculated based on the distance between observed y values and the ϵ boundary.

Find α^+ and α^- so that:

$$\sum_{i=1}^L (\alpha_i^+ - \alpha_j^-) t_i - \epsilon \sum_{i=1}^L (\alpha_i^+ - \alpha_j^-) - \frac{1}{2} \sum_{i,j} (\alpha_i^+ - \alpha_j^-) (\alpha_i^+ - \alpha_j^-) \phi(x_i) \cdot \phi(x_j) \quad (3.4)$$

Is maximized.

Where α^+ and α^- are lagrange multipliers.

Calculate

$$w = \sum_{i=1}^L (\alpha_i^+ - \alpha_j^-) \phi(x_i) \quad (3.5)$$

Determine the set of support vectors S by getting the indices i where $0 < \alpha < C$ and $\xi_i = 0$. All observations inside the epsilon tube have lagrange multipliers $\alpha_i^+ = 0$ and $\alpha_j^- = 0$. If either α_i^+ or α_j^- is not zero, then the corresponding observation is called a support vector. The property, α , of a trained SVM model stores $(\alpha_i^+ - \alpha_j^-)$ which is the difference between two lagrange multipliers. b and x_m are stored in the properties Bias and Support Vectors respectively.

The calculate

$$b = \frac{1}{N_s} \sum_{s \in S} [t_i - \epsilon - \sum_{m=1}^L [(\alpha_i^+ - \alpha_j^-) \phi(x_i) \cdot \phi(x_m)]] \quad (3.6)$$

Each new point is evaluated by

$$y' = \sum_{i=1}^L (\alpha_i^+ - \alpha_j^-) \phi(x_i) \cdot \phi(x') + b \quad (3.7)$$

Implementing a SVM comes down to selecting the variable w and b that the training data could be described by, α is the lagrange multiplier. Summarily, the goal is to find a function y that deviates from y_n by a value no greater than

This algorithm was used to predict both porosity in Chapter Four and for the uncertainty reduction across log, core and well test discussed in Chapter Six. The former SVM porosity model used a dataset comprising 446 data points from Northern, Central and Southern North Sea sandstone reservoirs comprising Unconsolidated, Semi-consolidated and Consolidated sandstones, 435 training sets, 11 validation sets and separate data for testing while the latter SVM model used a total of 114 datasets which was randomly divided into 105 training set and 9 validation sets while a separate data form another field was used for testing its ability to generalise.

3.6.5.3.4 Ensemble Models

Several base models can be combined to form a better meta-model. This involved placing different machine learning models over one another and their predictions are passed to the top layer which then makes decisions based on these base predictions. Gradient Boosting model and Neural

Networks models were built separately after which they were combined to form a stacked Ensemble which gave a better prediction. This algorithms was used to develop the total water saturation, irreducible water saturation and the residual fluid saturation models with great accuracy. The dataset for the Stacked Ensemble water saturation model was randomly divided using stratified technique into 425 training sets, 24 validation set and 22 test set. The irreducible water saturation had 290 datasets randomly divided into 243 training set, 16 validation set and 31 testing set. The residual oil saturation model used 243 datasets again randomly divided into 195 training set, 16 validation set and 32 test sets. A second level of deep learning algorithm (metalearner) was then trained to learn optimal combinations on the level-one models.

Chapter Four

4.1 Uncertainty Reduction in Porosity Usage

4.1.1 Introduction

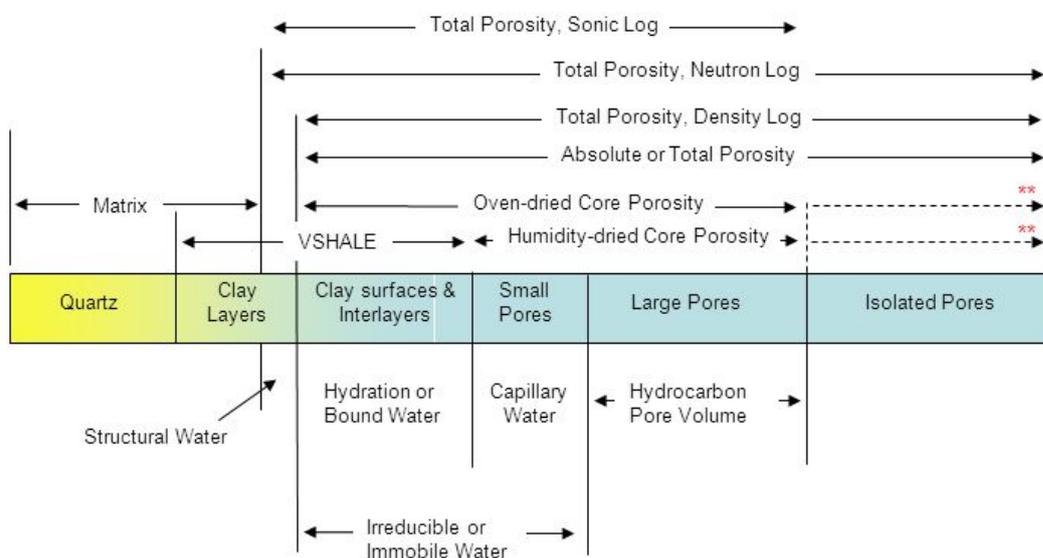
One of the most essential rock properties for reservoir evaluation is porosity. Despite its importance, its determination has lots of uncertainty. The density log has been regarded by several authors as the best source for porosity from logs but it gives different results if different fluid types are assumed. Same goes for Raymer and Wyllie if different lithology and compaction levels are uncertain or assumed. While cores on the other hand may not be available due to unconsolidation or secondary porosity such as vugs or natural fractures hence porosity will have to be determined from well logs. These logs do not measure porosity directly but have a greater radius of investigation than cores. Haldorsen (1986) presented that the volume of rock investigated by a sonic log is 162 times the volume of a one inch diameter core plug especially for large spacing between source and receivers getting undamaged measurements beyond the borehole wall.

Therefore log data if appropriately used could provide a more accurate prediction of porosity. Porosity from different data sources should not give different results and although their discrepancy especially for neutron and density, is useful for the identification of fluid type and contact, the goal of obtaining accurate measurement of porosity should not also be forgotten. The differences in these sources of porosity must be reconciled and accounted for. Rightly so since by design, neutron and density tools measure formation largely filled with mud filtrate and having residual hydrocarbon saturation as it assumes that most of the hydrocarbon in the sensitive volume measured by these tools are moved by invasion in low to moderate formations (Gaynard and Poupon, 1968; Segesman and Liu, 1971; Schlumberger, 1972). In practical terms, the uncertainty from neutron, sonic, density and even core measurements produce a scatter observed in the crossplots between them (Figure 2.11). It seems right to say that the problem of accurately estimating reserve (which is an industry problem) is very sensitive to the uncertainty in porosity as shown in Figure 2.13.

The differences in core and log porosity in general must also be taken into consideration (Figure 4.1) since the former is thought to provide effective porosity values depending on pore types present, clay content and method of cleaning and drying. This is so because helium porosity from cores (used in this work) should exclude unconnected pores and clay-bound water. It should also be remembered that core conditions are different from that of the reservoir. On the otherhand,

total porosity (which includes microporosity, fracture porosity and bound water associated with clays) is obtained from neutron, density although isolated pores are excluded in sonic total porosity. Needless to say that caved and irregular holes present an interesting problem in density log interpretation thus might not give actual porosity reading. Hence the interpretation of averaging the neutron and the density porosity with respect to sonic transit time may lead to wrong conclusions. (James, Jordan and Campbell, 1986). Spears (2006) also presented evidence that this averaging for gas bearing zones at Erha field in the Niger Delta were way different from the available conventional core porosity range. Practically, uncertainties in neutron, sonic, density and core porosity produce scatter observed in the cross plots between them (Figure 2.11).

The Chapter compares results of existing porosity models, new empirical model and machine learning specifically Support Vector Machine Regression (SVM) using data from Northern, Central and Southern North Sea. Log data such as neutron, sonic, and density were explored using core porosity, saturation or permeability (as the case may be) as the dependent variables to build the model. Their predictive power in unconsolidated, semi-consolidated and consolidated sandstones was investigated. It further went on to discuss fluid saturation and consequently permeability. There is need to reduce the uncertainty and confusion in the prediction of these parameters. This becomes crucial since they are sensitive to one another and to other parameters.



** If sample is completely disaggregated

Figure 4. 1 Log and Core porosity measurements (after Eslinger and Pevear, 1988)

Table 4. 1 Typical ranges in porosity of common rock types (modified after Freeze and Cherry, 1979; McWaorter and Sunada, 1977)

Lithology	Porosity %	Mean Porosity
Sandstone	5.0 – 50.0	25.0
Shale	10.0 – 30.0	20.0
Clay	34.0 – 51.0	42.0
Karst Limestone	7.0 – 56.0	30.0
Dolomite	0.0 – 20.0	10.0
Siltstone	21.0 – 41.0	35.0
Weathered granite	34.0 – 57.0	45.0
Weathered gabbro	42.0 – 45.0	43.0
Gravel	24.0 – 36.0	28.0
Fractured basalt	5.0 – 50.0	25.5
Schist	4.0 – 49.0	38.0

4.1.2 Methodology

The existing porosity models (Equations 2.1-2.4) are better for some sandstone types and not so great for others. Also knowing for sure the type of fluid can be a challenge as different fluids gives different porosity results. This inconsistency creates uncertainty. Three different approaches were developed for porosity modelling of the North Sea (a) Improving on Raymer-Hunt-Gardner (RHG) and Wyllie Time Average (WTA) models (referred to as Raymer and Wyllie in this work for convenience) to obtain their local and regional mean using sonic log (b) Develop a completely empirical but modified Schlumberger density model to reduce the uncertainty caused by fluid density (c) Machine learning technique using Support Vector Regression allowing the model to learn directly from a combination of these well logs. The different porosity model results were then analysed and compared. In addition to these, to further assess the accuracy of these models, they were applied to three sandstone datasets i.e unconsolidated, semi-consolidated and consolidated types. Core porosity was used as dependent variable hence their depths were ofcourse accurately tied to log depths (Figure 4.2). Important to note that the purpose of (a) is to get regional and local constants for RHG and WTA for the NorthSea, (b) is to empirically modify the Schlumberger density porosity model such that it is not affected by the uncertainty of the fluid type and (c) which is actually the main focus, is a machine learning model that can accurately predict for all sandstone types.

The deviation effects of the different logs were examined by comparing their models with actual data. The default coefficients of both Raymer and Wyllie were perturbed to obtain unique values both locally and regionally thus preserving the original form of these equations as well improving their accuracy. The models predict porosity in the mesoscale and especially the machine learning model can be used in all sandstone types whether consolidated, semi-consolidated or

unconsolidated. It can also be used where core porosity could not be obtained and when the fluid type whether water, oil or gas is uncertain.

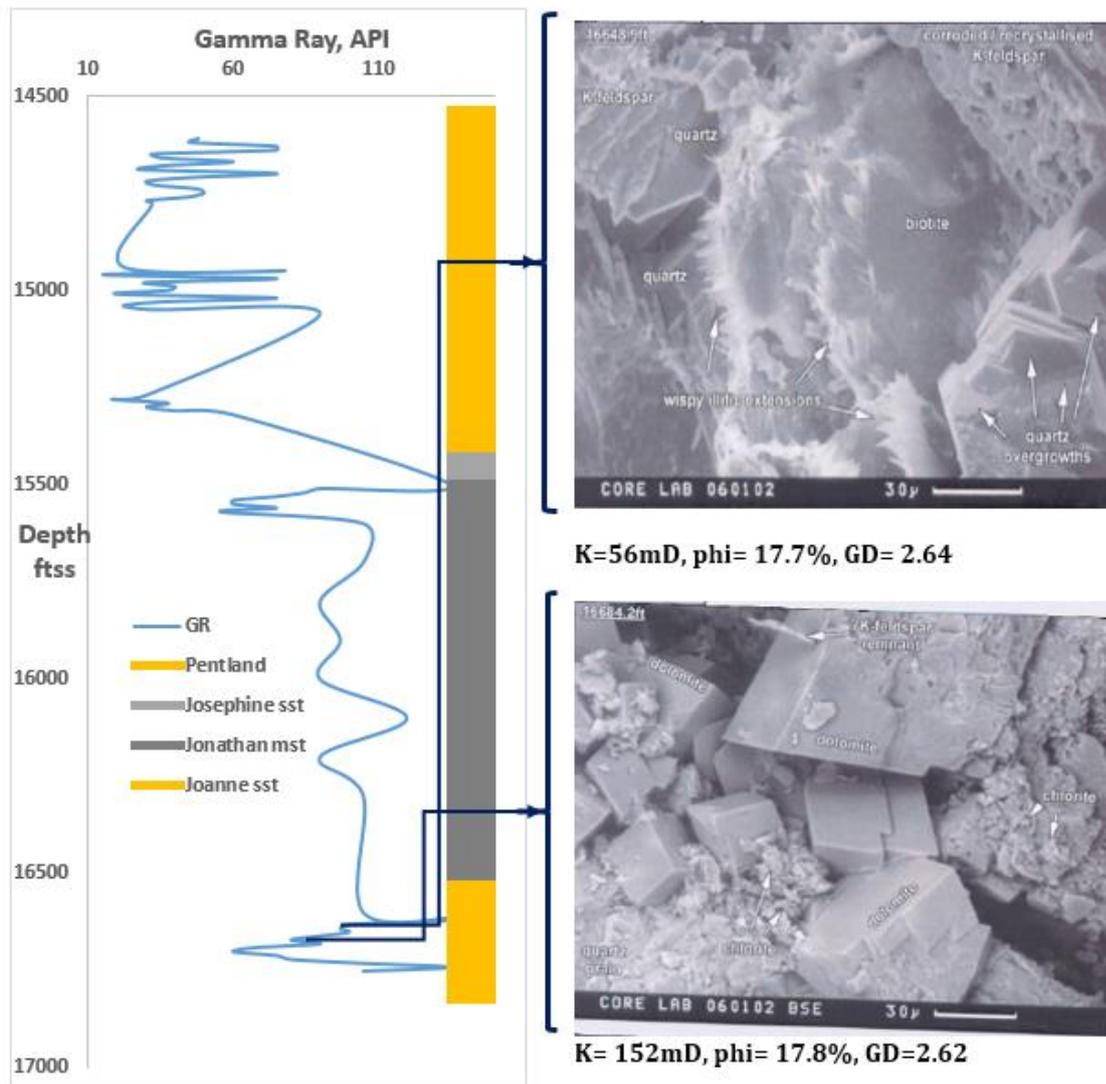


Figure 4. 2 Gamma ray showing lithology with photomicrograph of core taken from Joanne sandstone Formation indicating permeability, porosity and grain density.

4.1.2.1 Uncertainty in Porosity Determination

As shown in Figure 2.11, there is discrepancy in porosity values from the different sources. Different density porosity values are also obtained if water, oil or gas are assumed (Figure 4.3). Same goes for sonic log as Raymer gives different results if consolidated or unconsolidated formation is assumed (Figure 4.4). Therefore for porosity models to be accurate, the fluid type (whether water, oil or gas) and the degree of consolidation (loose or indurated) of the formation should be known to reduce uncertainty.

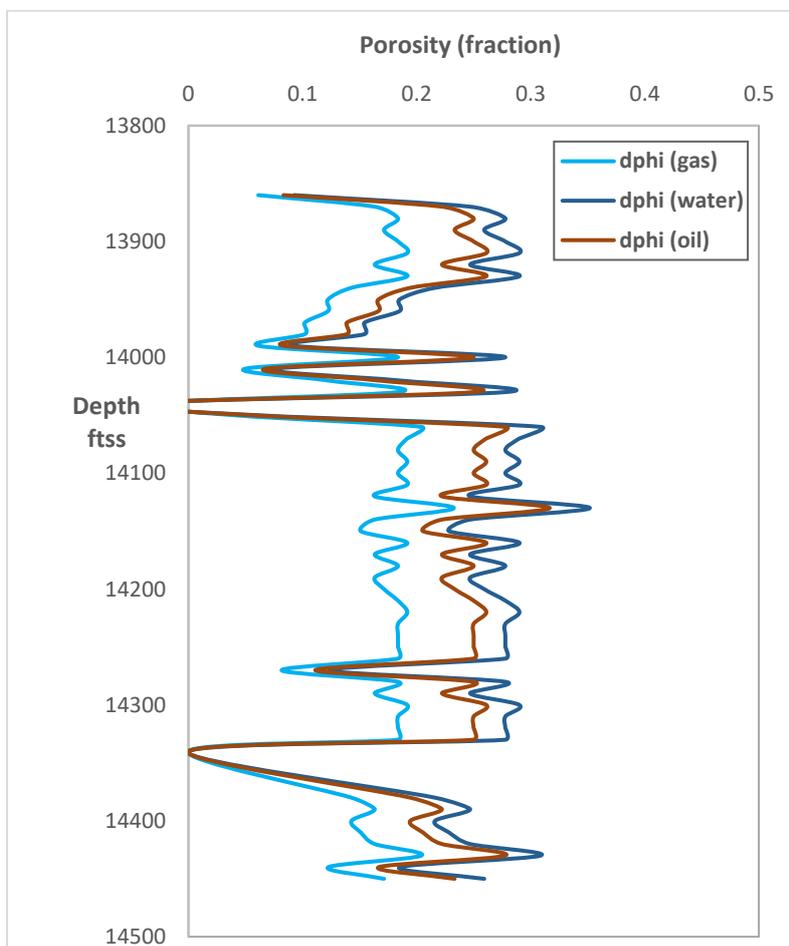


Figure 4. 3 Fluid type uncertainty in porosity determination from density logs (dphi) from Equation 2.4 (Schlumberger 1974). The effect of the fluid type uncertainty for oil, water or gas is obvious in porosity.

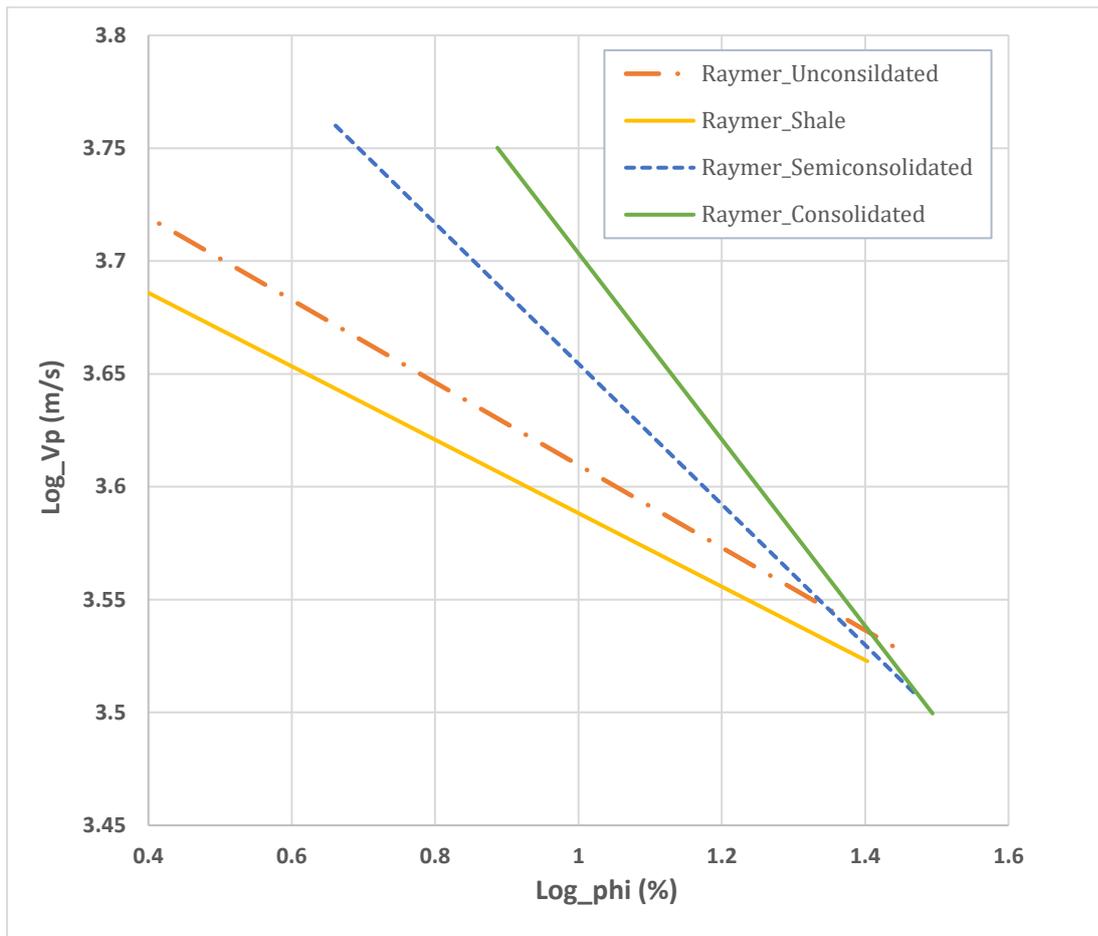


Figure 4. 4 Lithology uncertainty on Raymer-Hunt-Gardner porosity from Equations 2.2 & 2.3 using sonic log. Different results obtained if consolidated, semi—consolidated, unconsolidated or even a different rock type (shale) was considered. If the level of induration is not known for sure, porosity calculation from these models will lead to error.

4.1.2.2 Support Vector Regression Modelling (SVR)

This is the machine learning technique (mentioned in Section 4.1.2 above) for modelling porosity in addition to the modified Raymer, Wyllie and Schlumberger models. The data was first cleaned to ensure no missing numbers are in any row or column. If such were found they were removed since just a few rows (< 4 rows) had this issue. A median transformation would have been applied if such points were many. The relationship between the independent (nphi, rhob and sonic) and dependent variables gave very visual trends that was used for input data selection (Figure 4.5). Thus bulk density (rhob) was not used for the prediction. A non-probabilistic support vector regression model was developed to learn the non-linear function using a linear learning machine mapping from the input space into a high dimensional feature space. More explanation of this technique has already been given in Section 3.6.5.3.2. The dimensionality of the feature space

does not influence the parameters controlling the capacity of the system. The dataset comprising 446 data points from Northern, Central and Southern North Sea sandstone reservoirs comprising unconsolidated, semi-consolidated and consolidated sandstones, was randomly divided into 435 training sets and 11 validation sets. Separate datasets of Unconsolidated, Semi-consolidated and Consolidated sandstones were used for testing the developed SVM model. Neutron porosity (nphi) and sonic data were used as the independent (input) variables while core porosity data was used as the dependent variable (as this is a supervised learning technique). The radial basis kernel function was used for the training and prediction using a gamma value of 0.2 (a parameter needed by kernels except linear kernel), with a constant of the regularization term in the Lagrangian formulation (cost) of 100 and an epsilon in the insensitive loss function set as 0.1. As a reminder, gamma is an hyperparameter that controls the trade-off between error due to variance and bias in the SVR model. The small gamma value was used to try to avoid overfitting as possible. Also, the cost is a hyperparameter that controls how much we penalize our slacked variables and if we need more penalty, we just increase the value of the cost. Slacked variables allow us to relax the constraints. The radial basis was also used as the preferred kernel function due to its applicability in non-linear modelling. The tuning function was then applied for the SVM (grid search) using 10-fold cross validation. Values of epsilon between 0 and 1 using steps of 0.01 (i.e. 101 values of epsilon) and cost function with exponential steps of 2. Summarily 101 values of epsilon and 8 values of cost function, thus 808 models were tested to get the best. After tuning, the best epsilon and cost were 0.15 and 128 respectively (Figure 4.6). This model was used in

other chapters to compute porosity especially where needed but not present.

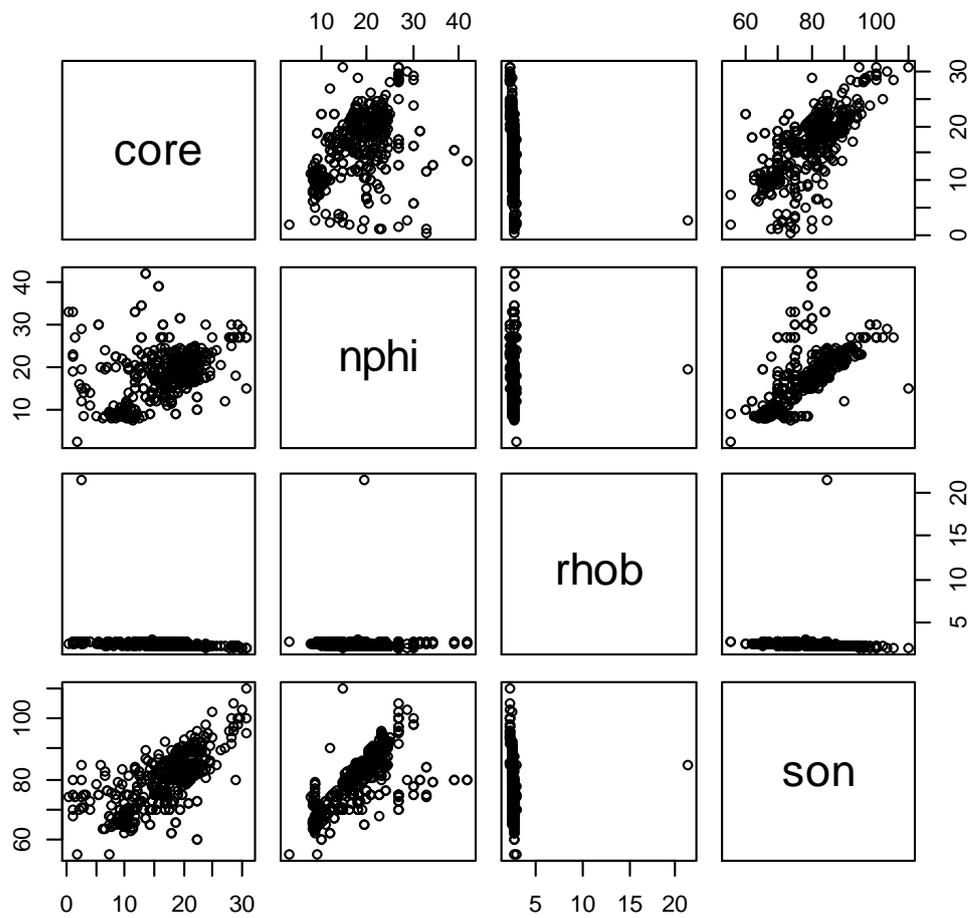


Figure 4. 5 Feature engineering showing the relationship between the different data types such as nphi (neutron porosity), rhob (bulk density), son (sonic) considered and to the independent variable (core) at a glance.

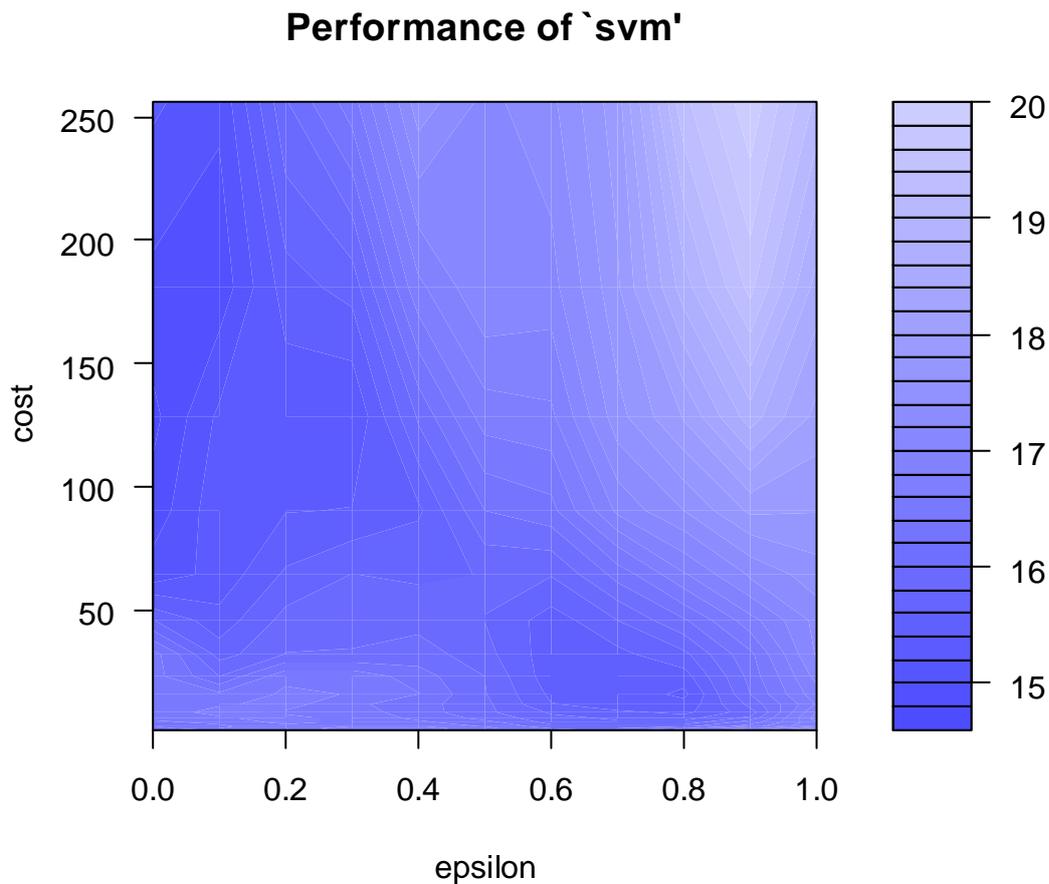


Figure 4. 6 Porosity SVR model performance

Colour coding was used to show the performance of the various models. Darker areas imply better accuracy indicating that the best values for epsilon lie between 0 and 0.2. Although one can further zoom in on this interval and retune using lower steps but had to stop to avoid overfitting. The legend to the right shows the Mean Squared Error (MSE) of the prediction. Root mean squared error was therefore reduced from 0.1481 to 0.0541 (equivalent to RMSE of 0.3848 - 0.2326).

Other machine learning algorithms tried include neural networks with RMSE value of 0.3121, Gradient Boosting Model had RMSE of 0.2993, Random Forest gave 0.3235 and Stacked Ensemble with best RMSE value of 0.2598. These are good RMSE values but the Support Vector Models gave the best value of 0.2326 hence it was taken forward as the machine learning model of choice in the instance to model porosity.

Table 4.2 shows the different combinations of the training set used. After 97%, the root mean squared error began to increase at 98% hence 97% was used for the analysis. This combination was used for the unconsolidated, semi-consolidated and consolidated sandstones.

Table 4.2: Different combinations of training data sets showing the best performance at 97%.

S/N	Proportion of Training Set (%)	RMSE
1	60	2.5631
2	70	0.9453
3	80	0.6001
4	90	0.4222
5	95	0.3879
6	97	0.2326
7	98	0.2594

4.1.3 Results and Discussion

4.1.3.1 North Sea Porosity Modelling

In addition to developing local and regional coefficients for Raymer and Wyllie, a modified Schlumberger model (Equation 4.7) for density porosity was formulated to deal with the fluid uncertainty as described in section 4.1.2.1. The developed model is completely empirical emanating from data. The comparison was made with Schlumberger model as shown in Figure 4.7. Equation (4.7) which was empirically fitted with data was modified from Schlumberger Porosity model.

$$\phi_D = d * \left(\frac{\rho_{\log} - \rho_{ma}}{\rho_{\log}} \right) \quad (4.7)$$

Where $d = 1.22$ for the North Sea wells considered as shown in Table 4.3.

Table 4. 3 Distribution of best fit (BF) coefficients and their associated errors for Default Raymer (DR), Default Wyllie (DW) and Modified Schlumberger (MS) models (Equation 4.7)

Field	Raymer		Wyllie			Mod Schlum			
	Best fit	NRMSE	NRMSE	Best fit	NRMSE	NRMSE	Best fit	NRMSE	NRMSE
	BF	DR		BF	DW		BF	MS	
1	0.5	2.89	3.97	0.83	2.76	3.9	1.2	0.68	1.36
2	0.64	7.62	7.68	1.2	8.44	9.84	1.24	1.1	3.2
3	0.66	0.38	0.42	0.85	0.73	0.85	1.25	0.45	0.49
4	0.55	0.97	1.47	0.95	1.63	1.67	1.33	1.36	1.41
5	0.46	4.76	7.82	0.82	5.28	6.5	1	8.67	12.48
6	0.61	0.55	0.67	0.9	0.96	1.5	1.2	0.4	2.11
7	0.6	0.32	0.33	0.98	0.31	0.32	1.2	0.26	2.08
8	0.55	0.7	1.71	0.8	0.78	2.77	1.25	1.65	2.76
9	0.4	4.64	18.33	0.65	4.34	18.92	1.25	3.79	5.89
10	0.5	3	5.55	0.8	3.12	5.49	1.28	3.43	4.89
Regional Mean	0.55	2.58	4.80	0.88	2.84	5.18	1.22	2.18	3.67

The error for default coefficients and local best fit for the different wells are summarized in Table 4.3. Local mean is the average porosity obtained from the individual fields numbered 1-10 of the same Table. Regional mean values were then obtained from the average of all the local means considered. These results were obtained from the default and locally derived values in the well. There is considerable variation in the porosities estimated for Raymer, Wyllie and modified Schlumberger model using their default values. RHG (Equation 2.2 & 2.3), WTA (Equation 2.1) and the Modified Schlumberger (Equation 4.7) have coefficients, hence coefficient was used along side their root mean squared error. For Raymer, the best coefficients ranges from 0.4 up to 0.66 while for Wyllie, they range from 0.8 to 1.15. The coefficients of the modified Schlumberger model, range from 0.65 to 1.28. The calculated regional mean coefficient values of 0.55 for Raymer instead of default of 0.625, 0.88 for Wyllie instead default of 1 and 1.22 for Modified Schlumberger Density Model instead of a default of 1 (Equations 2.1-2.4). Figure 4.7 is the graphical representation of Table 4.3 where the ten fields are shown in the abscissa. The errors generated while using the best fit and default coefficients for Raymer, Wyllie and Schlumberger are shown in the ordinate.

Figure 4.8 shows an improved accuracy (determined from the range of errors associated with the local values) of 1-74% for Raymer, a lower value means that not much difference between the default and the best fit. For Wyllie, the improved efficiency ranged from about 4-77% and about 3-88% for the modified Schlumberger model. In Field 9 for example, the normalised root mean squared error for default Raymer was 18.33 but this error dropped to 4.64 when calibrated locally.

It is clear from Table 4.3 that the default Raymer and Wyllie were quite different from their derived forms as their fitting coefficients decreases or increases with respect to the default values. The data specific coefficients reduced the uncertainty in these estimates. Regionally, there was a 12% decrease from the original Raymer model coefficient, a 12% decrease in the Wyllie's coefficient which was originally 1 for consolidated sandstones and 22% increase for Schlumberger coefficient.

Although the Schlumberger density model has been described by several authors as the most accurate for porosity determination, it gave an error of 3.67 regionally but this reduced by about 40% when the modified Schlumberger model was used. It is also important to note that more often than not these inaccuracies in prediction are always in terms of over-prediction rather than under-prediction.

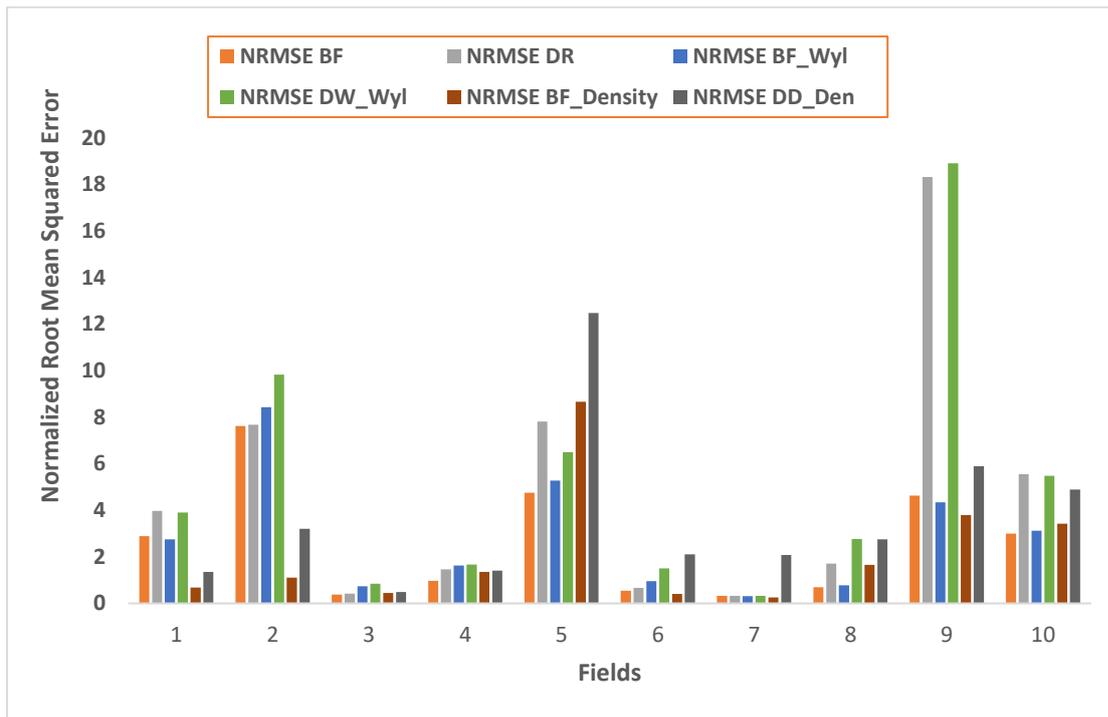


Figure 4. 7 Errors associated with estimation of porosities from sonic using derived best fit for Raymer, Wyllie and Density

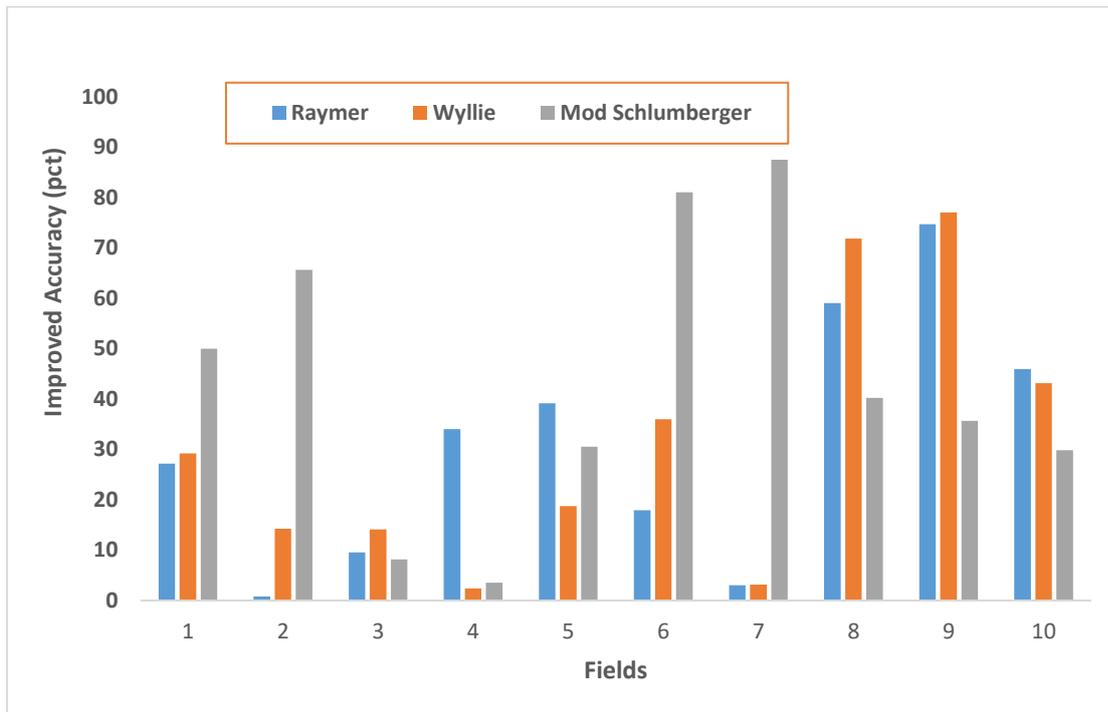


Figure 4. 8 Improvement accuracy of estimated based on new coefficients over the default Raymer, Wyllie and Schlumberger Density model coefficients.

The performance of the models were examined for the unconsolidated, semi-consolidated and consolidated case described in the following section.

4.1.3.1.1 Unconsolidated Sandstones

Galley sandstone Formation located in Bowmore Field (15/24a-9), is a loose and friable sandstone formation, moderately well sorted, with poor inferred porosity. The volumetric clay content is from zero to 0.5. It is an over-pressured sand underlain by the Claymore Sandstone with Piper sandstone further down which is normally pressured and condensate liquids of 20.8 BCF of gas. Estimated proved reserves of 4.6 MMBBL of crude oil. The bulk of the field occupies the culmination of a structural terrace on the downthrown part of the E-W fault forming the Northern margin of the Witch Ground Graben with the east, west and south having a dip closure while the North has a fault closure.

The line is drawn at 45° from both the predicted and actual axis to show the level of correlation between them. The line passing through the data would mean a good level of correlation between the predicted and actual values. Figure 4.9 shows the effectiveness of density logs in unconsolidated sandstones. Although the default Schlumberger model performed less effectively, the modified Schlumberger model (Equation 4.7) did better. The SVM model made very good predictions in this geological setting as it was able to map the non-linear relationship between the dataset using a linear function in a higher dimensional feature space.

Figure 4.10 also shows that default Raymer and Wyllie are less effective in predicting porosity in unconsolidated sandstone formations as have been pointed out by several authors (Dvorkin and Nur, 1996). Amongst other things, this might be caused by the uncertainty in accurate determination of compaction factor for Wyllie.

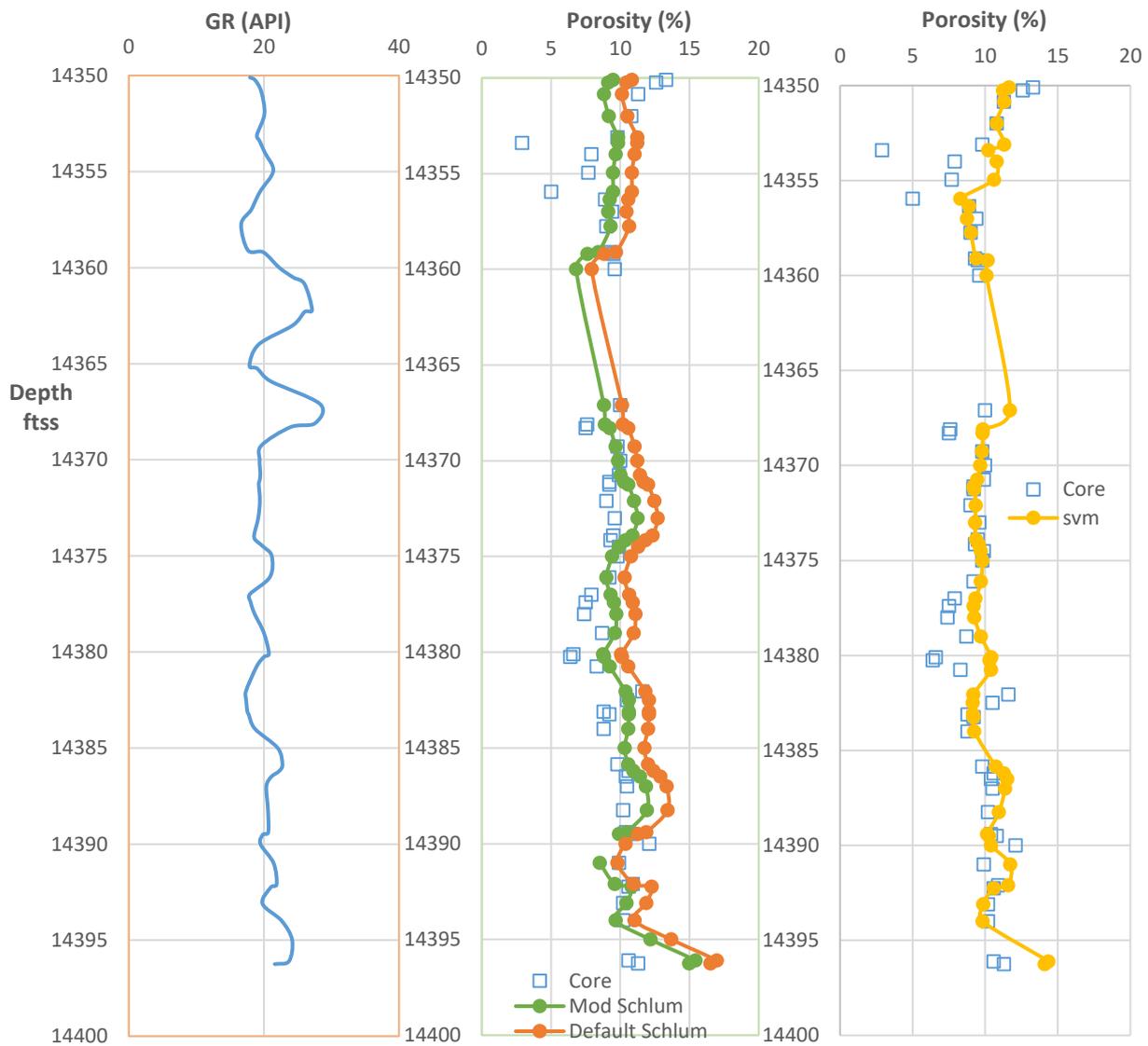


Figure 4. 9 Porosity prediction using default Schlumberger, modified Schlumberger, and Support Vector Machine Regression models for the unconsolidated Galley sandstone turbidites.

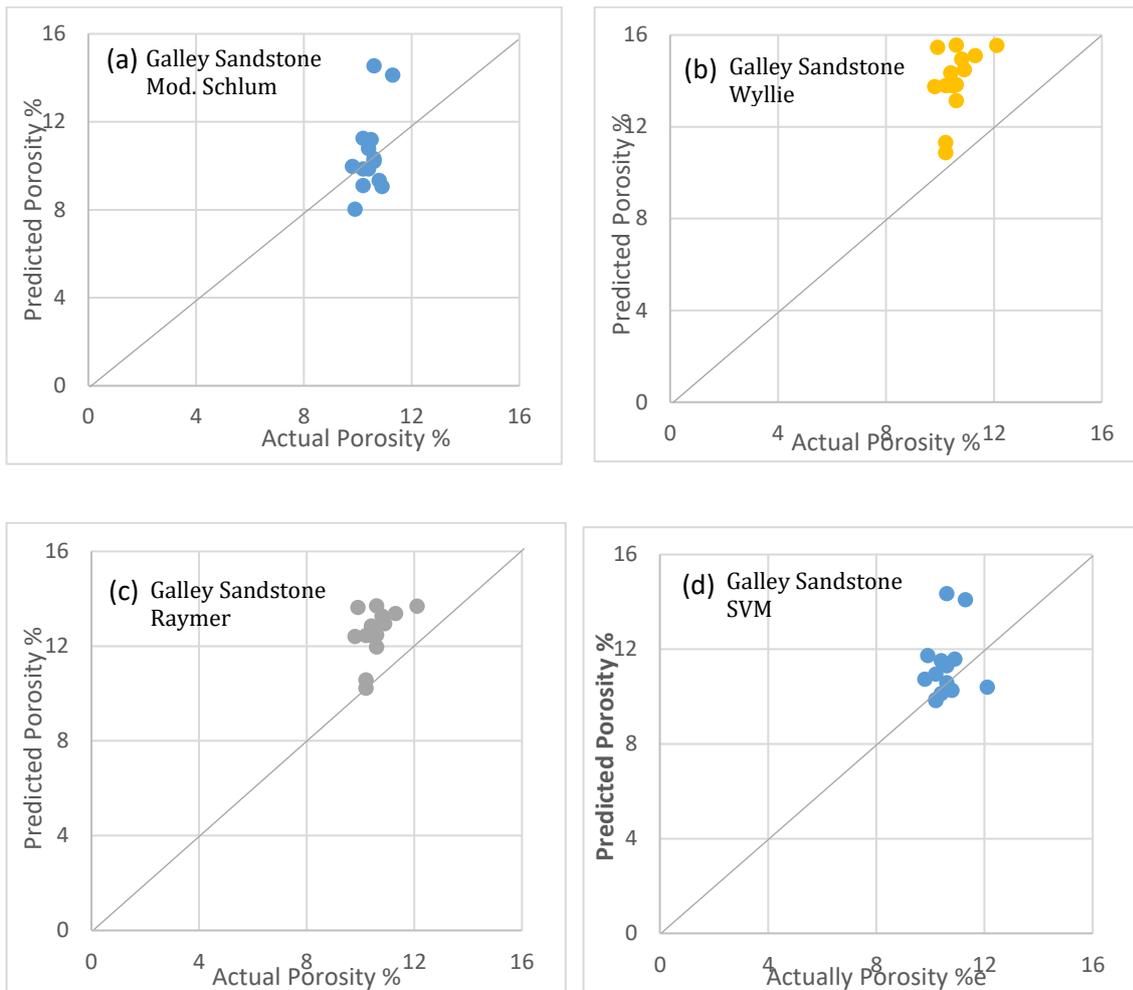


Figure 4. 10 Comparison between predicted and actual core porosity for the unconsolidated Galley Sandstone formation.

4.1.3.1.2 Semi-Consolidated Sandstone

Captain Sandstone member of the Carrack Formation in Blake Field (13/24a-6) is dominated by massive, structureless and minor parallel stratified sandstones with friable quartz grains and argillaceous intervals. They have been interpreted to be mainly deposited from sand rich high density turbidity currents hence are relatively shale free. It has an erosive basal contact with the underlying Vauxhall Formation shales. Formation pressures are from 2321.1-2441.6 psia. Average grain size ranges from very fine to moderate and sorting ranges from poor to moderate well sorted with an average of moderately sorted.

Figure 4.11 shows a good correlation between the Modified Schlumberger model and the actual porosity. Again demonstrated that there is need for correction of the Schlumberger density model for it to be appropriate for porosity determination. The SVR model gave very excellent predictions of porosity in this case as well. The prediction was better than the developed empirical model in this semi-consolidated case as well. Again showing the power of machine learning to accurately determine reservoir parameters. The majority of the porosity seems to be effective between 20-30% and can result in very high permeabilities. Like most reservoirs, this macroporosity and pore connectivity alongside total cement proportion will play a huge part in its character. This porosity will enhance vertical sweep in the reservoir since permeability decreases with depth.

Figure 4.12 shows the correlation between their predicted values and the actual porosity. Raymer and Wyllie this time gave better predictions than in the unconsolidated example. It seems that Raymer gave better prediction than Wyllie even in this case where porosity seems to be effective thus the general notion that Wyllie gives effective porosity is not accurate.

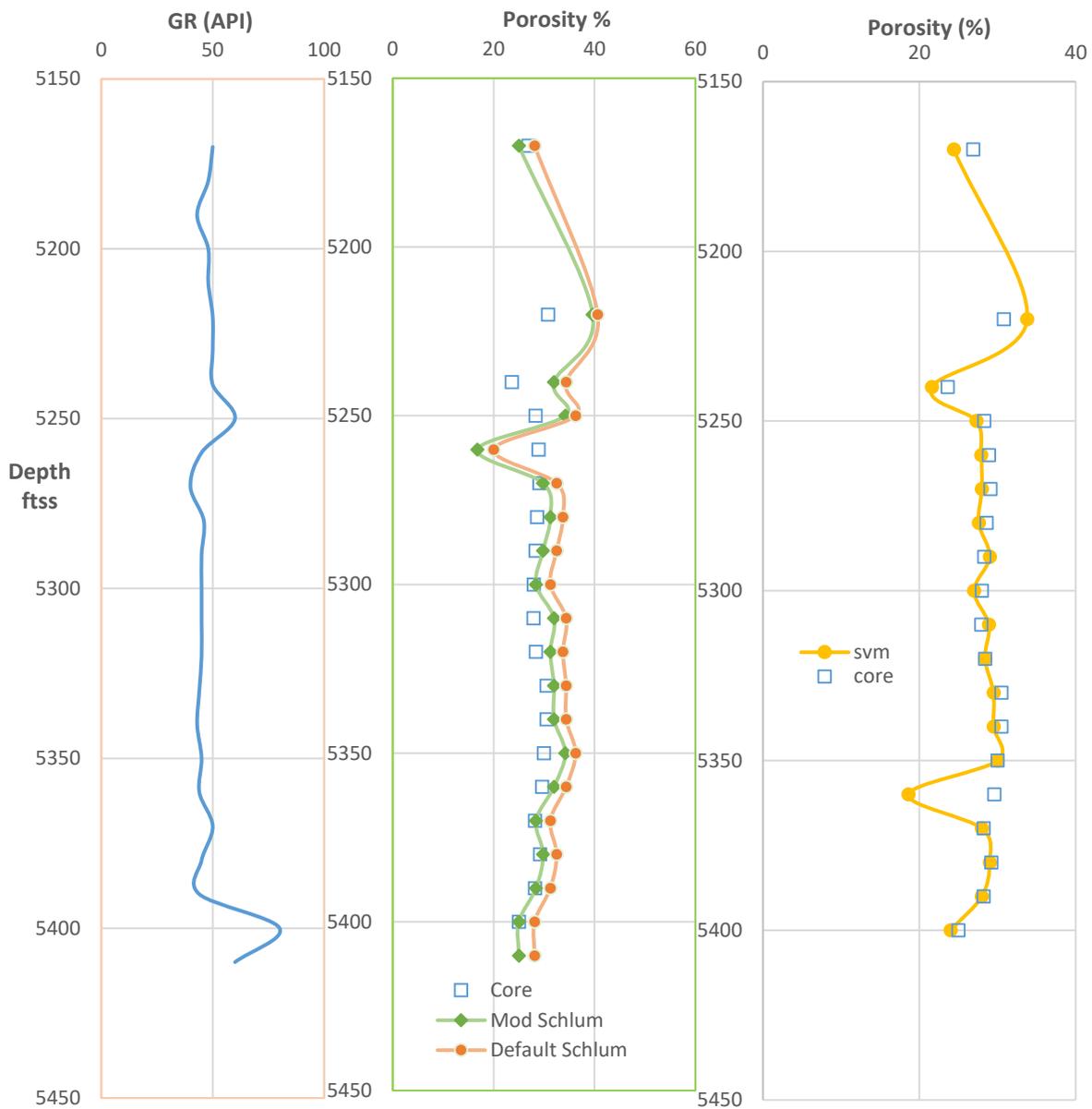


Figure 4. 11 Showing the different porosity models alongside gamma ray for semi consolidated Captain Sandstone Member., Carrack Formation, Blake Field, 13/24a-6.

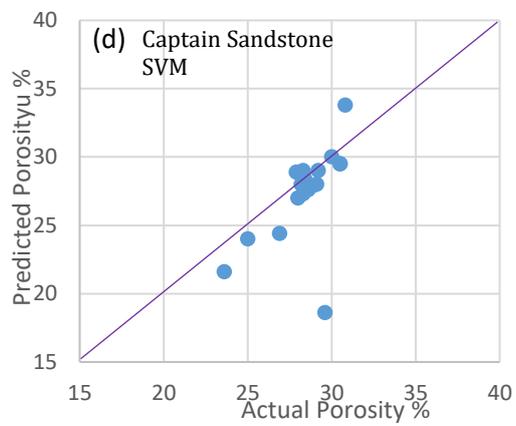
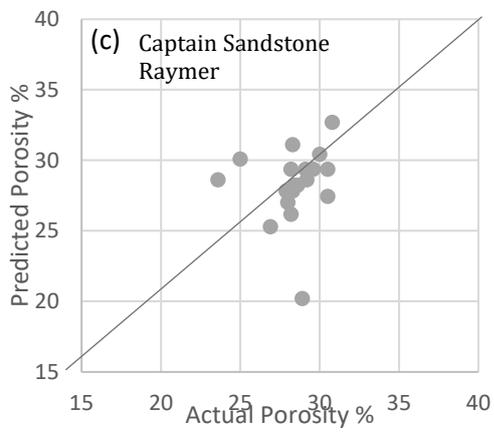
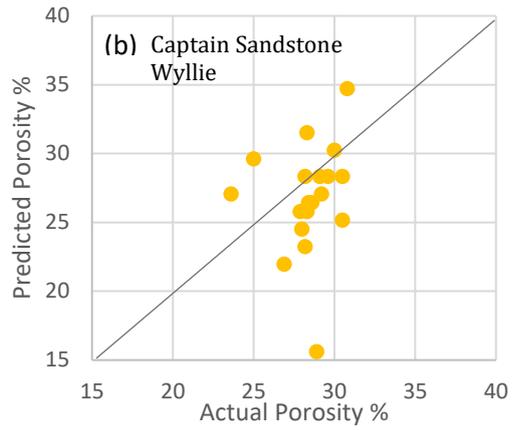
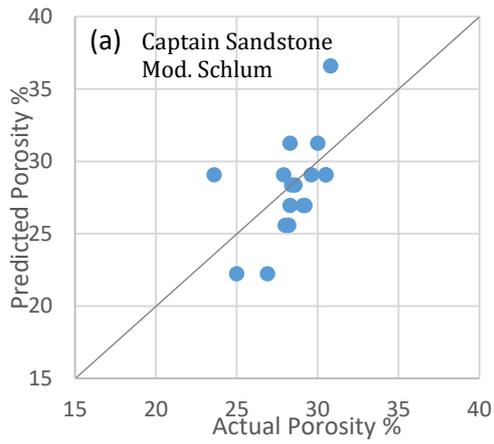


Figure 4. 12 Comparison between predicted and actual core porosity for the Semi-consolidated Captain Sandstone formation, Carrack Formation, Blake Field, 13/24a-6.

4.1.3.1.3 Consolidated Sandstones

The Dogger Reservoir, Upper Brent Formation is an indurated, well sorted, laminated sandstone with poor to good porosity. This Field is located in the Northern part of North Sea (3/9a-1). Significantly has both horizontal and vertical permeability. The Kimmeridge Clay overlies it. The Broom, Rannoch, Etive, Ness and Tarbert units are genetically related to those of Lyell, Hutton and others.

Figure 4.13 shows modified Schlumberger making very good predictions in consolidated sandstone formations. No doubt that these sandstones, compaction is a main factor influencing its porosity. Most of the porosity lies between 20-30%. Again the Default Schlumberger model prediction shows that if used the way it is (which is the case most times) could lead to poor porosity results and thus consequently affect other properties like water saturation, permeability and even reserve estimation. The SVM model on the otherhand, also made good predictions.

Figure 4.14 clearly illustrates the improvement of Raymer and Wyllie model predictions in this consolidated example. Compared to less indurated rock formation discussed in previous sections, the 45° line cuts through most of the data indicating its correlation with the core data. The default Schlumberger model actually performed the worst in this instant.

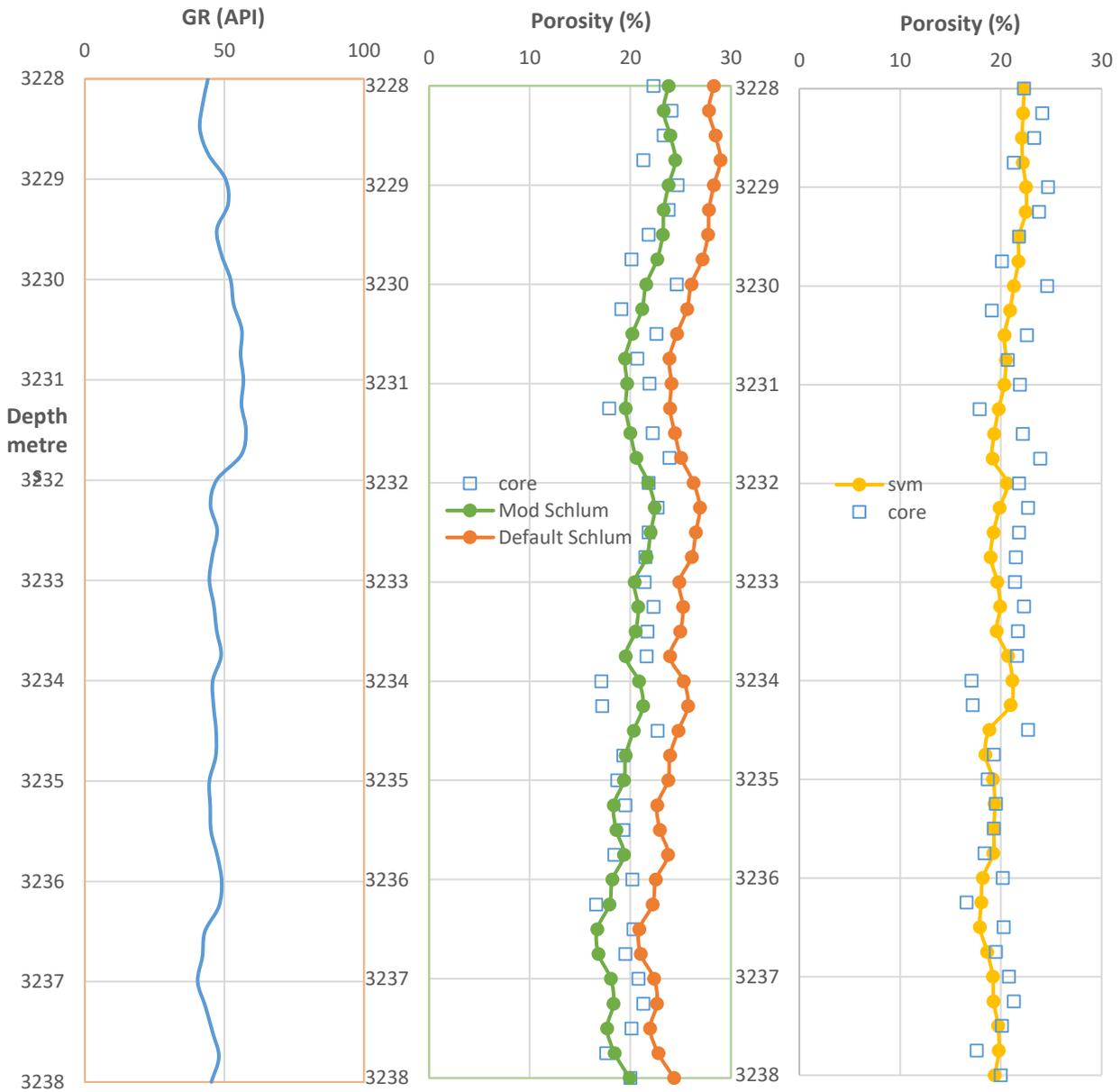


Figure 4. 13 Plot of Gamma Ray, Modified Schlumberger, Default Schlumberger for consolidated Dogger Reservoir, Brent Formation, Northern North Sea (3/9a-1).

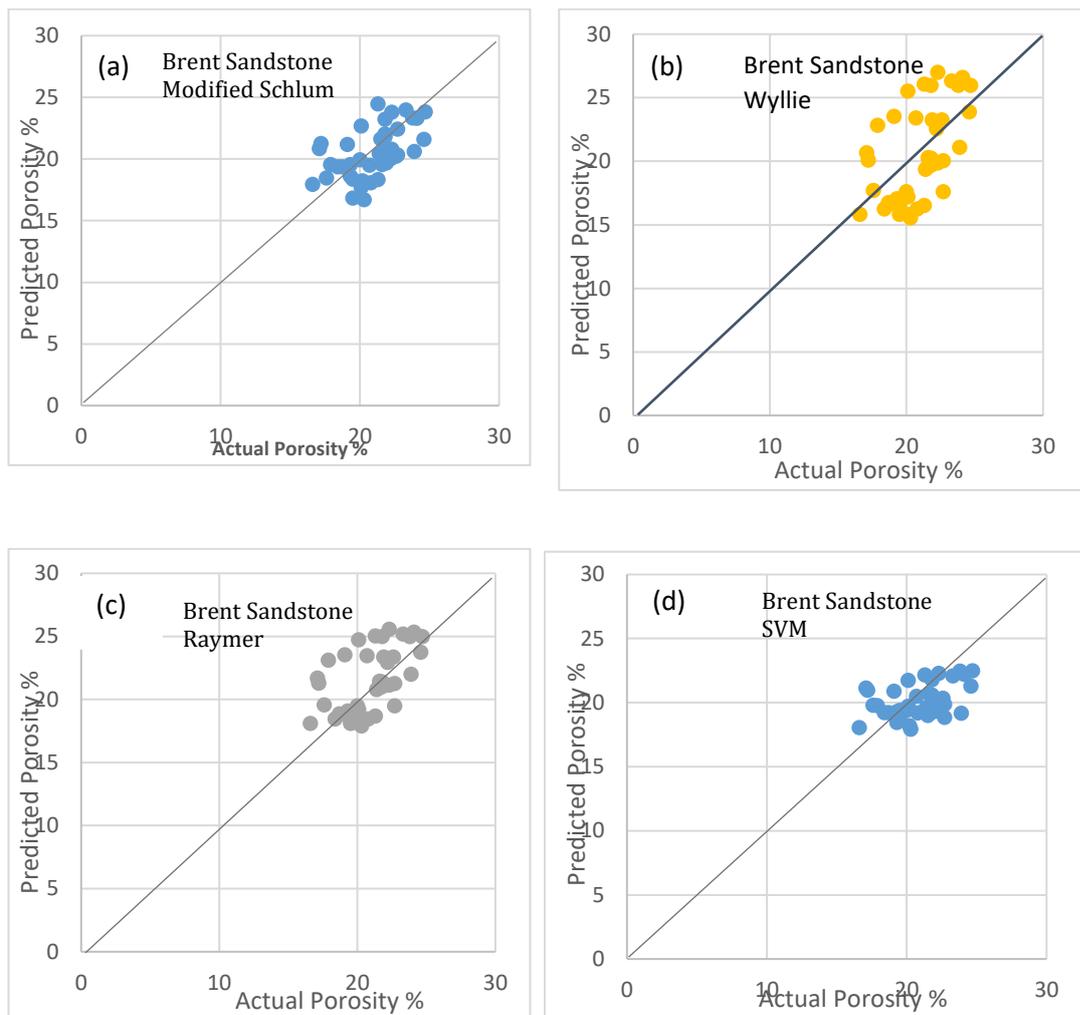


Figure 4. 14 Comparison between predicted and actual core porosity for the Consolidated Captain Sandstone formation, Carrack Formation, Blake Field, 3/09a-1.

4.1.3.2 Effect on Reserve and Production Forecast

The different predictions were used to estimate oil originally in place assuming there are no uncertainties in other parameters like reservoir area, pay thickness, water saturation, formation volume factor used for its computation. Table 4.4 shows the different results from the different models considered compared to the base value. Table 4.5 shows the error associated with these disparity showing default Schlumberger, Wyllie and Raymer could lead to errors of up to 17%, 11% and 14% for unconsolidated formation, up to 5%, 4% and 4% in semi-consolidated formations while lesser values of error of about 2%, 3% and 3% respectively in consolidated sandstone formations. The newly developed modified Schlumberger and support vector machine models overall gave errors of less than 5% in reserve estimation thus reducing the uncertainty.

Table 4. 4 Porosity uncertainty in reserve estimation (in MMSTB) assuming other parameters have no uncertainty.

Lithologic Description	Actual	Modified Schlum	Schlumberger	Raymer	Wyllie	svm
Unconsolidated	149	139	175	166	170	152
Semi-consolidated	450	440	472	434	430	445
Consolidated	331	324	325	338	340	329

Table 4. 5 Percentage error in reserve estimation arising from porosity uncertainty assuming other parameters have no uncertainty.

Lithologic Description	Modified Schlum (%)	Schlumberger (%)	Raymer (%)	Wyllie (%)	Svm (%)
Unconsolidated	7	17	11	14	2
Semi-consolidated	2	5	4	4	1
Consolidated	2	2	2	3	1

4.1.4 Summary and Conclusion

Raymer and Wyllie models for the North Sea were formulated as well as Modified Schlumberger model thus reducing the uncertainty associated with its prediction. Support Vector Machine Regression was also used to make predictions and found to perform well in unconsolidated, semi-consolidated and consolidated sandstones investigated. The chapter becomes very relevant as later chapters deals with permeability or water saturation with some wells not having core porosity data hence knowledge of local, regional and support vector machine regression were comfortably used to determine this with greater level of confidence.

The results show the default Schlumberger model exaggerating predictions while default Raymer and Wyllie are not suitable for porosity determination in unconsolidated formations. Their accuracy in sandstone formations seems to be related to consolidation. SVM performed best in all sandstone types considered. It is also important to note that more often than not, these inaccuracies in prediction are always in terms of over-prediction rather than under-prediction in most of the cases examined as can be seen by the upward lying of most points above the 45° line on the actual versus predicted cross plots. Thus the assumption that Wyllie gives effective porosity seems not to be true. Local and regional mean values for Raymer, Wyllie and Modified Schlumberger were then determined giving values of 0.55, 0.88 and 1.22 respectively for the North Sea.

Application of the different sources for reserve estimation in fields of known reserve value gave seemingly different reserve figures as well. The commonly used Schlumberger, Raymer and Wyllie Time Average could lead to errors of up to 17% in unconsolidated sandstones, about 5% in semi-consolidated sandstones and about 3% in consolidated sandstone reservoirs.

4.2 Solving the Fluid Saturation Problem Using Stacked Ensemble and Empirical Methods

4.2.1 Introduction

Water production control and identifying pay zones with high irreducible water saturation are important formation evaluation issues which can result to delayed decisions on completion and cause additional expenses on field management. High water saturation is not much of a problem if huge fraction of the water is immovable (Oraby and Eubanks, 1997). Their accuracy impacts on hydrocarbon saturation. Complexities in the determination of water saturation emanates due to the fact that not only are there different approaches for its determination but also typically complicated that they all produce different results. This may result to remarkable differences due to their sensitivity in reserve and Net Present Value estimation as have been discussed in Chapter Two. Technical teams are posed with the challenge of obtaining a consistent technique or model hence they try to resolve the differences among the water saturation results obtained with the different procedures in order to arrive at its best calculation and distribution (vertically and areally) throughout the reservoir.

Again, reliable identification of water free hydrocarbon producing zone is a difficult task and reservoir analysts in general are sceptical about the production of water free hydrocarbon at water saturation greater than 50% leading to unwillingness to carry out a well test on such zones especially in the presence of less uncertain intervals. By-passed zones are common with the calculation of reserves using only reservoir with $< 50\%$ water saturation and assuming that those with $> 50\%$ water saturation contribute little or nothing to reserves. The domain of the constant range relative permeability curve depends on the combination of irreducible water saturation and residual oil saturation. Statistically, the irreducible water saturation falls between 10-20% and 2-5% for residual oil saturation (Calhoun, 1953; Amyx, 1960). As can be seen from Figure 4.15, 1-20% seems to be the most sensitive range of this parameter as it pertains to its relationship with permeability and porosity. Again a small increase in water saturation could result in a significant reduction of oil relative permeability obviously because relative permeability depends on phase saturation.

Each existing technique used in water saturation determination (Table 4.6) has weaknesses such as selecting from a myriad of models with a variety of input parameters (Just as Herrick and Kennedy (2009) pointed out on the quagmire of these equations actually compounding the problem instead of relieving the confusion of its determination), conversion from laboratory to

reservoir conditions, time taken for laboratory measurements to reach equilibrium, application to the lowest part of transition zones where the brine phase is mobile, limited amount of cores due to expense, inaccurate procedures for parameter determination.

Table 4. 6 Some water saturation models for shaly sands with R_t as the true resistivity, water resistivity (R_w), formation factor (F) and resistivity of shale (R_{sh}).

Models	Formula
Poupon et al. 1954	$\frac{1}{R_t} = \frac{(1 - V_{sh})S_w^2}{F \cdot R_w} + \frac{V_{sh}}{R_{sh}}$
Simandoux (1963)	$\frac{1}{R_t} = \frac{S_w^2}{F \cdot R_w} + \frac{V_{sh}}{R_{sh}}$
Schlumberger (1972)	$\frac{1}{R_t} = \frac{S_w^2}{F(1 - V_{sh})R_w} + \frac{V_{sh}S_w}{R_{sh}}$
Hossin (1960)	$\frac{1}{R_t} = \frac{S_w^2}{F \cdot R_w} + \frac{V_{sh}^2}{R_{sh}}$
Poupon and Leveaux (1971)	$\frac{1}{R_t} = \frac{S_w^2}{F \cdot R_w} + \sqrt{\frac{V_{sh}^{2-V_{sh}}}{FR_w R_{sh}}} S_w^2 + \frac{V_{sh}^{2-V_{sh}} S_w^2}{R_{sh}}$
Clavier et al (1977)	$\frac{1}{R_t} = \frac{S_w^2}{F_o R_w} + \frac{(C_{bw} - C_w)V_Q Q_v S_w}{F_o}$
Waxman and Smits (1968)	$\frac{1}{R_t} = \frac{S_w^2}{F^* \cdot R_w} + \frac{BQ_v S_w}{F^*}$

Where R_t is true resistivity of the pristine rock formation and not the flushed zone (R_{xo} not used here), V_{sh} is volume of shale, R_w is water resistivity, R_{sh} is shale resistivity, F is formation factor, Q_v is counter ion concentration (meq/gm), B is equivalent conductance of clay cation, C_{bw} is bulk water conductivity, C_w is water conductivity.

Nuclear Magnetic Resonance (NMR) has shown capability to estimate reservoir fluid saturation. Although the physical concept of the NMR interpretation has a generic nature, practical experience with NMR logs in about twenty (20) North Sea oil and gas wells indicated the existing need for cores to investigate potential effects of changes in wettability by oil based mud invasion especially in multiphase flow environments and also to determine BVI cut-offs (T2 spectra calibration). Again NMR parameters are difficult to acquire due to vertical resolution, depth of investigation and disturbing environmental/borehole effects. On the other hand, maintaining reservoir overburden pressure (NOB), preserving core wettability, time, expense and unconsolidated reservoirs are major concerns for accurate special core analysis.

Several workers like Al-Bulushi (2009) used neural networks for water saturation determination from both well logs and core parameters. Mardi et al., (2012) used four different structures to make predictions of water saturation from log data after which comparison was made with dual water model and found to perform better. Helle and Bhatt (2002) used committee machines comprising neural networks for the prediction of reservoir fluids using well log data. The best nine experts were selected from 20 trained experts and averaged to obtain final predictions. Kamel and Mabrouk (2002) both developed empirical and regression methods for the prediction of total water saturation. Kenari and Syamsiah (2013) also developed committee machine for the prediction of water saturation using well log data. They also applied pruning steps to make the models better. Most of them focussed on total water saturation and not much has been done with respect to irreducible water and residual oil saturation.

In this work a Stacked Ensemble model (Section 3.6.5.3.4) in which the predictions that are generated by using different learning algorithms in this case Gradient Boosting Model (GBM) and Neural Networks (NN) as inputs in a second level learning algorithm and the stack is then used to predict total water, irreducible water and residual oil saturation. More details of the Stacked Ensemble models used for the prediction of fluid saturation is explained in Section 3.6.5.3.4. In addition to this, a new empirical model was also developed for irreducible water saturation using the factors affecting the parameter.

4.2.2 Methodology

The methodology involves the development of both empirical and an ensemble of machine learning models for connate and irreducible water as well as residual oil saturations using North Sea wells. Details are given in the following sections.

4.2.2.1 Machine Learning (Stacked Ensemble) Approach

Stacked Ensembles combines several individual models (base models) into one model (meta model) in order to get improvement in prediction performance. The machine learning approach involves using conventional well logs which are always available to estimate connate and irreducible water saturation as well as residual oil saturation. For the total water saturation model, a total of 471 datasets with four inputs such as Gamma Ray (GR), Resistivity (ILD), Neutron (NPHI) and Bulk Density (RHOB) while core Sw was used as the output. The dataset was randomly divided using stratified technique into 425 training sets, 24 validation set and 22 test set. The irreducible water saturation had 290 datasets randomly divided into 243 training set, 16 validation set and 31 testing set. The residual oil saturation model used 243 datasets again randomly divided into 195 training set, 16 validation set and 32 test sets. As already mentioned, the data comprises wireline logs and core which were normalized lying between 0 and 1 to ensure that the input variables were independent of their measurement units. Several sensitivity analysis were undertaken by varying some parameters such as the number of hidden layers, learning algorithms, transfer functions and scaling methods. A design matrix was done to specify both the predictors and target variables after which the data was divided into training, validation and test sets. An ensemble approach that combined the best machine learning algorithms for the data after testing several algorithms was adopted using a technique known as stacking in bid to reduce prediction variance.

Two levels of learning were used. The first set of learners (base learners) involved different algorithms like Gradient Boosted Model (GBM) and Neural Networks (NN). From Figure 4.16, both models did not correlate hence they have strengths in different aspects of the data. They were trained using the training dataset and then a 5-fold cross validation was performed on each learner. These predicted results were combined to form a matrix which together the original response vector is called the level-one data. A second level of deep learning algorithm (metalearner) was then trained to learn optimal combinations on the level-one data. The resulting ensemble model comprises the base learning models and the metalearning model. It is imperative to note that Stacked Ensembles only work with cross validation. In selecting the models for the ensemble, it was ensured that individual models fulfil a particular accuracy criteria and that the

base model are not highly correlatable (Figure 4.15) with each one capturing different aspects of the data. The different types of ensemble models are discussed in detail in Chapter 3.

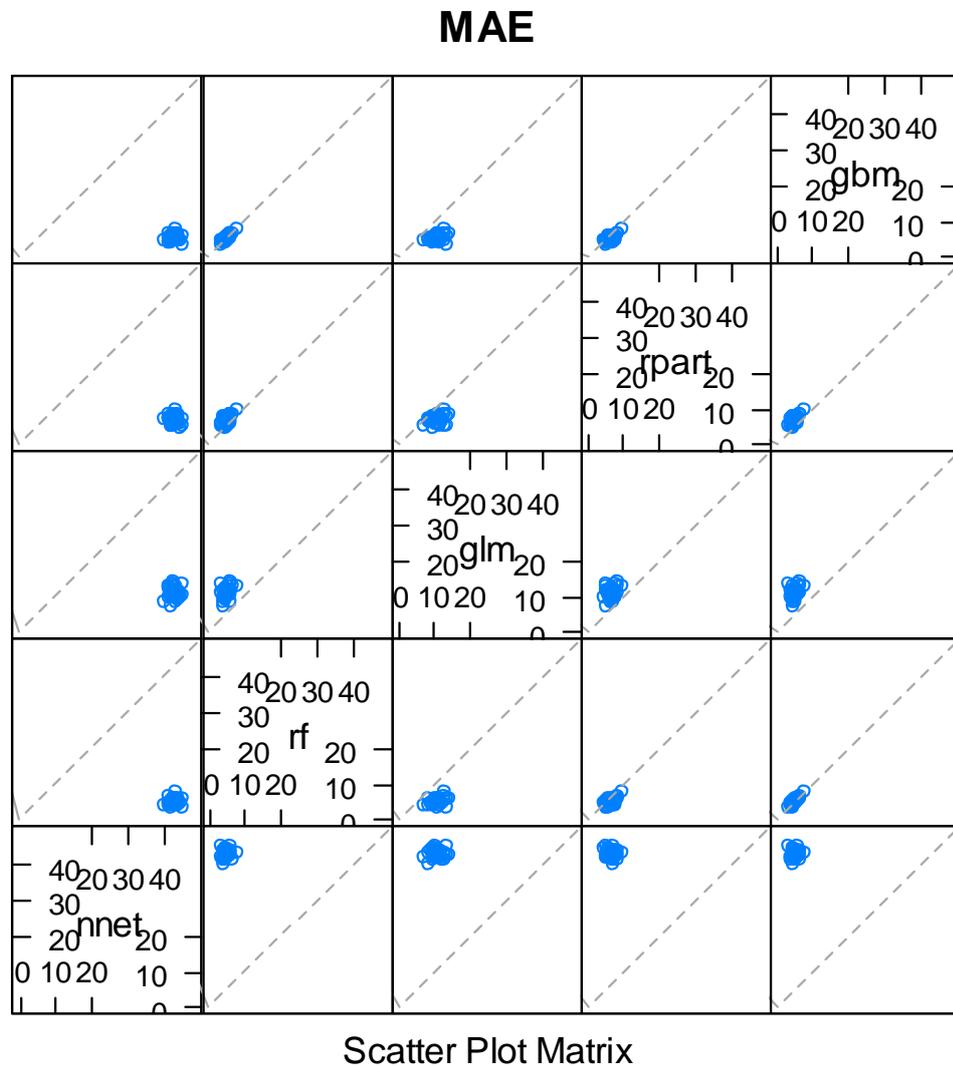


Figure 4.15 Correlation between the different base algorithms. The least correlatable i.e blue plots farther away from the 45° line were selected. GBM, Neuralnet and Random Forest where selected as base models for the Stacked Ensemble saturation models.

4.2.2.2 Empirical Model Approach

Empirical model was also systematically developed for the estimation of irreducible water saturation to be valid for each product of rock that forms the reservoir. It was entirely computed from influencing parameters such as specific surface area to pore volume ratio, volume of shale and flow zone indicator hence obtaining the property as continuous curve across reservoir intervals from well logs is discussed (Figure 2.11). Results were then compared to actual core data.

4.2.3 Results and Discussion

Excellent results were obtained from testing on North Sea reservoirs showing good correlation with measured data. Model construction was done using both the training and validation sets while the test set data was used for out of sample prediction capability. This section is divided into different sub-sections representing total water saturation, irreducible water saturation and residual oil saturation.

4.2.3.1 Ensemble Total Water Saturation Model

The ensemble model was chosen to predict the water saturation in both a Northern and Central North Sea reservoir which was then compared with commonly used water saturation models like Archie, Simandoux, Schlumberger, Indonesian. Hosin. Figure 4.16 and 4.17 give the results of the above models alongside the ensemble model and actual core data. All models appear to have performed well probably since the reservoir cut by the well is a relatively cleaner sand. Simandoux gave a root mean squared error value of 0.0349, Schlumberger had values of 0.0427, Indonesia gave values of 0.0648, Hosin had values of 0.0711, Archie had 0.0712 while the stacked ensemble model gave the lowest error value of 0.0136. Figure 4.18 and 4.19 show that the reservoir seems to be more heterogenous as seen from the gamma ray behaviour at that well location. Simandoux had a root mean squared error of 0.1477, 0.1357 for Schlumberger, 0.1975 for Indonesia, 0.2265 for Hosin, for Archie it was 0.2282 while Stacked Ensemble gave a value 0.054. Modeling at depths of about 15250-15350ft, 15450-15560ft, representing areas of high shaliness, the disparity between the different models is more. Archie is suited for clean sands while the other shaly sand models that seem to have done well in Figure 4.16, have not done well in the more shaly sand formation shown in Figure 4.18. This inconsistency in accuracy raises the question of which one should be used for water saturation estimation thus increasing the

uncertainty. As have been seen from both figures, the trained water saturation ensemble model performed well in both cases and this consistency answers the above question that machine learning models can be trusted to give accurate predictions aligning to the 45° line which shows the correlation between the predicted and actual values.

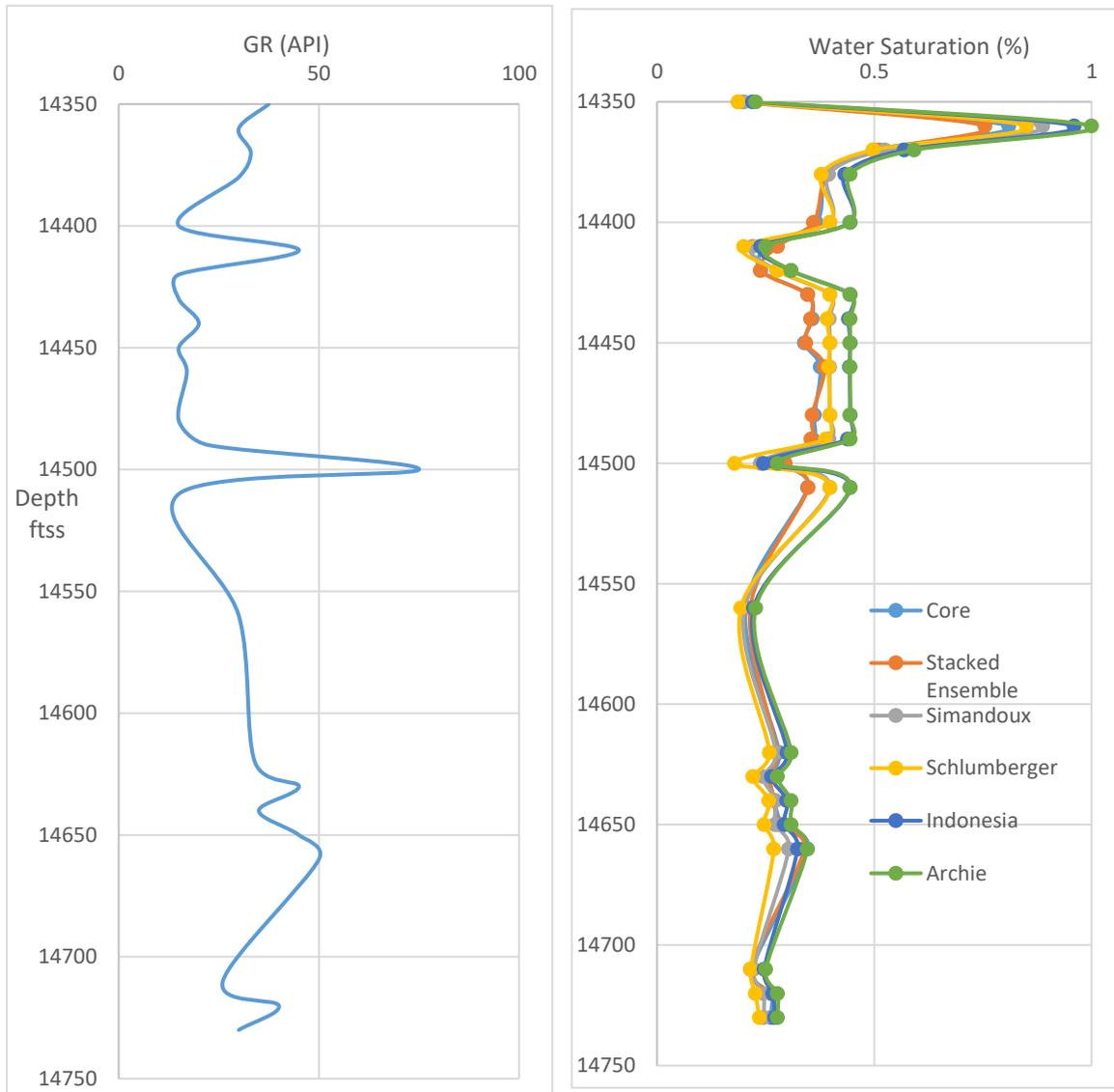


Figure 4.16 Stacked Ensemble, Archie, Schlumberger, Simandoux, Indonesian and Core alongside Gamma Ray for the Brae field, Northern North Sea.

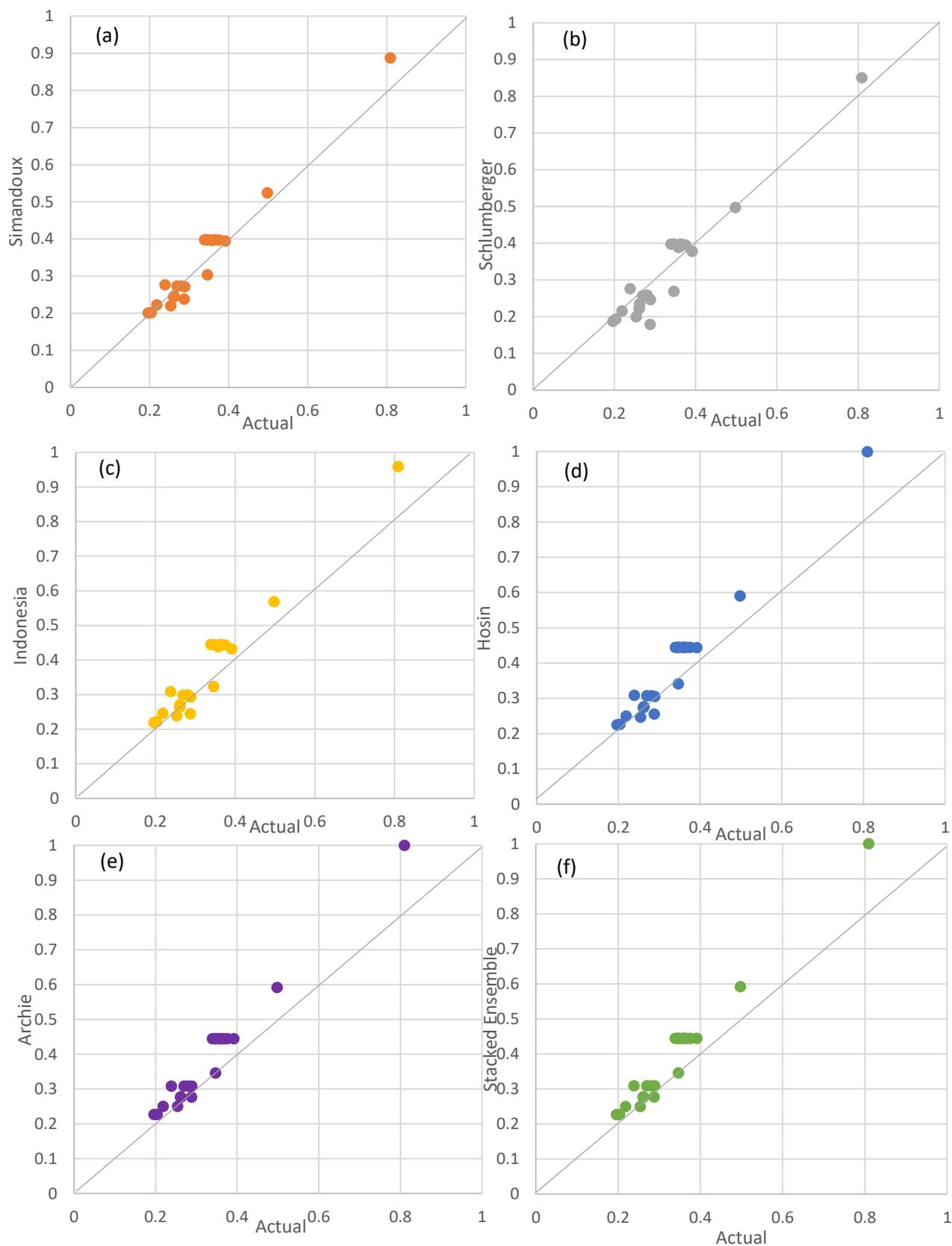


Figure 4.17 Cross plot of actual core against (a) Simandoux with rmse of 0.0349 (b) Schlumberger with rmse of 0.0427 (c) Indonesia gave rmse of 0.0648 (d) Hosin had rmse of 0.0711 (e) Archie gave rmse of 0.0712 and (f) Stacked Ensemble had 0.0136

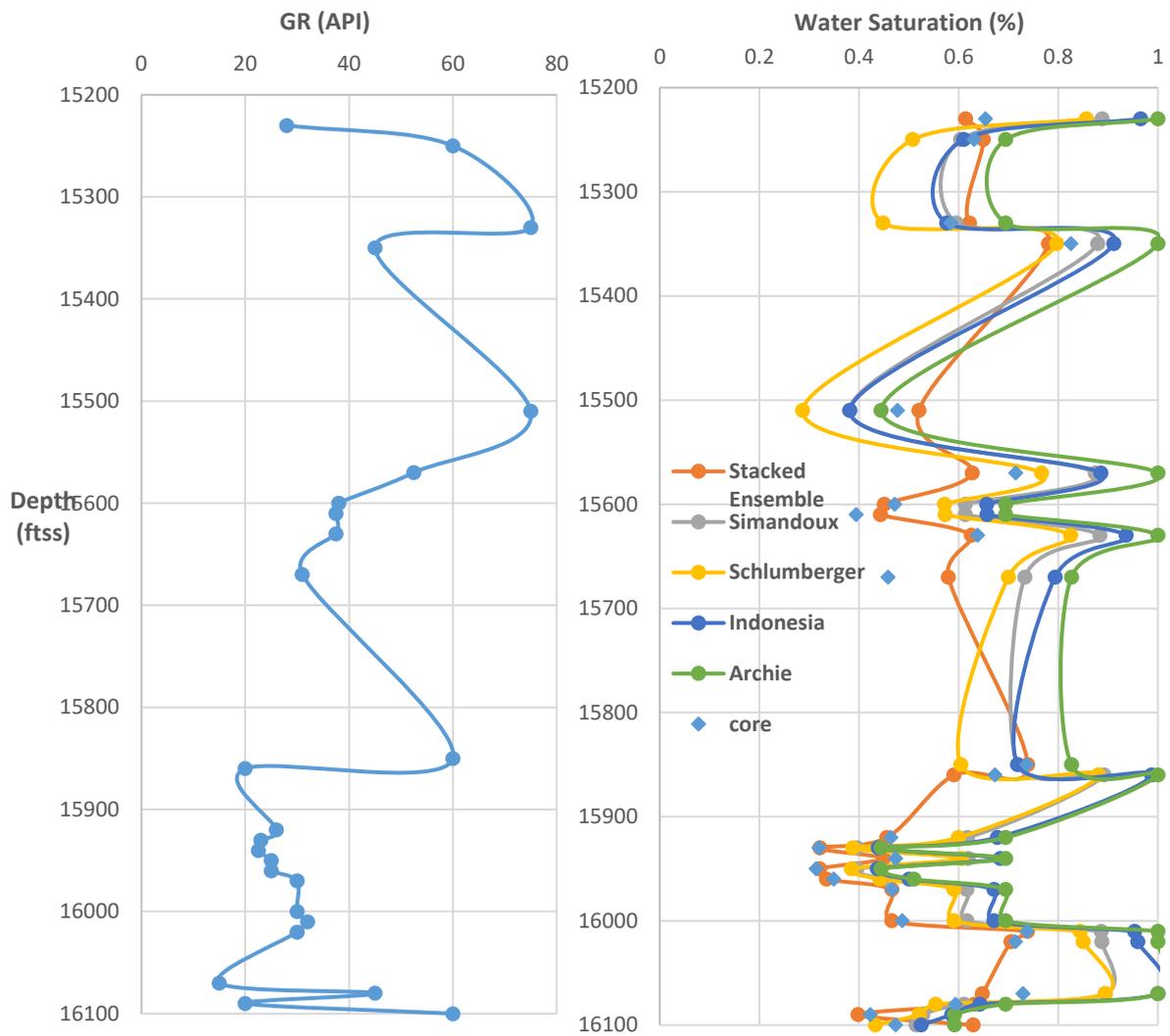


Figure 4.18 Stacked Ensemble, Archie, Schlumberger, Simandoux, Indonesia and Core alongside Gamma Ray for the Brae field, Central North Sea.

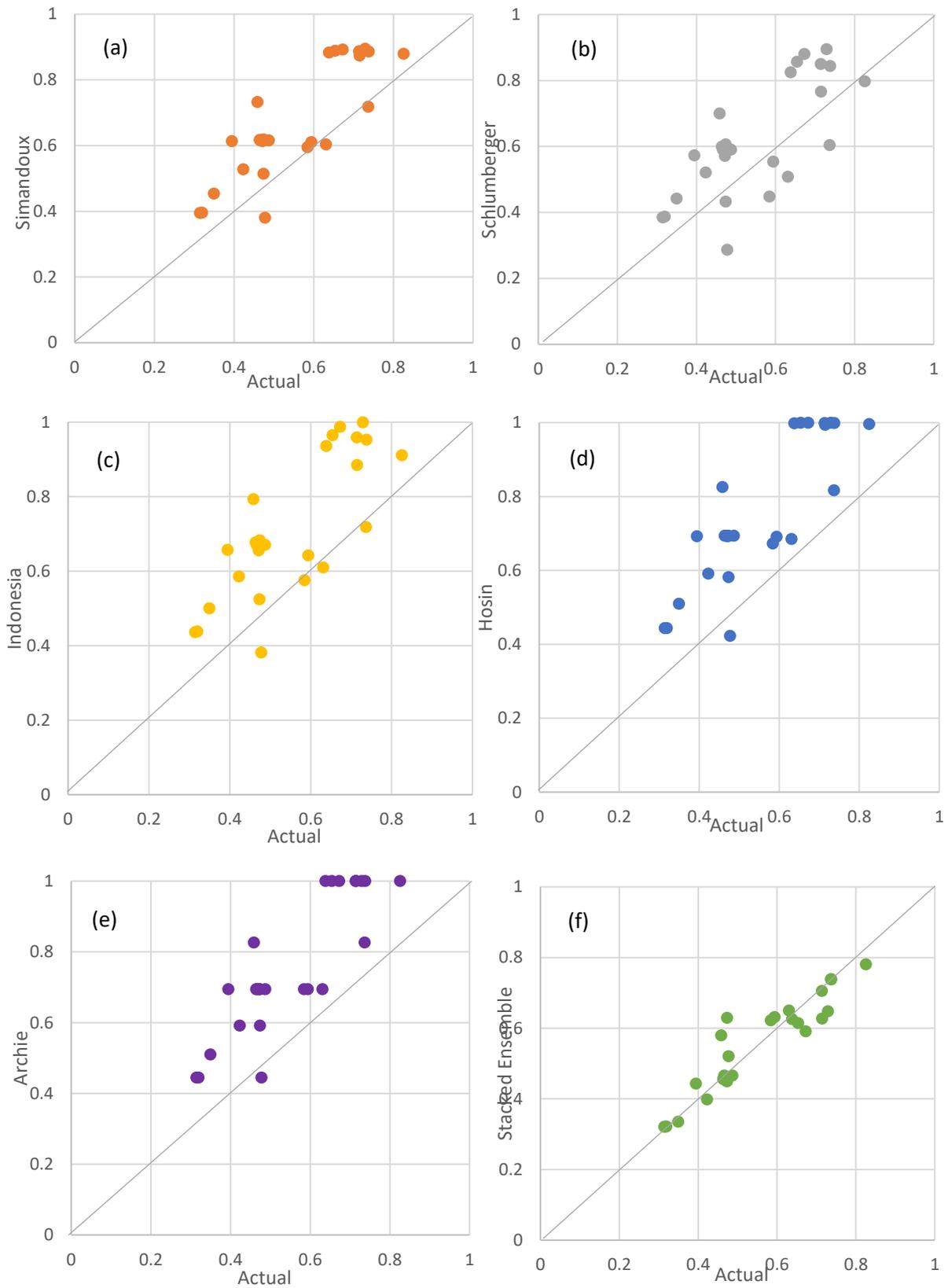


Figure 4. 19 Cross plot of actual against (a) Simandoux with 0.1477 (b) Schlumberger had 0.1357 (c) Indonesia had 0.1975 (d) Hosin gave 0.2265 (e) Archie gave 0.2282 and (f) Stacked Ensemble had 0.054 values of rmse.

4.2.3.2 Empirical and Ensemble Swi Models

The results of both the comprehensive but simple irreducible water saturation and the ensemble of regression models to reduce the uncertainty associated with its prediction is discussed. Equation 4.16 represents a non-linear regression on volume of shale, specific surface area and flow zone indicator to obtain the empirical model for irreducible water saturation (Figure 4.20). Figure 4.21 shows the sensitivities of the different parameters in the model in Equation 4.18.

$$\log(S_{WI}) = 0.2483 * \log(V_{sh}) - 0.0766 * \log(FZI) + 0.4352 * \log(S_p) \quad (4.17)$$

hence

$$S_{wi} = \frac{V_{sh}^{0.25} S_p^{0.44}}{FZI^{0.08}} \quad (4.18)$$

where V_{sh} is in percent, S_p in m^{-1} , FZI in μm .

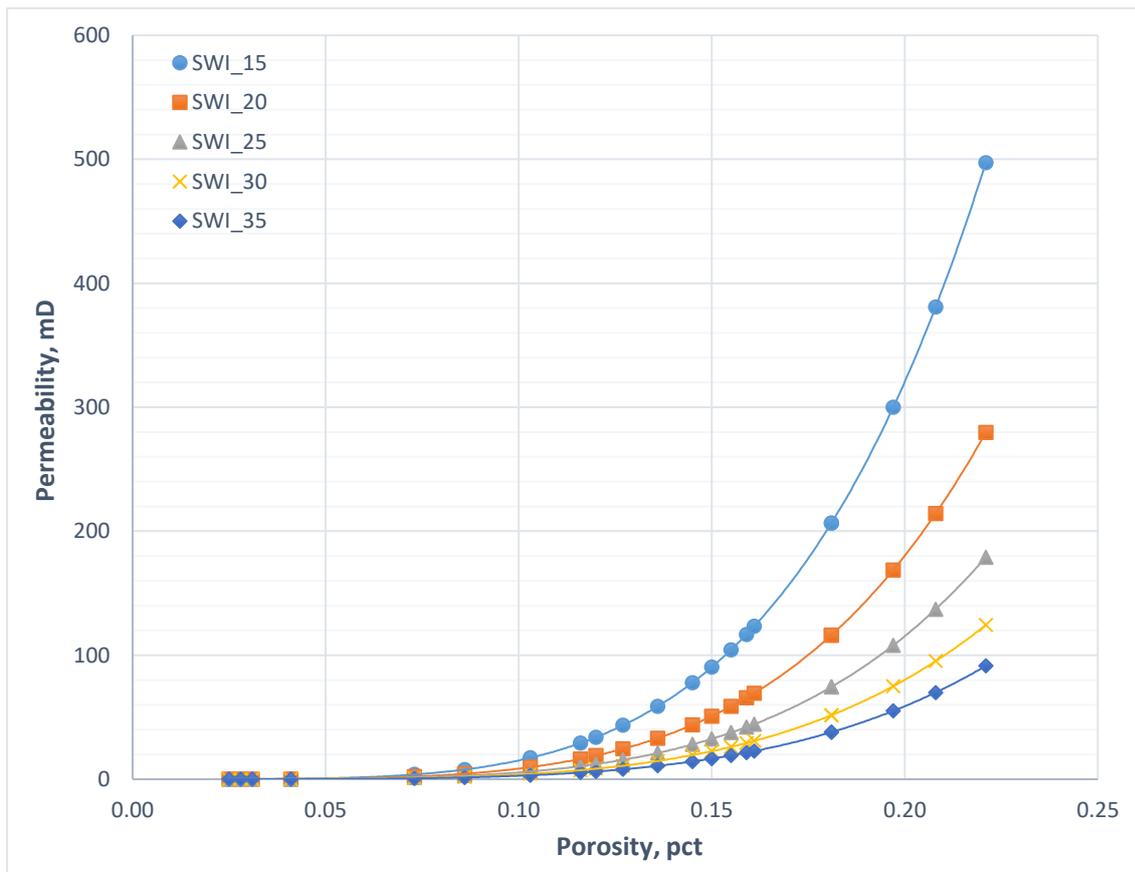


Figure 4. 20 Showing the influence of irreducible water saturation on permeability with each point as a measurement at a depth.

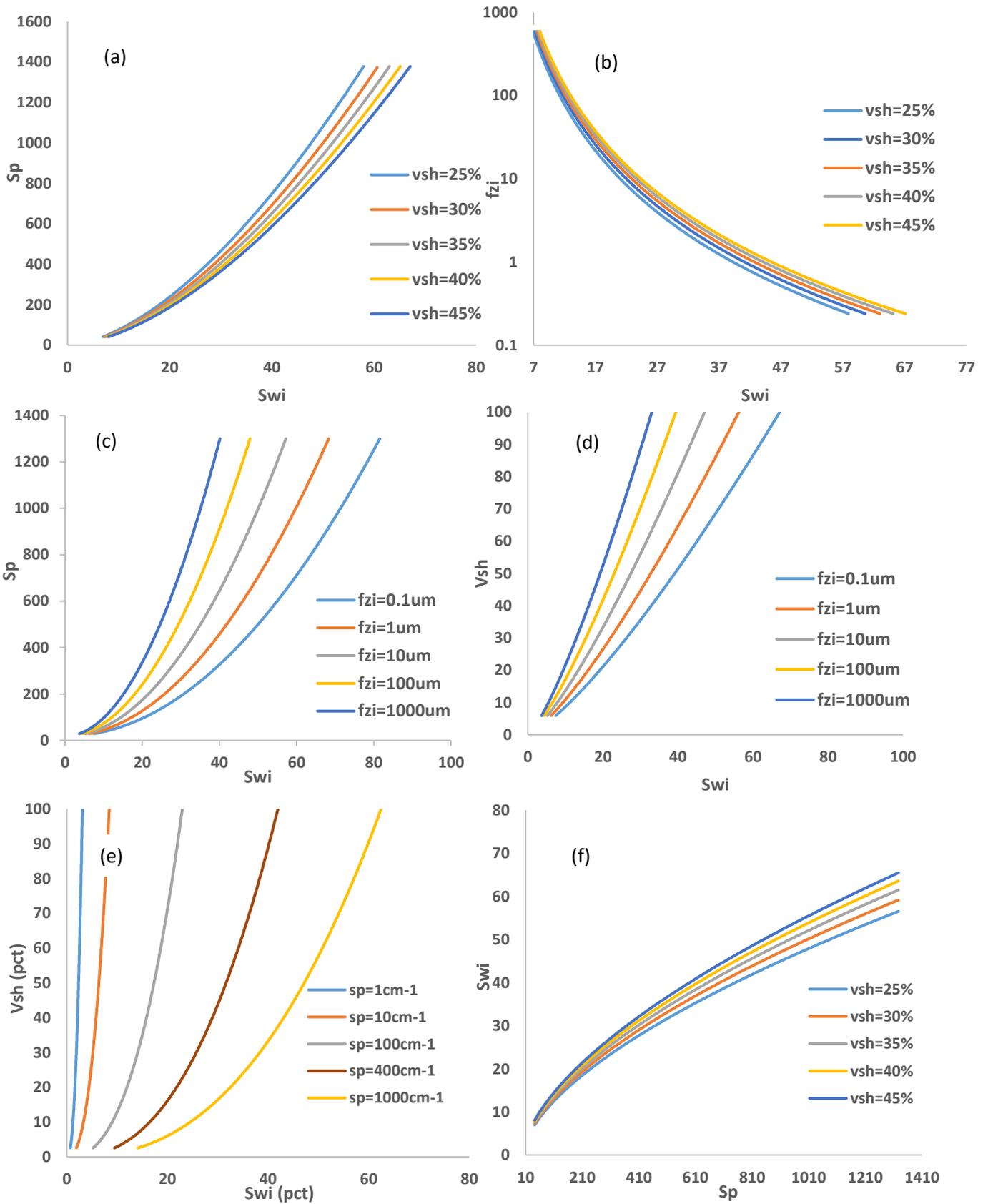


Figure 4. 21 Sensitivity of different parameters used in the empirical model (Equation 4.18) i.e volume of shale, specific surface, FZI and irreducible water saturation relationship.

Figure 4.20 demonstrated the importance of irreducible water saturation to the accurate determination of permeability. The inverse relationship between permeability and irreducible water saturation means that errors made at lower values of irreducible water have a higher impact on permeability as shown in the above figure. Again, these are synthetic porosity datasets used to compute permeability using different values of irreducible water saturation of 15%, 20%, 25%, 30%, 35% thus illustrating the importance of irreducible water saturation and the sensitivity of permeability to it. As irreducible water lies mainly within the 10-25% range, it is pertinent to get it as accurately as possible hence the need for more accurate machine learning models. Figure 4.21 shows the sensitivity of the different parameters used in the empirical model.

Figure 4.22 shows a comparison of the base learners (GBM, Neural Networks, Random Forest), metalearner (Stacked Ensemble) and the empirical model to the actual core data for the validation set. Again, closeness and alignment to the 45° line gives an indication to how correlatable both axis i.e actual and model, are to each other hence the $y=x$ fit line not used. It is clear to see that all models did well with RMSE values of 0.0603 for GBM, Neural Networks gave 0.0691, Random Forest had 0.0619, the Stacked Ensemble model gave 0.0581 while the empirical equation gave 0.071. Figure 4.23 shows the different models against depth alongside gamma ray. The gamma ray shows an intercalation of sand and shale lithology with both coarsening and fining upwards sequences. Areas with relatively high amount of shale seems to have higher irreducible water saturation as shown hence the inclusion of volume of shale as a parameter in the developed empirical model. The gamma ray response can give a good indication of irreducible water saturation and should be used as an index for the parameter in sandstone reservoirs. For Figure 4.24 and 4.25, the Gradient Boosting Model gave an RMS error of 0.0626, Neural networks had 0.0619, Random Forest gave 0.063, Stacked Ensemble gave values of 0.0459 while the Empirical model gave values of 0.0773.

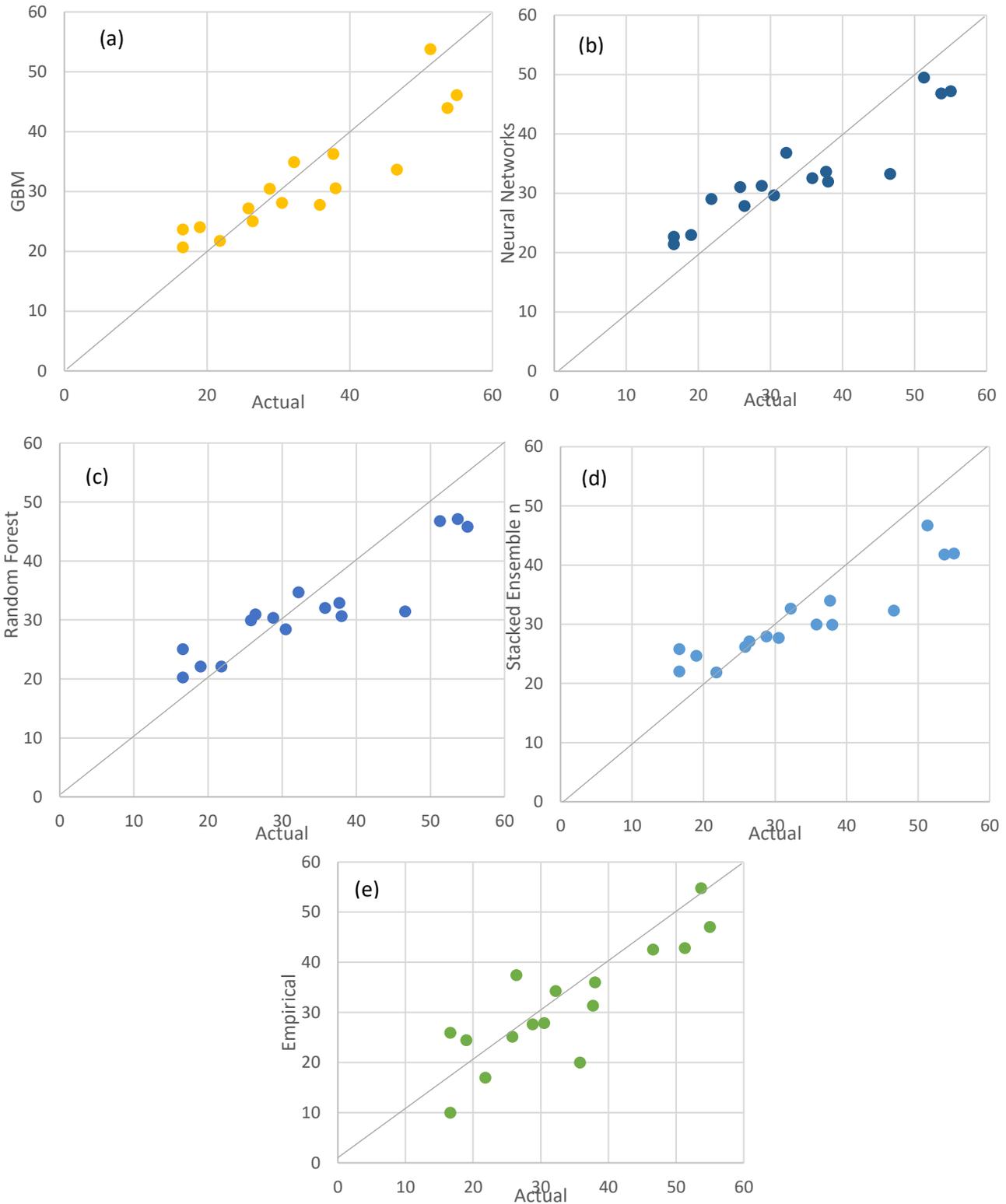


Figure 4. 22 Comparison between (a) GBM with rmse of 0.0603 (b) Neural Network had rmse of 0.0691 (c) Random Forest gave rmse of 0.0619 (d) Stacked Ensemble gave rmse of 0.0581 (e) Empirical predicted gave rmse value of 0.071 with actual (core) for the Brent validation set, Hutton Field, Northern North Sea.

Generally for all the test cases considered, the different machine learning models gave good predictions of irreducible water saturation demonstrating the power of this technique. Furthermore the Stacked Ensemble model's ability to combine the different capabilities of the individual base models to form a stronger irreducible water model was responsible for it giving the best performance of them all for this property. This is again illustrated with a plot of the different models against depth alongside gamma ray log (Figure 4.25). Figure 4.26 shows resistivity of the formation as the most sensitive for the GBM generation. Most of the irreducible water saturation values lie mainly between 10 to 50% unlike total water saturation which has values of up to 100% in some reservoirs especially when Archie is used.

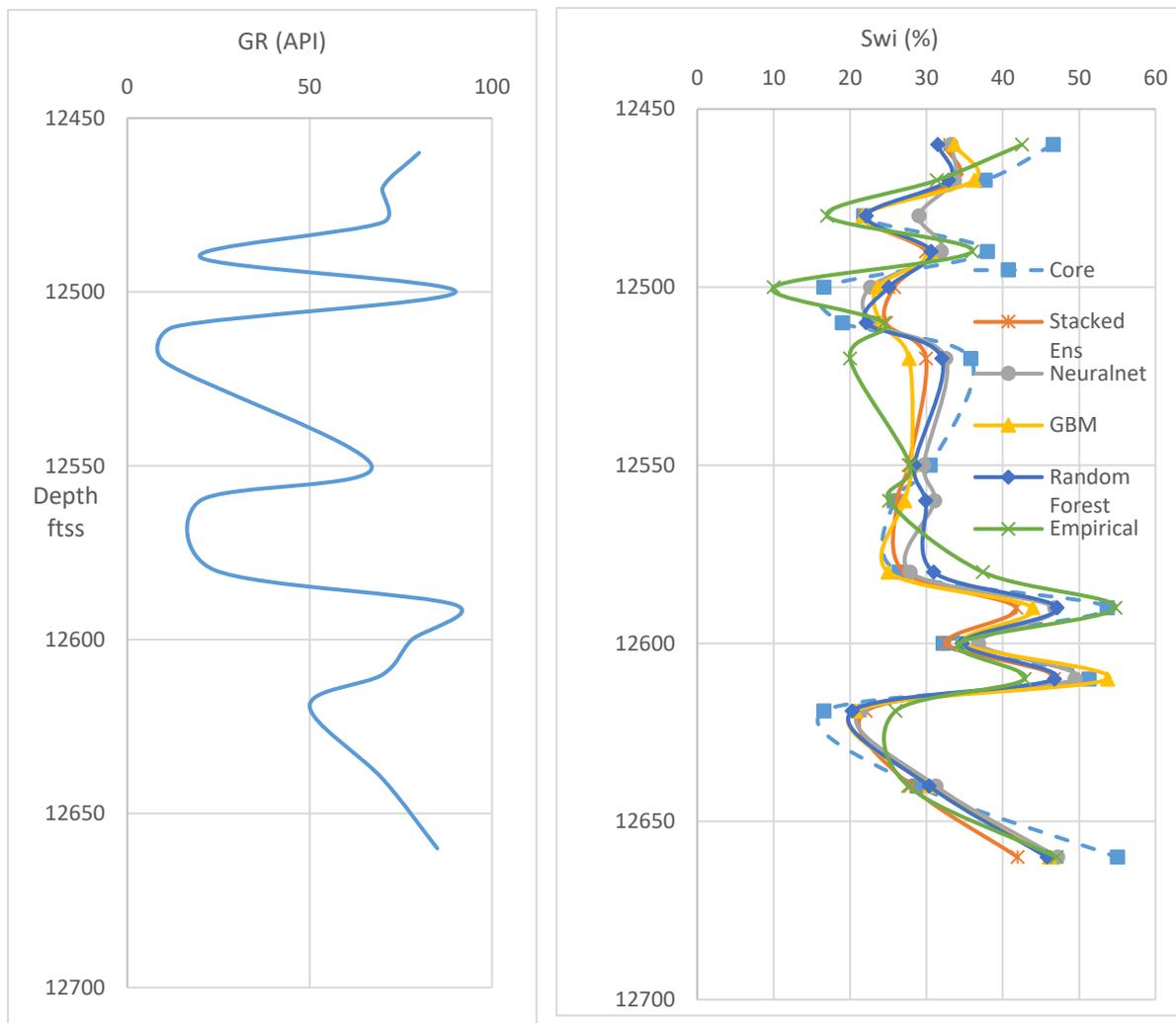


Figure 4. 23 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM, Random Forest, and Empirical Models for the Brent Reservoir, Hutton Field, Northern North Sea.

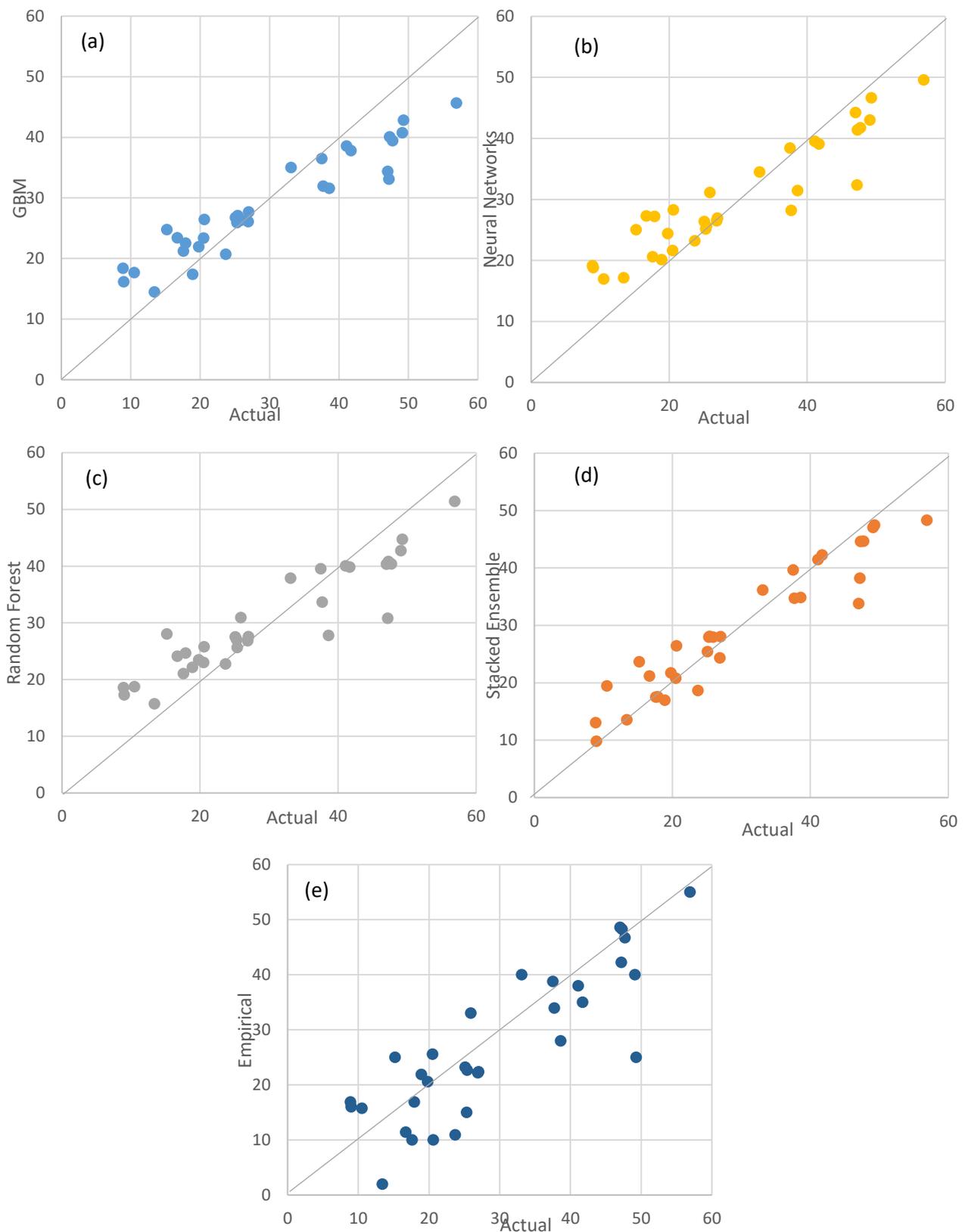


Figure 4. 24 Comparison between (a) GBM with rmse of 0.0626 (b) Neural Network had rmse of 0.0619 (c) Random Forest gave 0.063 for rmse (d) Stacked Ensemble had rmse of 0.0459 and (e) Empirical gave 0.0773 values for rmse with actual for the Test set, Hutton Field, Northern North Sea

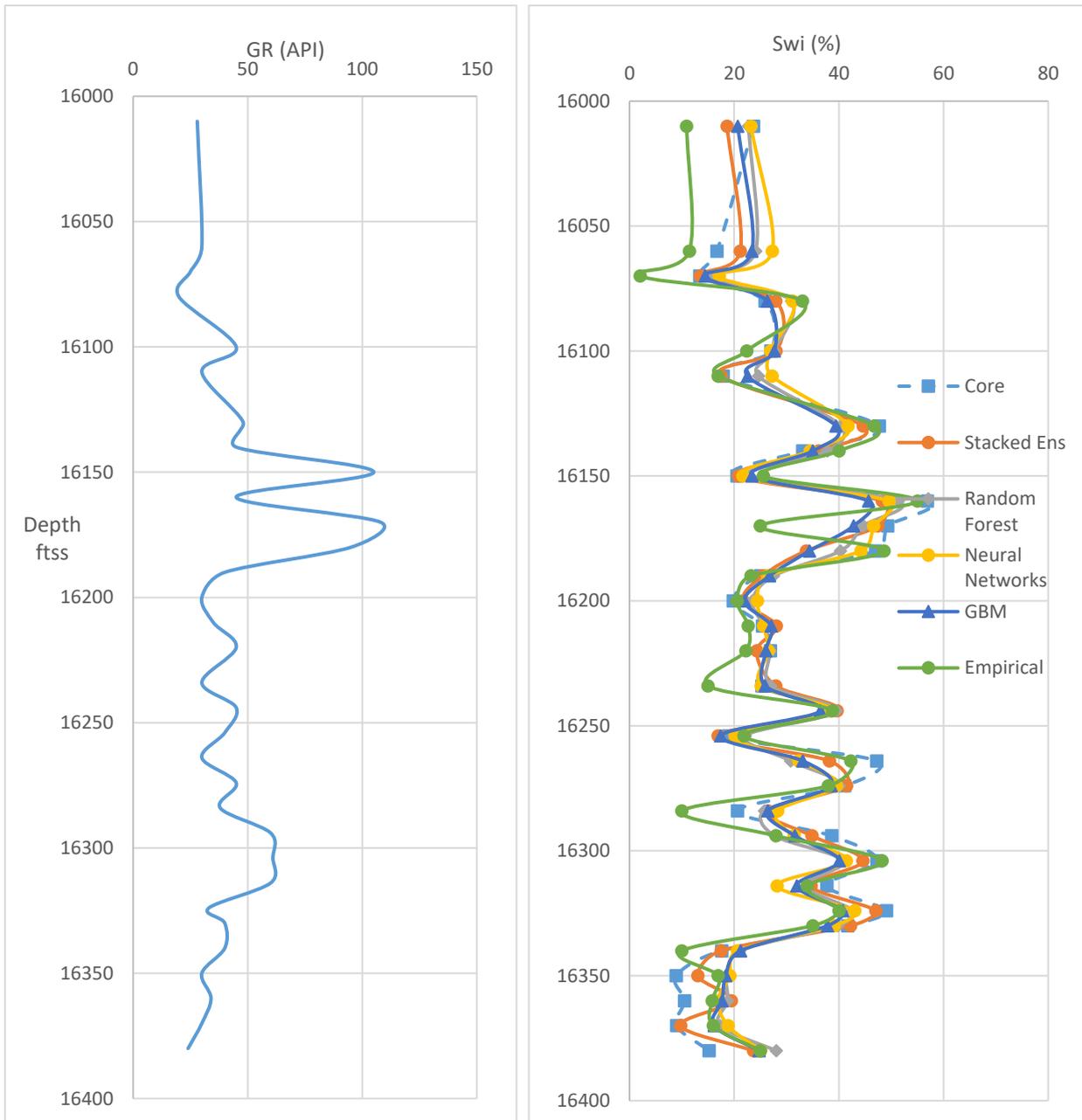


Figure 4. 25 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM, Random Forest, and Empirical Models for the Brent Reservoir, Hutton Field, Northern North Sea.

Figure 4.26 illustrates the sensitivity of the variables used for the irreducible water saturation Stacked Ensemble model showing the deep induction log to be most sensitive followed by porosity.

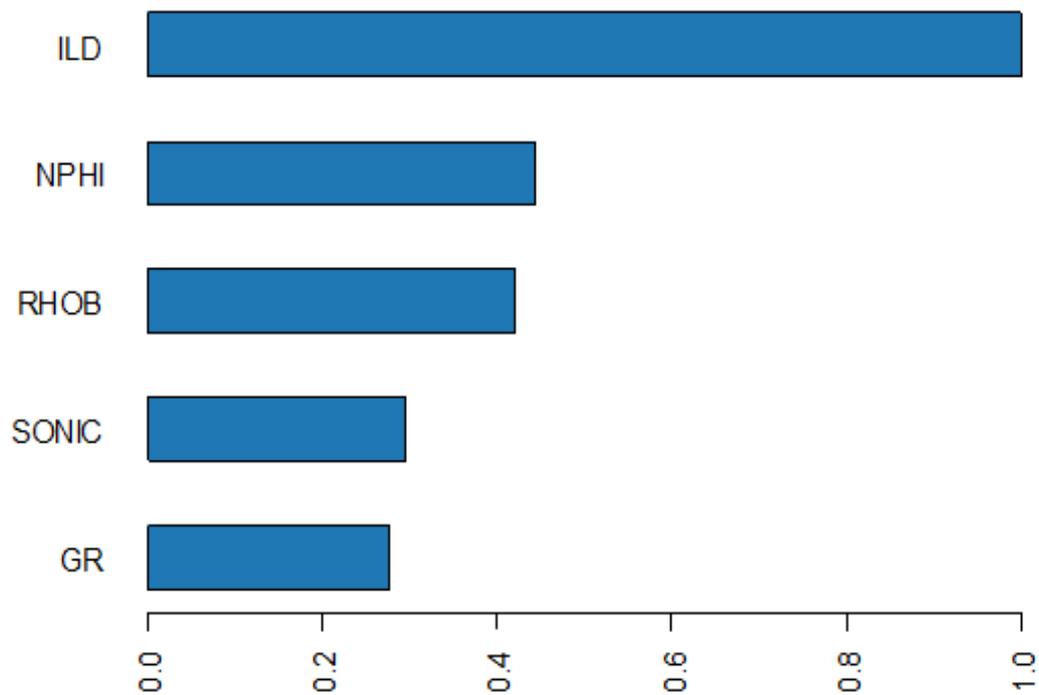


Figure 4. 26 Parameter sensitivity of the machine learning model and training scoring history.

4.2.3.3 Ensemble Residual Saturation Model

The importance of residual oil saturation to reservoir engineers cannot be over emphasized. The property is affected by wettability (Schneider and Owens, 1982) and heterogeneity (Sherborne et al. 1967). Also important to point out that there are four types of residual oil saturation after a water flooding and these are: oil ganglia trapped at pore throats mainly observed in water wet rocks. Its mobilization can be estimated in terms of capillary number (Stegemeier 1974; 1977); the second is oil locked in rock crevices and dead end pores associated with both oil and water systems; there is also the thin film coating rock surfaces observed in oil wet rocks while the fourth is residual oil from waterflooding in rocks with heterogeneity on a small scale (Chuh and Gary, 2008; Wreath, 1989). It is logical to say that the presence of one type of residual oil does not exclude the presence of the other in a reservoir. A machine learning approach has been adopted

for the determination of this property so as to help in accurate planning of enhanced oil recovery projects and give an accurate saturation value of any other phase(s) present in the reservoir.

Figure 4.27 shows Gradient Boosting Model giving values of 0.0611, Neuralnet has 0.0678, Random Forest gave 0.0632 while the Stacked Ensemble had values of 0.0539 for RMSE. It again shows that combining several models rightly, improves the accuracy of predictions. Figure 4.28 is a plot of these models with depth of the pay zone. Overall machine learning gave predictions within 5% error although the Ensemble model performed best. Accurately determining S_{or} will help to assess technical feasibility and profitability of oil and gas development projects.

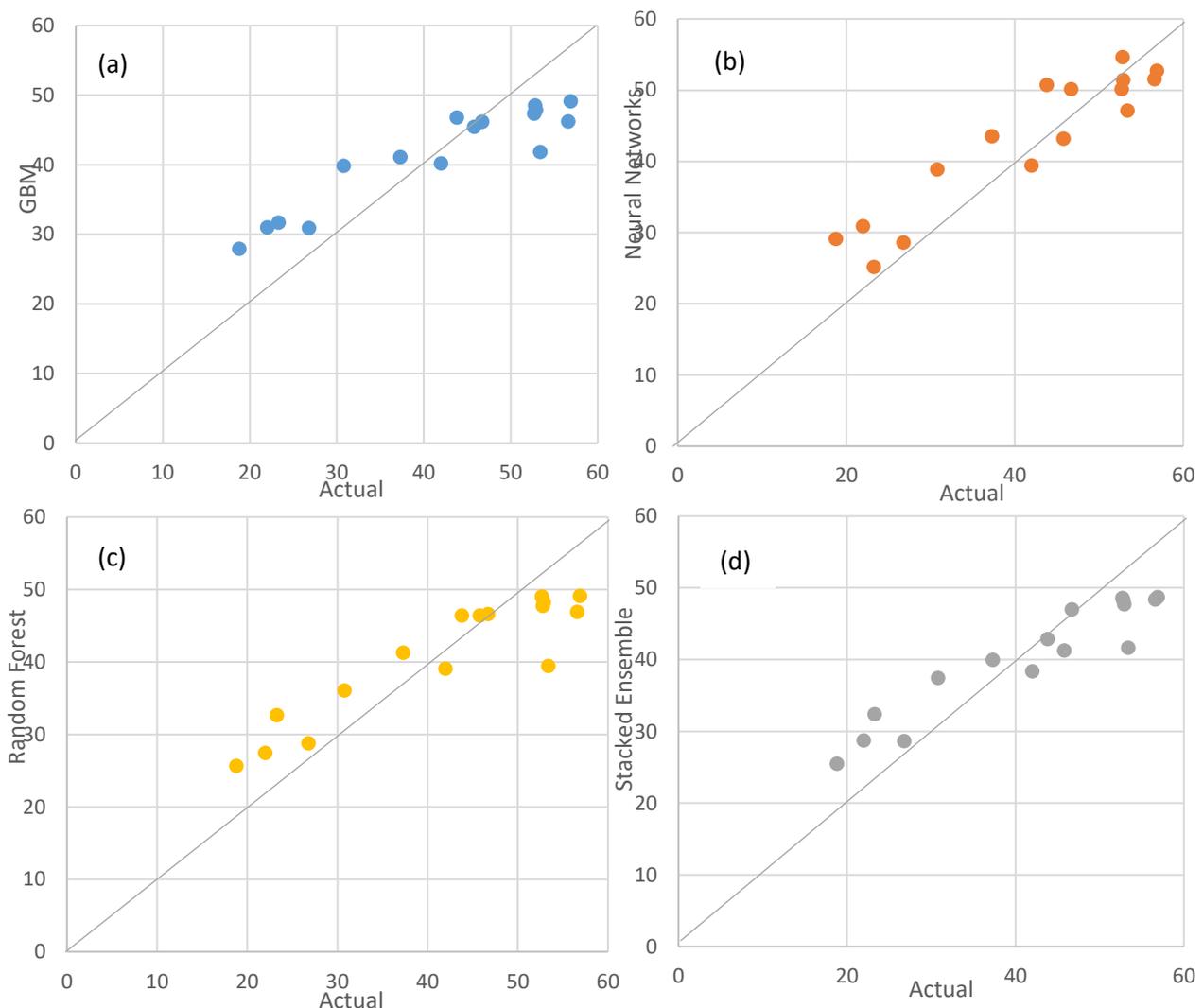


Figure 4. 15 Comparison between (a) GBM gave rmse values of 0.0611 (b) Neural Network had rmse of 0.0678 (c) Random Forest gave rmsee of 0.0632 and (d) Stacked Ensemble had rmse values of 0.0539 with actual for the validation set, Hutton Field, Northern North Sea.

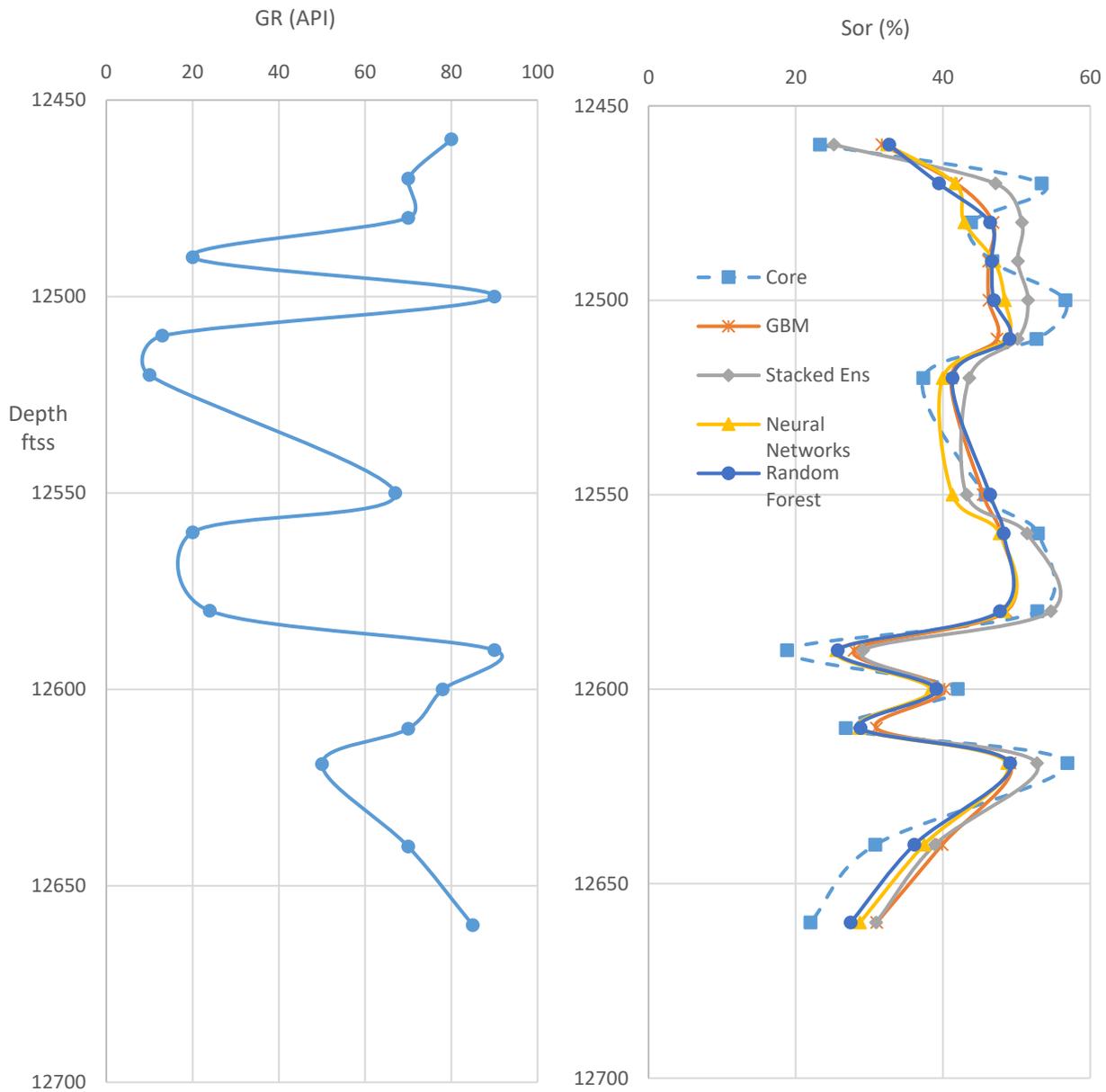


Figure 4. 16 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM and Random Forest models for the Brent Reservoir, Hutton Field, Northern North Sea.

Figure 4.29 describes prediction error for the test set. The GBM model gave values of 0.0694, for Neuralnet it was 0.0556, Random Forest had 0.0607 while Stacked Ensemble gave values of 0.0568 for Root Mean Square Error. Figure 4.30 shows a more vertical less heterogenous reservoir than Figure 4.28 as indicated by the gamma ray log. This supervised learning algorithm was able to use the optimal combination of these base models to make a better prediction.

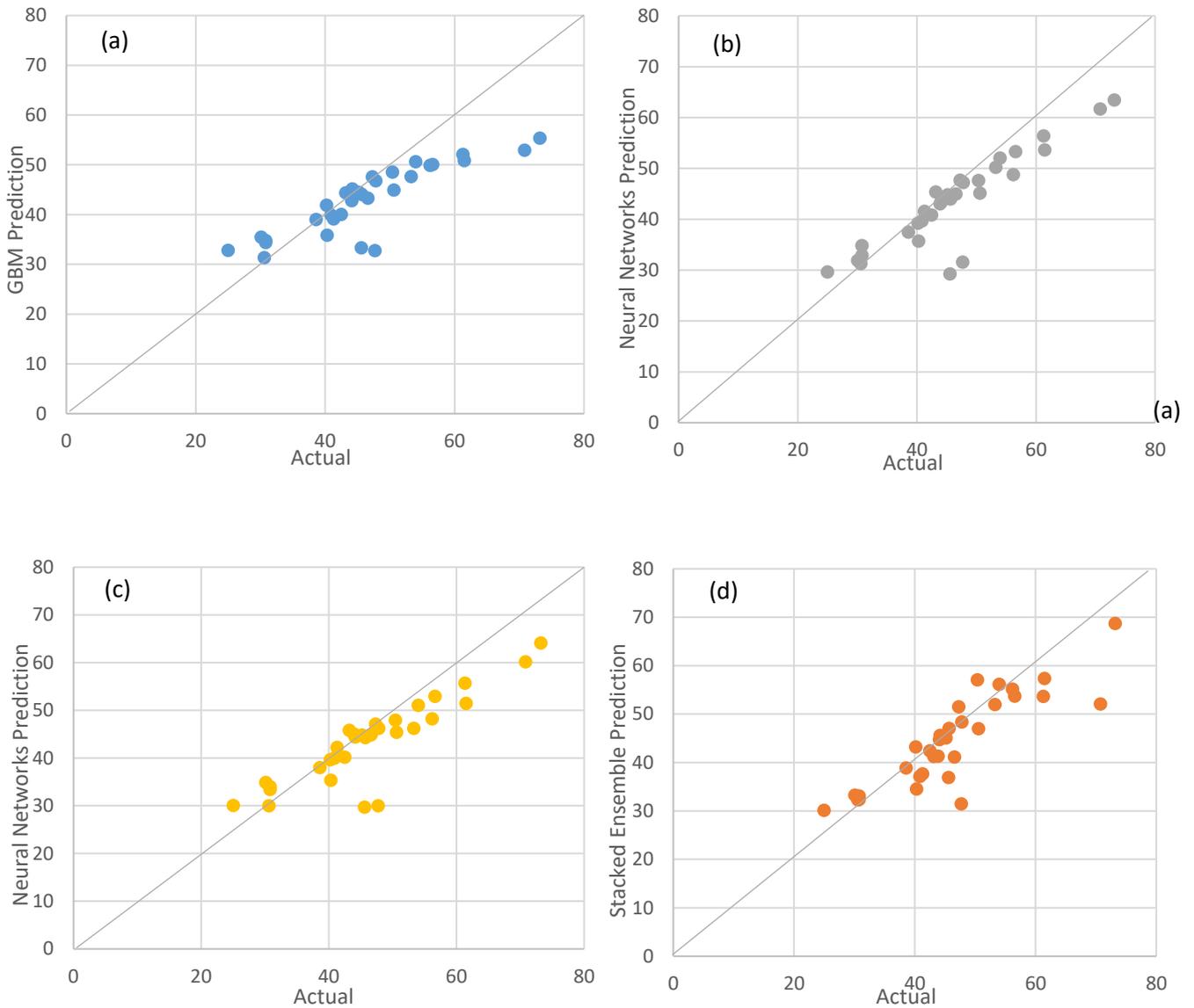


Figure 4. 17 Comparison between (a) GBM gave rmse of 0.0694 (b) Neural Network had rmse of 0.556 (c) Random Forest had rmse of 0.0607 and (d) Stacked Ensemble had rmse of 0.0568 with actual for the validation set, Hutton Field, Northern North Sea.

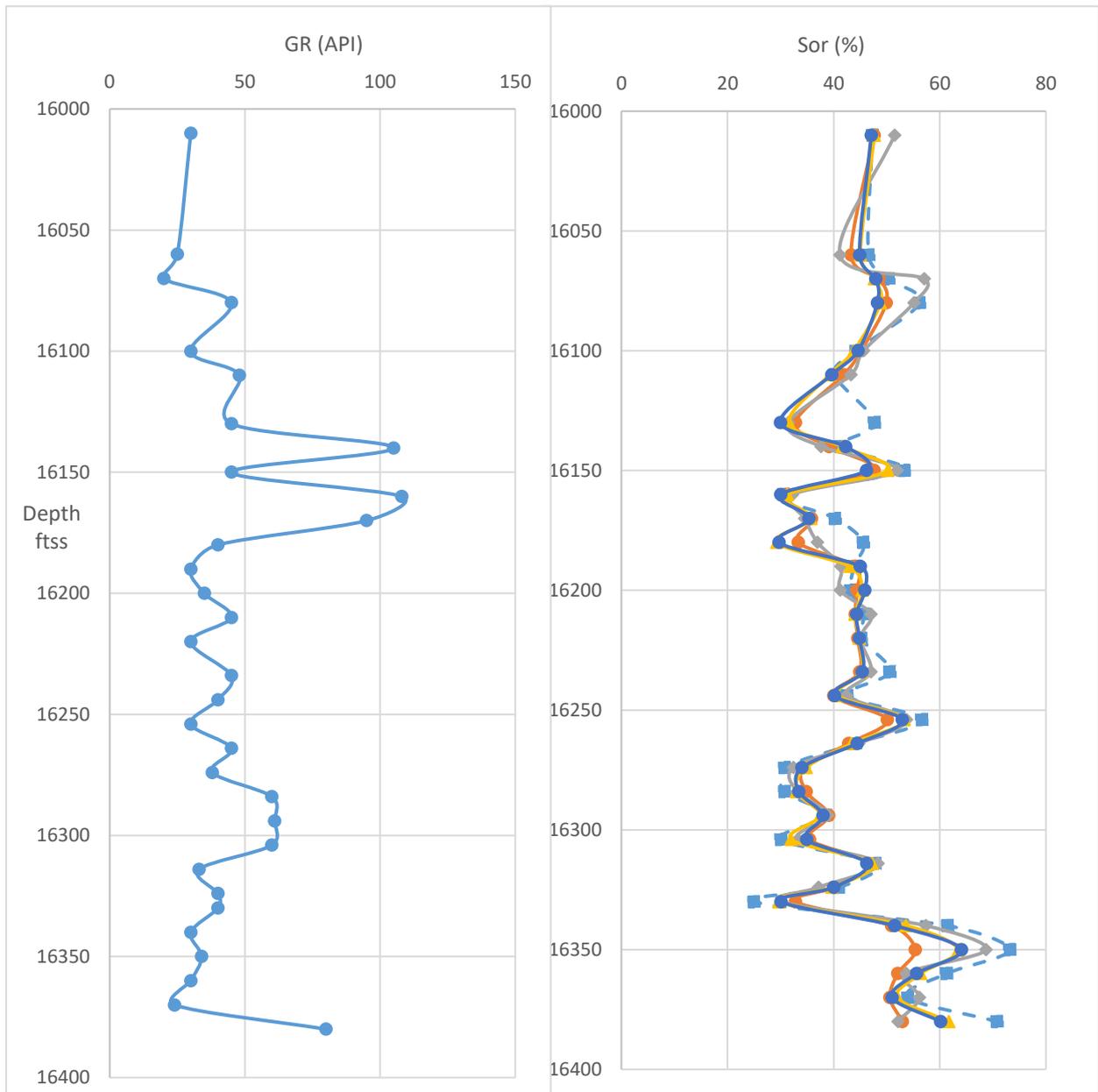


Figure 4. 18 Plot of Gamma Ray, Core, Stacked Ensemble, Neural Networks, GBM and Random Forest models for the Brent Reservoir, Hutton Field, Northern North Sea.

4.3 Real-Time Permeability from Logs

4.3.1 Introduction

Relative permeability is the most important property of porous media to carry out reservoir prognosis in a multiphase situation (Delshad and Pope 1989; Yuqi and Dacun 2004) and therefore needs to be as accurate and readily accessible as possible. Theoretically, it is the ratio of effective and absolute permeability. It is useful for the determination of reservoir productivity, effective mobility, wettability, fluid injection for EOR, late-life depressurization, gas condensate depletion with aquifer influx, injectivity, gas trapping, free water surface, residual fluid saturations, temporary gas storage amongst others (Figure 4.31). It is well known that a significant variation in relative permeability data can have a huge impact on a macroscopic scale.

The oil and gas industry has a need for easily available and reliable relative permeability data, expense reduction on experiments and a more general model for the parameter judging by the pitfalls pointed out by several researchers after testing the existing two and three phase relative permeability models (Table 4.6). Such workers like Fayers-Matthews (1984) and Juanes et al. (2006) after testing non-wetting relative permeability interpolation models such as Baker and Stone's I and II, against Saraf et al. (1982), Schneider and Owens (1970), Saraf and Fatt (1967) and Corey et al. (1956) experimental data, presented the same conclusion that they give similar results for high oil saturations but are different as it tends towards residual oil saturation. Manjnath and Honarpour (1984) concluded that corey gives higher values for non-wetting phase relative permeability after comparing against Donaldson and Dean data.

Based on the assumption that water and gas relative permeability depends only on their saturation and not on that of other phases, Delshad and Pope (1989) concluded after a comparative study of 7 relative permeability models that Baker and Pope performed better but also stated the need for better models. Siddiqui et al. (1999) found Wyllie-Gardner and Honarpour to yield consistently better results at experimental condition after testing 10 relative permeability models. Al-Fattah (2009) found Honarpour regression model to be the best after comparing with 5 other models and also developed his own regression model. Since the coefficients of these regression models are not generalized, they are not suitable for real time applications.

Furthermore, for wetting phase relative permeability in consolidated media, Li and Horne (2006) showed that the Purcell model best fits the experimental data in the cases studied by them provided the measured capillary pressure curve had the same residual saturation as the relative

permeability curve which is sometimes not the case. Saraf and McCaffery (1985) could not recommend a best model due to scarcity of three phase relative permeability data. The different relative permeability correlations have limitations and assumptions which no doubt have implications thus increasing the uncertainty in reservoir simulation studies hence the need for a more generalized model.

Therefore, the purpose of this study is to implement a generalised Deep Learning model for the prediction of relative permeability accounting for reservoir depletion, saturation and phase changes with time. Most of the reviewed models (Equation 4.19) are static but Deep Neural Networks if appropriately tuned can capture the transients faster and more accurately throughout the reservoir life while also getting better as more data becomes available with time (Arigbe et al., 2018). Training can be done offline and the trained networks are suitable for on-board generation of descent relative permeability profiles as their computation requires a modest CPU effort hence not a concern to real time application. A separate analysis was also done for absolute permeability with the aim of demonstrating their ability to model North Sea sandstone formations and thereby reducing the uncertainty in its determination from these existing models like Timur, Tixier, Morris and Biggs, Coates, Dual water. Generalized empirical models were also generated from Coates and Dumanoir model since they offer the flexibility interms of the relationship between permeability and irreducible water saturation.

Wyllie and Rose type models

$$k = C * \frac{\varphi^x}{S_{wi}^y} \quad (4.19)$$

Where φ is the porosity, S_{wi} is irreducible water saturation and C, x and y are constants.

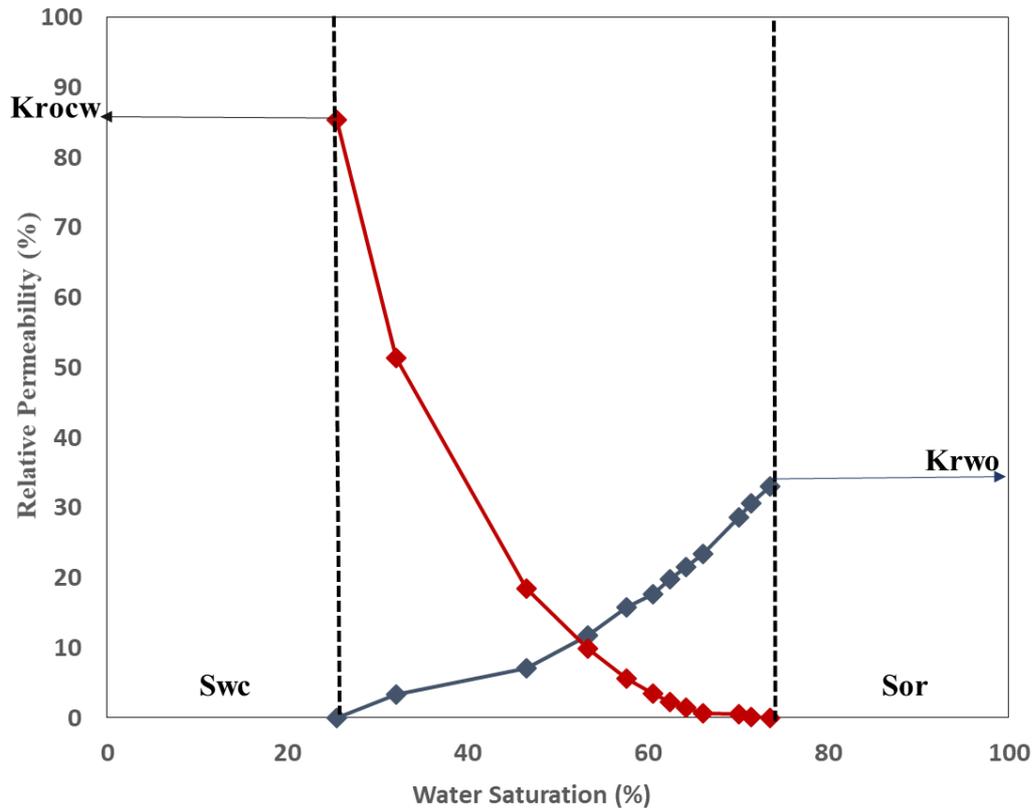


Figure 4. 19 Schematic of oil-water relative permeability curve

Table 4.7 shows the assumptions and applications of the commonly used two and three phase relative permeability models. It is a fact that relative permeability is dependent on saturation history as well as saturation. The direction of saturation change which could be imbibition or drainage referring to increase or decrease in wetting phase saturation respectively. The models shown in the table are based on the idea that the nonwetting phase is partly immobile and partly mobile during saturation change in the imbibition direction as well as the assumption that the amount of entrapment at any saturation could be derived from the connection between non wetting phase saturations in the drainage direction and residual saturations when imbibition is finished.

The weighting factors $(S_w - S_{wc})$, $(S_g - S_{gr})$ and $(S_o - S_{or})$ appears in both the numerator and denominator must be positive and at worse zero. L'Hopital's rule has to be applied to either saturation to get their proper limit, in the indeterminate case where both the numerator and denominator are zero. L'Hopital's rule which is often applied to provide limits of indeterminate form and converting it to a mathematical expression which can then be easily evaluated by substitution. The values of K_{row} and K_{rog} were obtained from the oil/water and gas/oil relative permeability data interpolated at the actual oil and gas saturation respectively.

Table 4. 7 Assumptions and application of the commonly used two and three phase relative permeability correlations

MODEL	CORRELATION	PHYSICS	ASSUMPTIONS	APPLICATION WINDOW
Corey et al. (1956)	$K_{ro} = \left(\frac{S_o - S_{or}}{1 - S_{or}} \right)^{\frac{2+3\lambda}{\lambda}}$ $K_{rg} = \left(\frac{1 - S_o}{1 - S_{or}} \right)^2 \left[1 - \left(\frac{S_o - S_{or}}{1 - S_{or}} \right) \right]^{\frac{2+\lambda}{\lambda}}$	An extension of Purcell (1941) and Burdine (1953) which is based on the mean hydraulic radius concept of Kozeny-Carman (bundle of capillaries model) for each pore size in a rock with large variety of pores and tortuosity expressed in terms of fluid saturation.	$K_{ro} \propto$ oil pore area and saturation of water and gas phases Relative permeability of wetting and non-wetting phase independent of saturation of other phases.	Requires a single suite of K_{rg}/K_{ro} data at constant S_w to calculate K_{rg} and K_{ro} for all saturations Not flexible to force end points of isoperms to match measured data. Applies only to well-sorted homogenous rocks.
Wyllie (1951)	$K_{rw} = \left(\frac{S_w - S_{wc}}{1 - S_{wc}} \right)^4$ $K_{rg} = \frac{S_g^2 [(1 - S_{wc})^2 - (S_w + S_o - S_{wc})^2]}{(1 - S_{wc})^4}$ $K_{ro} = \frac{S_o^3 (2S_w + S_o - 2S_{wc})}{(1 - S_{wc})^4}$	Based on bundle of capillaries cut and rejoined along their axis with related entrapment of the wetting phase.	Considers irreducible water as part of the rock matrix.	Applied when water saturation is at irreducible level.
Stone (1970)	$K_{ro} = S_o^* \left(\frac{K_{row}}{1 - S_w^*} \right) \left(\frac{K_{rog}}{1 - S_g^*} \right) = S_o^* \beta_w \beta_g$	Based on the channel flow theory which states that in any flow channel, only one fluid is mobile hence the basis for β_w and β_g is tied to the notion of identical microscopic fluid distribution around a two-phase interface	Assumes gas/oil displacement at zero water saturation Makes use of a probability model to estimate three phase relative permeability.	Experimental evidence shows that S_{om} (irreducible residual oil saturation) should be in the range $0.25 * S_{wc}$ to $0.5 * S_{wc}$
Stone (1973)	$K_{ro} = (K_{row} + K_{rw})(K_{rog} + K_{rg}) - (K_{rw} + K_{rg})$	Permeability to oil at irreducible water saturation and zero gas saturation.	Assumes total permeability is the product of total water/oil and gas/oil permeabilities Uses a revised probability model for better fit with experimental data.	K_{row} and K_{rog} must be unity at irreducible water saturation for accurate estimation of two-phase relative permeability as third phase saturation tends to zero.

<p>Honarpour et al. (1982)</p>	<p>Water wet</p> $K_{rw} = 0.035388 \left(\frac{S_w - S_{wc}}{1 - S_{wc} - S_{orw}} \right) - 0.0108074 \left(\frac{S_w - S_{orw}}{1 - S_{wc} - S_{orw}} \right)^{2.9} + 0.56556(S_w)^{3.6}(S_w - S_{wc})$ <p>Any wettability</p> $K_{ro} = 0.76067 \left[\frac{(S_o/1 - S_{wc}) - S_{or}}{1 - S_{orw}} \right]^{1.8} \left[\frac{S_o - S_{orw}}{1 - S_{wc} - S_{orw}} \right]^{2.0} + 2.6318\phi(1 - S_{orw})(S_o - S_{orw})$ $K_{rg} = 1.1072 \left(\frac{S_g - S_{gr}}{1 - S_{wc}} \right)^2 K_{rgo} + 2.7794 S_{org} \left(\frac{S_g - S_{gr}}{1 - S_{wc}} \right) K_{rgro}$	<p>Based on proposed empirical relationships describing experimentally determined permeabilities.</p>	<p>Assumes normally distributed variables</p>	<p>New constant will have to be developed for other areas to have a good fit.</p>
<p>Parker et al. (1987)</p>	$K_{ro} = (\bar{S}_t - \bar{S}_w)^{1/2} \left[(1 - \bar{S}_w^{1/m})^m - (1 - \bar{S}_t^{1/m})^m \right]^2$	<p>Based on relative permeability, saturation-fluid pressure functional relationships with a flow channel distribution model in two or three phase flow subject to monotonic saturation path and to estimate effective mean fluid conducting pore dimensions.</p>	<p>Wettability takes the water > oil > gas sequence. Irreducible fluid saturation is independent of fluid properties or saturation history No Gas/water contact occurs in the three phase region until the level where oil exists as discontinuous bolbs or pendular rings</p>	<p>Limited to cases where a satisfactory fit to the two-phase data is provided by the fitting equations $m = 1 - 1/n$</p>
<p>Baker (1988)</p>	$K_{ro} = \frac{(S_w - S_{wc})K_{row} + (S_g - S_{gr})K_{rog}}{(S_w - S_{wc}) + (S_g - S_{gr})}$ $K_{rw} = \frac{(S_o - S_{or})K_{rwo} + (S_g - S_{gr})K_{rwg}}{(S_o - S_{or}) + (S_g - S_{gr})}$ $K_{rg} = \frac{(S_o - S_{or})K_{rgo} + (S_w - S_{wc})K_{rgw}}{(S_o - S_{or}) + (S_w - S_{wc})}$	<p>As the saturation of a phase tends to zero, that of the other two-phase will dominate.</p>	<p>The end points of the three phase relative permeability isoperms coincide with the two-phase relative permeability data.</p>	<p>Weighting factors $(S_w - S_{wc})$ and $(S_g - S_{gr})$ must be both positive</p>
<p>Frode (2005)</p>	$K_{row} = K_{ro}^x \frac{(1 - S_{wn})^{L_w^o}}{(1 - S_{wn})^{L_w^o} + E_w^o S_{wn}^{T_w^o}}$ $K_{rw} = K_{rw}^o \frac{S_{wn}^{L_w^o}}{S_{wn}^{L_w^o} + E_w^o S_{wn}^{T_w^o}}$	<p>Based on the mean hydraulic radius concept of Kozeny-Carman (bundle of capillaries model)</p>	<p>Assumes that the whole spectrum of the relative permeability curve can be captured with the L, E, T parameters.</p>	<p>It exhibits enough flexibility to reconcile the entire spectrum of experimental data.</p>

	$S_{wn} = \frac{S_w - S_{wi}}{1 - S_{wi} - S_{orw}}$			
--	--	--	--	--

4.3.2 Methodology

The most commonly available factors influencing relative permeability such as porosity, ϕ , viscosity, μ , permeability, k , saturation, S_w , together with Baker and Wyllie parameter combinations were used as inputs for the network. Baker gave correlation coefficients of 0.96 and 0.86 while Wyllie has correlation coefficients of 0.91 and 0.89 for Corey and Leverett-Lewis datasets respectively (Table 4.8). The table was compiled by Baker 1988 using different models (left column labels) on existing datasets (top row label). Modified by added the last row (Wyllie) to it.

Ten (10) sets of water-oil relative permeability data with 132 data points from a NorthSea field with four-fifths used as training set and one-fifth as validation set. Another set of water-oil relative permeability data from a separate field were used as the testing set after data wrangling and normalization. A seed value was set to ensure the repeatability of the model. An optimised number of hidden layers (100) was used to reduce the need for feature engineering. The best cross validation result in a 5-fold arrangement was automatically used to train the DNN models until convergence using Nesterov accelerated gradient descent (which minimize their cost function). The rectifier activation function were used in the DNN modelling to increase the nonlinearity of the model, significantly reduce the difficulty in learning, improve accuracy and can accept noise (Equation 4.20). This allows for effective training of the network on large and complex datasets making it helpful for real time applications.

$$f(x) = \max(0, x + Y) \quad (4.20)$$

Where $Y \sim \mathcal{N}(0, \sigma(x))$ is the Gaussian noise applied to the rectifier linear unit.

They were then validated and tested to check the generalization and stability of the models for out of training sample applications. The aim is infer relative permeability data for core samples where no flood experiments have been carried out based on existing data.

Table 4. 8 Comparison of relative permeability models (vertical) with different datasets (horizontal) using correlation coefficient (Modified after Baker 1988)

DATA	COREY	LEVERETT AND LEWIS	REID	SNELL	SARAF ET AL	HOSAIN	GUCK ERT
STONE I	0.97	0.76	0.90	0.57	0.82	0.85	0.48
STONE 11	0.77	0.75	0.87	0.75	0.68	0.33	0.50
AZIZ AND SETARRI	0.8	0.75	0.95	0.75	0.74	0.9	0.48
COREY	0.88	0.83	0.89	0.48	0.50	0.74	0.6
BAKER	0.96	0.86	0.88	0.58	0.9	0.84	0.57
NAAR AND WYGAL	0.74	0.67	0.78	0.50	0.55	0.54	0.50
PARKER	0.85	0.73	0.88	0.56	0.87	0.93	0.52
LAND	0.93	0.8	0.89	0.50	0.66	0.74	0.55
WYLLIE	0.91	0.89	-	-	-	-	-

The developed Deep Neural Networks model was further applied to predict other experimental data carried out based on Buckley-Leverett (1942) frontal advance theory (Figure 4.32) and Welge (1952) method for average water saturation behind the water front. Permeability is a tensor (time dependent), hence relative permeability was obtained from saturation change with time that is, the developed Deep Learning model made predictions of relative permeability as a function of time from the saturation history.

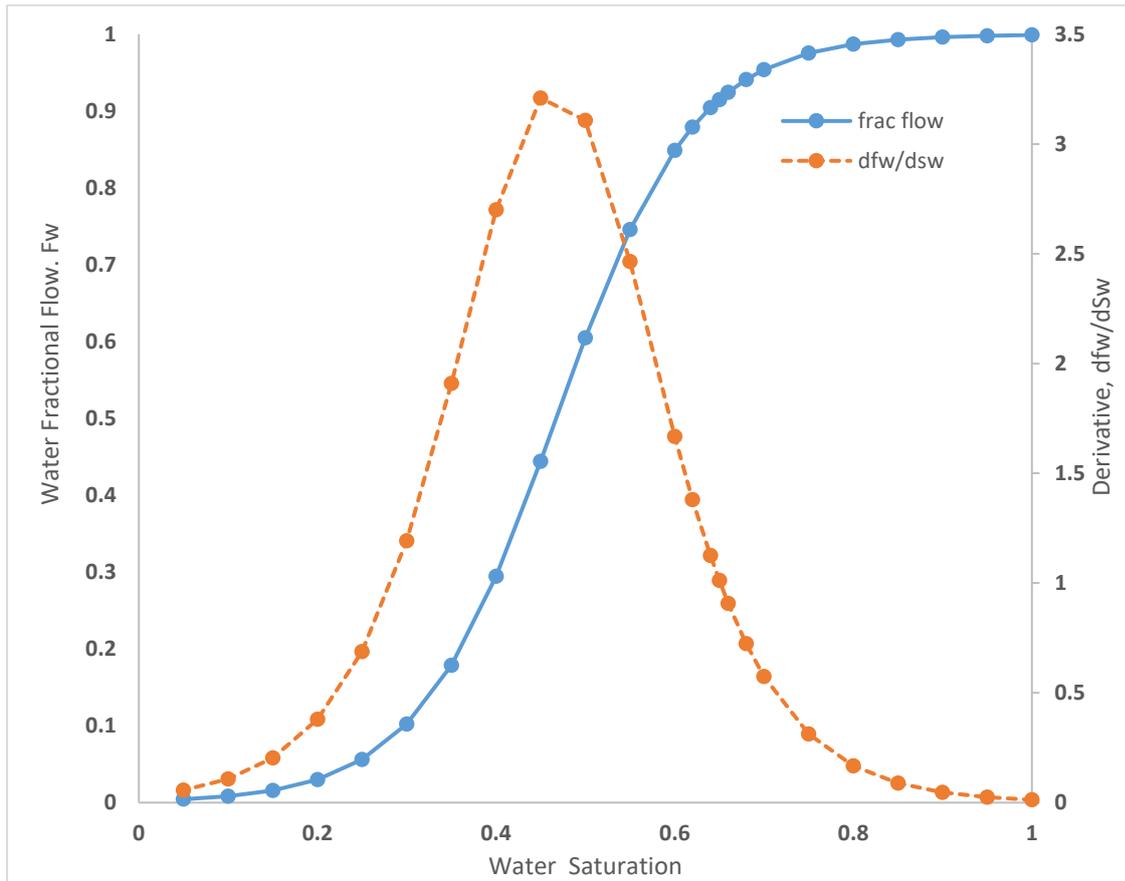


Figure 4. 20 Water fractional flow curve with its derivative for the field considered.

4.3.2.1 Deep Neural Network

DNN is a feed-forward, artificial neural network with more than one layer of hidden units between its inputs and outputs (Figure 4.33). The ability of the model to transfer to a new context and not over-fit to a specific context (generalization) was addressed using cross validation which is described in detail below. All networks were trained until convergence with Nesterov accelerated gradient descent which also minimizes the cost function (which measures the performance of a machine learning for the data). In addition, both λ_1 and λ_2 regularization (Equation 4.21) were used to add stability and improve the generalization of the model.

Mathematically,

$$J(\theta) = \frac{1}{2} \sum_{i=1}^n (\theta^T x^{(i)} - y^{(i)})^2 + \lambda \sum_{j=1}^p \theta_j^2 \quad (4.21)$$

Where x are inputs, θ are parameters, J is the regularized objective function, λ is the tuning parameter which decides how much the flexibility of our model should be penalized represented as $\mathbf{l1}$ or $\mathbf{l2}$ (preferred to $\mathbf{l0}$ for convexity reasons). Regularization helps to reduce the variance in

predicting test data without increasing bias thereby reducing overfitting. DNN are well suited for modelling systems with complex relationships between input and output (Burke, 1992; Hubick, 1992) which is what is obtainable in natural earth systems. In such cases with no prior knowledge of the nature of non-linearity, traditional regression analysis is not adequate (Gardner and Dorling, 1998). It has been successfully applied to real time speech recognition, computer vision, optimal space craft landing etc.

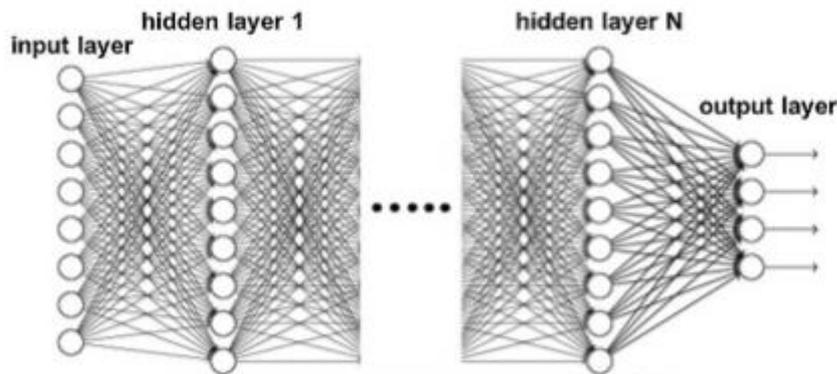


Figure 4. 21 Deep Neural Network model architecture showing input, hidden and output layers (Lee et al. 2017).

Again the scale issue is a major concern for most algorithms. We construct bigger neural networks and continually train them with more and more data (which is the case for real time applications) their performance continues to increase. Performance refers to the prediction accuracy of the model for new data not used in the training set. This is generally not the case for other machine learning algorithms whose performance reaches a plateau (Figure 4.34).

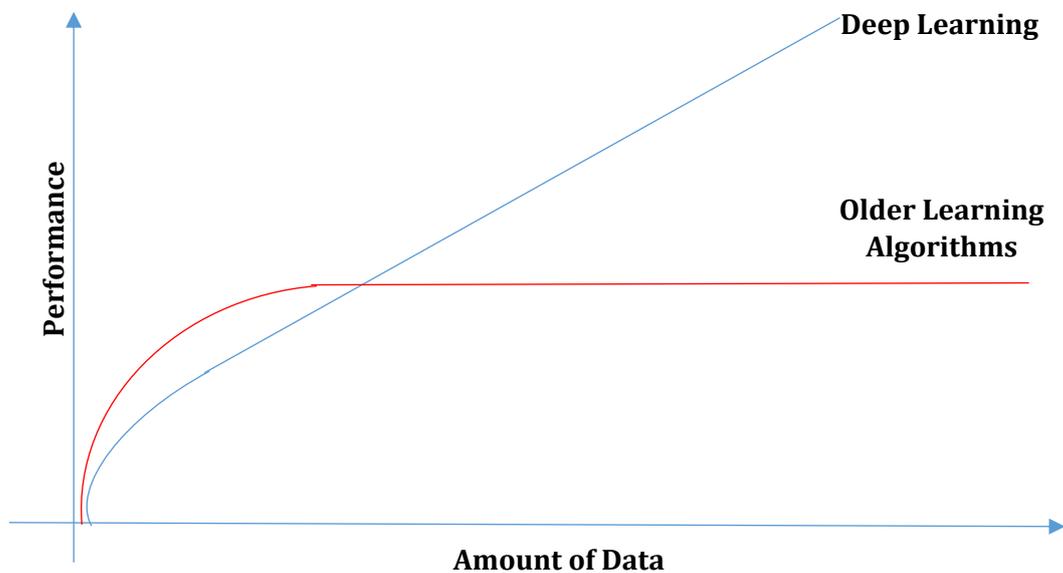


Figure 4. 22 How data science techniques scale with amount of data (Andrew, 2013)

4.3.2.2 Cross Validation

Overfitting which is the single major problem of prediction when independent datasets is used was reduced through cross validation by estimating out of sample error rate for the predictive functions built to ensure generalisation. Other issues like variable selection, choice of prediction function and parameters and comparison of different predictors were also addressed. A 5-fold cross validation technique was used to split the data set into training and test set, build a model on the training set, evaluate on the test set and then repeat and average the errors estimated. A weight decay was chosen to improve the generalization of the model by suppressing any irrelevant component of the weight vector while solving the learning problem with the smallest vector. This also suppresses some of the effects of static noise on the target if chosen correctly.

4.3.3 Results and Discussion

4.3.3.1 Absolute Permeability Modelling of the North Sea

Figures 4.35, 4.36 & 4.37 are cumulative curves showing the permeability distribution for Northern, Central and Southern North Sea. Different permeability datasets from sandstone reservoirs in the North Sea were plotted into a cumulative distribution curve to show the distribution of permeability in the different regions of the North Sea. It appears that the median permeability range in the Northern North Sea is between 100 to 500mD (Figure 4.35), the Southern North Sea has a median permeability of 0.1 to 1.0mD (Figure 4.36) while it is between 1 and 10mD in the Central North Sea. Generally absolute permeability in the North Sea ranges from 0.0001mD to about 3692mD. The median range of values for the Northern North Sea appears higher than both for Central and Southern North Sea. This distribution gives a better understanding of how varied the datasets from the North Sea is. The median values are the mid points in the distribution of permeability in Northern, Central and Southern North Sea.

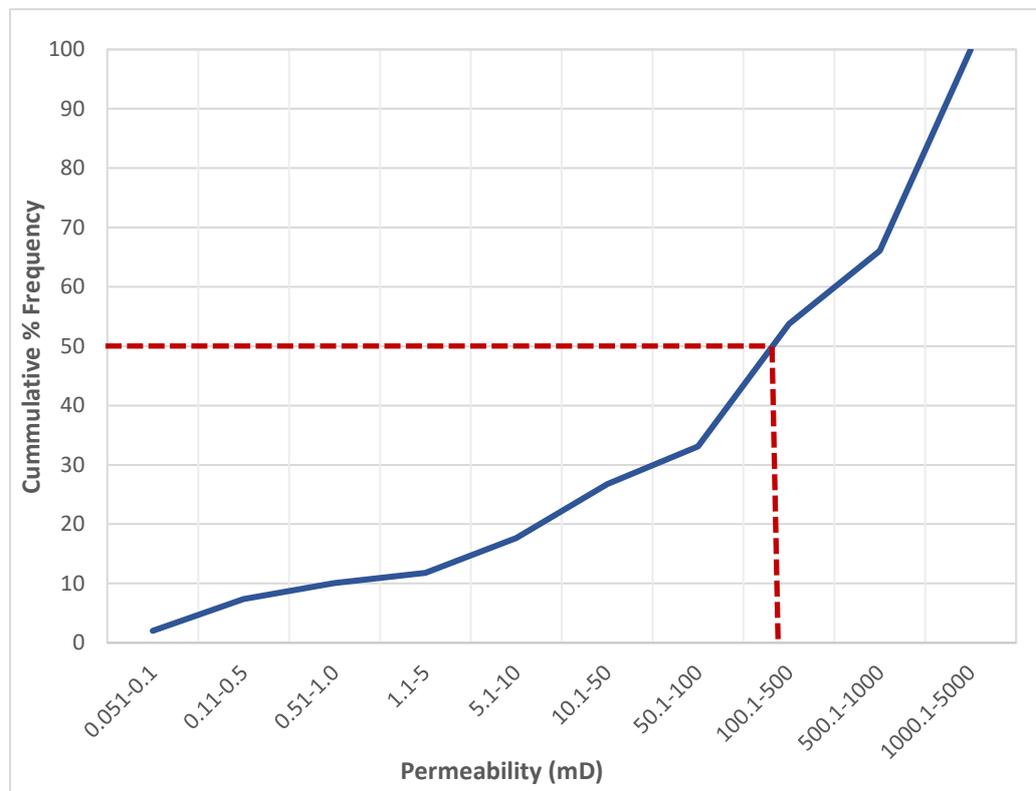


Figure 4. 23 Permeability distribution in Northern North Sea sandstone reservoirs.

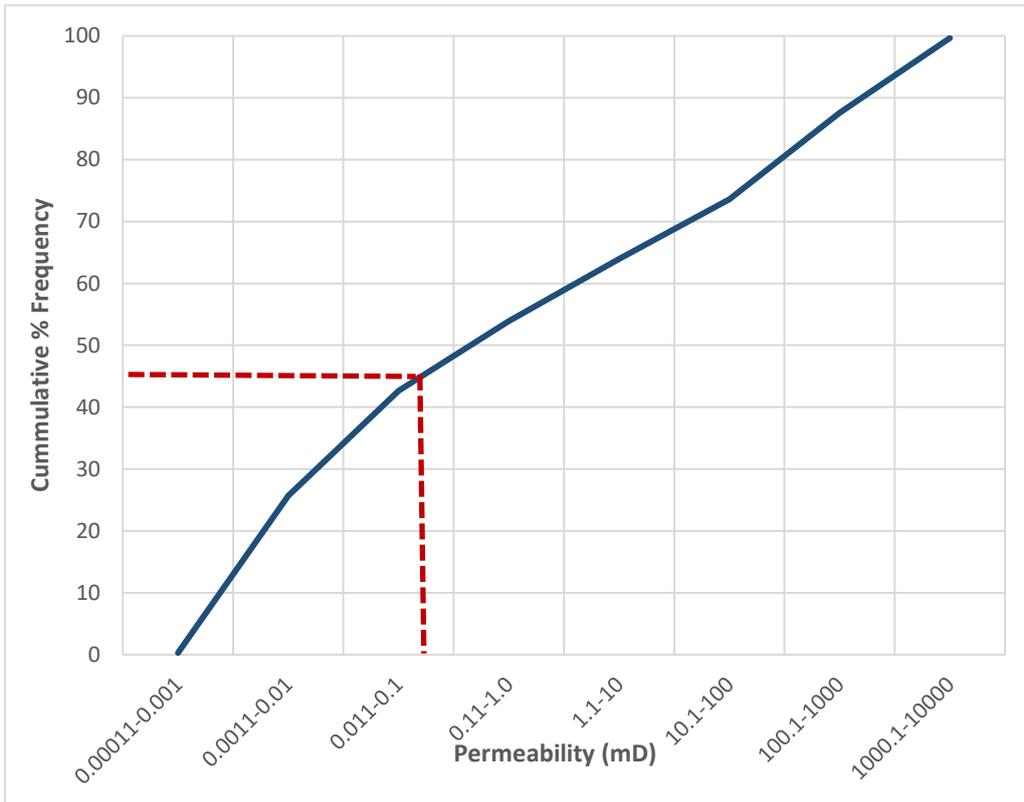


Figure 4. 24 Permeability distribution in Southern North Sea sandstone reservoirs.

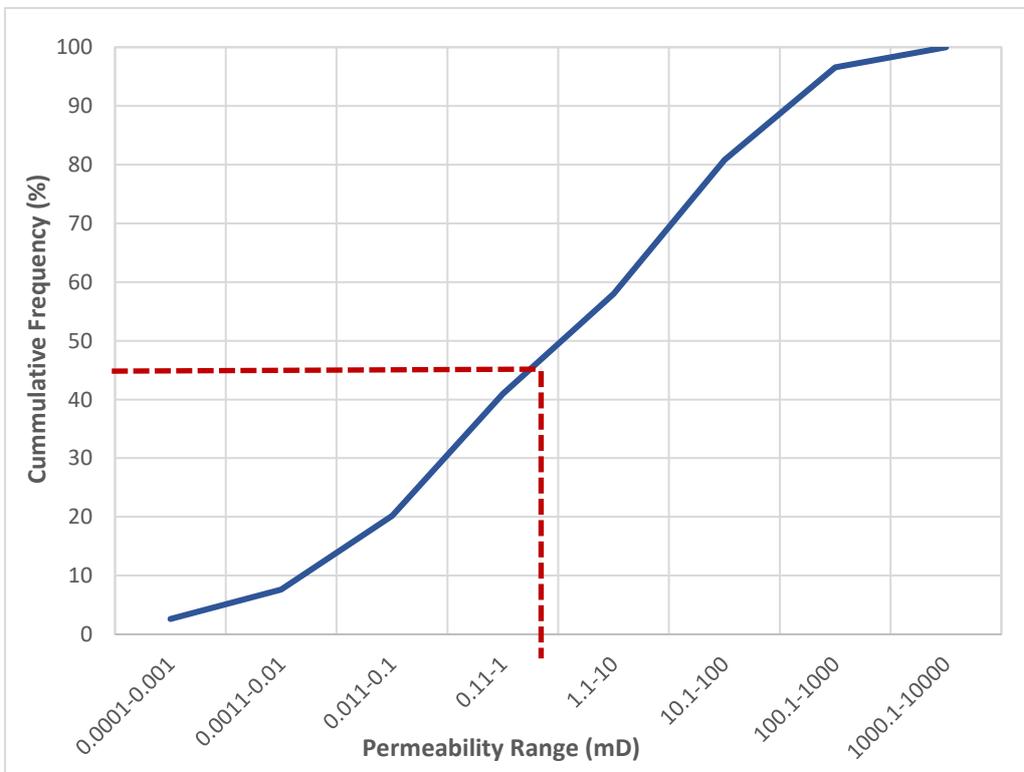


Figure 4. 25 Permeability distribution in Central North Sea sandstone reservoir.

The different absolute permeability models have been dealt with in Chapter Two. Several of the most commonly used absolute permeability models like Timur, Tixier, Morris and Biggs, Coates were modelled for the North Sea as shown in Figures 4.38-4.39. As can be seen from the figure below, none of the existing models perfectly models the North Sea data although Timur gave the closest. Ahmed et al. (1991) also pointed out that the Wyllie-type models generally suffer from the difficulty of getting zero permeability as irreducible water saturation approaches 100% and when porosity approaches zero hence they might disregard irreducible water approaching 100% and honour only the porosity limit for permeability approximation that are zero. Since the Coates and Dumanoir (1973) free fluid model allows zero permeability at zero porosity and 100% irreducible water saturation, it was modified to model absolute permeability for North Sea sandstone reservoirs. The original Coates and Dumanoir (1973) model is shown in Equation 4.25 while the Modified Coates and Dumanoir model for the North Sea is given by Equation 4.26 although more data is needed to get a more confident model (it is not the model in Coates original paper but Equation 4.22 to 4.25 are the original models).

Timur 1968

$$k = 0.136 \frac{\phi^{4.4}}{S_{wi}^2} \quad (4.22)$$

Tixier 1949

$$K^{1/2} = 250 \frac{\phi^3}{S_{wi}} \quad (4.23)$$

Kozeny-Carman 1927

$$k = A_1 \frac{\phi^3}{S_o^2(1 - \phi)^2} \quad (4.24)$$

Coates and Dumanoir (1973)

$$K^{0.5} = 70 * \frac{\phi^2(1 - S_{wi})}{S_{wi}} \quad (4.25)$$

Modified Coates and Dumanoir

$$K = 0.004 * \frac{\phi^{2.7}(1 - S_{wi})}{S_{wi}} \quad (4.26)$$

Figure 4.38 illustrates the original Coates model for the North Sea while Figure 4.39 shows the good fit of the modified Coates model in logarithmic porosity scales.

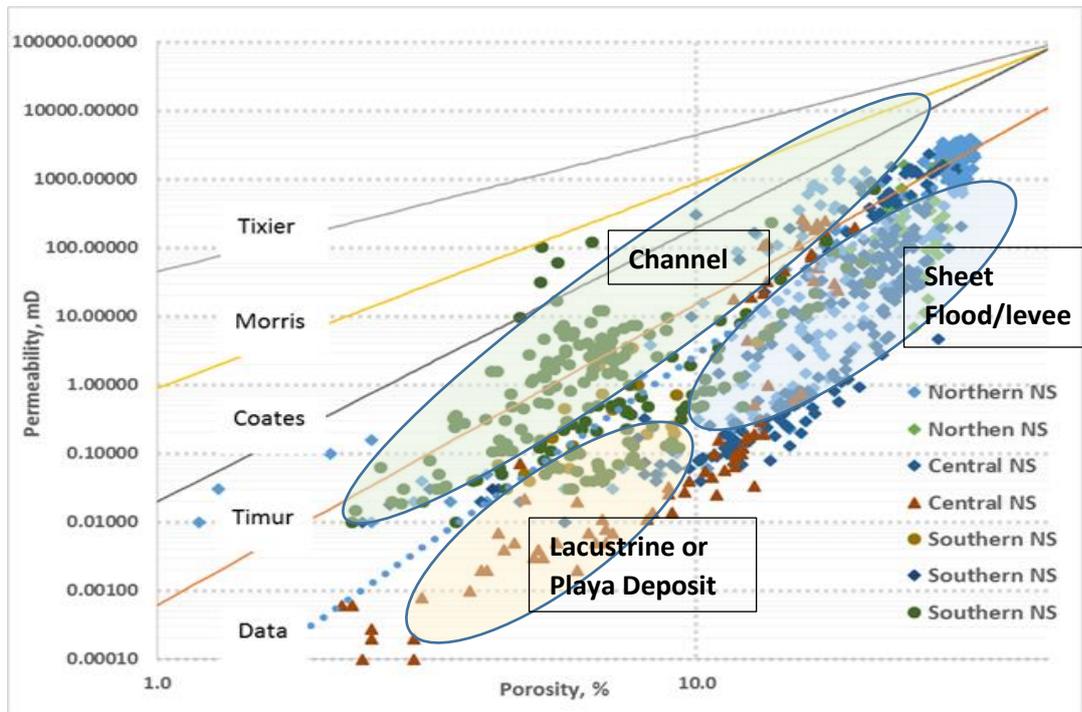


Figure 4. 26 Distribution of permeabilities across the Northern, Central and Southern North Sea indicating their different depositional environments

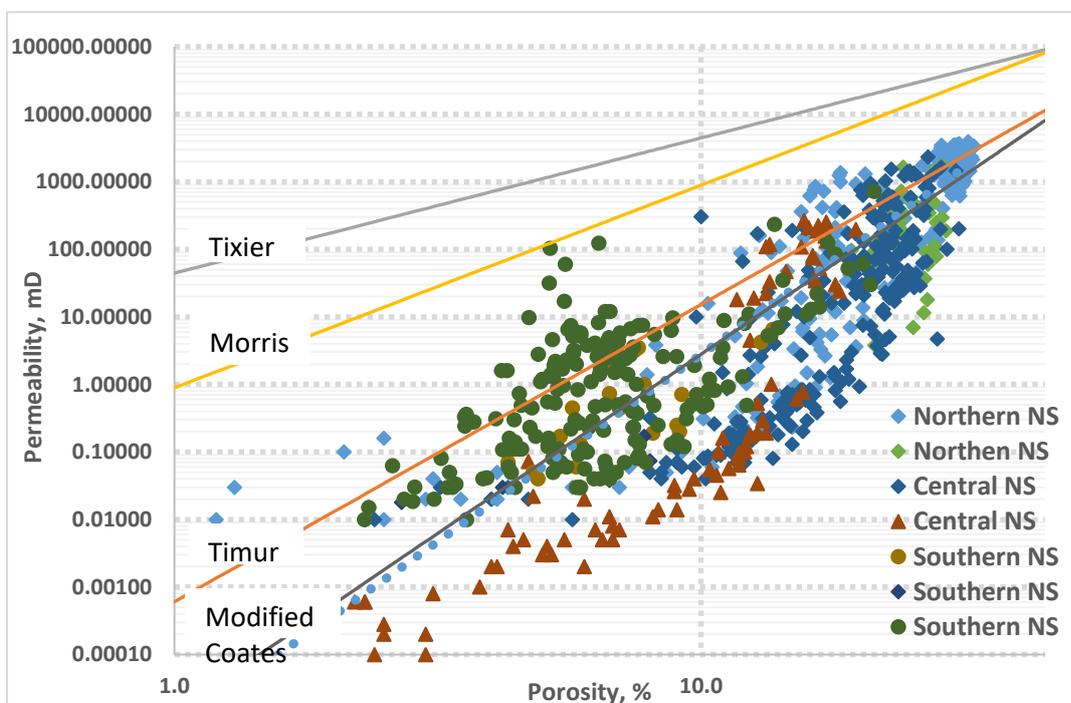


Figure 4. 27 Tixier, Morris, Coates, Timur modelled over North Sea data. Broken line is obtained from the data and passes through the centre.

4.3.3.2 Oil-Water Relative Permeability Prediction

Deep neural networks model have been validated using separate out of sample datasets not used for the training. The good agreement between experimental data and DNN's model predictions indicates that the complex, transient, non-linear behaviour of reservoir fluids can be effectively modelled as their saturation and phase changes with time.

Figures 4.28, 4.29 & 4.30 give a comparison between actual experimental values and model predictions using neural networks without cross validation, neural networks with cross validation and the deep neural networks. The objective here was to see how Deep Learning out performs ordinary networks on new data. These cross plots show the extent of agreement between the laboratory and predicted values. A perfect agreement means all points lie on the 45° line on the plot. For the testing set drawn from a different field from the training set, the Deep Neural Networks for both the wetting and non-wetting phase relative permeability (Figure 4.28 b&d) gives very close values to the perfect correlation line in all data points compared to the other models. Figure 4.40 a&c representing Neural Networks without cross validation, gave an RMSE value of 0.2484 and 0.0767 while Neural net with cross validation gave an RMSE of 0.0624 and 0.0765 (Figure 4.29 a&c). The Deep Neural Net gave an RMSE value of 0.2517 and 0.065 (Figure 4.30 a&c) for both wetting and non-wetting relative permeability. It is clear that all the models did well for the validation set although the deep neural networks performed better than the other two models. The different models were then shown new data from a separate field to see how they performed. For the test set (which is an out of sample dataset) obtained from a different field, the RMSE for neural network without cross validation is 0.9996 and 0.8483 (Figure 4.28 b&d), 0.2295 and 0.8022 with cross validation (Figure 4.29 b&d) while DNNs gave 0.0759 and 0.15 (Figure 4.30 b&d) for wetting and non- wetting relative permeability respectively. The deep learning model used the fourth cross validation model which happen to be the best for the wetting phase with a correlation coefficient of about 97% (Table 4.9) and the lowest error of 0.0014 while the second cross validation model was used for the non-wetting phase relative permeability having 96% correlation coefficient and the lowest error value of 0.0010 (Table 4.10).

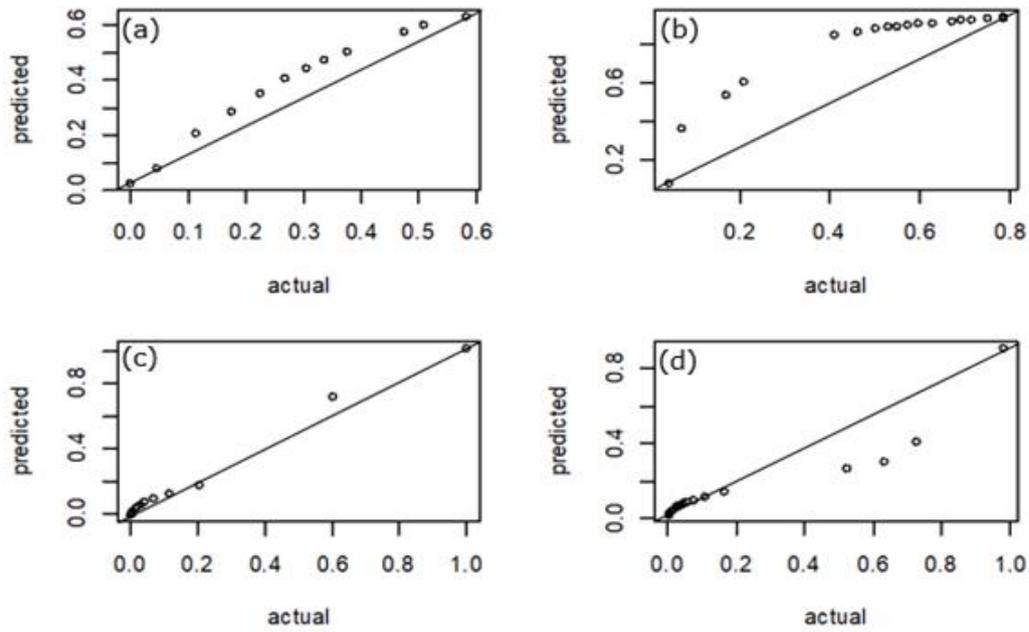


Figure 4. 28 Actual vs predicted value for neural networks without cross validation with (a) wetting phase relative permeability for validation set with an error of 0.2484 (b) wetting phase relative permeability for test set with error of 0.9996 (c) non-wetting relative permeability for validation set with error of 0.0767 (d) non-wetting relative permeability for the test set with error of 0.8483.

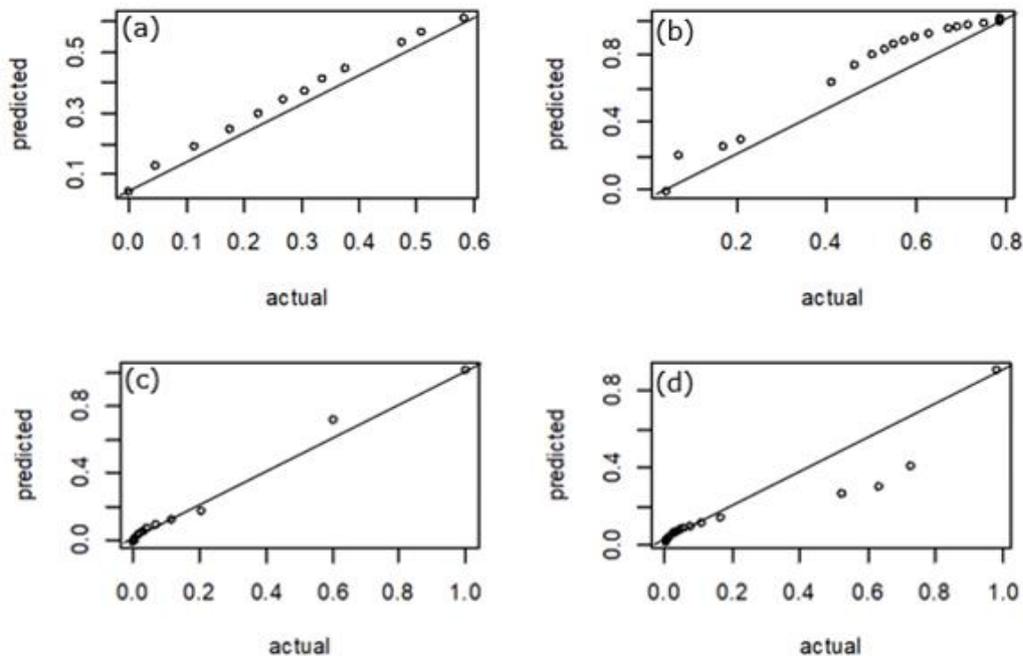


Figure 4. 29 Actual vs predicted value for neural networks with cross validation with (a) wetting phase relative permeability for validation set without an error of 0.0624 (b) wetting phase relative permeability for test set haing an error of 0.2295 (c) non-wetting relative permeability for validation set with an error of 0.0765 (d) non-wetting relative permeability for the test set with an error of 0.8022.

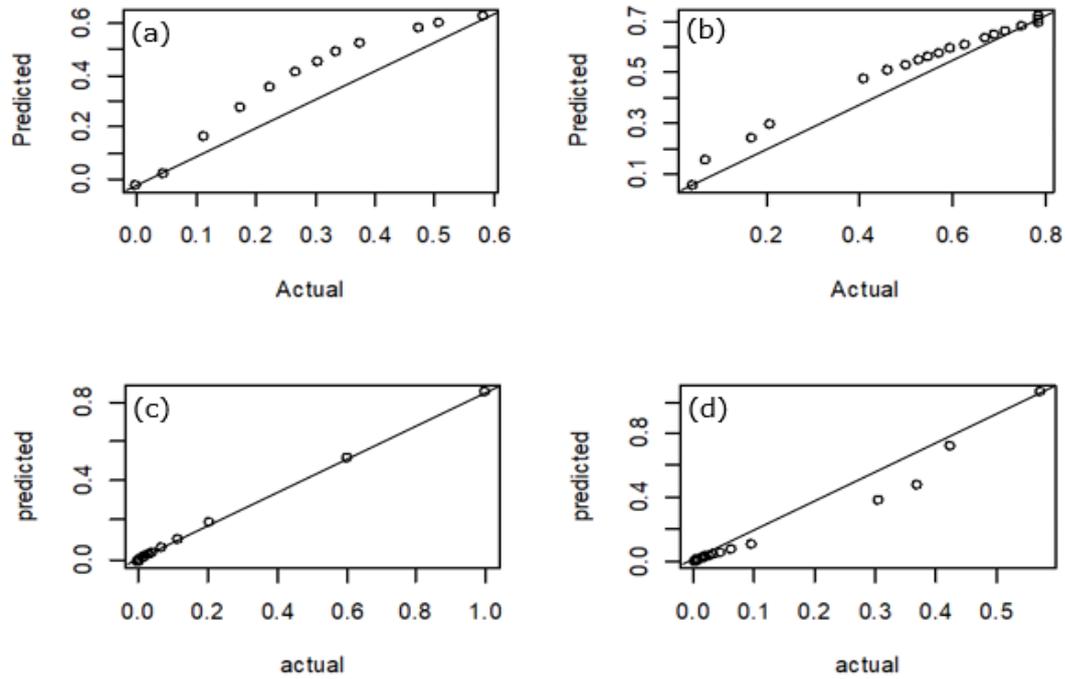


Figure 4. 30 Actual vs predicted value for deep neural networks with (a) wetting phase relative permeability for validation set with an error of 0.2517 (b) wetting phase relative permeability for test set with an error of 0.0759 (c) non-wetting relative permeability for validation set with an error of 0.065 (d) non-wetting relative permeability for the test set having an error of 0.15.

Table 4. 9 Accuracy of the deep learning model for the wetting phase

	mean	sd	5-Fold Cross Validation Results				
			1	2	3	4	5
mae	0.0489	0.0068	0.0558	0.0477	0.0612	0.0330	0.0468
mrd	0.0052	0.0022	0.0053	0.0047	0.0108	0.0014	0.0038
mse	0.0052	0.0022	0.0053	0.0047	0.0108	0.0014	0.0038
r2	0.9259	0.0186	0.9121	0.9086	0.9018	0.9745	0.9325
rd	0.0052	0.0022	0.0053	0.0047	0.0108	0.0014	0.0038
rmse	0.0689	0.0150	0.0728	0.0684	0.1037	0.0380	0.0615
rmsle	0.0541	0.0130	0.0509	0.0558	0.0854	0.0277	0.0509

Table 4. 10 Accuracy of the deep learning model for the non-wetting phase

	mean	sd	5-Fold Cross Validation Results				
			1	2	3	4	5
mae	0.0470	0.0109	0.0633	0.0395	0.0593	0.0209	0.0521
mrd	0.0052	0.0019	0.0065	0.0038	0.0089	0.0010	0.0060
mse	0.0052	0.0019	0.0065	0.0038	0.0089	0.0010	0.0060
r2	0.9214	0.0217	0.8800	0.9636	0.9099	0.9043	0.9492
rd	0.0052	0.0019	0.0065	0.0038	0.0089	0.0010	0.0060
rmse	0.0690	0.0153	0.0805	0.0619	0.0941	0.0309	0.0774
rmsle	0.0489	0.0090	0.0641	0.0466	0.0578	0.0266	0.0492

Figures 4.31 and 4.32 display the trend comparing the different models using the standard relationship between saturation and relative permeability. The Deep Learning model clearly outperforms the other models giving better predictions for both the wetting and non-wetting phases. Measurement error which causes input values to differ if the same example is presented to the network more than once is evident in the data. This limits the accuracy of generalization irrespective of the volume of the training set. The Deep Neural Networks model deeply understands the fundamental pattern of the data thus able to give reasonable predictions than ordinary networks and empirical models (Figures 4.31 and 4.32). The curves show that significant changes in the saturation of other phases has large effect on the wetting phase ability to flow as observed from the less flattening of the water relative permeability curve and vice versa for the flattened curve. Although this flattening behaviour is usual in the secondary drainage and imbibition cycles but mainly in the wetting phase when flow is mainly through small pore networks. Again, the curve flattening of the oil relative permeability curve could from experience be from brine sensitivity and high rates causing particle movements resulting to formation damage.

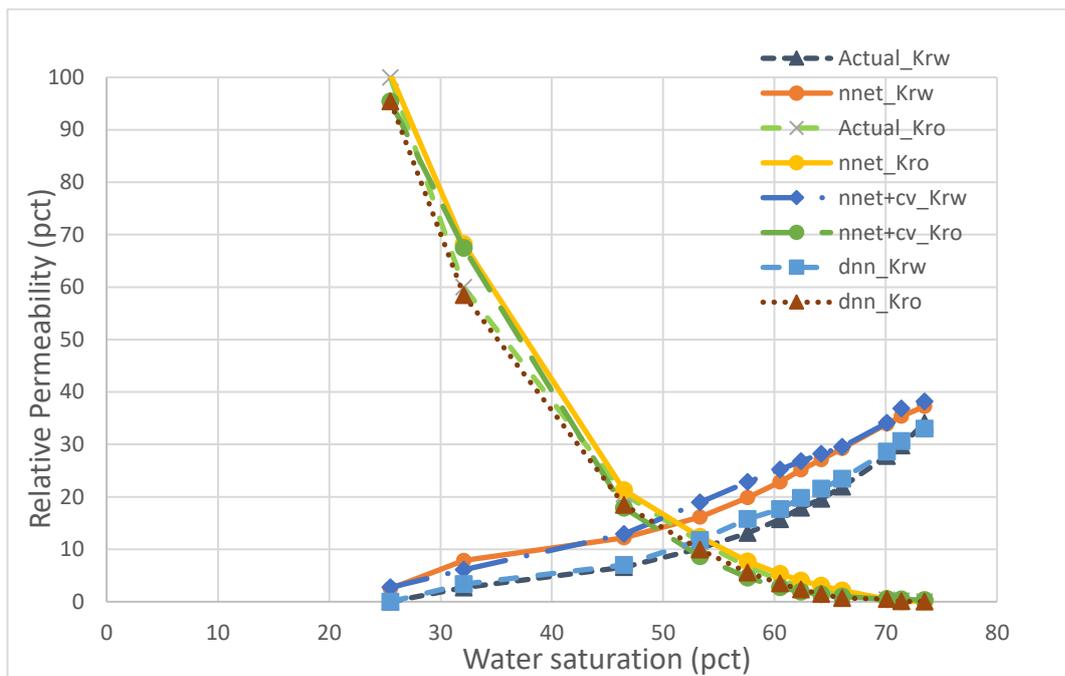


Figure 4. 31 Experimental and predicted relative permeability models using neural network with and without cross validation and deep neural networks on the validation set. The neural network model with cross validation (cv) partitioned the dataset into 5-fold and then trained and tested the model using the different folds.

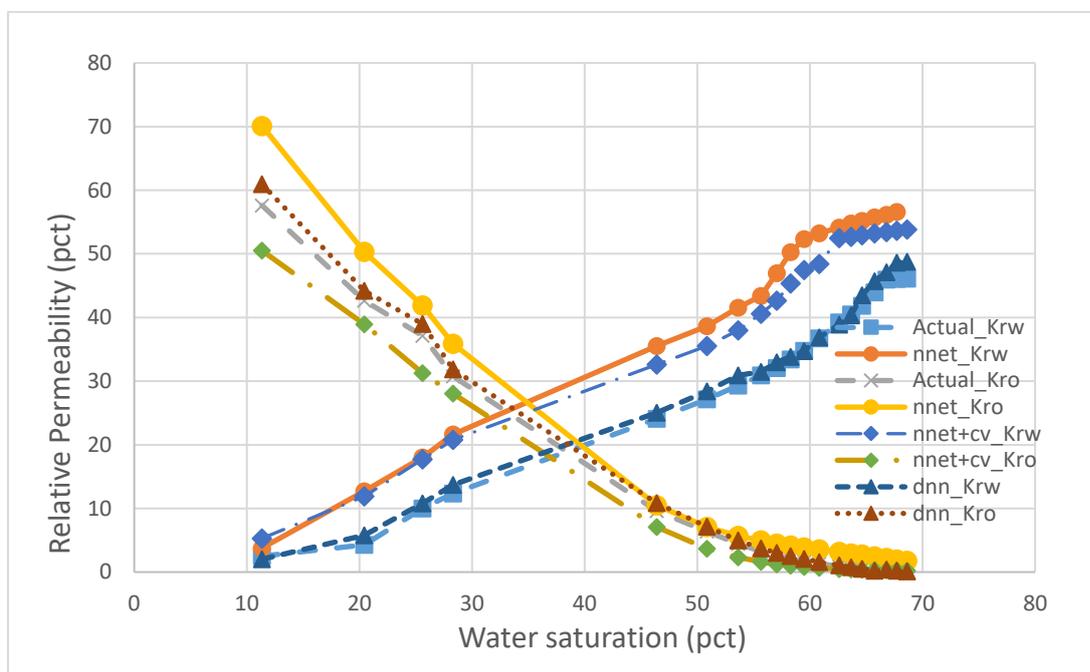


Figure 4. 32 Experimental (actual) and predicted relative permeability models using neural network (both with and without cross validation) and Deep Neural Networks on the out of sample test set (Stafjord reservoir). Cross validation (cv) involved in the network helped to improve its accuracy for out of sample datasets.

Figure 4.33 and 4.34 compares the Deep Neural Network model with commonly used empirical relative permeability models like Baker, Wyllie, Honarpour, Stones, Corey, Parker. The Baker equation is used to extrapolate oil-water and oil-gas relative permeabilities to the three phase case. Despite the fact that some of these models were developed using lots of datasets way more than the amount used for training the Deep Neural Networks, it still outperformed them showing that it is more able to capture the transients and eddies in real time scenarios due to its ability to regularize and generalize using its robust parameters as discussed earlier.

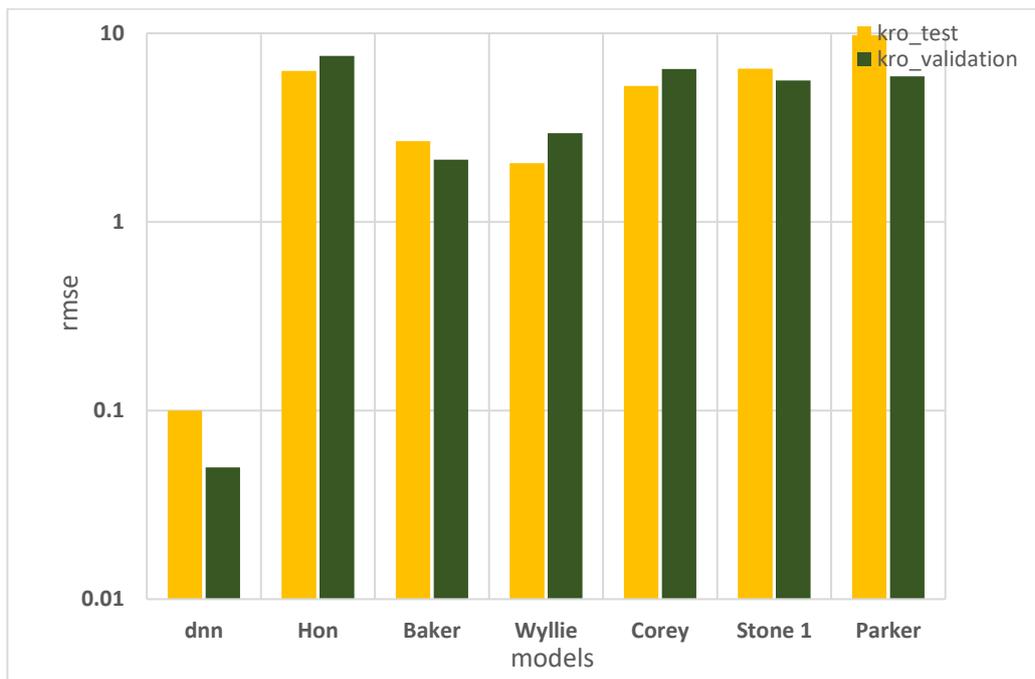


Figure 4. 33 Comparison of Wyllie, Corey, Parker, Stone, Baker, Honarpour, Deep Neural Networks for the Brent reservoir, North Sea. The DNN gave better prediction than the existing models for this validation set. Corey's λ , taken to be 2 and Parker's n parameter.

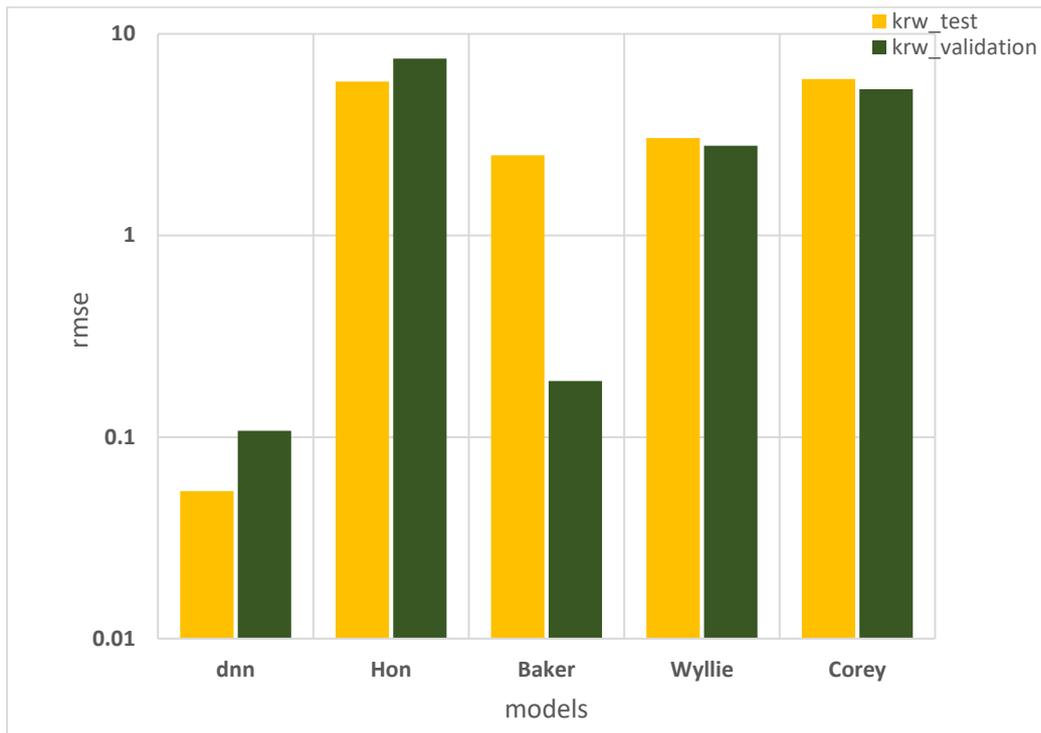


Figure 4. 34 Comparison of Wyllie, Corey, Baker, Honarpour, Deep Neural Networks models for the Stratjford reservoir, NorthSea.

Figures 4.35 and 4.36 corroborates the earlier observation that the Deep Learning model predicts better compared to most of the relative permeability models used in reservoir modelling software. It is important to note here that the empirical models have a problem of generalization especially as every reservoir is unique. Again, the assumptions associated with their formulation might not be practically true in all cases but this reservoir uniqueness or generalization is captured by the Deep Learning model bearing in mind that it will perform even better as more real time data is added to the training set.

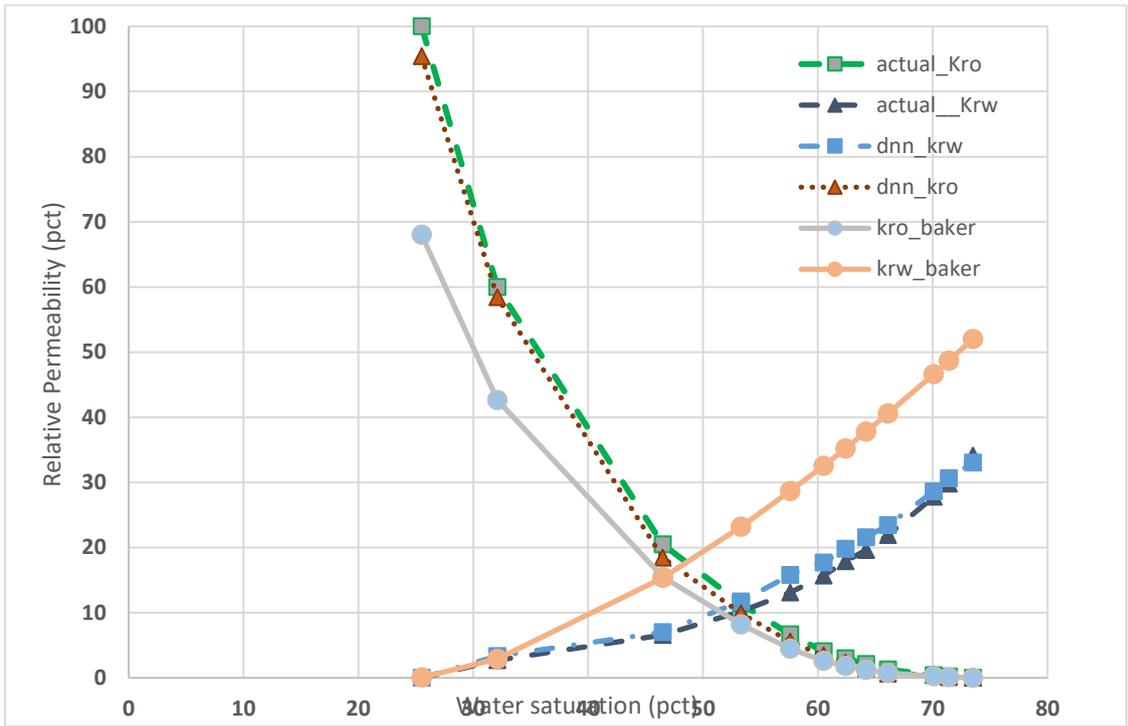


Figure 4. 35 Comparison of Deep Neural Networks and Baker with the measured wetting and non-wetting relative permeability models for the validation set (Brent reservoir).

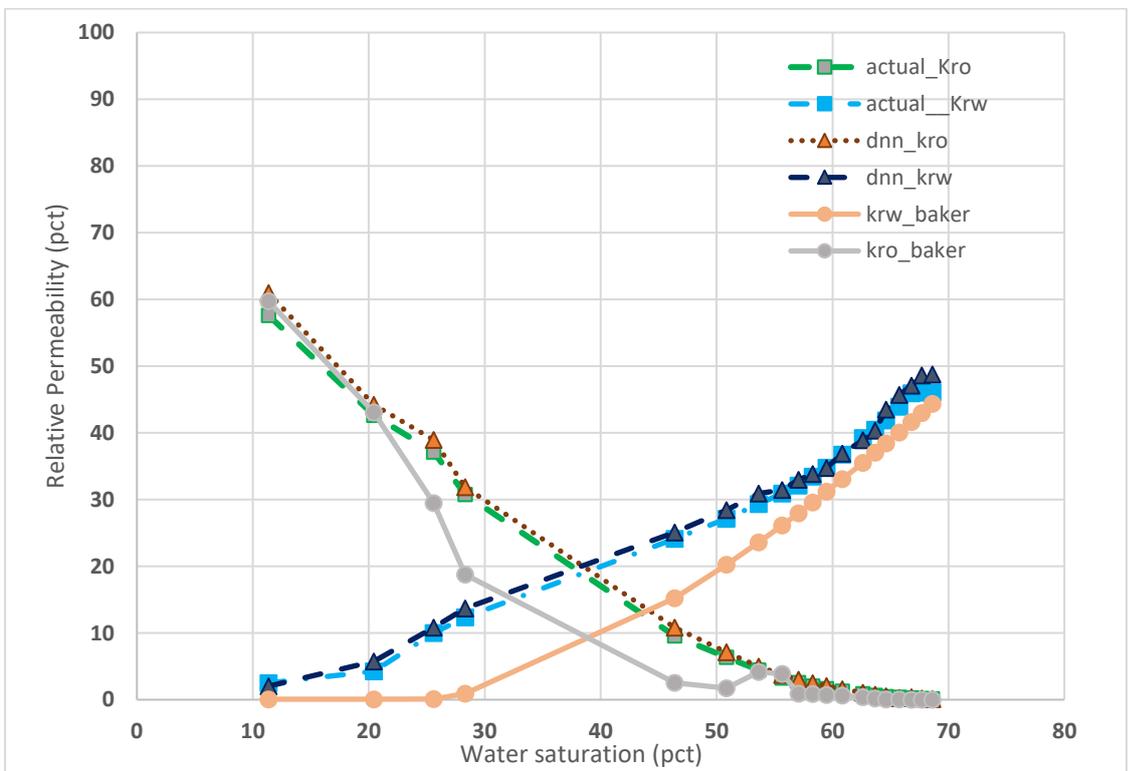


Figure 4. 36 Comparison of Deep Neural Networks and Baker with the wetting and non-wetting phase relative permeability models with for the test sets (Stratford reservoir). Baker was used since it performed best among the models compared.

Field Data of a water flooding operation were further used to test the capability of the deep neural network model for other real time applications. Figure 4.37 shows the amount of water injected at fixed time for flooding operation. The alphabets s-z represent the water flood fronts at fixed time of 30 to 240 days respectively. This becomes essential as water moves into the production well (water-oil interface moves upwards) due to greater water mobility than oil (Dake, 1983).

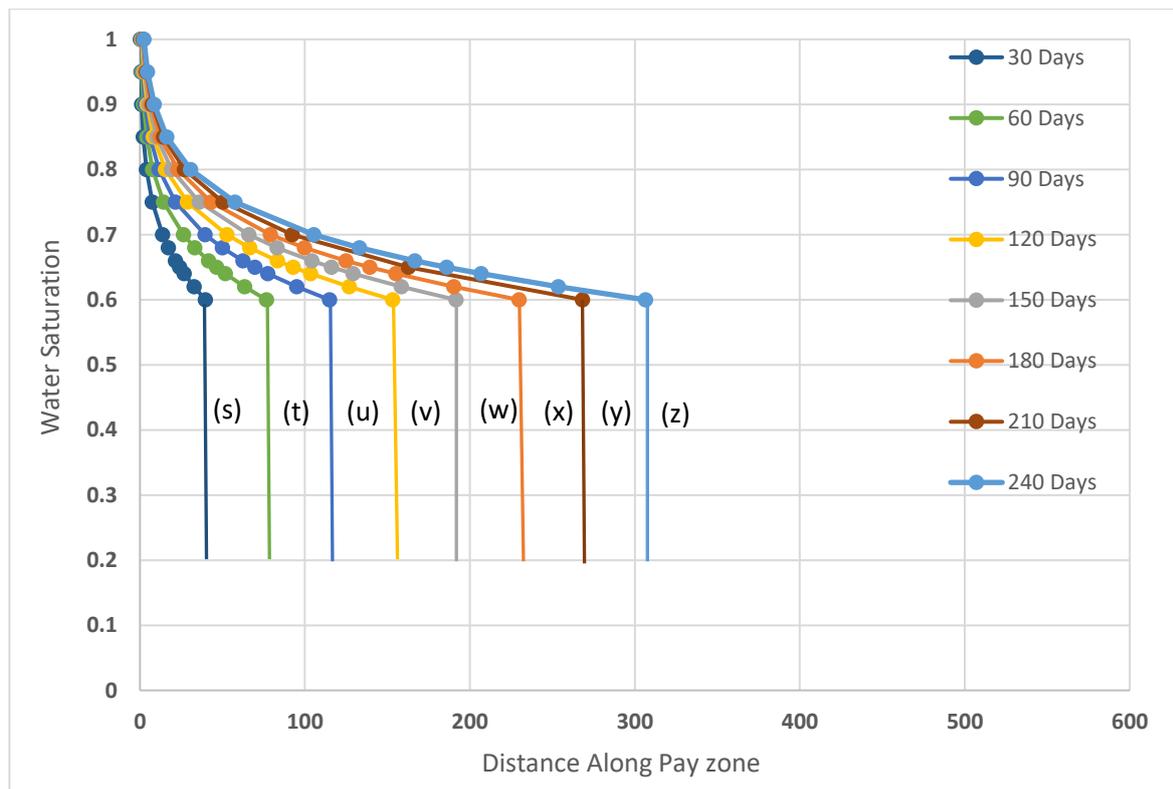


Figure 4. 37 Water saturation against distance along pay

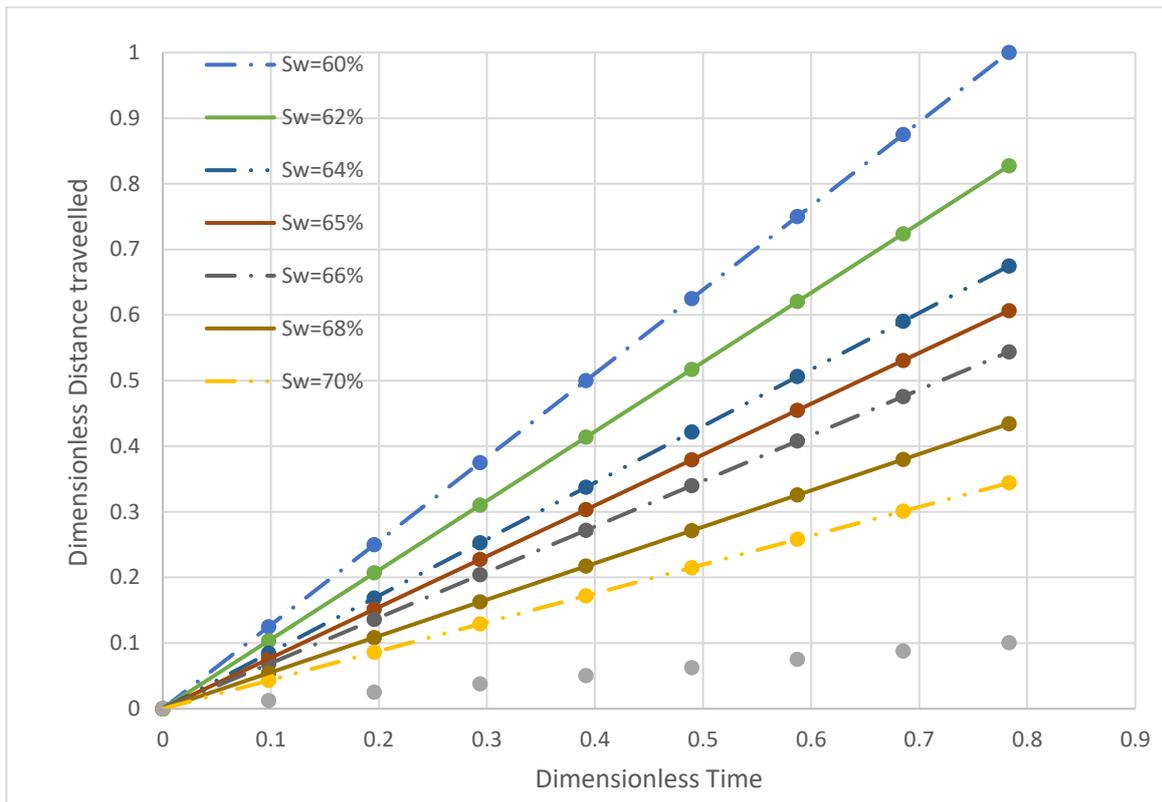


Figure 4.38 Dimensionless distance against dimensionless time describing saturation history

Based on the given depletion profile and history (Figures 4.37 & 4.38), the corresponding relative permeability was delineated for both the wetting and non-wetting phase (Figure 4.39 & 4.40) using the tested Deep Learning model. As the oil saturation drops, the water saturation increases leading to coning which is one of the major problems in reservoirs under water drive hence reservoir engineers seek ways to prevent or delay it. This is so vital especially since the cost of treating and disposing this water is very high.

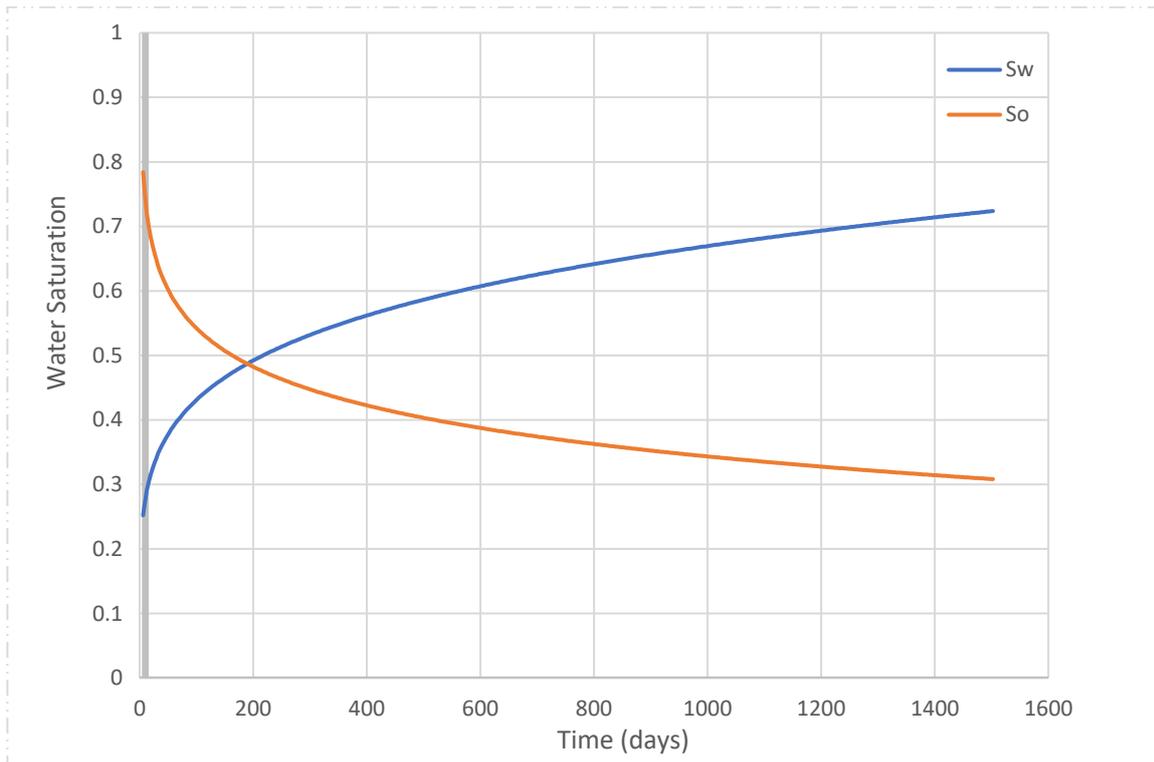


Figure 4. 39 Wetting and non-wetting phase saturation history of case field

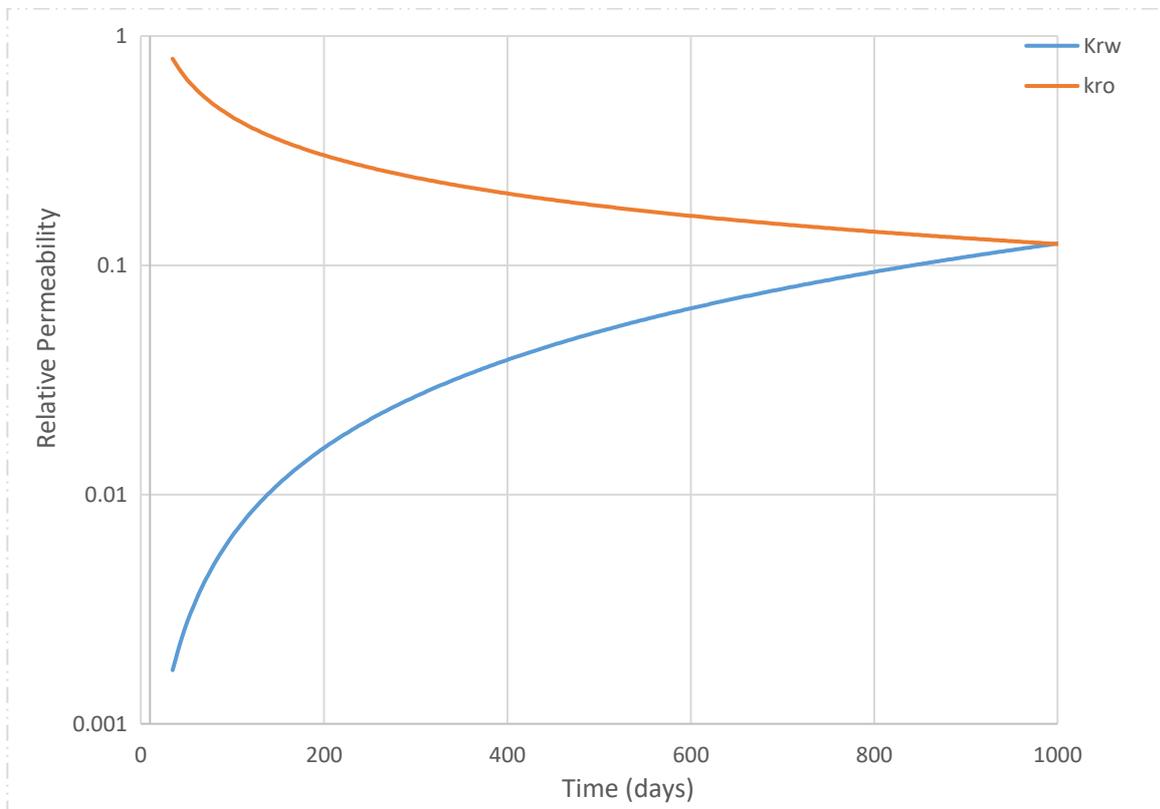


Figure 4. 40 Real time wetting and non-wetting phase relative permeability of a case field.

Figures 4.41 and 4.42 describe the relative importance (sensitivity) of the variables used for the wetting and non-wetting Deep Learning relative permeability models. The wetting phase model was more sensitive to its saturation and relatively less sensitive to that of the non-wetting phase while the non-wetting phase model was very sensitive to both its saturation and that of the wetting phase. Both models were also more sensitive to their own viscosities than the other. These models seem to obey the basic physics underlying relative permeability modelling. The least important variable still contributed above the median mark although in general, all variables show greater sensitivity in the non-wetting model than in the wetting relative permeability model. The $S_w - S_{wc}$ was the most sensitive of all the input parameters in both the wetting and non-wetting phase and immediately followed by S_w in the wetting-phase and by S_o in the non—wetting phase relative permeability.

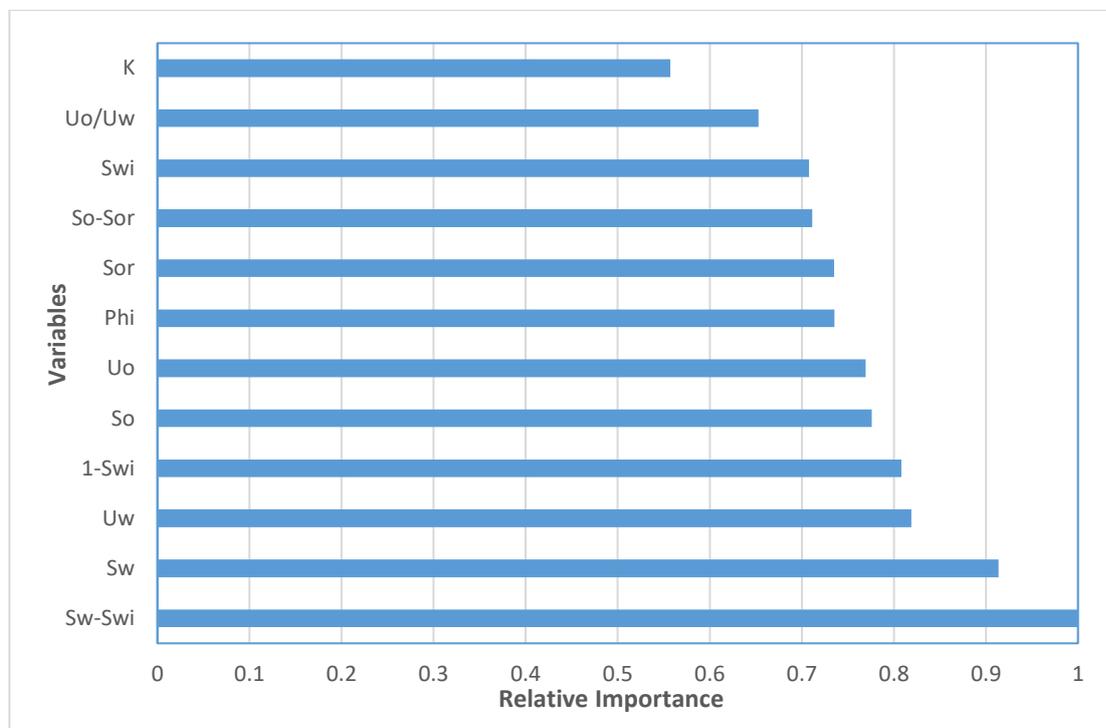


Figure 4. 41 Sensitivity analysis of individual variables used for building the wetting phase Deep Learning relative permeability model.

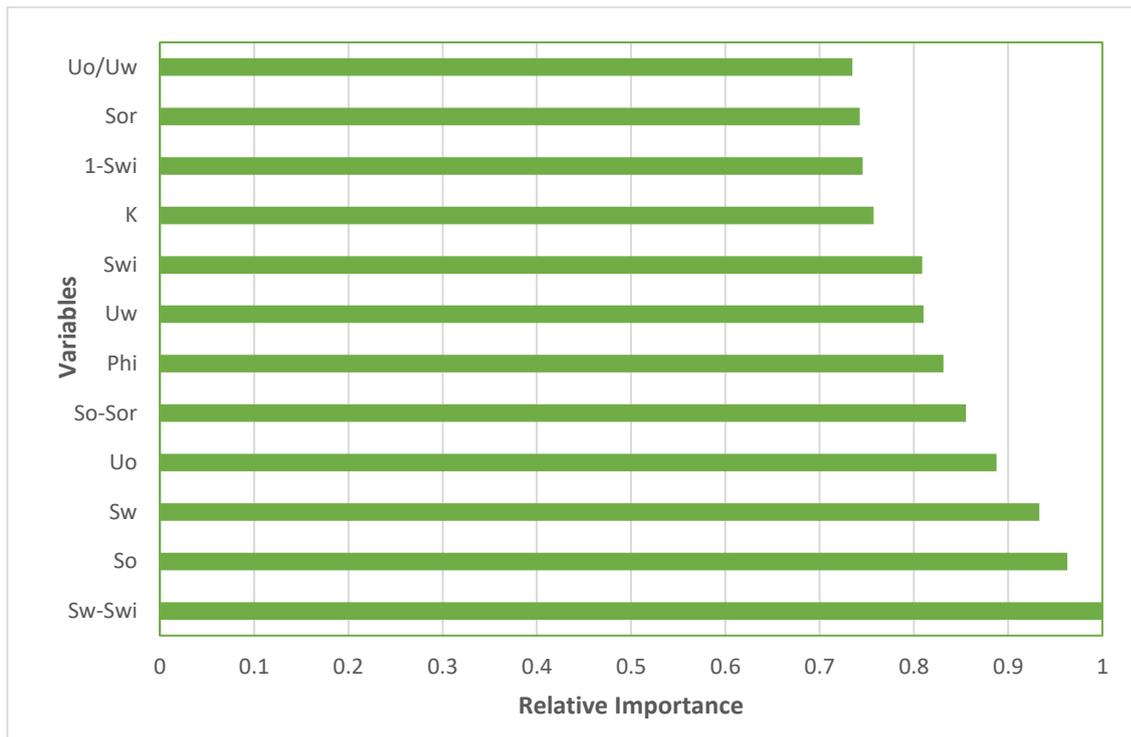


Figure 4. 42 Sensitivity analysis of individual variable used for building the non-wetting phase Deep Learning relative permeability model.

Table 4.11 shows the performance of the different variables combinations for both the wetting and non-wetting phase model. There was an increase in model performance when just the input parameters from the factors affecting relative permeability were increased for both phases shown from Cases 1 to 7. The error dropped from 0.1204 to 0.0481 for the wetting phase and from 0.1532 to 0.0691 for the non-wetting phase. In Cases 8 to 11, the input parameter now includes the functional links from Baker and Wyllie models. As can be seen the performance continued to increase as shown by the error reduction from 0.0481 to 0.0380 for the wetting phase and 0.0681 to 0.0619 for the non-wetting phase.

Table 4. 11 Sensitivity analysis showing the importance of the different to both water and oil relative permeabilities.

Cases	Input Variables		Model Metric (RMSE, fraction)	
	Factors affecting Relative Permeability	Functional Links (From Baker and Wyllie)	K_{rw}	K_{ro}
1	S_w, S_o		0.1204	0.1532
2	S_w, S_o, S_{wi}		0.1201	0.1057
3	S_w, S_o, S_{wi}, S_{or}		0.1153	0.0712
4	$S_w, S_o, S_{wi}, S_{or}, k$		0.0906	0.0698
5	$S_w, S_o, S_{wi}, S_{or}, k, phi$		0.0705	0.0671
6	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o$		0.0616	0.0691
7	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$		0.0481	0.0681
8	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc})$	0.0463	0.0667
9	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc}), (S_o - S_{or})$	0.0449	0.0652
10	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc}), (S_o - S_{or}), (1 - S_{wc})$	0.0508	0.0732
11	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc}), (S_o - S_{or}), (1 - S_{wc}), (\mu_o/\mu_w)$	0.0380	0.0619

4.3.6 Summary and Conclusion

A Deep Neural Networks methodology has been formulated for wetting and non-wetting phase relative permeability predictions taking into account phase and saturation changes hence its capability for real time applications. This work has the following conclusions:

1. Deep Neural Network has shown to be a good predictive and prescriptive tool for relative permeability. Its ability to generalize and regularize helped to stabilize and reduce the main problem of all predictive tools which is over fitting.
2. In an industry where big data is now available, Deep Learning can provide the platform to systematically forecast reservoir fluid and rock properties in order to drastically optimize the cost and time needed for laboratory experiments. Even with the amount of data used, the power of the Deep Neural Networks is evident in that it gave reasonable predictions which will dramatically improve if more data were available.
3. Different results were obtained from different relative permeability models for the same reservoir with some of the models giving better predictions at lower saturations but performs poorly at higher saturations and vice versa hence lots of uncertainty. Therefore, it is needful for practitioners to know the limitations of any correlation used for the prediction of wetting and non-wetting phase relative permeability.

Chapter Five

Well Test Uncertainty Analysis

5.1 Introduction

Well test analysis (sometimes referred to as Pressure Transient Analysis) is an important aspect of petroleum engineering as it seeks to give information about the reservoir on a wider scale. A transient well test is the only method which enables estimates of the permeability-thickness (conductivity) at a truly insitu condition representing an average which is of greater volume than core and log permeability. Several rock and fluid parameters are used as input variable for this analysis such as porosity, saturation, wellbore radius, pay thickness, compressibility, viscosity, flowrate, formation volume factor. As was shown in Chapter Two using Monte Carlo simulations to produce tornado plots, some of these properties are more sensitive than others and thus could have huge impact on the result of pressure transient analysis. Reaching a decision whether it is worth the time and money producing a reservoir at all, there is need to know reservoir permeability, initial pressure and boundary. The conductivity (kh) controls the speed of fluid flow into the well and thus it is useful for the design of number of wells and their spacing. Drilling a well costs companies over 10 million dollars and it is therefore vital to reduce the uncertainty associated with the property that controls this design.

Well test analysis has uncertainty associated with its different input parameters thus it is important to note that there is a limit to the level of detail that can be achieved from well testing techniques for reservoir description. This is because pressure transmission is an inherently diffusive process and therefore it is governed largely by average conditions rather than by local heterogeneities. Grader and Horne (1988) showed that it is possible to have a “hole” in the reservoir that is as large as the distance between the production well and observation well, without that “hole” making any discernible difference in an interference test. Though this observation looks discouraging at first, however it underlies the overall usefulness of well test analysis hence well tests can be interpreted to estimate bulk reservoir properties because they are insensitive to most local scale heterogeneities.

Since the SPE Monograph 5 (Earlougher 1977), one of the most significant breakthroughs in the well test analysis remains the development in the late 1970s and early 1980s of a general and systematic approach to the analysis of well tests. This approach unified the various independent techniques previously described in the literature and often gave opposing results (Ramey 1992)

into one methodology on the basis of signal theory (Jouanna and Fras 1979). This unified well test methodology considers the agreement of all diagnostics and plot used from identification of the interpretation model, verification to calculation of the interpretation model parameters. By considering well testing within the context of signal theory (Gringarten et al., 1979), it became easier to understand its scope and limitations like what type of result can realistically be obtained from well testing, what is the best method to obtain this result, how does well testing contribute to reservoir characterization compared to other methods. In signal theory, signal processing is schematically described below:

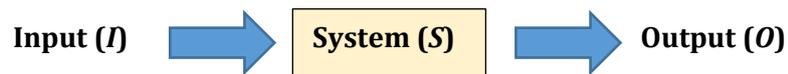


Figure 5. 1 Signal theory schematic for the reservoir (Gringarten 1985a)

Where I is the input signal usually flowrate applied to the reservoir system operator S to obtain an output signal O usually pressure in the case of pressure transient analysis. Ramey (1992) also gave an example to explain this theory assuming I is (1, 2 and 3), S is addition operator and O is 6. If both I and S are known (convolution), O has a unique solution which is also the case if S and O are known (deconvolution) since I has a unique answer. But if I and O are known (inverse problem), S has a non-unique solution as it can be addition operator (1+2+3) or multiplication (1×2×3) which is the case in well testing as we don't know what the system is. In practice, the inverse problem is solved during the identification of an interpretation model (Figures 5.2 and 5.3) which is a combination of the individual flow regimes components that dominate the flow period at different times. The emphasis of the integrated approach was on the well test "behaviour", which refers to the response of the well to changes in production conditions. As identification techniques become more powerful as with derivatives (Bourdet et al., 1983a, 1989; von Schroeter et al., 2001; Agarwal, 1989) and the resolution of measurements improves, the number of behaviour components that can be identified increases, resulting in more detailed interpretation models.

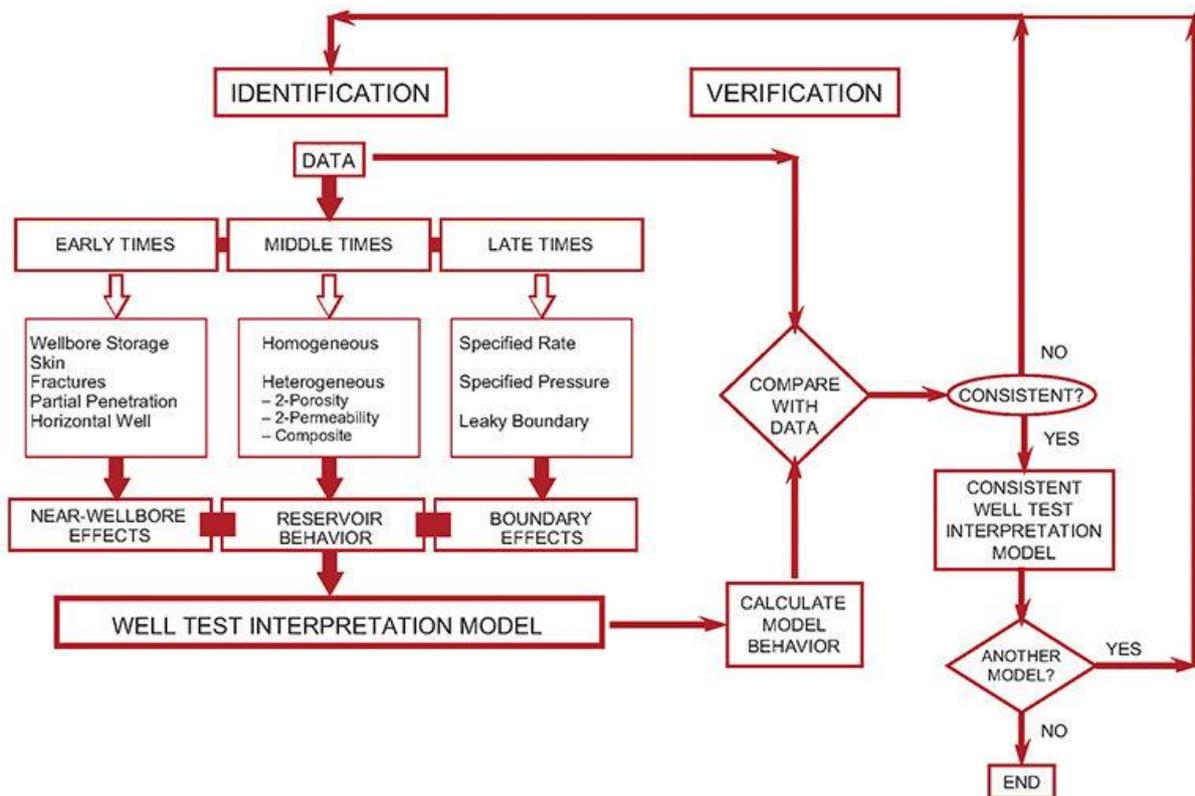


Figure 5. 2 Interpretation model identification process (Gringarten 2012)

Lorenz plot was used to get the dominant flow interval especially if different from the perforated interval. In other words, what percentage of the interval in contributing to flow which was then used as the pay thickness for the pressure transient analysis. The plot also gave an idea of the heterogeneity of the reservoir area cut by the well. Both buildup and drawdown were considered. The different sensitivities were also analysed

The purpose of this chapter is to accurately analyse well test data to obtain specifically permeability even though skin and well bore storage introduced by Van Everdingen and Hurst were also automatically determined by the analysis. The obtained well test permeability will then be collated with their respective well log and core permeabilities for the same interval in Chapter Six to build a database and model for real time calibrations. Important to note that no machine learning was considered in the chapter as this talks about well test analysis to primarily determine permeability for the sake of this work.

5.2 Methodology

5.2.1 Identification of the Interpretation Model (Inverse Problem)

Identifying the model, Σ , is the most important step of the analysis process. If the wrong model is selected then all reservoir parameters derived from the analysis will be incorrect. One must identify a model of the actual reservoir, S , say, the behaviour of which is identical to the behaviour of S (Figure 5.1). Identical behaviour in this case means that the observed output signal O obtained from the reservoir S and the output signal O' calculated from the model, exhibit the same qualitative characteristics (i.e., show similar shapes):

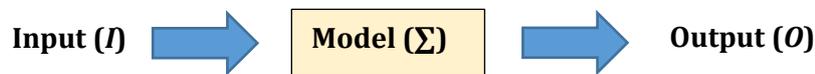


Figure 5. 3 Signal theory schematic for the chosen model (Gringarten 1985a)

Finding Σ which is the model implies solving the inverse problem which requires an identification or pattern recognition process. The solution is not unique. The degree of non-uniqueness increases with the complexity of the reservoir behaviour and decreases with the amount of information available on the well and reservoir being tested. The problem of non-uniqueness is now well recognized in the oil industry. It is the main reason for the increase of the use of stochastic modelling techniques, which aim at providing alternative equi-probable representations of the reservoir to capture the uncertainty associated with predictions (Hewett 1986; Suro-Perez et al., 1991).

Identification has improved tremendously with the development of a stable algorithm for deconvolution (von Schroeter and Hollaender, 2001). By converting pressure at a variable rate into pressure at constant rate, the process of deconvolution transforms a test into a single drawdown having the same duration as the test thus increasing the amount of data that can be analysed with conventional analyses. The gain is clearly greater in long tests, such as with permanent downhole pressure gauges, in which the total test duration of the longest flow period at constant rate. Deconvolution is also useful in short tests such as DSTs because it increases the radius of investigation and enables the differentiation between true test behaviour and artifacts of the derivation calculation.

With the model now known at this stage of the interpretation process, the problem to be solved is a direct problem. Based on the fact that the solution of the direct problem is unique, there is a unique set of model parameter values that can provide a best fit with the observed data. Therefore, once the interpretation model is selected, the reservoir parameters corresponding to that model are defined uniquely and the numerical values of these parameters are independent of the method used to calculate them. So the results must be the same whether parameters are calculated by the use of straight lines, log-log type-curve matching or nonlinear regression techniques (Rosa and Horne 1983) with the only acceptable difference being those caused by the differences in resolution of the various techniques. This was not universally understood before the development of the integrated methodology because the different methods gave different results.

5.2.2 Verification of the Interpretation Model

As already stated, the identification step involves solving the non-uniqueness problem, thus the interpretation model was verified. Consistency checks were made among other characteristics inferred by the model and the corresponding known information from the actual reservoir and measured data. If the model Σ satisfies all the checks, it is regarded as “consistent” and represents a valid solution to the problem but if it fails any checks, it is regarded as invalid. The interpretation process must be repeated to identify all possible consistent models which can then be ranked on the basis of decreasing probability and possibly a new well test designed to confirm the most probable model.

5.2.3 Calculation of the Interpretation Model Parameters (Direct Problem)

Once the interpretation model has been identified, its response must be generated, either analytically or numerically, and the parameters of the model must be adjusted until the model gives the same quantitative response as the actual reservoir and also the same qualitative response (e.g. the same shape), a condition that controlled the selection of the model in the first place. The adjusted numerical values of the model parameters are then said to represent the values of the corresponding reservoir parameters.

5.2.4 Pressure Buildup Analysis

Most of the pressure transient analysis where done with buildup data. The analysis describes the buildup of pressure in a well with time after a well has been shut in. Because the buildup of

wellbore pressure will follow a particular trend, it was possible to not only determine permeability around the wellbore but also to use a Non-Parametric Gaussian Process Regression method to reconstruct any missing data using flowrate or temperature and vice versa. Other uses of buildup test include determination of the presence and distance of fault, any interference between producing wells and reservoir boundary especially where there is no strong water drive. The technique require that the well was first produced at a constant rate Q_o STB/day for a flowing time t_p days (it's in effect over the entire time, $t_p + \Delta t$), long enough to stabilize before shut in for a shutin period Δt from a flow rate of Q_o to 0. The resulting pressure curve was then analysed to obtain reservoir properties. The composite effect is gotten by the addition of the individual constant-rate solutions at the specific rate-time sequence shown by:

$$p_i - p_{ws} = (\Delta p)_{total} \quad (5.1)$$

$$= (\Delta p)_{due\ to\ (Q_o-0)} + (\Delta p)_{due\ to\ (0-Q_o)}$$

With p_i being initial reservoir pressure (psi) while p_{ws} is wellbore shut in pressure (psi)

$$(\Delta p)_{due\ to\ (Q_o-0)} \quad (5.2)$$

$$= \left[\frac{162.6(Q_o - 0)B_o\mu_o}{kh} \right] \left[\log \left(\frac{k(t_p + \Delta t)}{\phi\mu_o c_t r_w^2} \right) - 3.23 \right]$$

$$+ 0.87s \Big]$$

While

$$(\Delta p)_{due\ to\ (0-Q_o)} \quad (5.3)$$

$$= \left[\frac{162.6(0 - Q_o)B_o\mu_o}{kh} \right] \left[\log \left(\frac{k\Delta t}{\phi\mu_o c_t r_w^2} \right) - 3.23 \right]$$

$$+ 0.87s \Big]$$

Then

$$p_{ws} = p_i - \frac{162.6Q_o B_o \mu_o}{kh} \left[\log \left(\frac{t_p + \Delta t}{\Delta t} \right) \right] \quad (5.4)$$

The above equation is commonly referred to as Horner's equation and a semilog plot of p_{ws} against $\left(\frac{t_p + \Delta t}{\Delta t}\right)$ produces a straight line relationship with intercept p_i and slope m from which the conductivity kh is calculated.

$$m = \frac{162.6Q_o B_o \mu_o}{kh} \quad (5.5)$$

And

$$kh = \frac{162.6Q_o B_o \mu_o}{m} \quad (5.6)$$

5.2.5 Drawdown Analysis

Some of the analysis carried out involved a draw down test. This are series of pressure measurements done during a constant flow period following a sufficient shut in period which allows the pressure to stabilize in the entire formation. Although this analysis could detect pore volume reservoir heterogeneity within the well's drainage area, assess the degree of damage induced by drilling and completion activities, the focus here was to obtain the average permeability of the reservoir formation. Since a well flowing at a constant rate Q_o under unsteady condition has pressure behaving as if it exists in an infinitely acting reservoir, the pressure behaviour was described with Equation 5. Below:

$$p_{wf} = p_i - \frac{162.6Q_o B_o \mu_o}{kh} \left[\log \left(\frac{kt}{\phi \mu_o c_t r_w^2} \right) - 3.23 + 0.87s \right] \quad (5.7)$$

Just like Equation 5.4, this also produces a straight line on a semi-log paper with slope m and kh obtained as in Equation 5.6.

5.2.6 Superposition, Type Curves and Pressure Derivatives

The unified methodology that allows for the several curves to match was adopted. Superposition principle states that the sum of the individual solutions to the diffusivity equation is also a solution to the equation since in reality there are more than one well in a field. Horner method is the solution of superposition of one rate change. In a case of 3 wells for example:

$$\begin{aligned}
 p_i - p_{wf} &= (\Delta p)_{total \ drop \ at \ well \ 1} & (5.8) \\
 &= (\Delta p)_{drop \ due \ to \ well \ 1} + (\Delta p)_{drop \ due \ to \ well \ 2} \\
 &+ (\Delta p)_{drop \ due \ to \ well \ 3}
 \end{aligned}$$

With the pressure drop at well 1 due to its own production is represented by the log approximation to the Eigen (E_i) function solution given by Equation 5.7. The additional pressure drops at test well 1 due to wells 2 and 3 production was written in terms of the E_i function solution

$$\begin{aligned}
 (\Delta p)_{total \ drop \ at \ well \ 1} &= p_i - p_{wf} & (5.9) \\
 &= \frac{162.6Q_{o1}B_o\mu_o}{kh} \left[\log \left(\frac{kt}{\phi\mu_o c_t r_w^2} \right) - 3.23 + 0.87s \right] \\
 &- \left(\frac{70.6Q_{o2}B_o\mu_o}{kh} \right) E_i \left[-\frac{948\phi\mu_o c_t r_1^2}{kt} \right] \\
 &- \left(\frac{70.6Q_{o3}B_o\mu_o}{kh} \right) E_i \left[-\frac{948\phi\mu_o c_t r_2^2}{kt} \right]
 \end{aligned}$$

Where Q_{o1} , Q_{o2} , Q_{o3} are the producing rates of well 1,2,3 respectively.

Finding the theoretical type curve Σ that matches the actual system response S from the test well and the reservoir on a log-log plot when subjected to changes in pressure or production rates was ensured. These type curves introduced by Agarwal, (1980) used dimensionless variables (made dimensionless by multiplying by a group of constants with opposite dimensions) to eliminate units and parameters in diffusion problems since they are plots of theoretical solutions to transient and pseudo steady state flow equations. Dimensionless time t_D and pressure p_D are shown below:

$$t_D = 0.0002637 \frac{k}{\phi \mu c_t r_w^2} \Delta t = A \Delta t \quad (5.10)$$

$$p_D = \frac{kh}{141.2qB\mu} \Delta p = B \Delta p \quad (5.11)$$

Where A and B are the groups that make a variable dimensionless. These type curves allowed for the identification of flow regimes during early (ETR) and middle time regions (MTR) for example a unit slope indicates wellbore storage effect.

The problem of similarity of shapes (Tiab and Kumar, 1980 and Bourdet et al., 1983) makes it difficult to find a match and thus the pressure derivative (Bourdet et al., 1983) is used to compliment this ensuring a double match. The derivative of p_D with respect to t_D/C_D with C_D being the Dimensionless wellbore storage, is given as:

$$p_D' = \frac{d(p_D)}{d(t_D/C_D)} = 1.0 \quad (5.12)$$

While for radial infinite acting flow period has pressure behaviour given by:

$$p_D = \frac{1}{2} [\ln(t_D/C_D) + 0.80907 + \ln(C_D e^{2s})] \quad (5.13)$$

Where s is the skin factor, hence differentiating with respect to t_D/C_D gives:

$$p_D' = \frac{d(p_D)}{d(t_D/C_D)} = \frac{1}{2} \left[\frac{1}{(t_D/C_D)} \right] \quad (5.14)$$

Thus on the log-log diagnostic plot, the unit slope and the half slope on the derivative were used to decipher flow regimes and calculate permeability.

5.3 Results and Discussion

5.3.1 Effective Flow Interval and Heterogeneity

Lorenz plot and Modified Lorenz plot were used to show what fraction of total storage is contributing to flow, EFI, (Pinisetti et al., 1998; Corbett et al., 1998), flow units (Gunter et al., 1997) as well as give an indication of the heterogeneity of the reservoir formation (Jensen et al., 1998; Jensen and Lake, 1988) which in turn gives an idea of the discrepancy one should expect from the different multiscale data sets. This is so crucial since well test permeability calculation is strongly influenced by formation thickness estimation (Corbett et al., 1996) as it is commonly taken as the total perforated interval (done on the basis of porosity which is a static property) which can be different from effective flow interval especially in a heterogeneous reservoir (e.g. channel sand). This can lead to incorrect results, moreover that the well test gives permeability thickness hence thickness has to be right. There is therefore a need to identify the interval likely to flow into the well. Both the ordered (Lorenz plot) and the unordered stratigraphic order (modified Lorenz plot) were used.

Figures 5.4-5.8 shows the Lorenz and Modified Lorenz curve for several wells from North Sea. EFI analysis like these should be done before carrying out well test analysis to reduce the uncertainty associated with accurate determination of its permeability prediction.

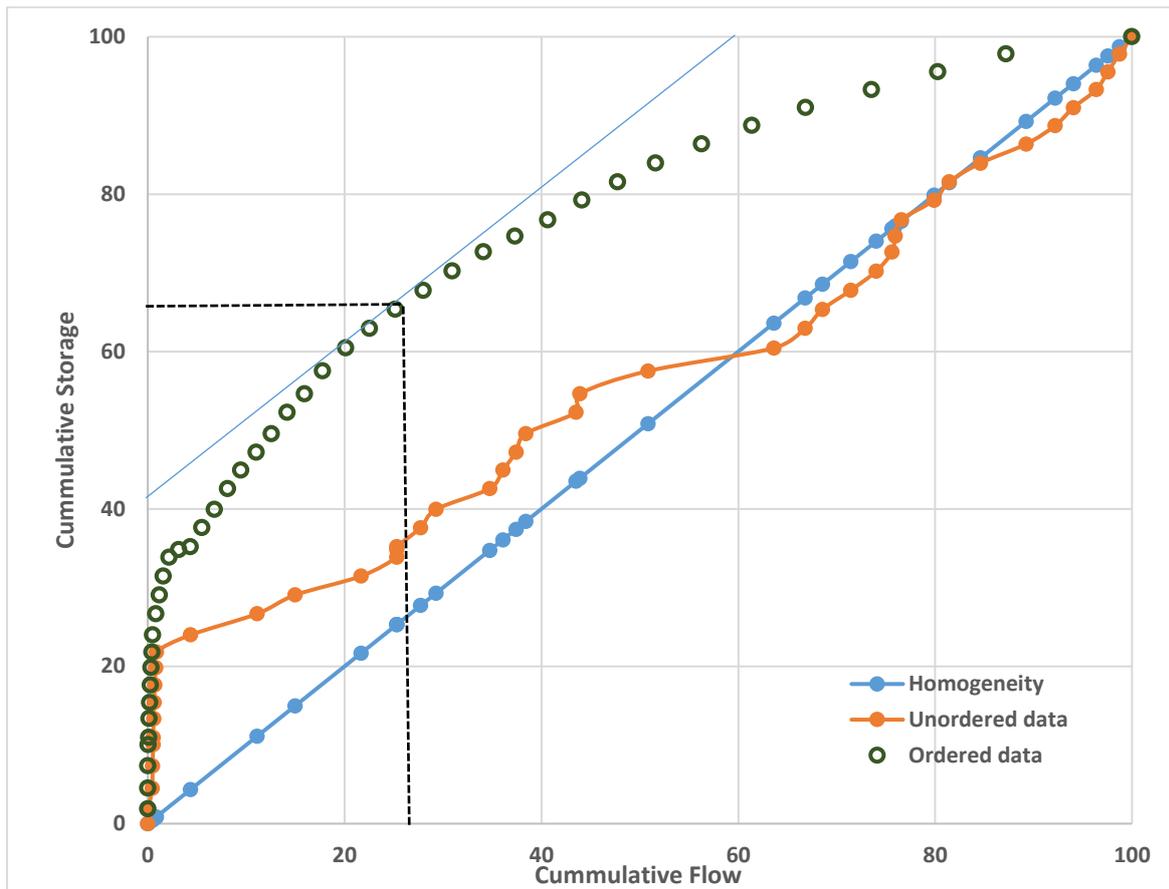


Figure 5. 4 Lorenz plot and modified Lorenz plot for well 1 sandstone formation. The Lorenz curve shows the effective flow interval through the intersection of the ordered plot with its tangent at 45° showing that about 70% of flow is coming from approximately 40% of storage. The unordered plot gives an indication of how heterogenous the formation is with respect to the homogeneity line as shown in legend.

The Lorenz plot shown above, displays not just the heterogeneity of the reservoir (Jensen et al., 1997; Jensen and Lake, 1988) but also the effective flow interval and not just considering the perforation height which could be misleading. Also some wells that have been perforated and not properly cleaned up prior to the well kill taken during any DST operation, raises concern that formation damage had taken place in the near wellbore region and was responsible for poor well performance. Bear in mind that the closer the values of flow and storage, the more homogenous the formation and the farther apart the values of flow and storage, the more heterogeneous the reservoir formation. Again, the zero line is taken as the well and the horizontal axis is the flow into the well hence reading is from the 100 mark on both axis (i.e. from reservoir to well) and not from the well to reservoir.

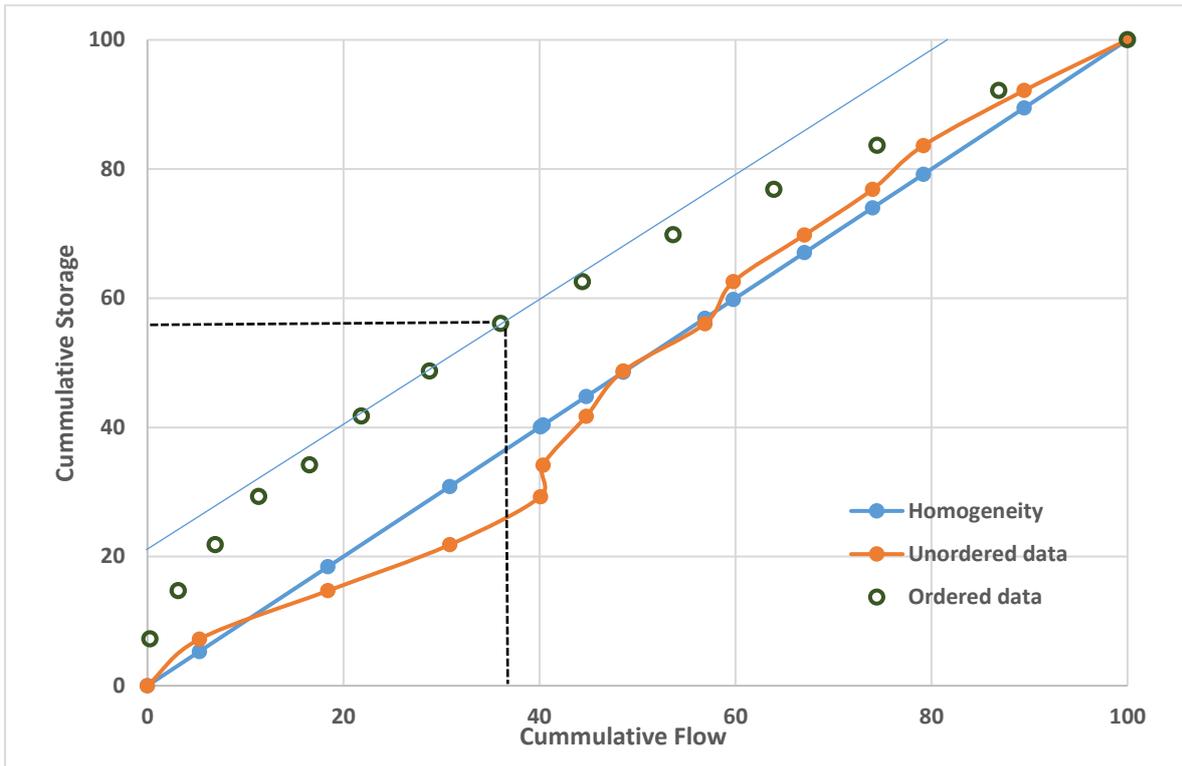


Figure 5.5 Lorenz plot and Modified Lorenz plot for well 4 sandstone reservoir. The plot shows that about 60% of flow is from approximately 45% of storage.

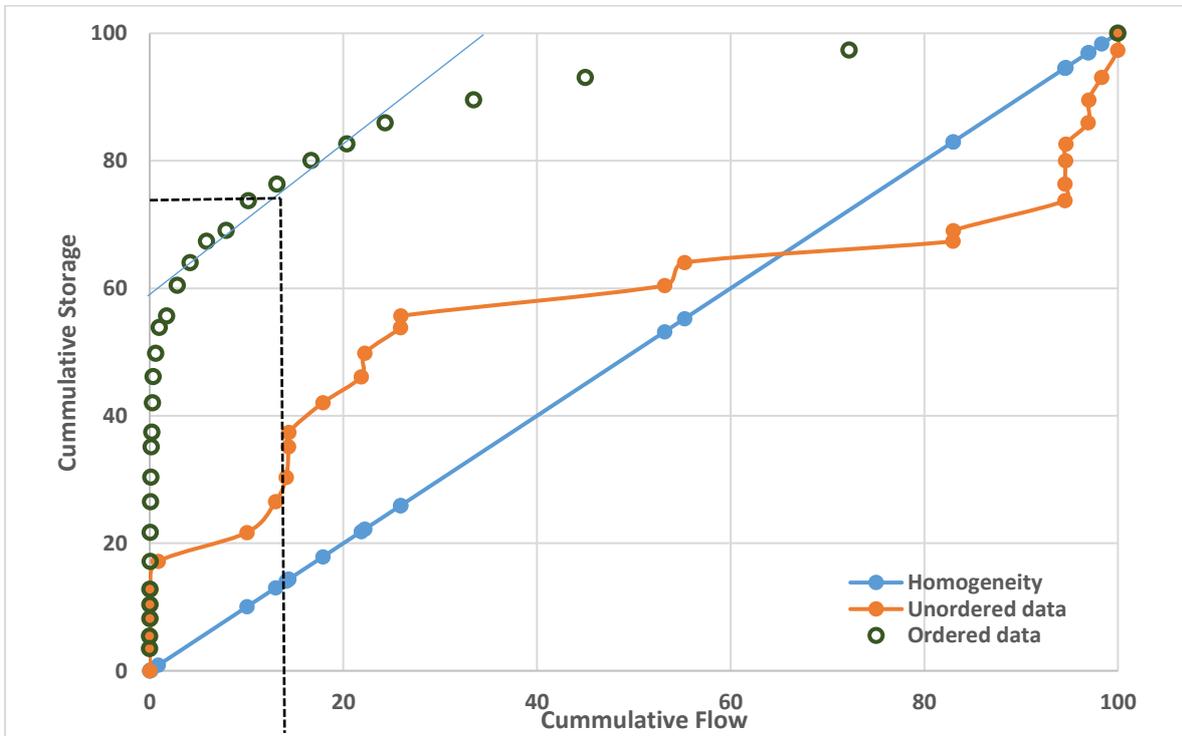


Figure 5.6 Lorenz plot and Modified Lorenz plot for well 3 sandstone reservoir. Here 85% of flow is from 25% of storage.

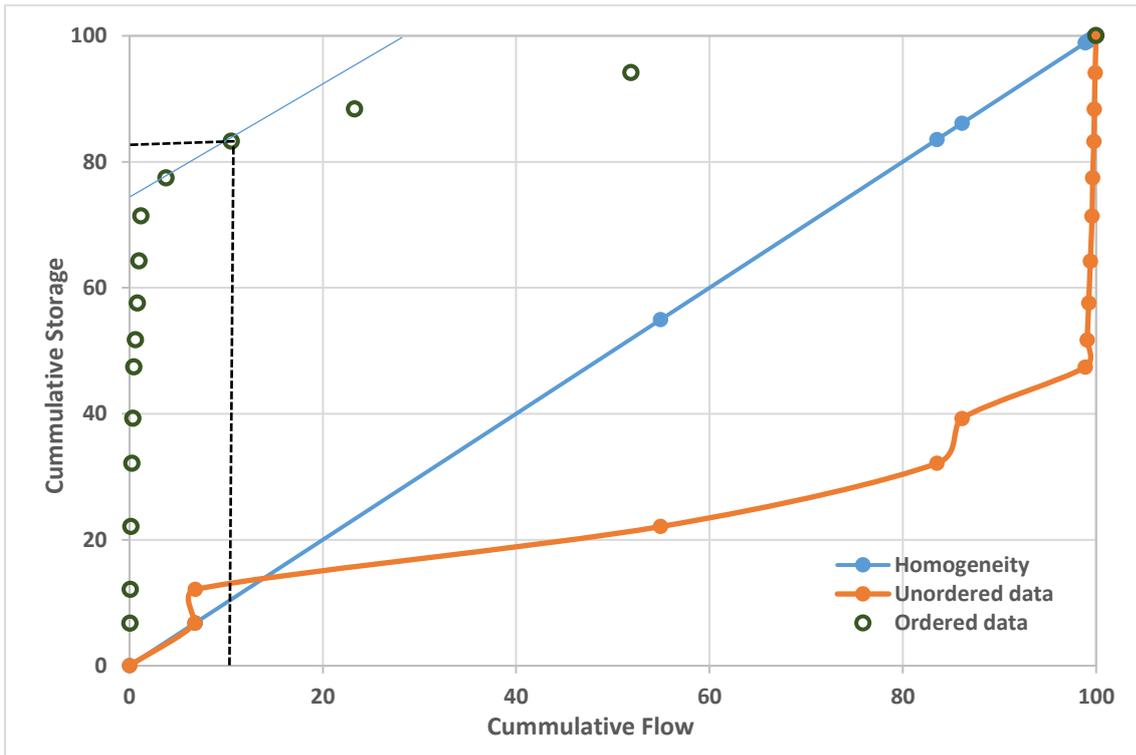


Figure 5.7 Lorenz plot and Modified Lorenz plot for well 2 sandstone reservoir. Here 90% of flow is from less than 20% of storage.

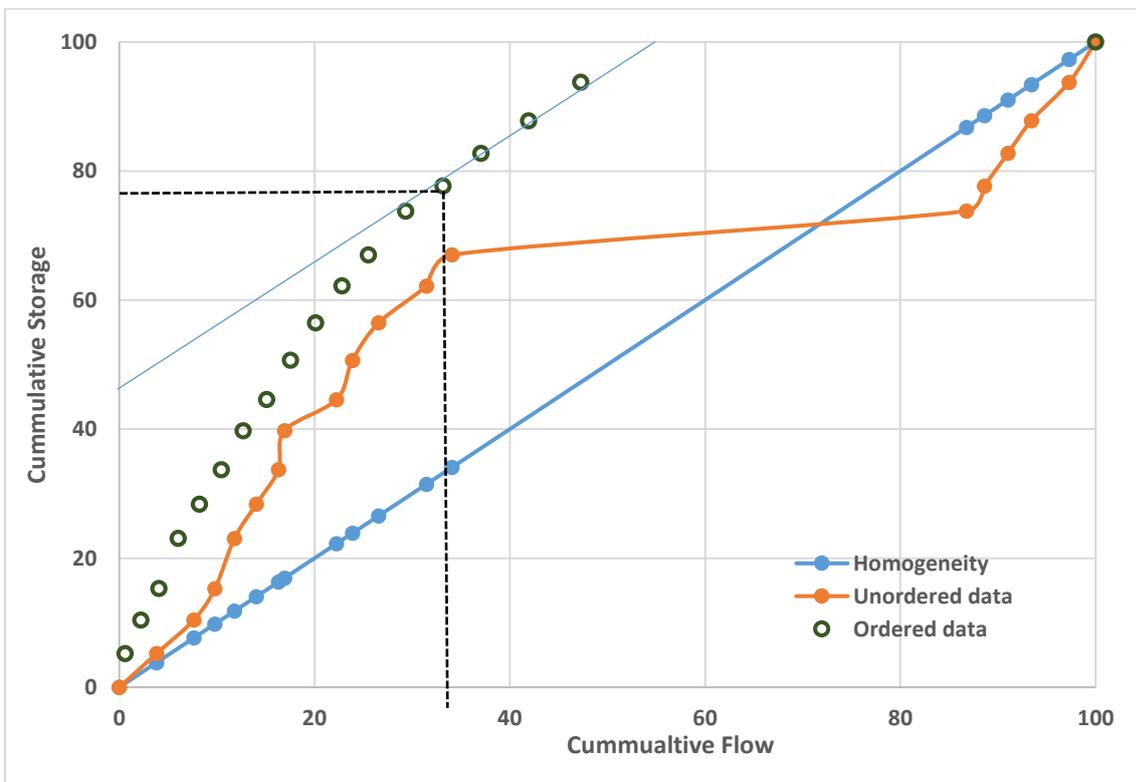


Figure 5.8 Lorenz plot and Modified Lorenz plot for well 6 sandstone reservoir. Here 70% of flow is from about 25% of storage.

5.3.2 Pressure Transient Analysis

Figure 5.9 is a history plot illustrating a 70-hour buildup test after an initial pressure drawdown of about 50 hours at constant rate of 1500 STB/D (although the entire drawdown period was for 85 hours). The green line is the observed pressure data from the reservoir system while the red line is the simulated bottomhole pressure from the model. The behaviour allows the identification of the applicable well test interpretation model, which controls the maximum number of parameters that can be obtained from a test and the meaning of those parameters.

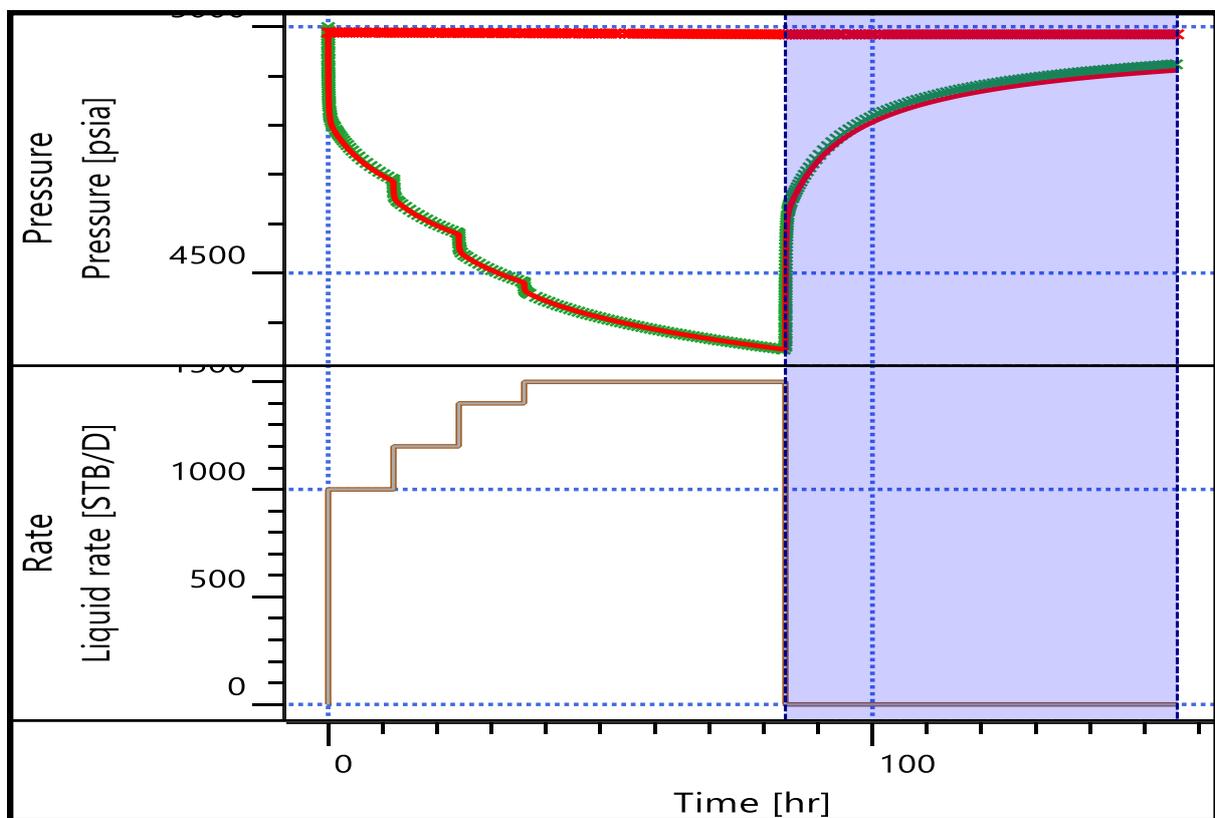


Figure 5. 9 History plot indicating pressure buildup analysis with observed data (green) matching the model (red).

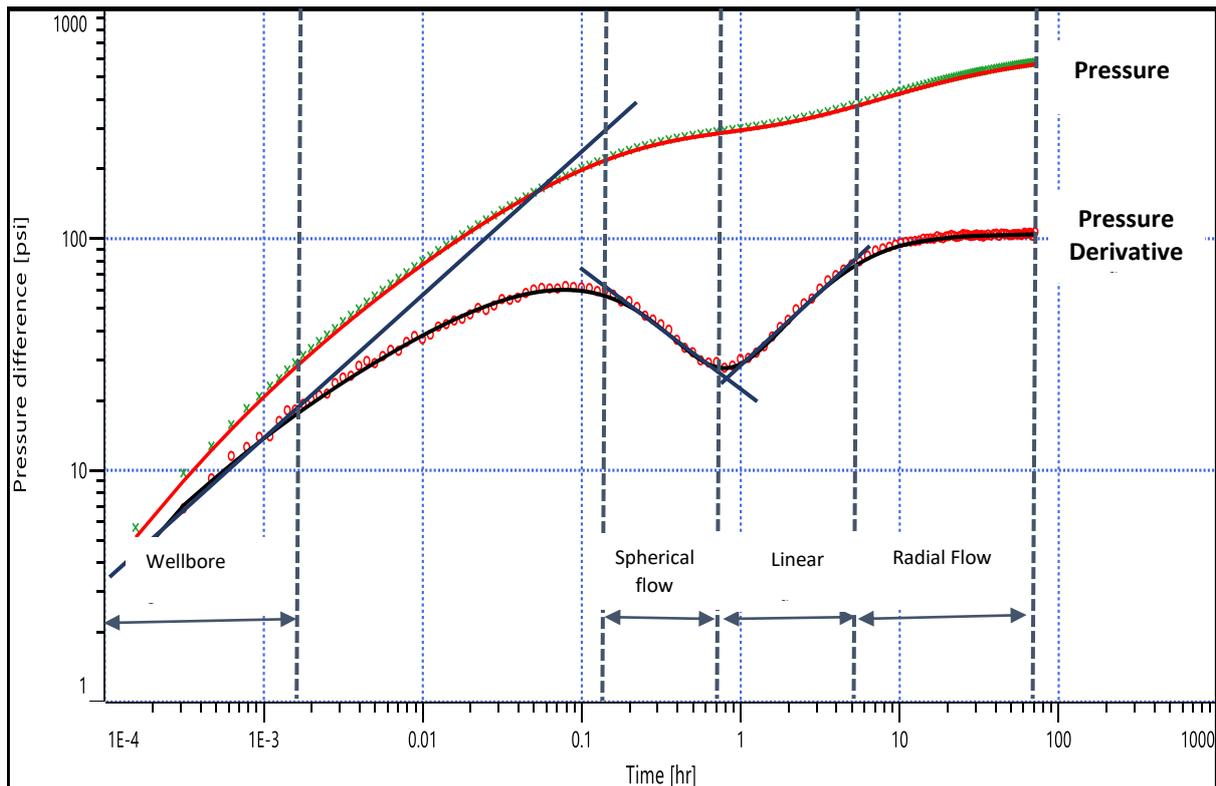


Figure 5. 10 Log-log diagnostic plot showing pressure change and derivative

Figure 5.10 describes a well with wellbore storage as well as skin in a homogenous reservoir with infinite conductivity fracture well model, dual-porosity pseudo-steady state reservoir model and infinite acting radial flow boundary model. At early time, the curves tend towards each other asymptotically with a slope equal to unity representing wellbore storage effect. Since it seems like the well is seeing other reservoirs (the possibility of an infinite conductivity fracture or well connected fracture network suggested by the strong slope of the derivative and the clear stabilization) and thus be cleaning up, the wellbore storage coefficient could drastically increase within a short space of time. When this early time wellbore storage effects is over, the constant sand face flowrate is then established. A double match of both pressure and its derivative ensured a higher level of confidence in the obtained results (Bourdet et al., 1983). Identification of boundary effects and analysis of heterogenous formations are deciphered from the derivative response since the skin coefficient has limited influence on it. Looks like the flow is not matrix dominated and this is supported by an evident $\frac{1}{2}$ slope trend in the pressure derivative plot (that indicate the presence of linear flow towards and into the conductive fractures). It does not appear as if the well is struggling to deliver rate. There is no evidence from the data of the wellbeing in a closed volume due the lack of a derivative roll-over and the system exhibits a radial flow behaviour representing an equivalent homogenous system composed of all producing elements.

The pressure response flattening during transition is generally difficult to spot on the log-log scale, thus in many instances, a semi-log scale has to be deployed for refining the pressure curve match. But with the derivative plot (Figure 5.10), the response heterogenous nature is clear thus eliminating the need for further adjustments with more plots although there must be consistency among all plots. The gain in sensitivity of the derivative technique is illustrated by this double porosity model.

The property of interest in these different well test analysis been permeability thickness of was found to give a value of 15mD which is the quotient of its conductivity and pay thickness. The heterogeneities were also amplified on the derivative with flow regimes having clear shapes and no doubt the overall quality of the interpretation improve by it.

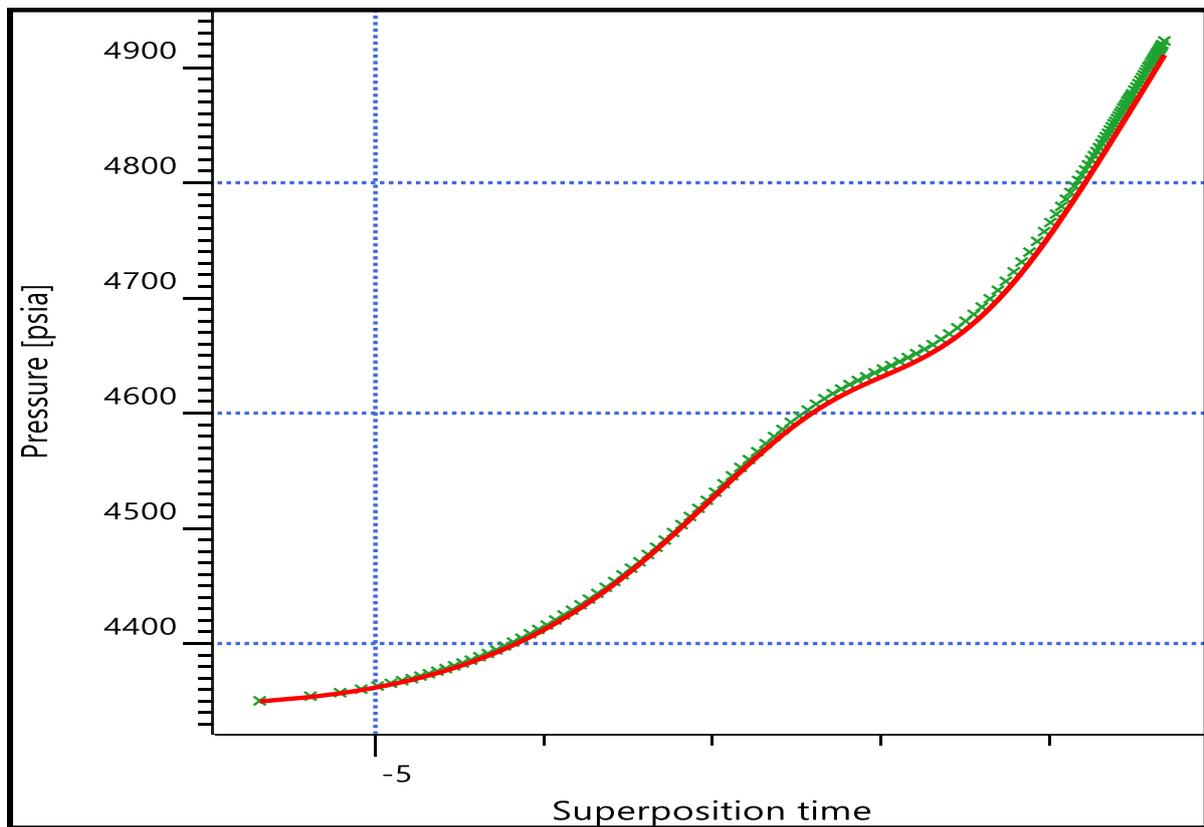


Figure 5. 11 Superposition plot of actual and simulated pressure

Figure 5.11 is consistent with the results from all other approaches such as the diagnostic and log-log plots mentioned earlier. No noticeable unusual jump which could indicate uncertainty in the rate history used for the analysis or that the well is seeing a higher pressure unit before it was shut-in. This kind of plot is justified as a global diagnosis with the necessary identification of all successive flow regimes present in the analysis is now part of a unified well testing methodology which can obviously help to provide more than just kh , skin and initial pressure. Figure 5.12 shows the numerical model defined based on the model type and the results obtained from the above mentioned diagnostic tools. This allowed us to create a model with the regular reservoir boundaries, faults and additional wells defined interactively.

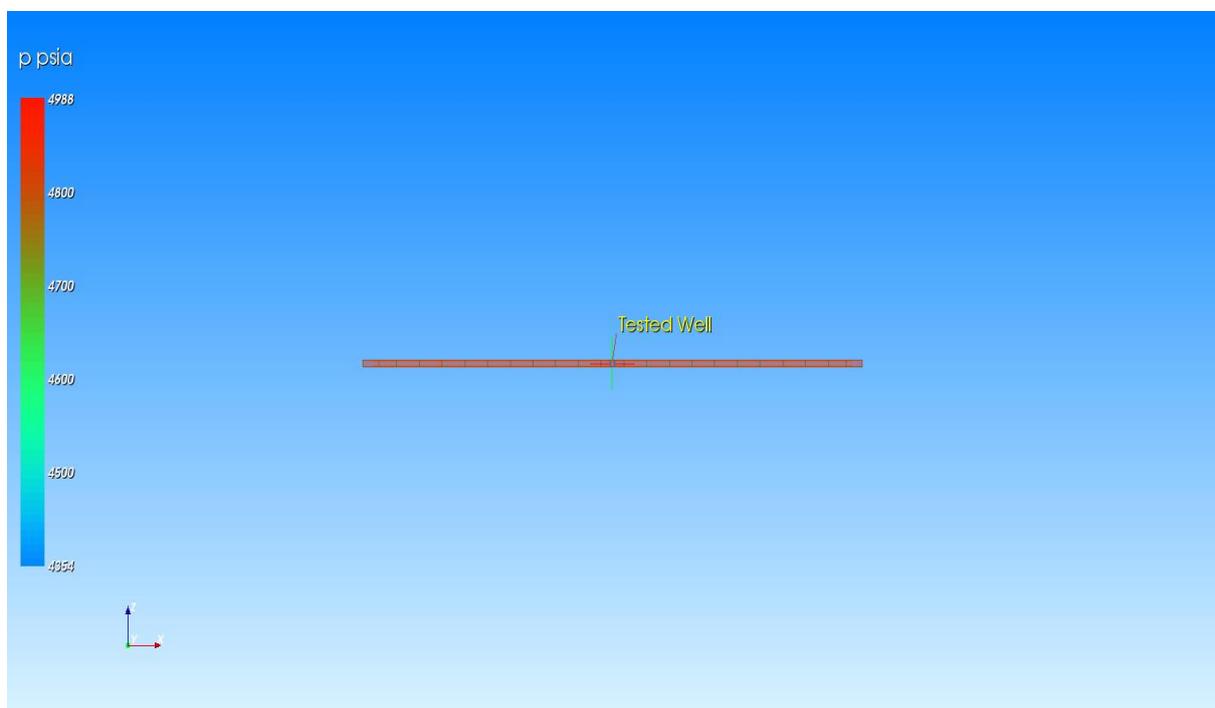


Figure 5. 12 Numerical modelling of the tested well

Figures 5.13-5.15 shows the sensitivities of some the most influential reservoir parameters on pressure change on log-log plots as also described by the Tornado chart in Figure 2.11. It is clear to see that permeability shows more sensitivity. This is in agreement with the Monte Carlo simulation carried out in Chapter Two. Porosity (values of 0.12, 0.20, 0.29, 0.37 and 0.46 were used with the base value been 0.23) had the least impact on the pressure change and its derivative. It is also very vital to use accurate value for the pay thickness as it can greatly impact permeability results (thick values of 50ft, 88ft, 120ft, 165ft and 200ft were used with the base value been 100ft). There are lots of uncertainties associated with their accurate determination as the pay thickness at the well may increase or decrease drastically away from the well. Thus we cannot rely on just one source of permeability hence Chapter Six builds a relationship based on the core, well log and well test permeabilities in order to capture both local and regional heterogeneities. More of this plot on another field can be found on Appendix A-7 to A-10.

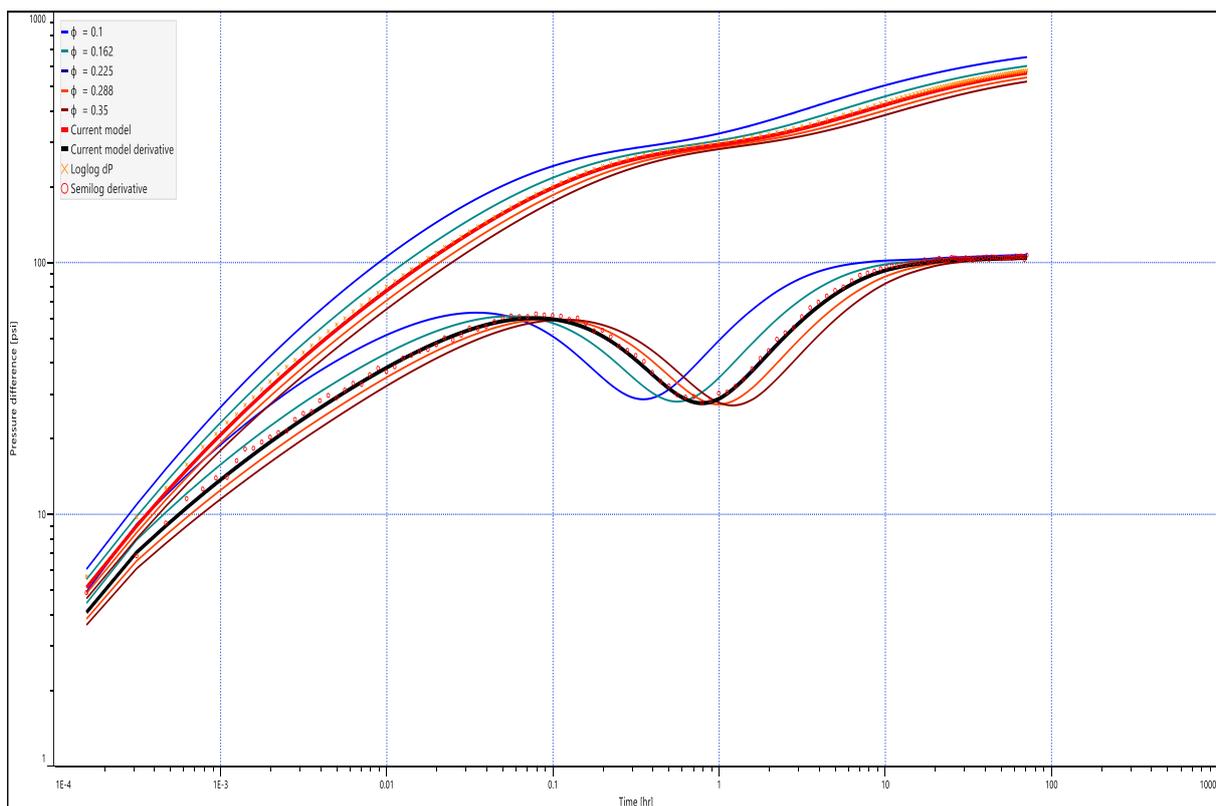


Figure 5. 13 Log-log plot showing model sensitivity to porosity

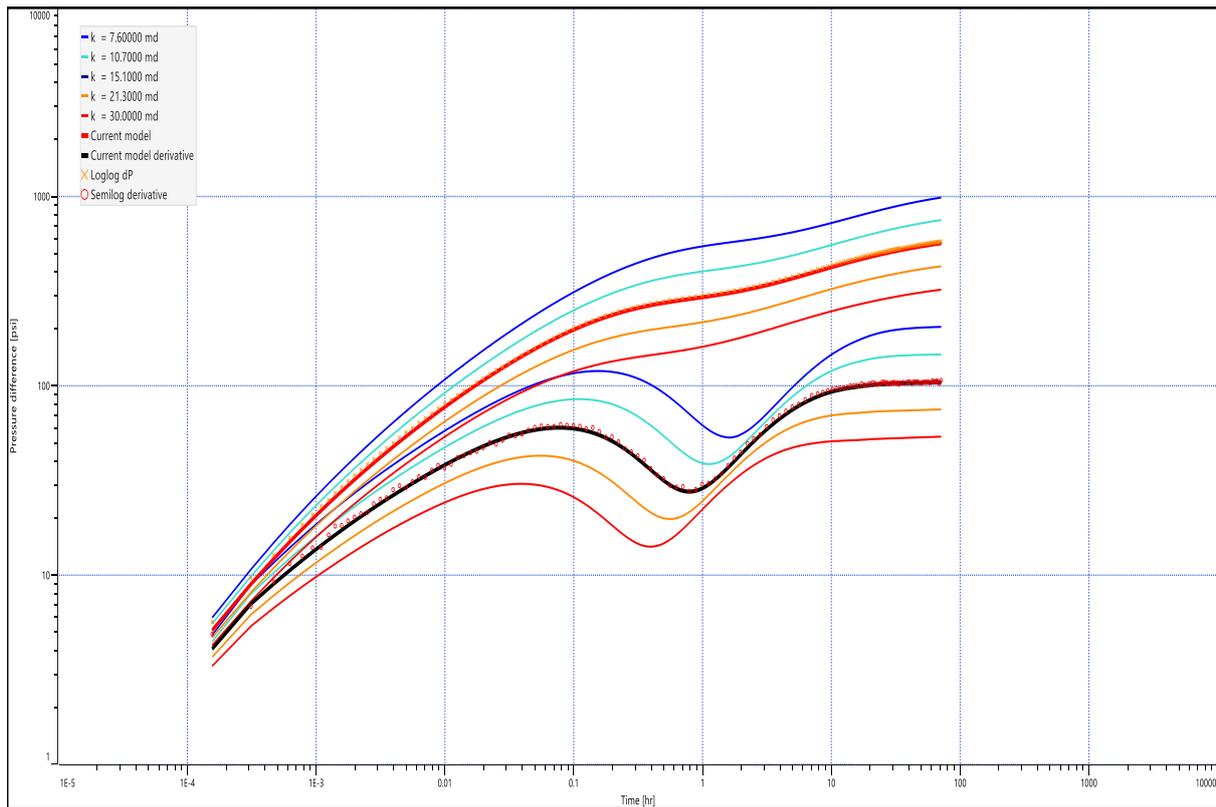


Figure 5. 14 Log-log plot showing model sensitivity to permeability

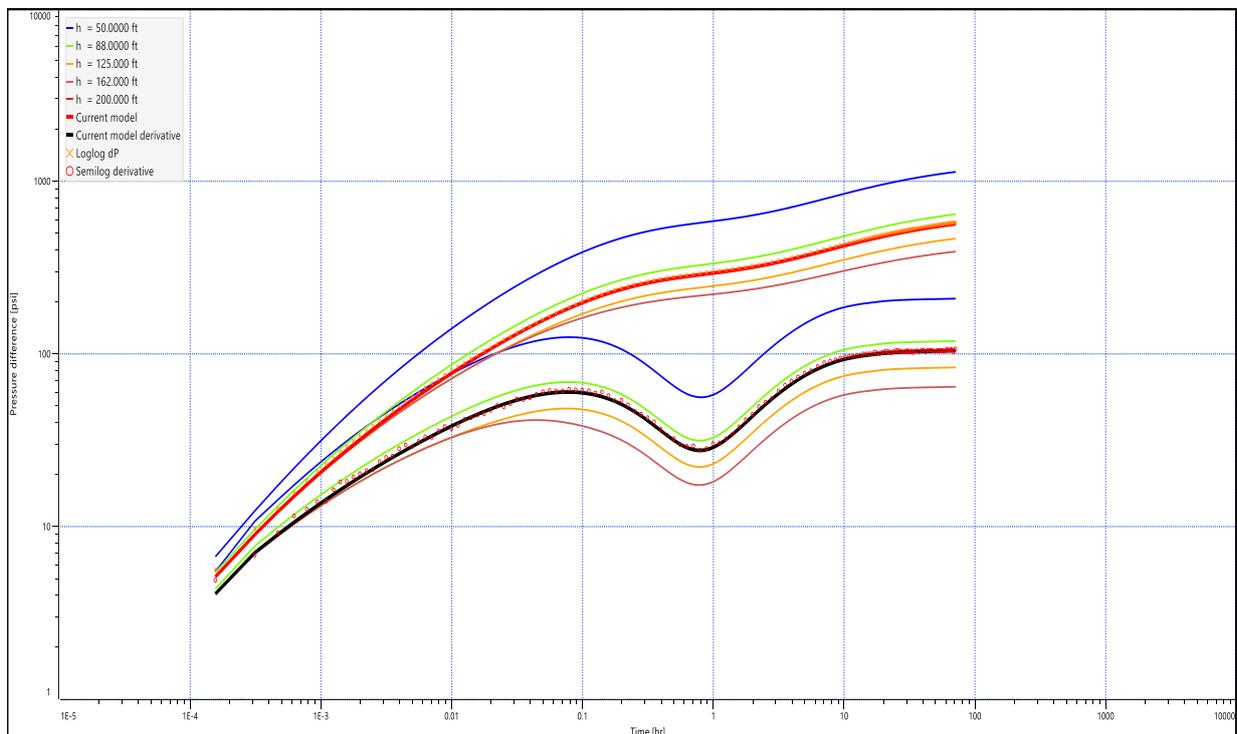


Figure 5. 15 Log-log plot showing model sensitivity to pay thickness

Figures 5.16-5.18 shows these sensitivities on both the drawdown and buildup pressures of the history plot. The disparity caused by porosity sensitivity is less than that produced by pay thickness and permeability on the stabilization pressure when the well was produced at a constant flowrate just before shut-in. At the beginning of constant production, changes in porosity of the reservoir to even up to 50% could not drastically affect the pressure stabilization before shut-in. Due to the uncertainty on pay thickness as already pointed out, its over prediction has less impact on both the stabilized pressure drawdown and buildup (Figure 5.18) as well as on the pressure derivative (Figure 5.15). The model used a thickness of 100ft but increasing it to 125ft, 162ft till 200ft generally had less of an impact than reducing it to 50ft. Although Zheng et al., (1996) estimated that a well not located at the average thickness of a varying channel sand body may lead to errors in well test permeability by 4-15%, it appears it is better to over-estimate pay thickness than underestimate it for relatively accurate prediction. Same conclusion can be drawn from Appendix A-11 to A-14 showing the sensitivity of porosity, permeability, pay thickness and total compressibility on another field. This comes handy as it is an obvious possible source of error or uncertainty as the most logical thing to use measurement done at the wellbore although the difficult but more accurate approach is to use the average pay tens or hundreds of metres away from the wellbore. There is somewhat of a trend for the well test permeability to be less than the core and well log permeability probably due to the fact that well test measures average reservoir permeability values unlike core and log which give permeability across the thickness of the reservoir and therefore were averaged for the comparison.

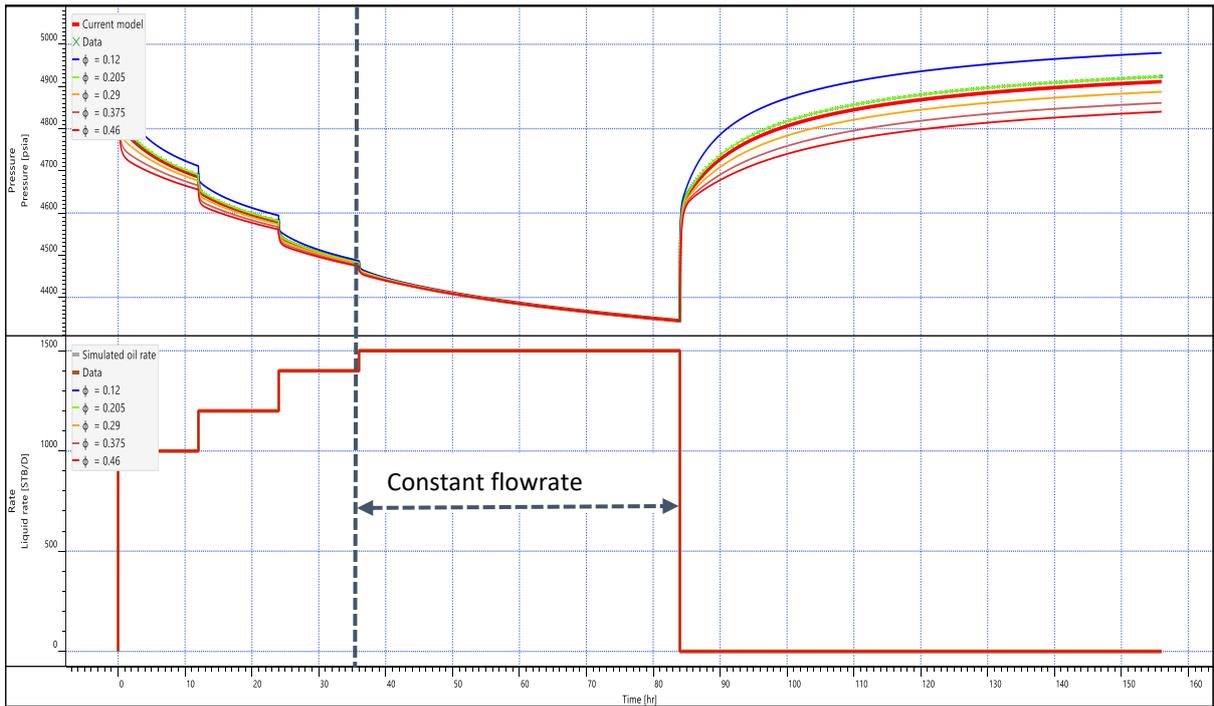


Figure 5. 16 Porosity sensitivity on history plot

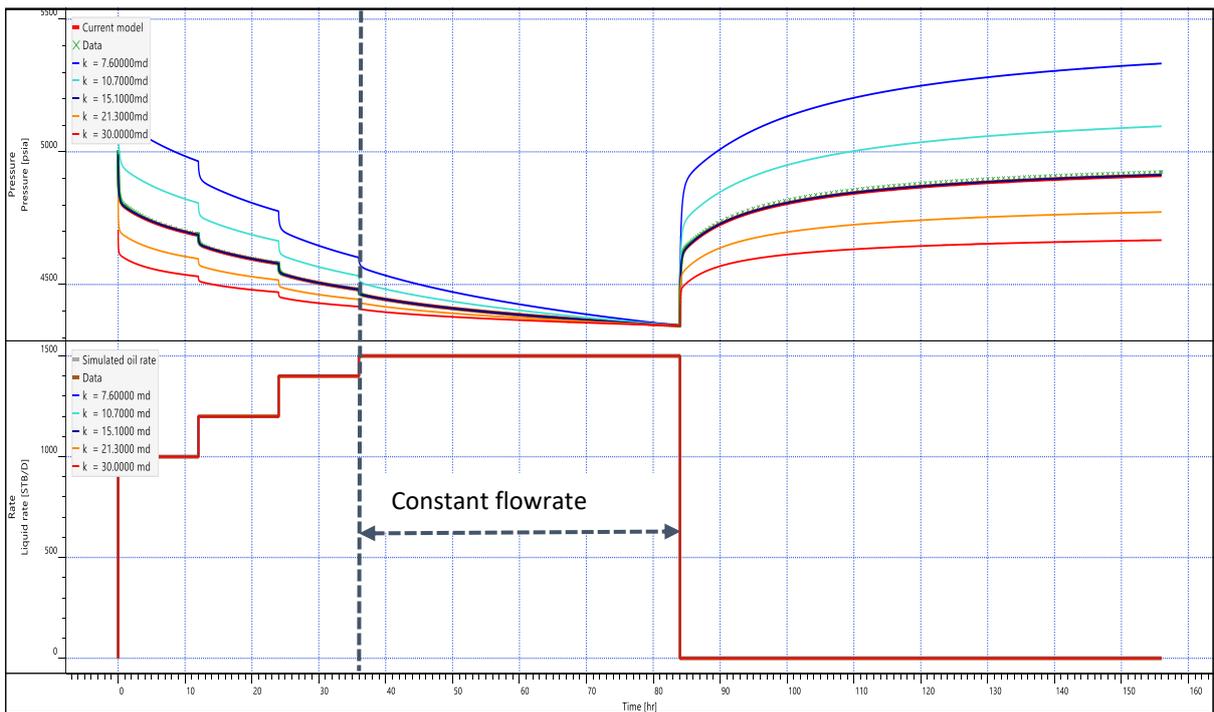


Figure 5. 17 Permeability sensitivity on history plot

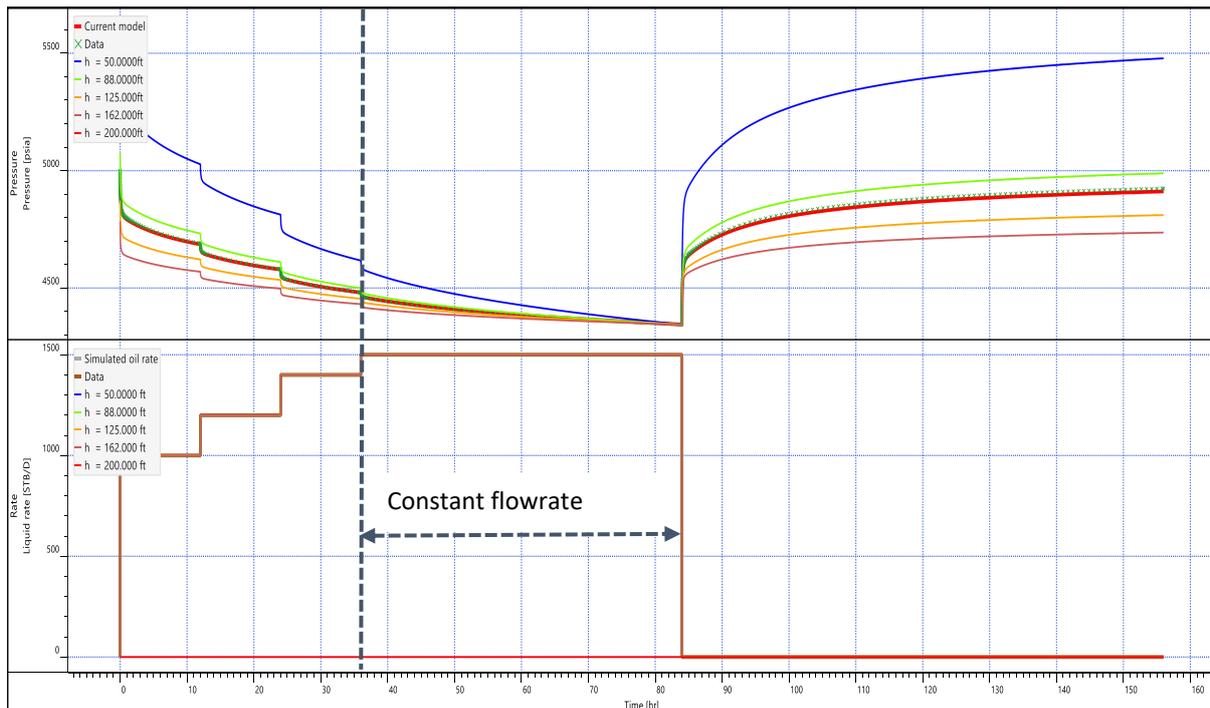


Figure 5. 18 Reservoir thickness sensitivity on history plot

5.4 Summary and Conclusion

This chapter described an in-depth well test analysis to obtain reservoir parameters with the focus on well test permeability to build a database alongside their already known core and well log permeability which has been used in Chapter Six to build a standard relationship between them thereby reducing uncertainty. The following conclusions were made:

- A 50% decrease in the pay thickness appears to cause more destabilization on constant and zero flowrate pressure drawdown, buildup respectively and its derivative than a 100% increase in pay thickness. Overall a high decrease in porosity, permeability and pay thickness was more detrimental than a high increase. With this been said, it is very important to accurately define Effective Flow Interval defined. Thus its calibration with other data sets such as core and logs is discussed in Chapter Six.
- Lorenz plot was used to determine the Effective Flow Interval as well as the heterogeneity of the reservoir formation.

Chapter Six

Log, Core, Test Across Scale Uncertainty Reduction (LCTUn)

6.1 Introduction

In heterogenous reservoirs, permeability is arguably the most important reservoir parameter as it is very crucial for field development. Evaluation of logs, core and test from several wells showed permeability across board to be different which also highlights the uncertainty in their interpretation in terms of scale, physics, environment, heterogeneity, formation damage, averaging issues etc. The question often asked is which source is the most accurate and can a relationship be developed between them in order to determine any of them i.e can we make good use of well logs and core permeability for example, to evaluate well test permeability with accuracy which is particularly important in the deep water offshore environment and marginal fields where the cost of DST's in every well in an entire field as well as in the different pay zones which the well penetrates, may not be justified.

Although Van der Post et al., (2015) presented a technique for estimating long term reservoir performance from log data having shown that core permeability over predicted reservoir productivity by almost 10-fold and also showed that it is not a good predictor of mid-term well performance, Corbett et al., (1998) carried out a study which gave insights into both core and well test integration and scaling of their measurements showing the need for improved geological, petrophysical and dynamic descriptions in fluvial systems. Appropriate log, core and test depth matching was also ensured including upscaling. Arithmetic, harmonic and geometric averages were determined. A database of core, logs and well test permeability from different sandstone reservoirs in the North Sea was developed. This comparison made before the permeability data is used in a reservoir model is a vital aspect of the reservoir description methodology as well as looking at the result from each one being used. Several machine learning algorithms were applied to build a relationship between these data since all the scales are relevant for robust characterization and the Support Vector Machine was found to be suitable. Model tuning was done in order to optimise its key parameters. Nomographic approach involving a numerical simulation run iteratively on a multiple non-linear regression model obtained from the dataset was also run.

This Chapter answered the above questions of which source is the most accurate and a relationship being developed between them. The gap between log, core and test was also closed

using machine learning and nomographic methods. From the smaller scale core and log permeabilities, the larger scale well test permeability was learnt using the machine learning technique thus reducing its uncertainty. The machine learning model was then coded into a dashboard containing the inputs for its training (Appendix B-4). Their relationship provides the bench mark to calibrate one against the other and also create the platform for real time reservoir properties prediction. The technology was applied to an independent dataset from Central North Sea deep offshore sandstone reservoir for the validation of these models with minimum tuning and thus effective for real time reservoir and production management.

6.2 Methodology

6.2.1 Scale Analysis

Important to state that the datasets considered have different scales (Figure 6.1). Almost all reservoirs are highly stratified. No matter the reservoir, as heterogeneity increases, permeability distribution tends towards its average value. Core probe permeameter measures small scale variations in permeability. Studies of this type have been done on cores (Hurst and Rosvoll, 1991; Corbett and Jensen, 1993) and also on the field (Goggin et al. 1988; Ringross et al. 1993a). Cross bedding which creates laminae is one of the most common characteristic structures of sandstones. Permeability contrast between these laminae are commonly in the range of 4:1 to 10:1 and may have a significant effect on oil recovery due to capillary trapping (Ringrose et al. 1993b). The size of these laminae captured by cores and logs ranges from millimetres to centimetre in thickness and from centimetre to decimetre scale in length. As this is small to be included in reservoir simulation models, their permeabilities were upscaled by averaging to provide effective permeability for large grid blocks of scales of 10metres or more (bearing in mind that well test gives average permeability as well). Several other reasons why core, log and well test permeability are different are shown in Table 6.1. Averaging is an upscaling method and upscaling can be first level where small set of laboratory data need to be interpreted at reservoir scale (Paterson et al., 1996) or second level meaning scaling properties from fine geological grid to coarse simulation grid. Figure 4.38 shows the porosity/permeability for both Northern, Southern and Central North Sea respectively. It is difficult to see a clear relationship between permeability and porosity from such a plot, since from a given porosity value, permeability may vary by several orders of magnitude. Hence there is need for careful upscaling through averaging so that predicted log permeability measurements can be related to larger scale permeability such

as well test. Again anomalous values were smoothed by upscaling while also retaining the original data.

Simple averages are used if flow is parallel or perpendicular to continuous layer (Marsily, 1986). These averages can be arithmetic, harmonic or geometric. For flows linear and parallel to the sedimentary layers and general laminar layering, arithmetic average should be used for effective permeability (Quintard and Whittaker, 1988). If permeability has a log linear relationship with porosity on a log scale, the upscaled values will lie slightly above this line. For n layers having permeabilities and thickness, it is described with equation below:

$$\bar{k}_{arith} = \frac{1}{n} \sum_{i=1}^n k_i \quad (6.1)$$

The harmonic is used if the flow is linear and perpendicular to the layers as it is equal to effective permeability. Mainly used for vertical permeability and hardly used for horizontal permeability except in steeply dipping or fractured systems. If permeability has a log linear relationship with porosity at log scale, harmonic averaged will lie below this line as the lower values control flow. It is represented below:

$$\bar{k}_{harm} = n \left(\sum_{i=1}^n \frac{1}{k_i} \right)^{-1} \quad (6.2)$$

Geometric average is used when there is no apparent preference for horizontal or vertical flow and there is no significant anisotropy for flow. Useful for well sorted sandstones with hardly any vertical barriers. If permeability has a log linear relationship with porosity at log scale, the upscaled values will lie approximately on this line trying to retain low and high values.

$$\bar{k}_{geom} = \left(\prod_{i=1}^n k_i \right)^{\frac{1}{n}} = \exp \left(\frac{1}{n} \sum_{i=1}^n \log_e(k_i) \right) \quad (6.3)$$

Measure of heterogeneity

$$\bar{k}_{harm} \leq \bar{k}_{geom} \leq \bar{k}_{arith}$$

Matheron (1967) stated that for steady-state linear flow, effective permeability will always lie between arithmetic and harmonic. If flow is at an angle to the layers, crossflow will occur i.e. flow may be induced perpendicular to the applied pressure gradient. In this work, arithmetic averaging was mostly applied as according to the principle of original horizontality most wind and water laid sediments were originally deposited horizontally before subsequent tectonic activity. For core data having vertical permeability measurements instead of horizontal permeability, harmonic averaging was used while in areas where reservoir tends towards homogeneity from their Lorenz plot, geometric averaging was used. Another term that is arguably used in place of averaging is upscaling. It is usually performed for every cell and properties in the coarse grid required in the reservoir dynamic flow simulation. Upscaling can hence be regarded as an averaging process where the static and dynamic properties of a fine scale model are approximated by that of a coarse scale model.

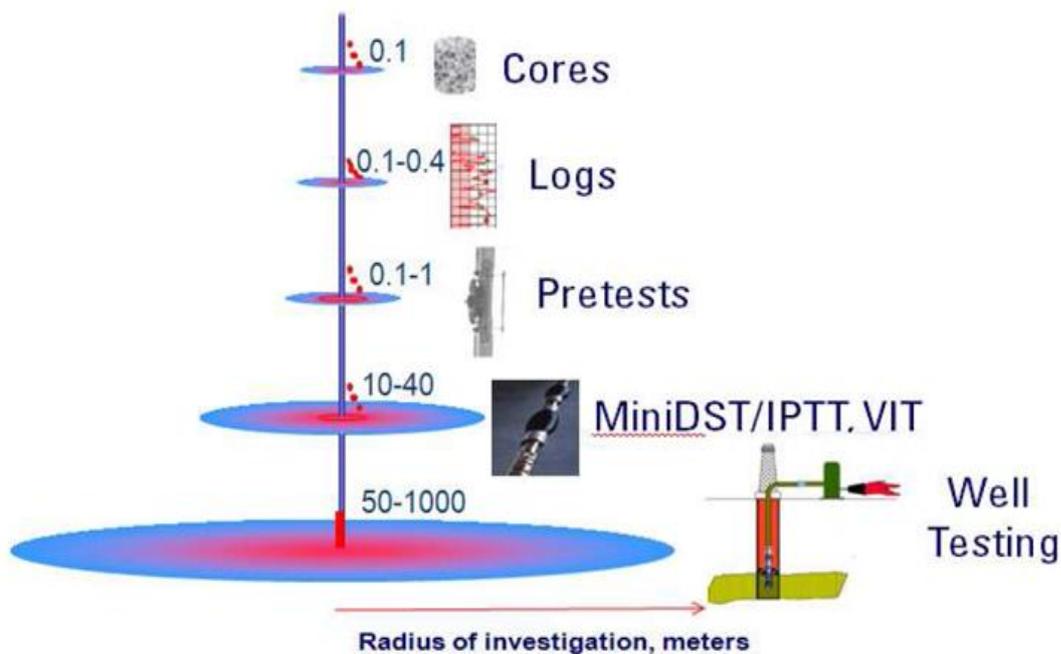


Figure 6. 1 Relative scales of permeability sources (Cantini et al., 2013)

6.2.2 Reconciling Permeability

Any agreement between core, log and well test happens if the zone of interest is relatively homogenous. The required additional uplift to match dynamic performance can be found in alternative mechanisms for distributing permeabilities away from the wellbore in the static model. Sequential Gaussian Simulation (SGS) methodology applied to build the static model produced significant lateral permeability changes from grid block to grid block in the model and

thus could degrade permeability by 5-15% compared to a more uniform layered permeability system. A variogram (type exponential) range of 5000 in both major and minor directions was used with an azimuth value of -45, a nugget of 0.1 and a sill of 1.0. There is definitely some variability in the oil and gas industry as to how these multiscale data should be expressed. In some instances, modification of the porosity-permeability transform by applying corrections and assignment of permeability in the static model enable permeability reconciliation between static and dynamic data. Important to note that there are different objectives for well testing. It could be to assess the degree of damage induced by the drilling and completion operation etc. Again the well test considered single phase flow for simplicity and due to the fact that only one fluid is flowing during the test hence viscosity is also assumed to be constant. Reconciliation was done having Table 6.1 in mind and it was necessary to separate between log and core permeability.

Table 6. 1 Factors to consider when reconciling permeability data (Dubrule and Haldorsen, 1986).

S/N	Factor	Effect
1	Core permeability not corrected for Klinkenberg effect	K_{core} too high
2	Core permeability not corrected for overburden pressure	K_{core} too high
3	Core permeability not corrected for initial water saturation influence	K_{core} too high
4	Core permeability not corrected for the effect of temperature	$K_{core} \neq k_{insitu}$
5	Core altered during sample recovery and preparation	$K_{core} \neq k_{insitu}$
6	Core samples does not contain natural fractures	$K_{core} \neq k_{PTA}$
7	Core samples have natural fractures	K_{core} usually too high
8	Unconsolidated core missing i.e not included in average permeability calculation	K_{core} too low
9	Different averages from different core averaging techniques	$K_{arith} > K_{geom} > K_{harm}$
10	PTA performed when multiple phases are flowing	$K_{PTA} < K_{core}$
11	Wrong choice of PTA flow model (Radial, linear, spherical, hemispherical)	K_{PTA} incorrect
12	The perforated interval for PTA and the cored interval do not match	$Kh_{PTA} \neq Kh_{core}$
13	When bedding, baffling and tortuosity reduce effective permeability	$K_{PTA} < K_{core}$

° K_{PTA} – Permeability from pressure transient analysis

° K_{core} – Permeability from core

A support vector machine algorithm using input features like $klog$, $kcore$, $(klog)^{1/2}$, $(kcore)^{1/2}$, $(klog)^2$, $(kcore)^2$ was developed to predict well test permeability after trying several other machine learning algorithms (i.e machine learning was used for reconciling the datasets). A total of 114 datasets was randomly divided into 105 training set and 9 test sets while a separate data from a different field not used for training the model was then used to validate the ability of the model to predict well test permeability. This proportion was chosen since the data was not much and therefore the training set got as much as possible. Training was done using radial basis kernel function. Since these non-linearity and strength of the SVM model are introduced by the kernel function, tuning of the kernel and other SVM parameters was done to achieve good results such as gamma needed by the kernel function, cost constant set as 1, tolerance of termination criterion set to 0.001, epsilon parameter set as 0.1, probability which is a logical parameter indicating whether the model should allow probability prediction was also set as true. Also the SVM does not require any assumptions about its functional form. Because the data is separated with maximum possible margin, the model becomes very robust and able to deal with incongruences like noisy test data and biased training data.

Summarily, the machine learning algorithm may be trained to recognise or correct for the true organisation of permeability based on the averages of the core and log data. The prediction is being corrected from the averages. Various averages were used depending on the stratigraphy at the well location as some contain layered reservoirs (arithmetic averaging was used) while it is more random at other well locations (geometric averaging was used). These were important features of the process.

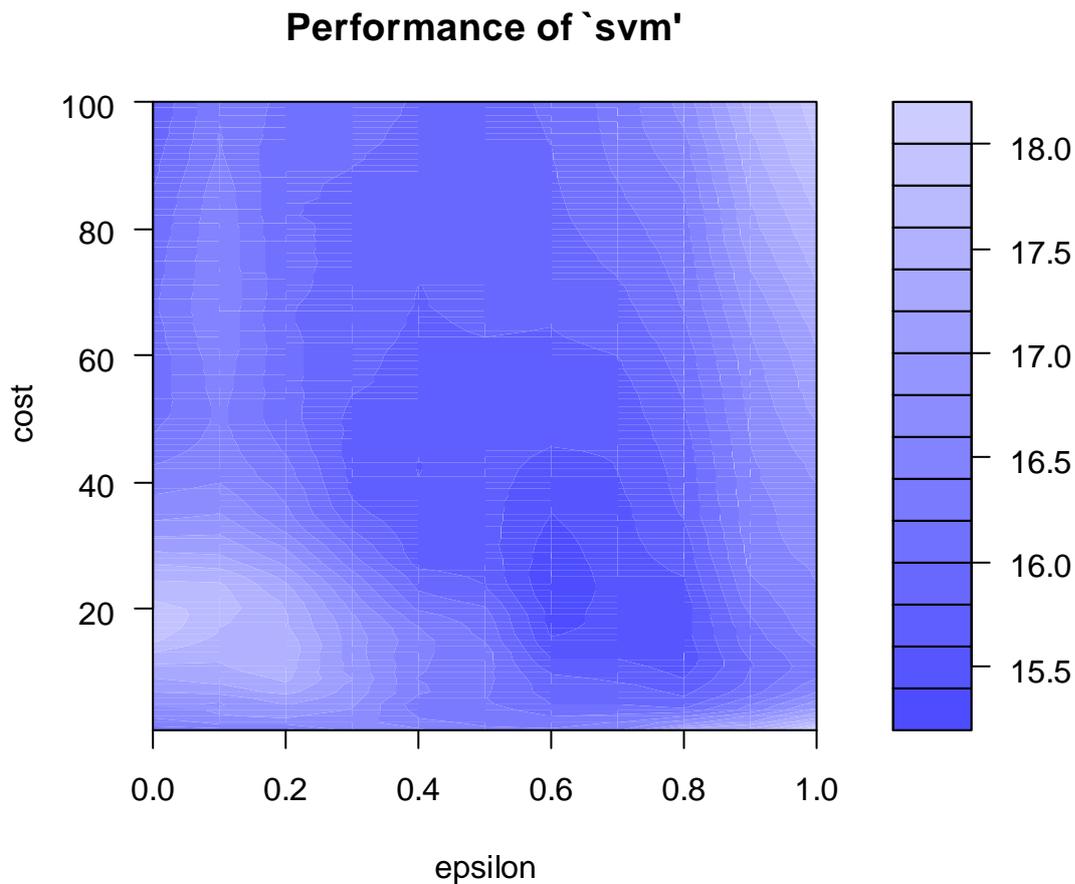


Figure 6. 2 Parameter tuning of the SVM model

Figure 6.2 involves searching for the best SVM model. The tune function used evaluated the performance of 1100 models for combinations of maximum allowable error of 0 to 1 with 0.1 steps and cost parameter of 1 to 100. The optimised model used has value epsilon and cost values of 0.6 and 24 respectively. The accuracy of the models are shown by the colour map with the legend to the right displaying the value of the mean squared error. This means that the darker areas have better accuracy.

Other methods like Generalized Linear Model (GLM), Gradient Boosting Model (GBM), Random Forest (RF), Deep Learning (DL), Stacked Ensemble models (SE) were tried as shown in Appendix B-3. None of these did well compared to the Support Vector Regression Model as they gave high root mean squared error values for the test set. The GLM gave an error value of 38.00, GBM gave error of 38.91, error value was 32.43 for Random Forest, 42.44 for Deep Learning and 47.93 for the Stacked Ensemble model. The Support Vector Regression Model gave relatively lower error values as seen in Section 6.3.2. This is not to say these models are bad, the point is that for this particular case their predictions were less accurate.

Also in terms of features used for the final Support Vector Machine model, the Table 6.2 below shows the effect of feature selection in the accuracy of the model. Using variants of core and log permeability as input features improved the accuracy of the model. With more data to train the model, it will definitely do better.

Table 6.2: Accuracy in terms of RMSE of the different features tried for the model.

S/N	FEATURES USED	RMSE
1	$klog, kcore$	68.12
2	$klog, kcore, (klog)^{1/2}, (kcore)^{1/2},$	17.68
3	$klog, kcore, (klog)^{1/2}, (kcore)^{1/2}, (klog)^2, (kcore)^2$	5.22

6.2.3 Depth Matching Multiscale

Lorenz plot which displays in a graphical form the relationship between permeability thickness (transmissivity assuming uniform viscosity) and porosity-thickness (storativity assuming uniform compressibility) using core and log measurements. It was used to ascertain the dominant flow interval and the degree of heterogeneity. Thus it provides a dynamic cutoff for well test permeability analysis. The production logging tool was also looked at where available to establish the effective pay contributing to the well test across the perforated interval. The PLT was run in some cases to ensure that good contribution was seen across the entire perforated interval and therefore representative data acquired from the test. Sandstone reservoirs can have variable thicknesses as most are water-laid comprising channel-levee (Figure 6.23) complexes stacked on top of each other.

6.3 Results and Discussion

Averaging is commonly used as a form of upscaling. Although Noetinger and Haas (1996) has detailed the appropriateness of the power average due to the problem of estimating its exponents. Table 6.3 describes the averaging techniques along side their well test permeability value showing areas where each of the averaging technique is closely related to the well test permeability.

Table 6. 3 Averages for upscaling core and log data for comparism with well test data

Arithmetic	Harmonic	Geometric	Log Perm	Well Test Perm
13.25	0.19	3.20	19.86	4.7
580.8	554.97	567.41	425.20	97
673.67	360.69	517.73	382.18	154
1742.96	128.28	731	188.32	158
2812	2059.39	2367.27	730.15	27.5
46.78	0.98	11.22	272.50	6.4

6.3.1 Scenarios of More Accurate Scale

Six (6) producing wells were considered in the reservoir with its pressure been maintained by injectors. Core, log and well test permeability were used to model the cumulative oil production rate. Figures 6.3-6.7 shows the different sources varying significantly leading to inconsistent estimate of well and/or reservoir performance. At well 6, core permeability appears to be the closest to observed production rate but performed very poorly in well 2. Log permeability performed best in well 5 while well test did best in well 4. From just the oil production rate, it appears difficult to point to one best source of permeability. Their discrepancy is huge in wells 6 and 5 but seems less in well 4. This discrepancy is a reflection of the degree of heterogeneity of the reservoir.

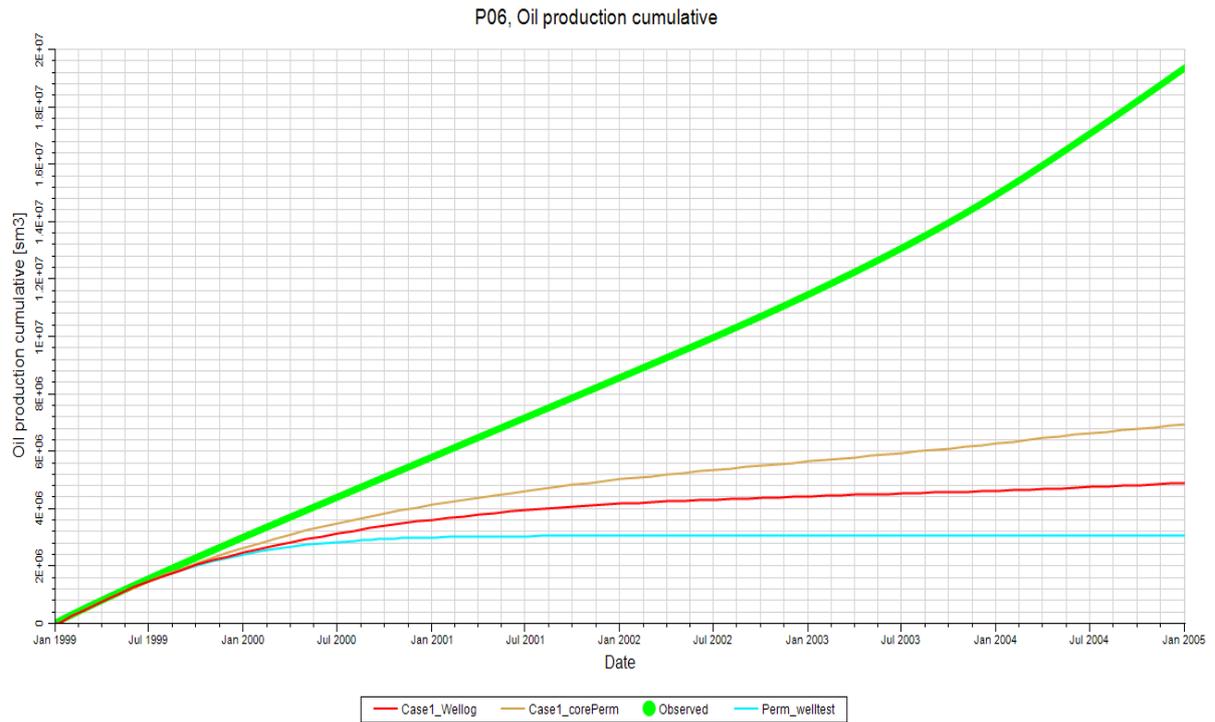


Figure 6. 3 Core, log, test permeabilities to model observed production rate for production well 6. Cases made for core (brown line), log (red line) and well test (blue line) permeability for predicting cumulative oil production. Core was the closest in this case to the actual.

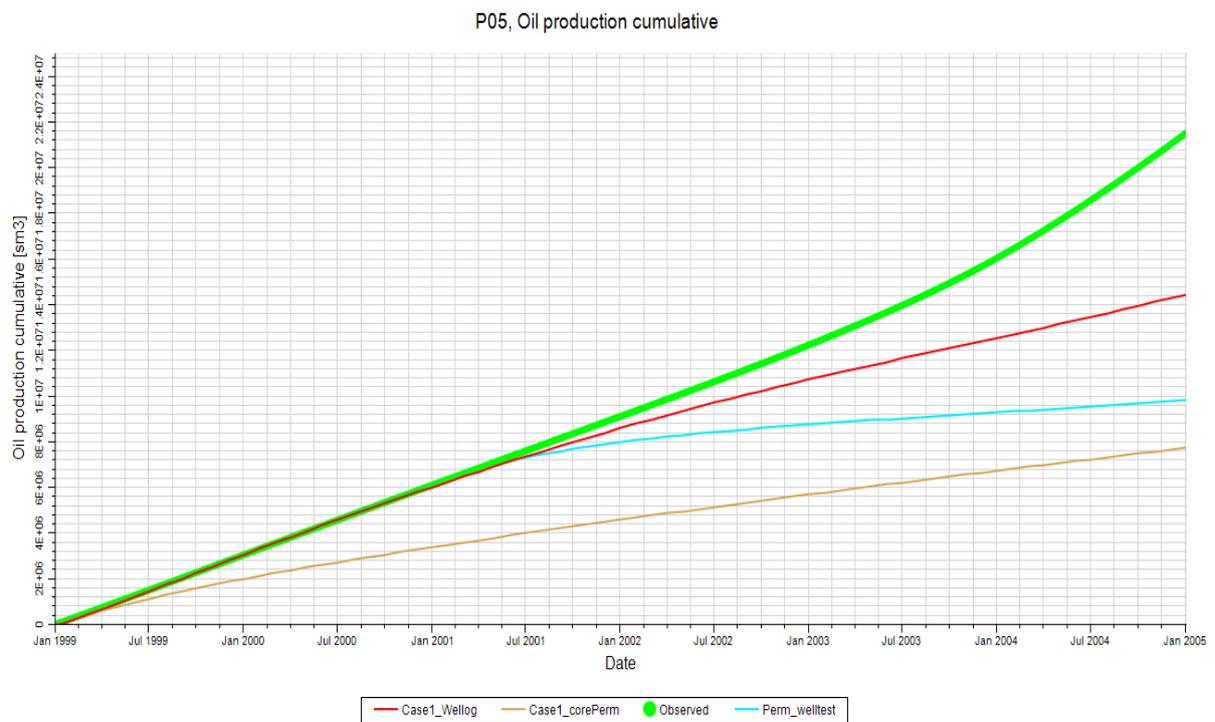


Figure 6. 4 Core, log, test permeabilities to model observed production rate for well 5

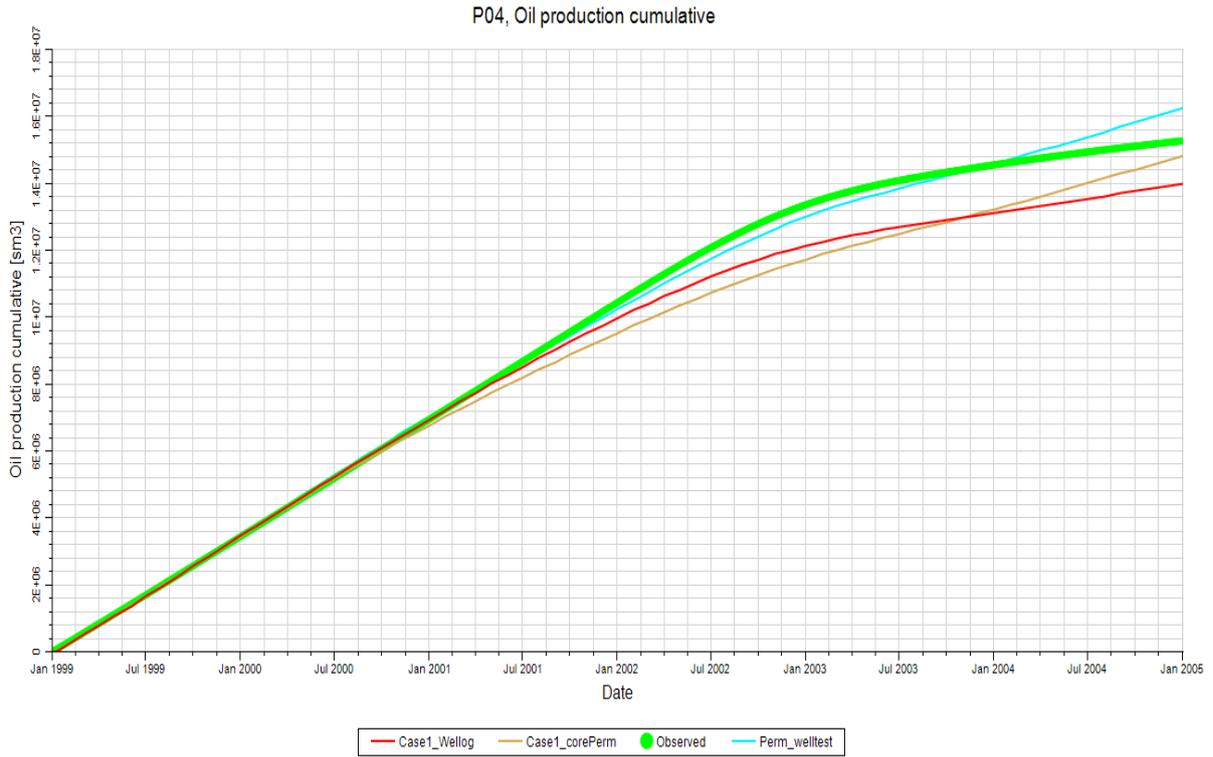


Figure 6. 5 Core, log, test permeabilities to model observed production rate for well 4

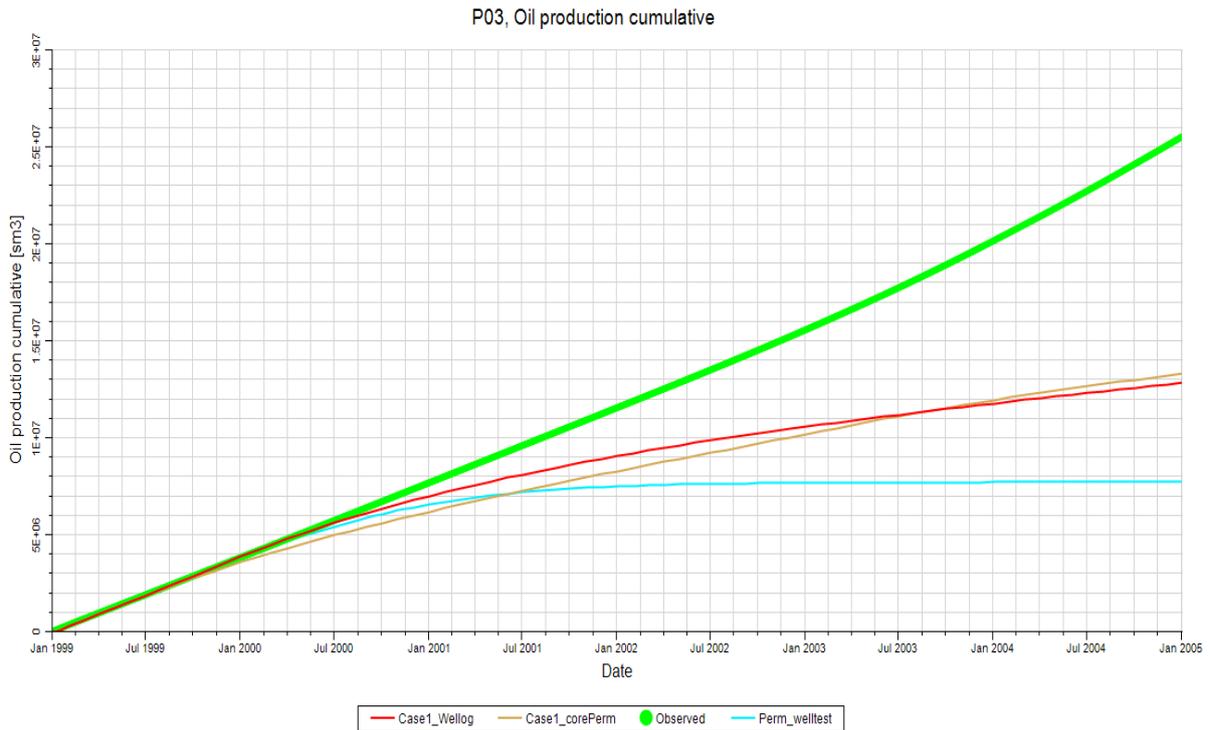


Figure 6. 6 Core, log, test permeabilities to model observed production rate in well 3.

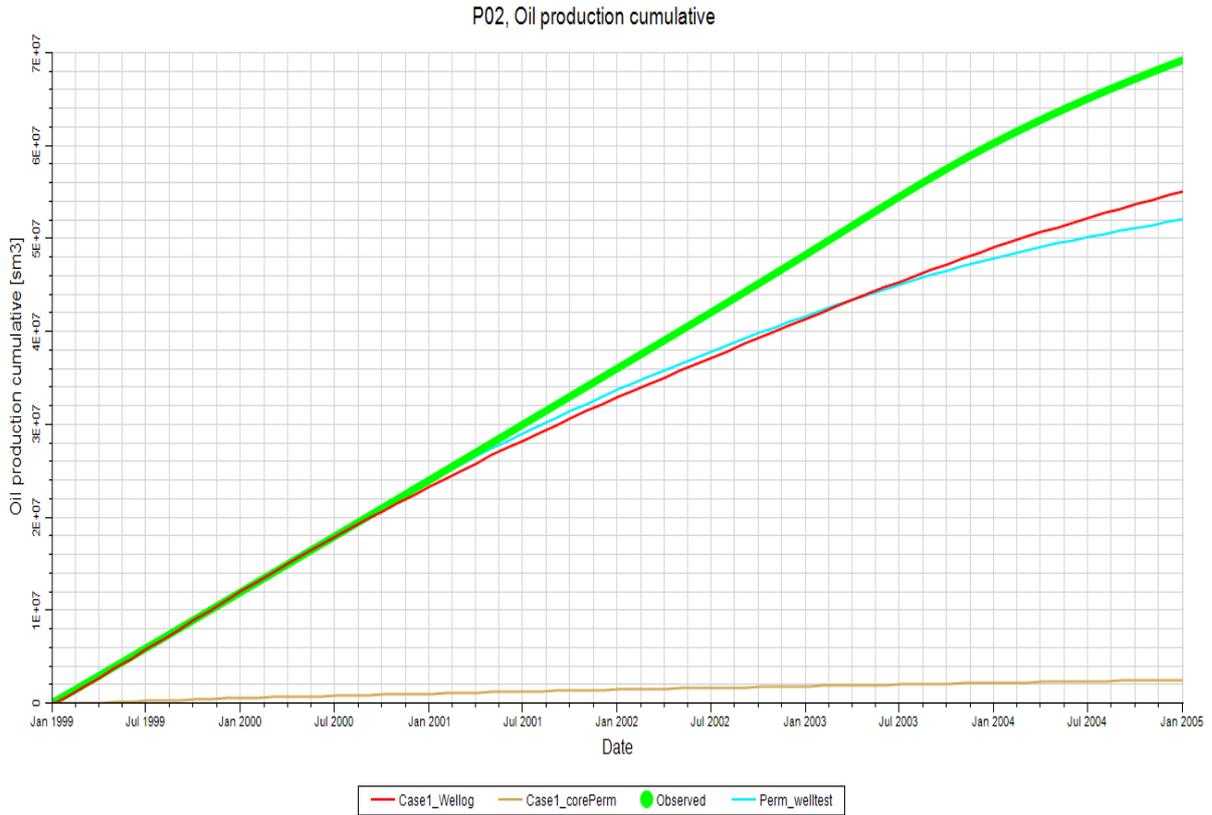


Figure 6. 7 Core, log, test permeabilities to model observed production rate for production well 2

These dataset were also used for the prediction of bottomhole pressure as shown in Figures 6.8-6.10. Well test permeability consistently gave better predictions compared to log and core permeability in all the wells considered. Cores on the otherhand, gave the worst prediction of bottomhole pressure probably because it has the least radius of investigation. Important to note that the flowing bottomhole pressure is also largely dependent on drawdown completion strategy, completion string size, tubing head pressures, choke size, pressure losses in the completion as well as multiphase fluid flow. These are all function of offtake performance. Again as this is the case of a reservoir that was already producing, well test permeability was better able to account for pressure gradient and hence bottomhole pressure.

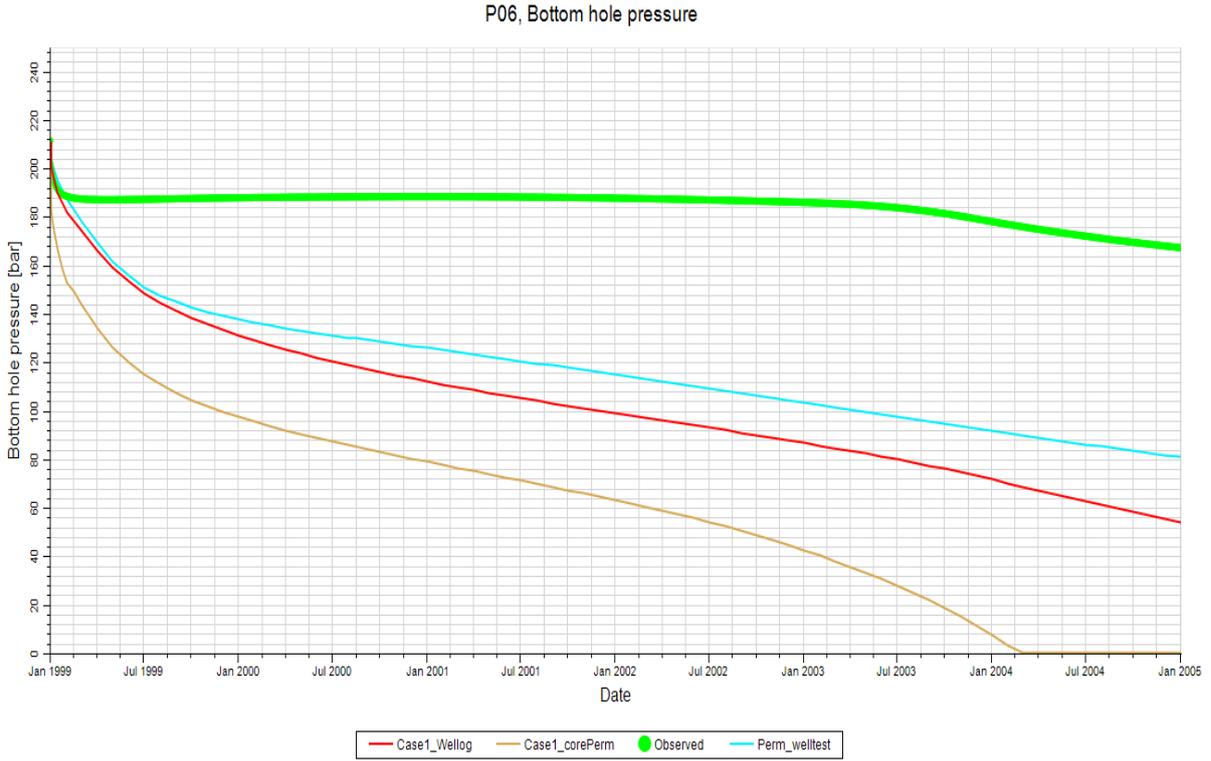


Figure 6. 8 Core, log and test permeability for estimating bottomhole pressure compared with observed bottom hole pressure for well 6.

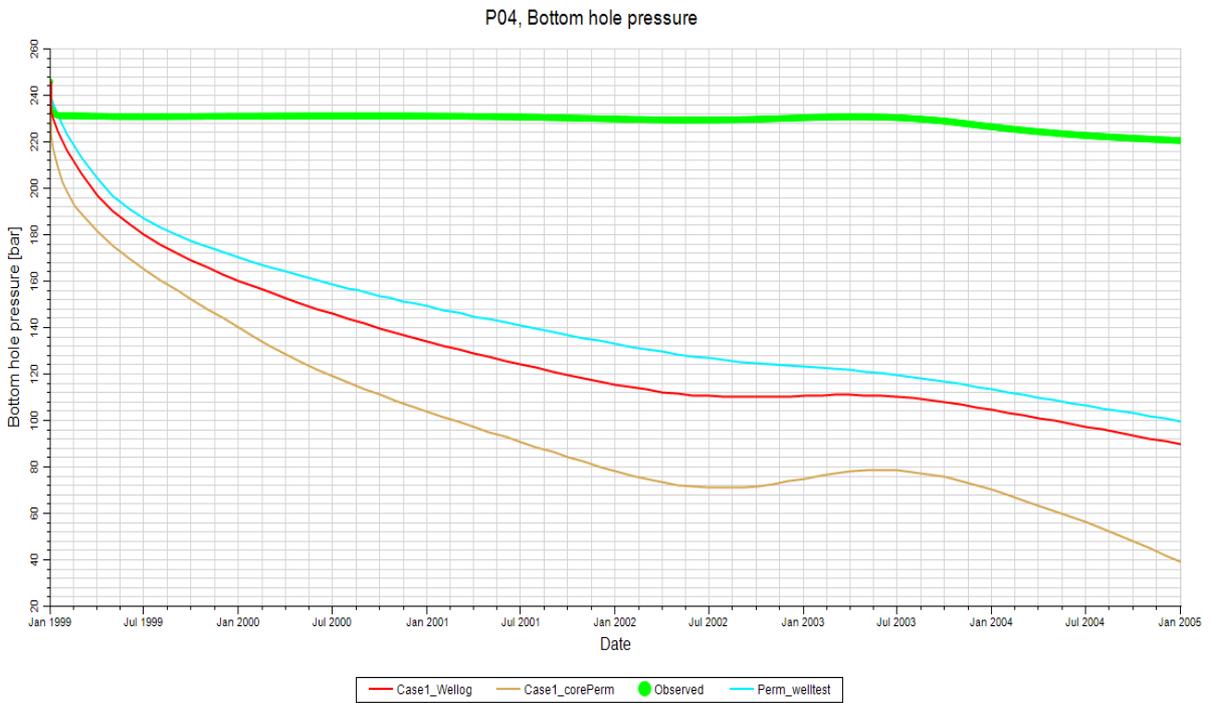


Figure 6. 9 Core, Log and test permeabilities for estimating bottom hole pressure compared with observed bottom hole pressure for well 4.

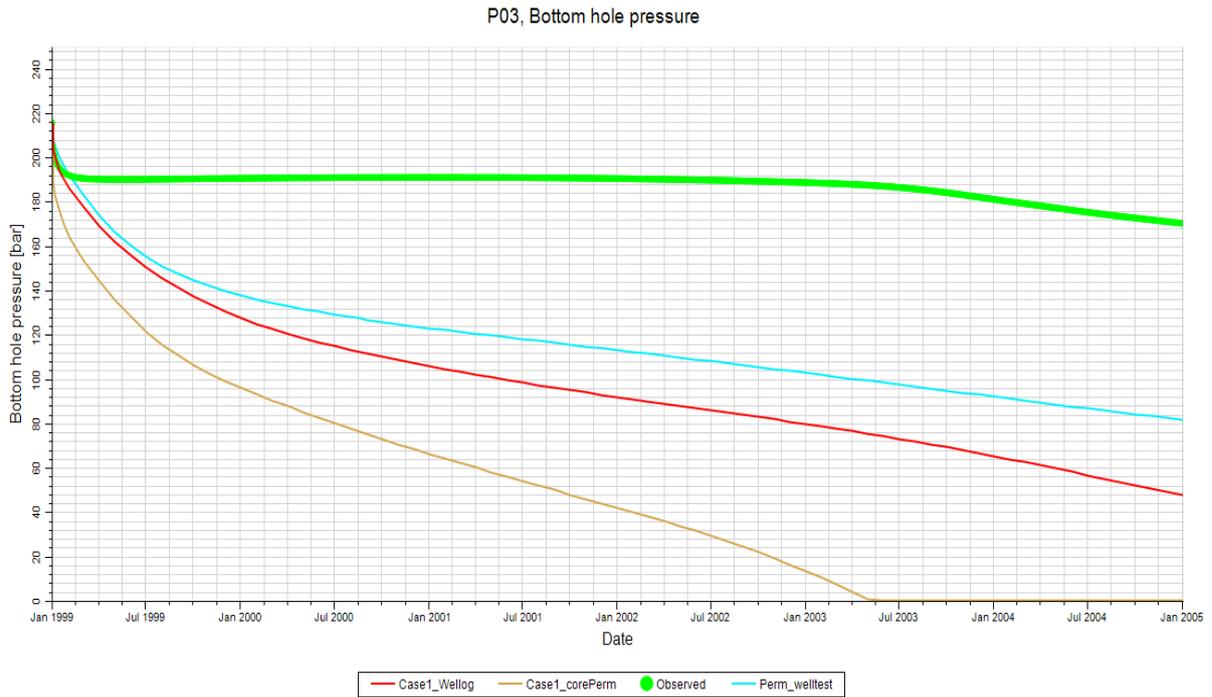


Figure 6. 10 Core, Log and test permeabilities for estimating bottomhole pressure compared with observed bottom hole pressure for well 3.

This work recognizes that the difference should not be ignored as all the data should be considered therefore be reconciled with each other (Figures 6.23 and 6.25) forming a relationship between them as well as with the geologic and reservoir models (Figures 6.11-6.19). Figures 6.11-6.13 shows the reservoir model using log permeability values in the I, j and k directions; Figures 6.14-6.16 shows the same reservoir models but this time using core permeability values in the i, j and k directions while Figures 6.17-6.19 represents the reservoir models using the well test permeability in the i, j and k directions. It is clear that different models are obtained when permeability is sourced using the different datasets. Getting an accurate representation of this variability and quantity is needful for building robust reservoir models. Figure 6.17-6.19 shows that well test lacks the fine detail provided by the core and log permeability. Hence well test permeability should be used but constrained with core and log permeability in areas where wells have been drilled in the reservoir.

A major reason for the observed difference between them is the scale of investigation of the different sources with log and core values at well-bore scale while the well test captures a much bigger portion of the reservoir with average of different flow values more heterogenous than seen at the former scales. Other reasons for the difference are shown in Table 6.1. This difference should not be ignored as all the data were instead reconciled with each other forming a relationship

between them as well with the geologic and reservoir models. core and log helps us to see the heterogeneity within the thickness and laterally as a result of the upscaling but the well test shows no variability along the thick just laterally. The cross section of the different wells as seen in Figure 6.20 showing the core permeability as demonstrates this.

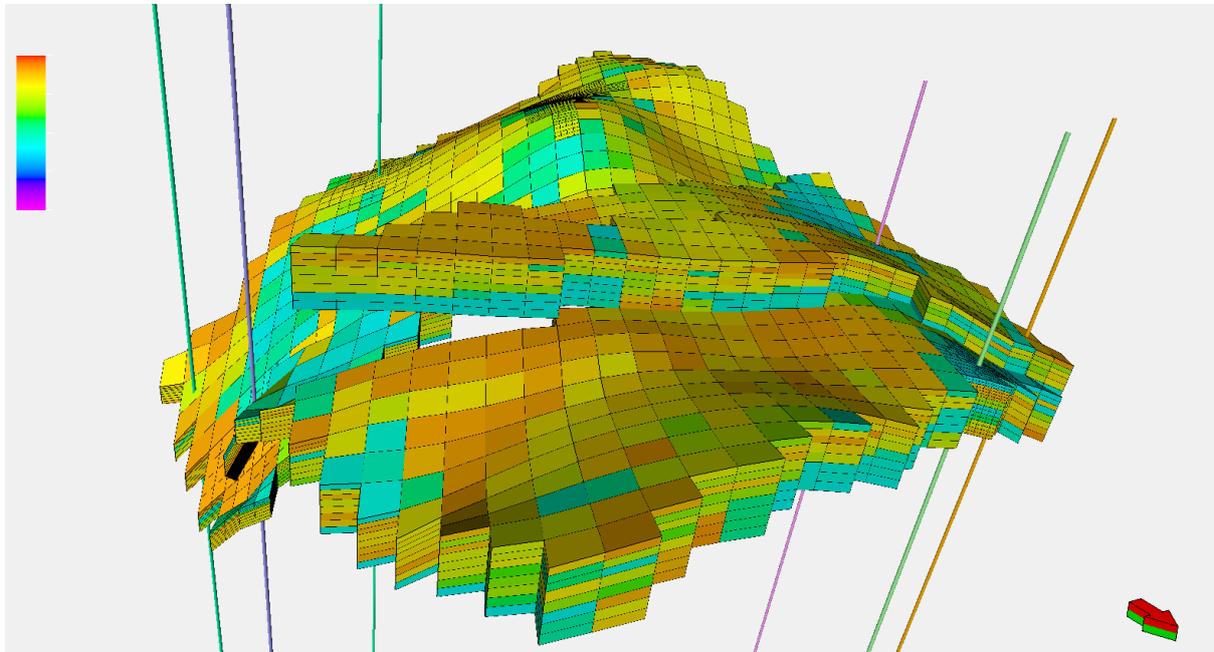


Figure 6. 11 Log Permeability distribution in the i-direction

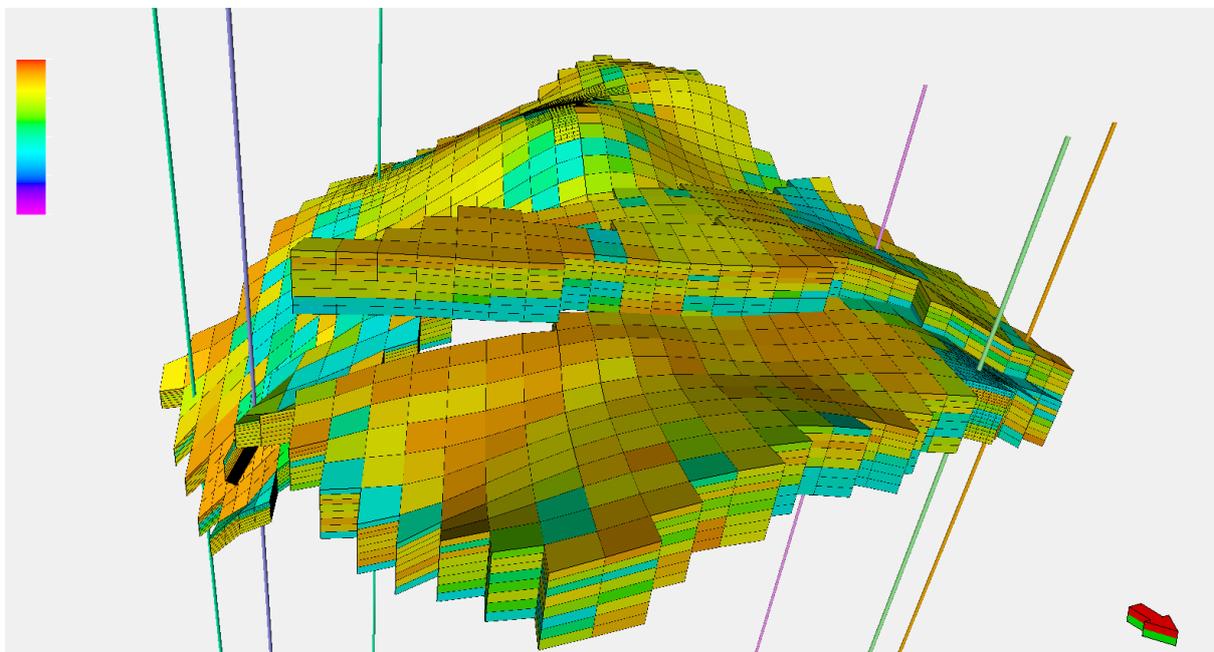


Figure 6. 12 Log permeability distribution in the j-direction

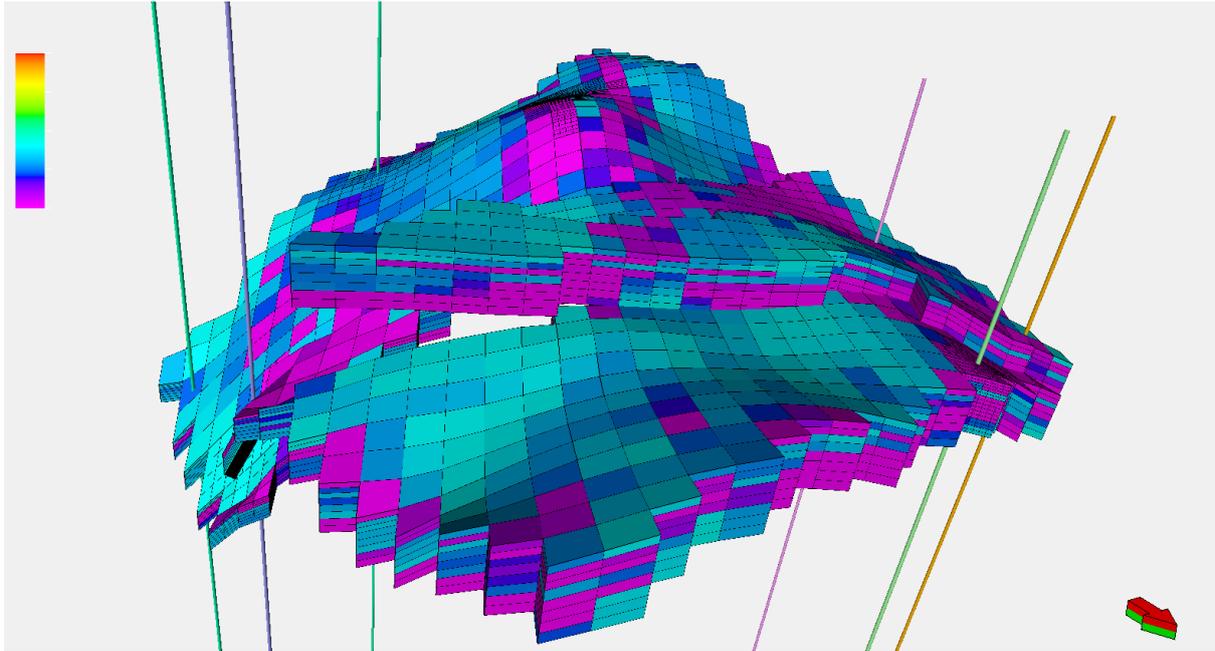


Figure 6. 13 Log permeability distribution in the k-direction

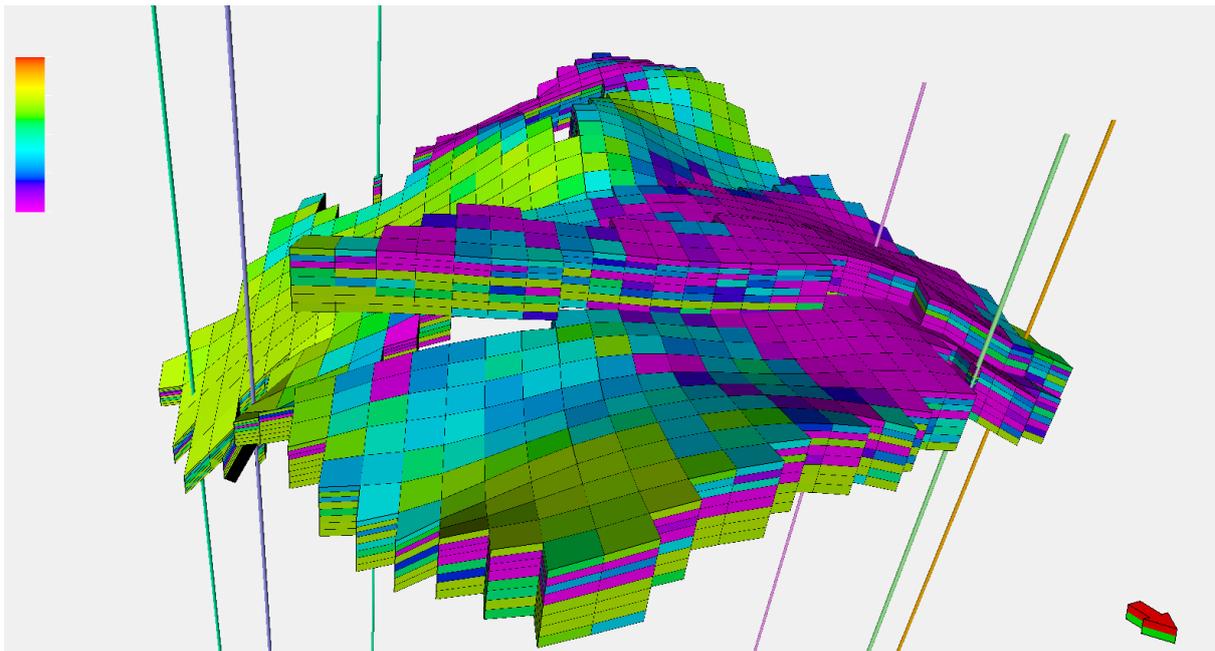


Figure 6. 14 Core permeability distribution in the i-direction

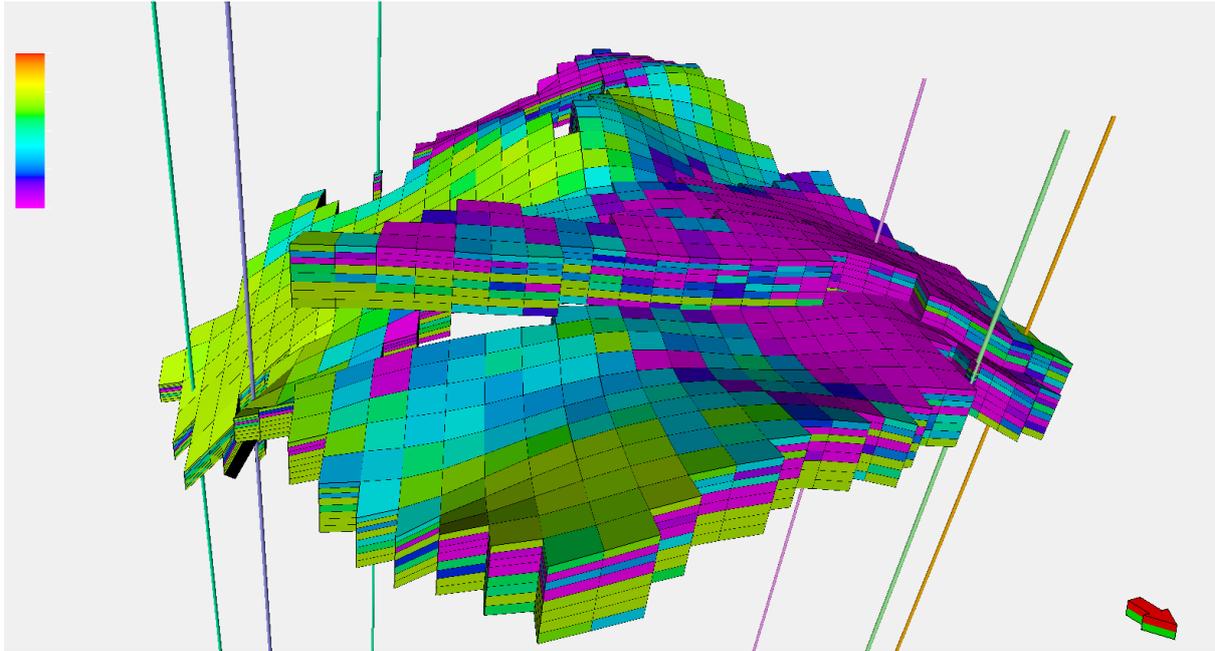


Figure 6. 15 Core permeability distribution in the j-direction

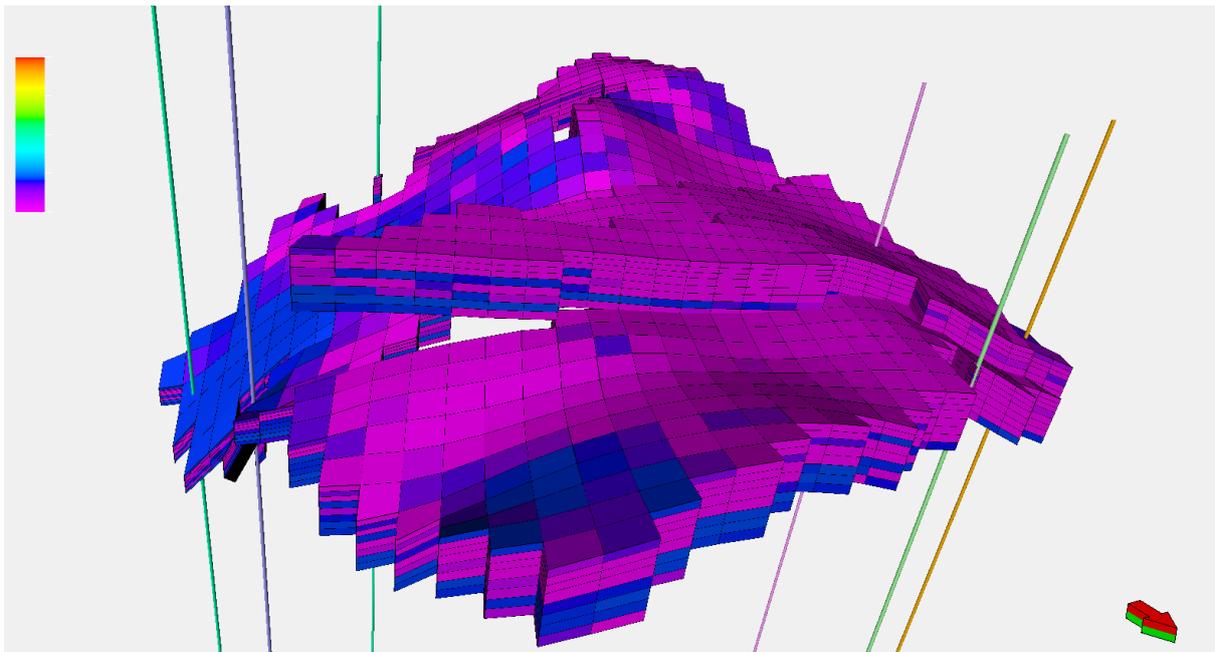


Figure 6. 16 Core permeability distribution in the k-direction

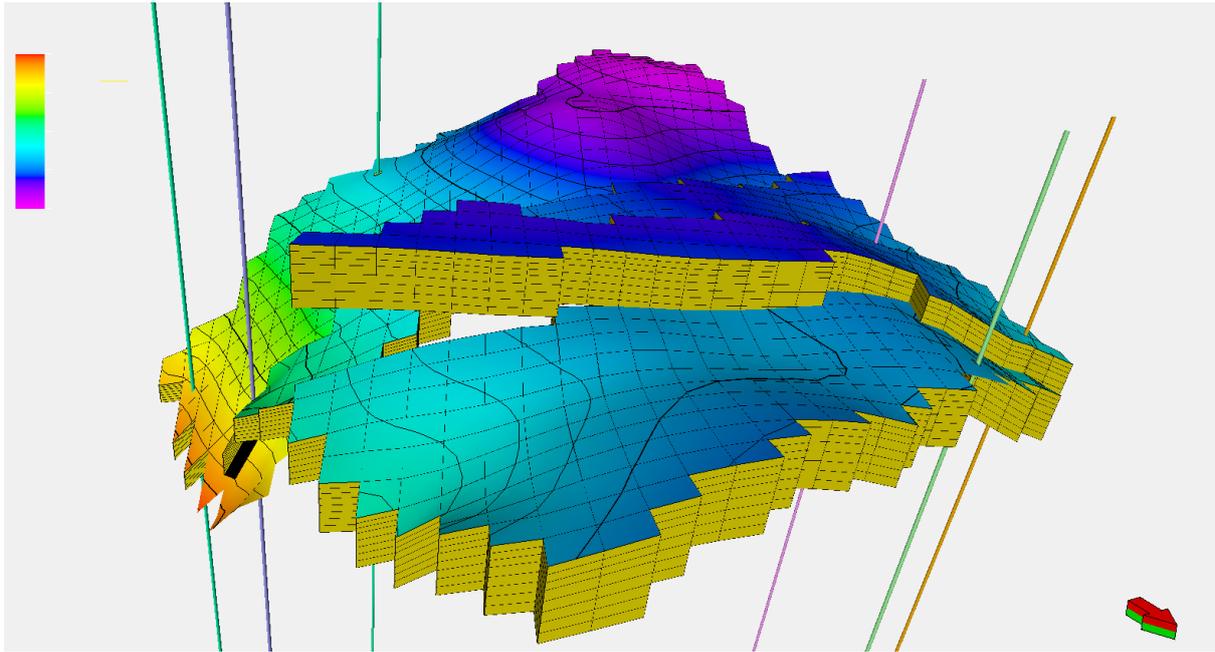


Figure 6. 179 Well test permeability distribution in the i-direction

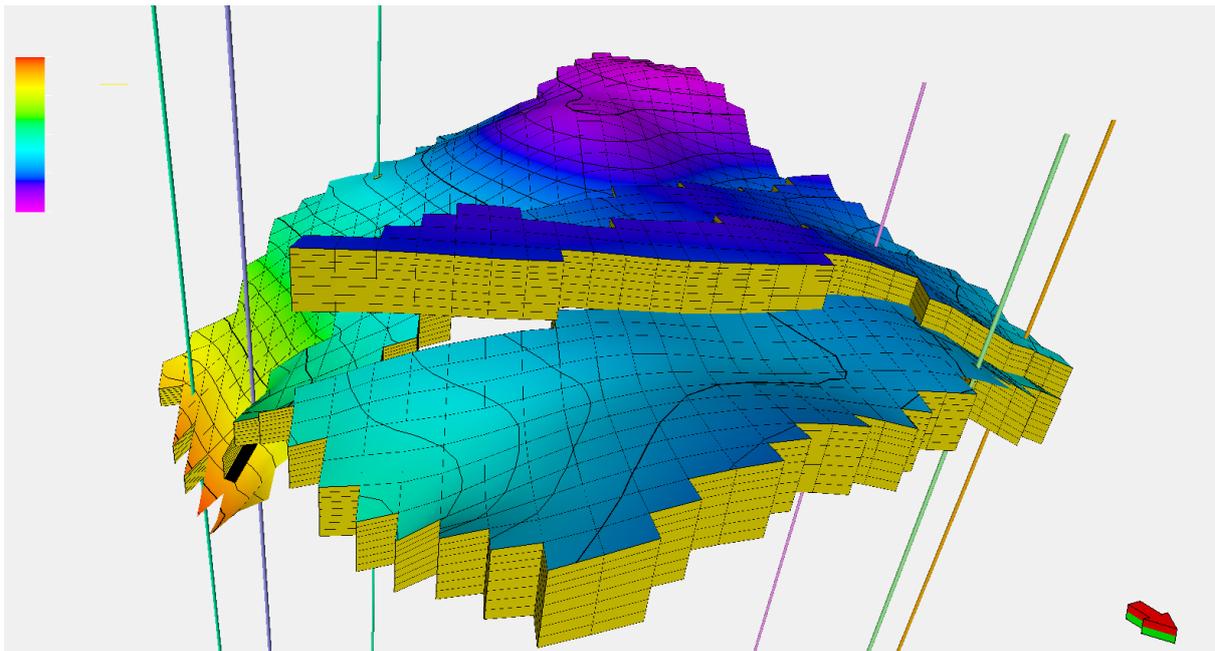


Figure 6. 18 Well test permeability distribution in the j-direction

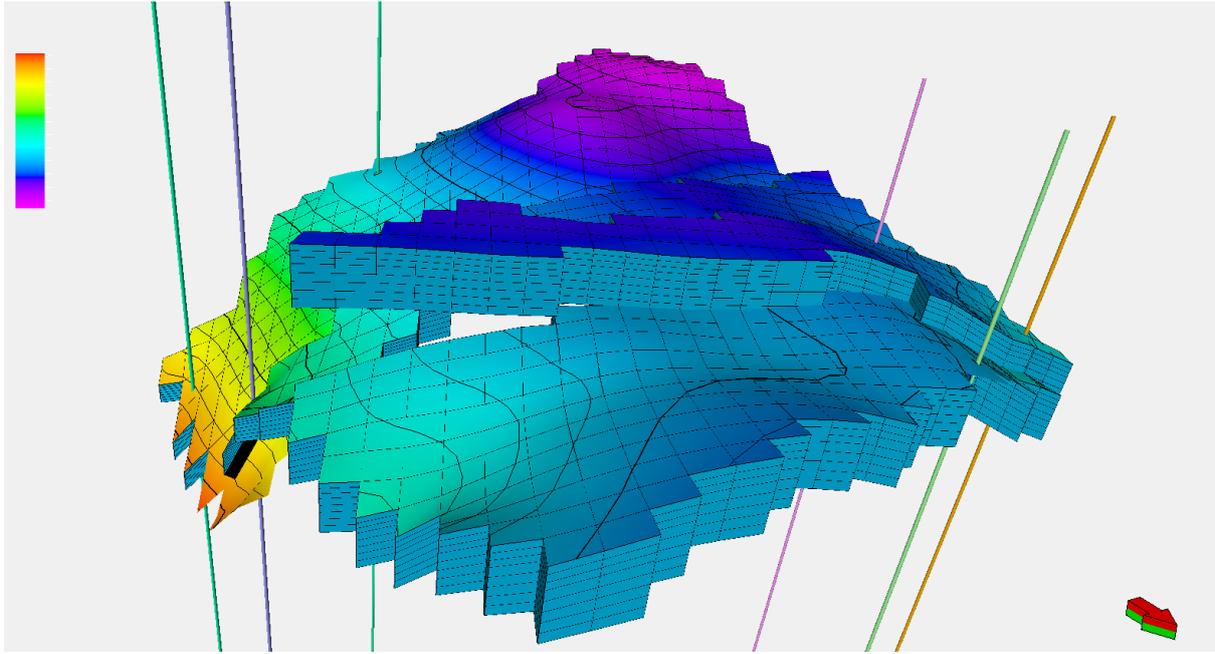


Figure 6. 19 Well test permeability distribution in the k-direction

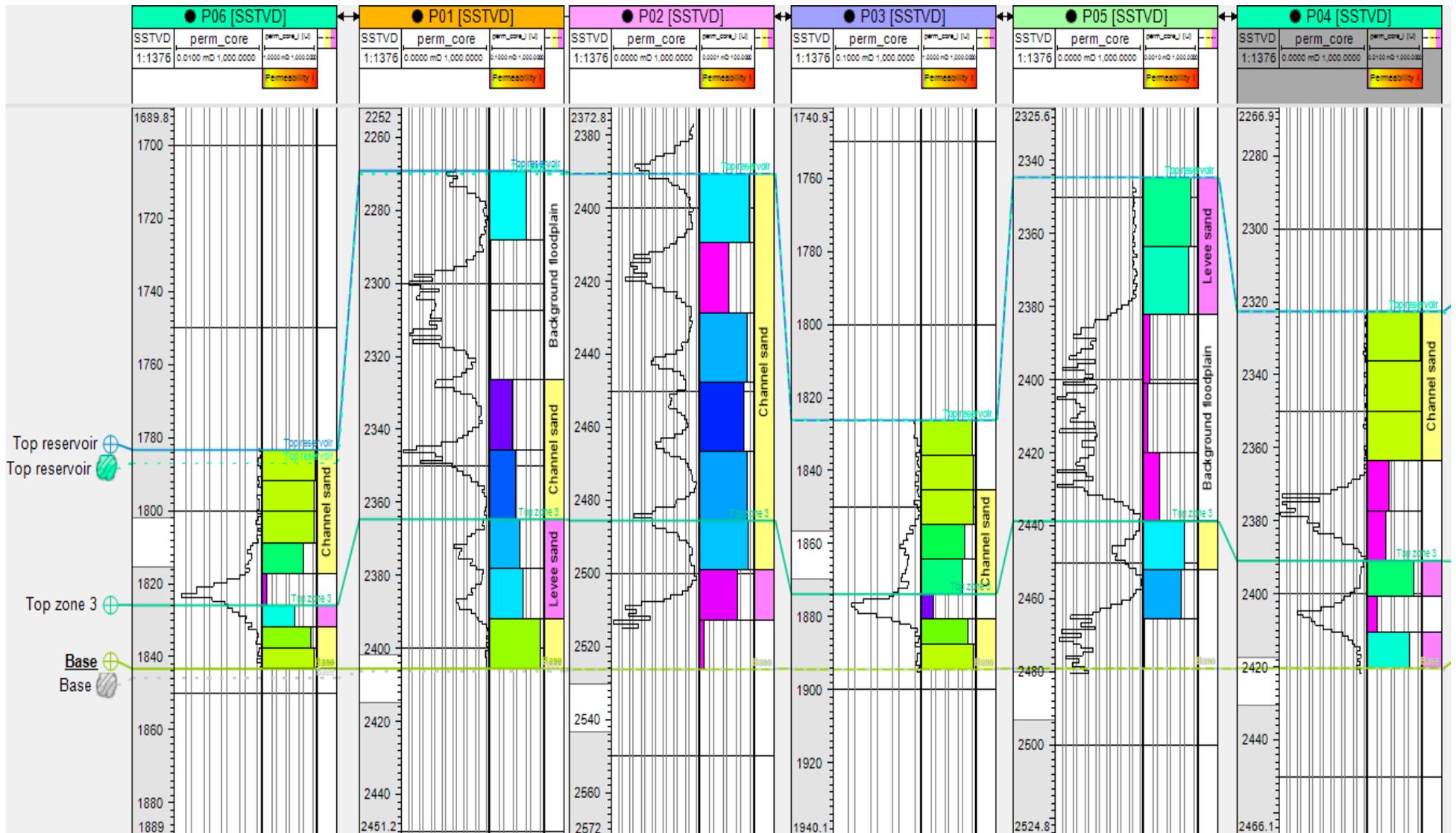


Figure 6. 20 Correlation of the pay zone across the different wells also showing variability along depth.

6.3.2 Log, Core and Test Permeability Relationship Using Machine Learning on North Sea Reservoir

Each source of permeability is vital at several stages in the life of a well. Thus having a relationship between the different sources in order to obtain any source can be useful whether in the well's early life where average horizontal effective permeability to oil or gas (from well test) is the major concern since it controls completion design and productivity of individual wells; or whether later in the life of the well where vertical permeability (from logs and core) is vital due to its influence on water and gas coning as well as productivity of both multilateral and horizontal wells; while both vertical and horizontal permeability distribution influences reservoir performance, amount and viability of secondary and tertiary recovery.

It is noteworthy to point out that well test measures effective permeability while core and log permeabilities are absolute. Log derived permeability depends on a larger scale than core scale permeability. In some fields, water saturation is very low ($< 10\%$) and therefore it is reasonable to expect that the effective permeability from well test will be close to the absolute permeability thus the relative permeability effects will be minor especially in the pre-production drill stem test. There is somewhat of a trend for the well test permeability to be less than the core and well log permeability. It is important to note that despite the fact that well test measures permeability thickness, it is more accurate to make the comparison with just the permeability instead of multiplying the log and core with thickness thereby increasing the uncertainty. A summary of the major causes of the difference especially of core and well test are highlighted in Table 6.1 above. Figure 6.21 is a flow chart of the step by step algorithms adopted for the machine learning process for reducing uncertainty across Log, Core and Test (LCTun SVM Model). In Chapter Four, we trained log data on core porosity, core fluid saturation, core permeability data so that the models generated are used to predict core scale permeability. In the same vein, the LCTun SVM model involved training core and log permeabilities on well test permeability data so core and logs are used to predict well test scale permeability.

Figure 6.22 and 6.23 is a crossplot of the support vector regression model against the actual well test permeability or the test cases considered with a root mean squared error value of 5.22 and 8.17 respectively. These indicate that making good use of well logs and core is sufficient to evaluate well test permeability and reservoir performance with accuracy which is particularly important in the deep water offshore environment where the cost of DST's in every well will not be justified. As pointed out by Potocki (2001), relative permeability effects associated with water saturation accounts for some of the difference between core, log and well test. He showed that

this difference becomes increasingly bigger as water saturation in the reservoir increase thus in chapter Four, water saturation and relative permeability techniques were developed.

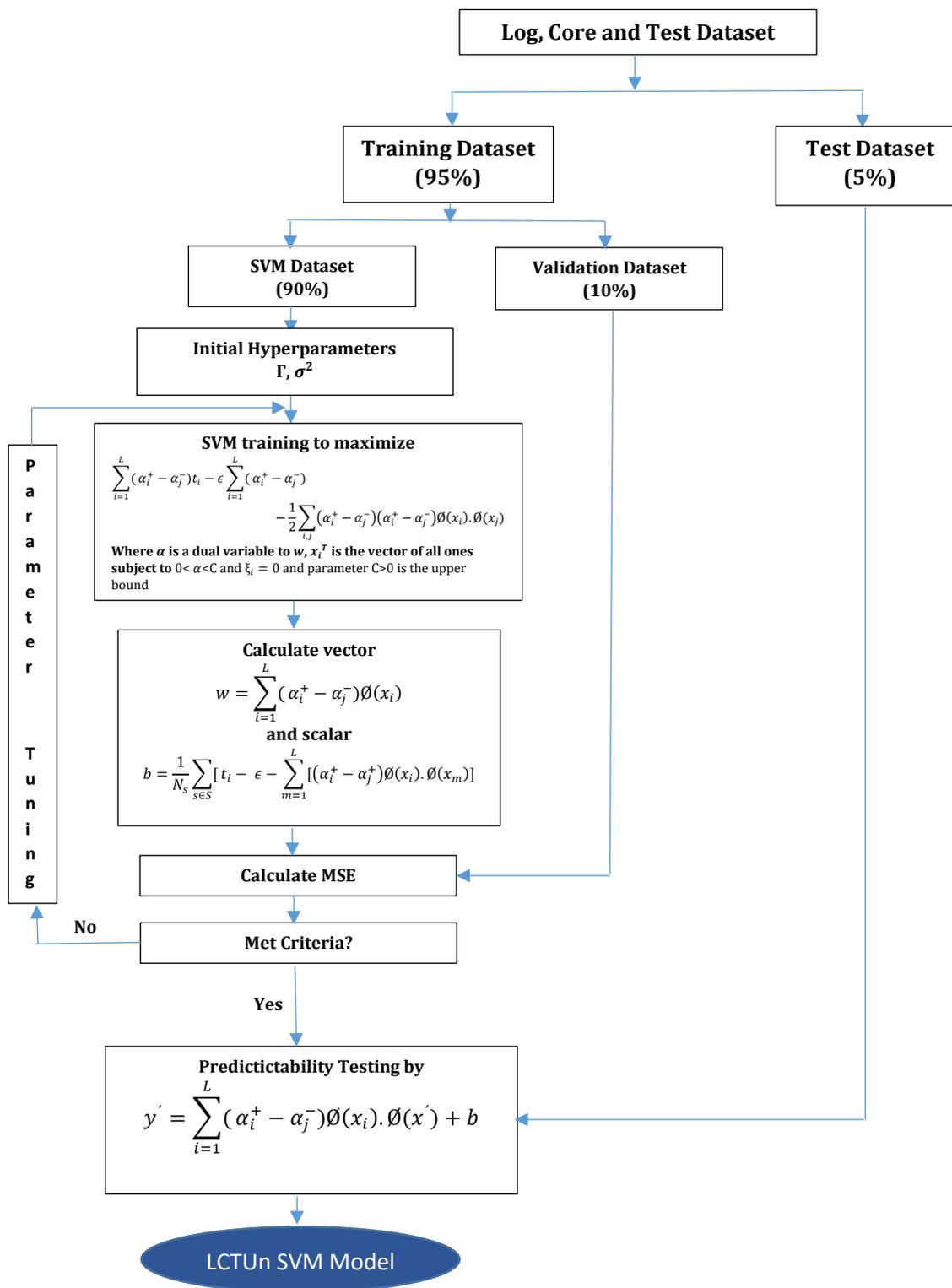


Figure 6. 21 Flow diagram showing Uncertainty Reduction Support Vector Regression Technique

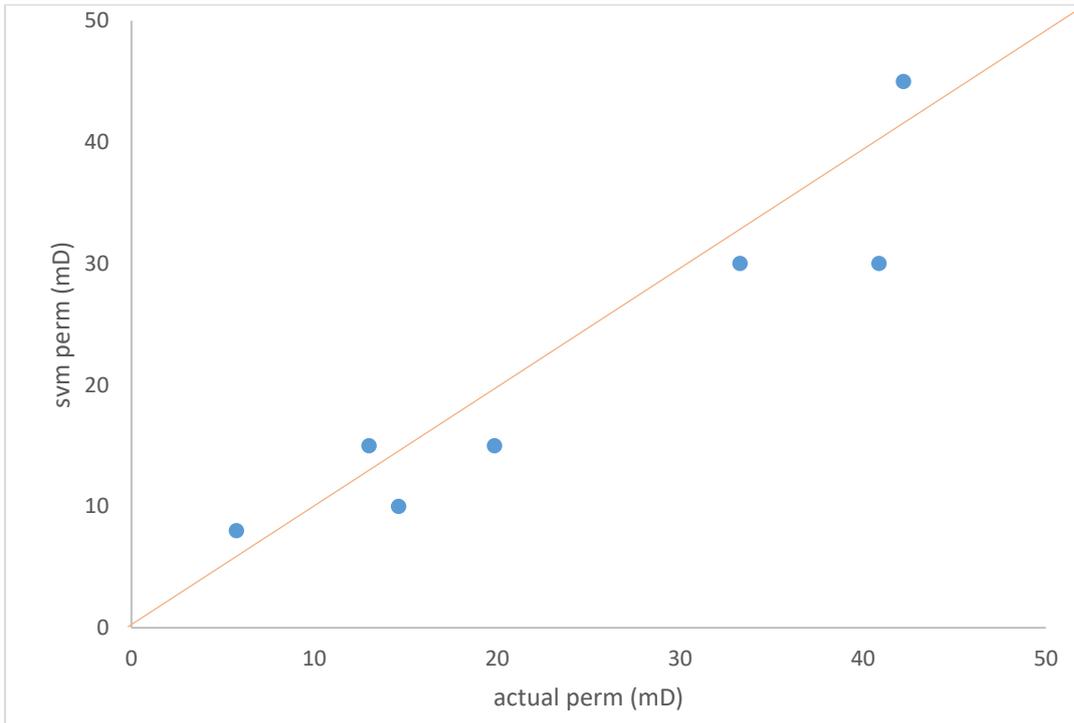


Figure 6. 22 Actual well test permeability versus SVM model predicted well test permeability (from core and logs permeability) on a randomly selected test set.

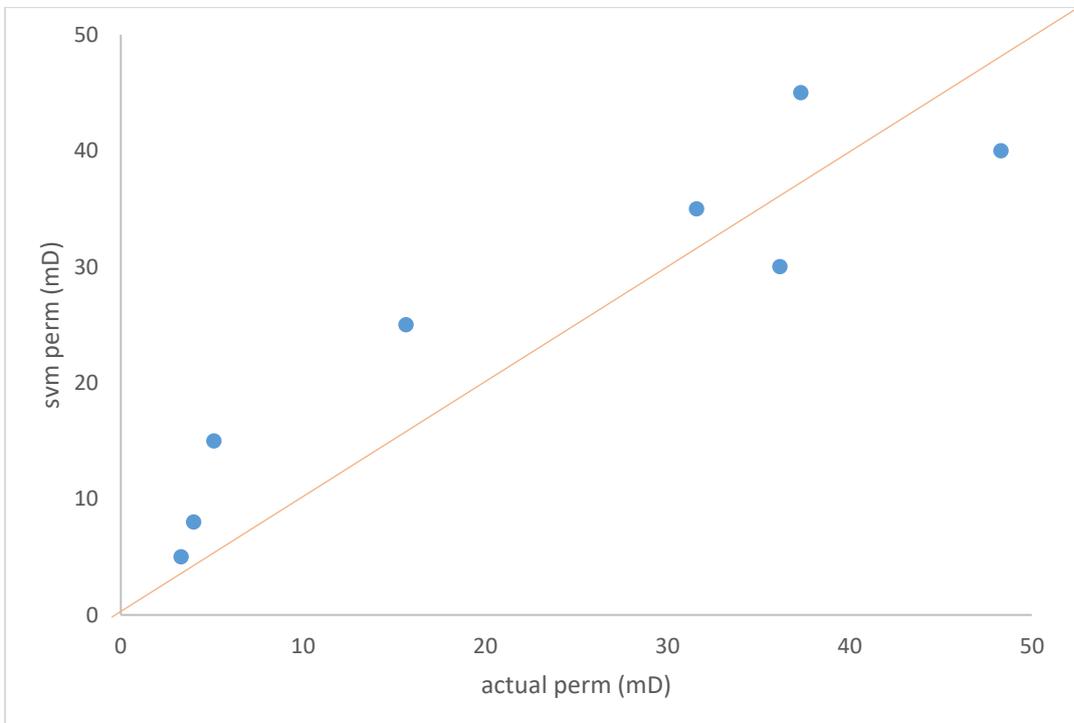


Figure 6. 23 Actual well test permeability versus SVM model predicted well test permeability (from core and logs permeability) on a second randomly selected test set.

Figure 6.24 is a dashboard of the Well Test SVM model (referred to as LCTun SVM model) built using log and core scale permeability data, their squared root and squares as input to the model as already discussed in Section 6.3.2. Parameters are appropriately entered in the slider input panel and this automatically computes the corresponding well test permeability using the support vector regression model already coded into the package. Considering additional features such as the variants of core and log permeability improved the prediction ability of the model (Table 6.2). The data tab gives the result of well test prediction made from $klog$, $kcore$, $(klog)^{1/2}$, $(kcore)^{1/2}$, $(klog)^2$, $(kcore)^2$. The package requires a PC with a windows operating system and R software package installed.

Well Test SVM Model

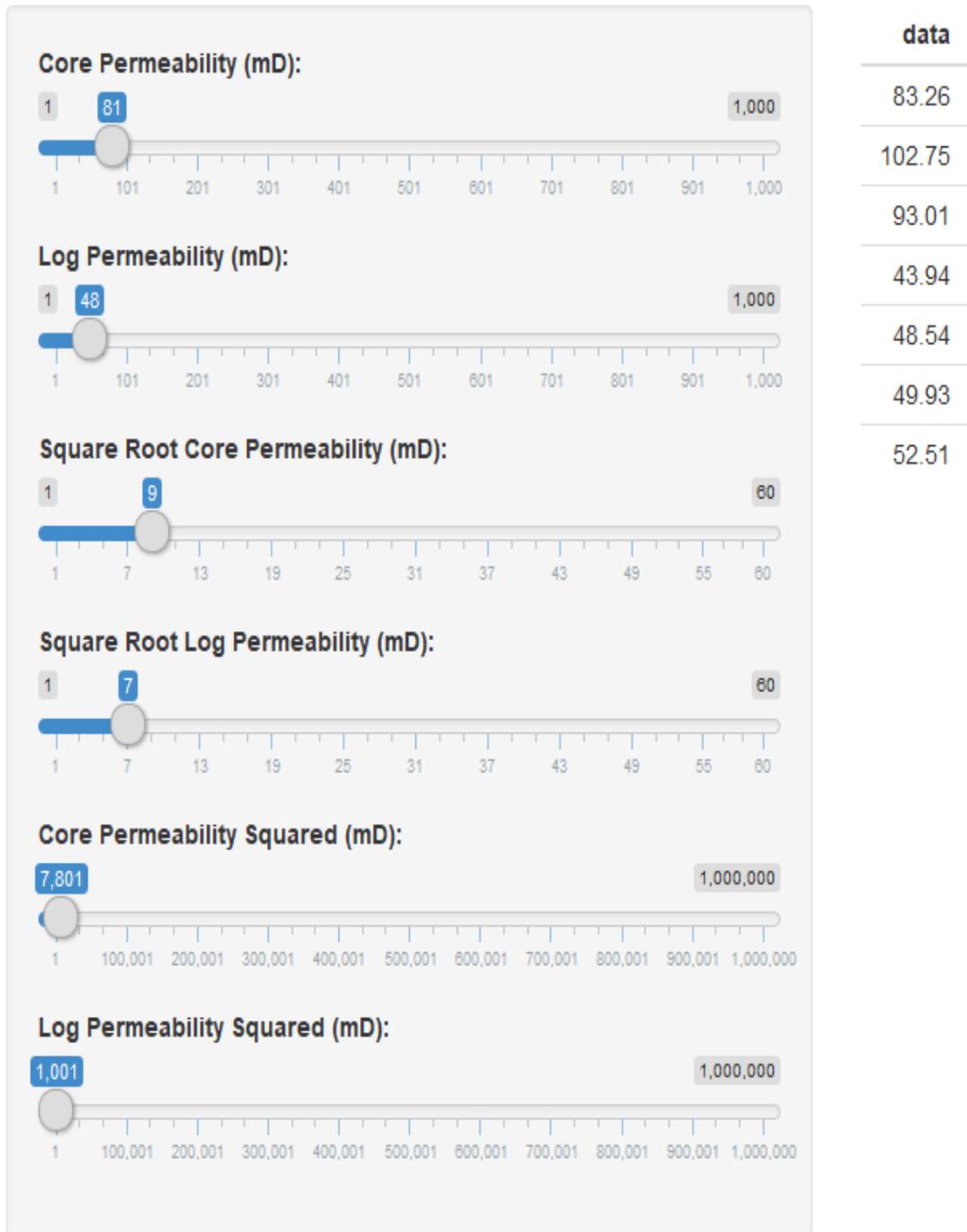


Figure 6. 24 Screen shot of well test permeability prediction dashboard result from core and log permeability.

6.3.3 Log, Core and Well Test Permeability Relationship Using Chart Method on North Sea Reservoir

A second approach using Nomographic technique was also used. Figure 6.25 is the result of a numerical simulation run iteratively on a multiple non-linear regression model obtained from the dataset on Matlab. No doubt models like this requires much more dataset to give more accuracy but the idea is to give a reasonable approximation. Similar to what the Schlumberger (1975) chart for porosity determination from its various sources like density, neutron and sonic does. The legend to the right gives the corresponding well test permeability from the known log and core permeabilities. This relationship from these different scales shows that generally well test permeability is lower than core and log permeability and since all scales matter, an idea of the more difficult to obtain is known. This shows a lot of promises in forecasting well test permeability with more data being made available from different fields around the world and not just from a few from the North Sea, which is the case here.

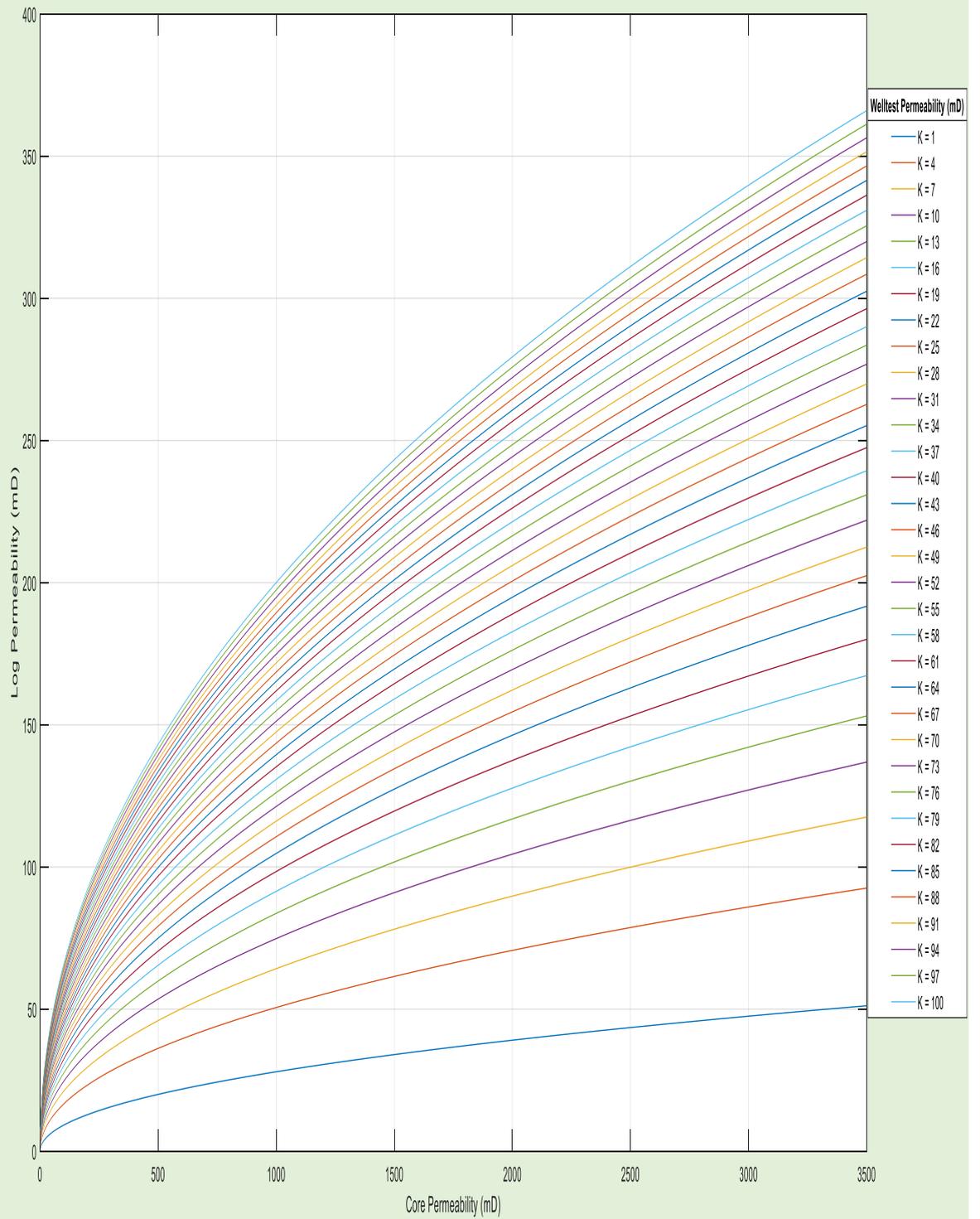


Figure 6. 25 Nomographic technique showing the relationship between core, log and well test permeability.

6.3.4 Machine Learning Model Validation

A separate North Sea field data not used for the training was used to validate the developed support vector machine model. The field is located in the Central Graben in the Central North Sea with the main reservoir been the Jurassic shallow marine shoreface Fulmar sandstone. Its upper sands appear to be of good quality (140ft thick, NTG =86%, porosity = 19%, $S_w = 17\%$) compared to the lower sands (125ft thick, NTG = 70%, porosity = 16%, and $S_w = 41\%$) as shown in Figure 6.26. The reservoir sections are extensively cored and DSTs also acquired. Some of the wells in this field are located in the horst caused by salt movement.

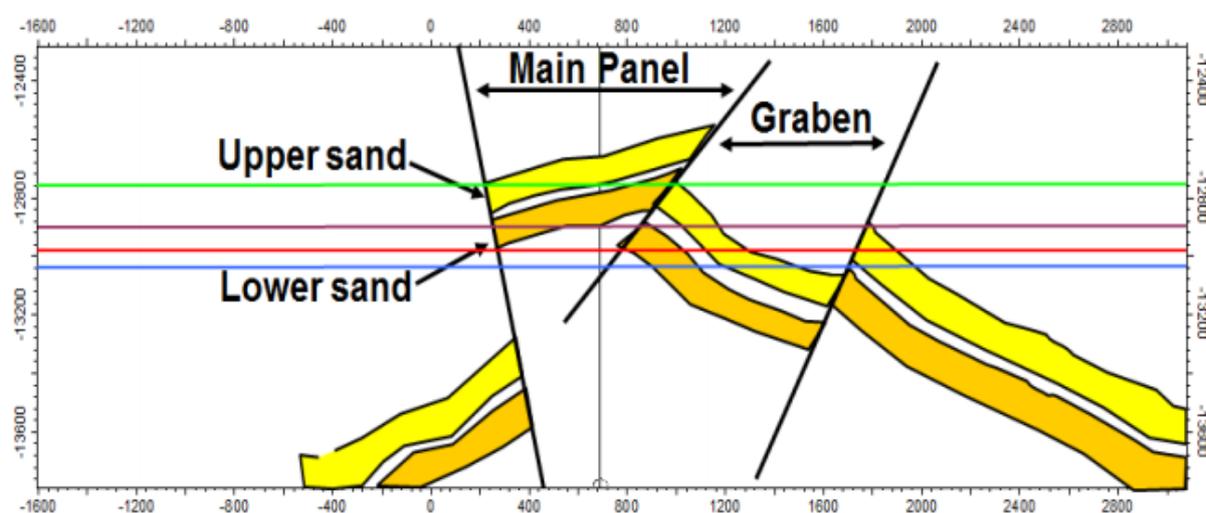


Figure 6. 26 Schematic showing the Upper and Lower sand penetrated by some of the wells used for validation.

Figure 6.27 shows the comparison between the actual well test permeability and the prediction made by the support vector machine model giving a root mean squared error value of 7.01. With more data and adequate parameter tuning, the quality of the model can be drastically improved to get real time predictions of the parameter from more readily available sources. No doubt more still has to be done in this area.

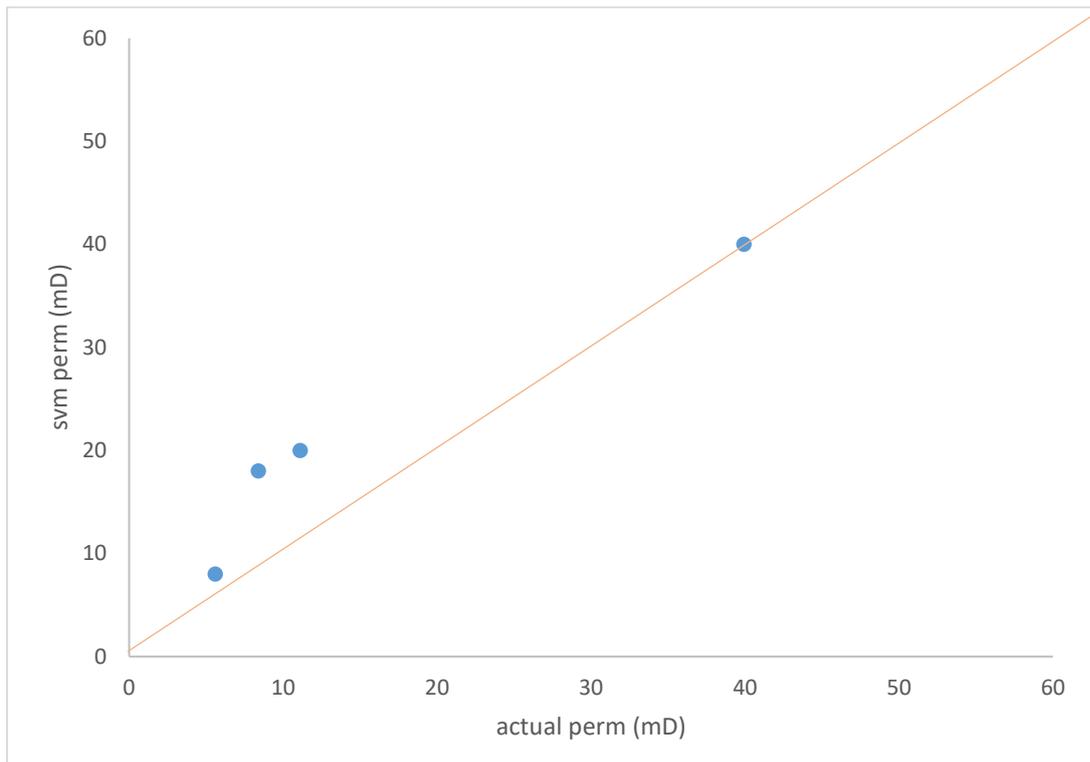


Figure 6. 27 Actual well test permeability versus model predicted well test permeability from core and logs permeability.

6.4 Summary and Conclusion

The multiscale database for permeability including core, log and well test for different sandstones reservoirs from the NorthSea have been built. These will provide a multidisciplinary data platform compilation which can be grown by adding more and more data in order to make predictions from big data more accurate.

It has been shown from using the techniques presented in this work that making good use of well logs and core is sufficient to evaluate well and reservoir performance with accuracy which is particularly important in the deep water offshore environment where the cost of DST's in every well in an entire field as well as in the different pay zones which the well penetrates, will not be justified. A relationship between the different dataset was learnt using appropriate machine learning algorithms.

Just like the Schlumberger (1975) chart for porosity from the different sources, a nomograph has been developed enabling a quick guide for permeability determination. More data still have to be gathered and then used to enhance the multiple nonlinear regression from which a numerical simulation was run to develop this model. This will no doubt improve the accuracy of the nomographic model

Using the different permeability source to forecast flow rate, it was hard to see which source is more accurate but for pressure forecast well test permeability appeared to be the most accurate of the different permeability sources considered.

Chapter Seven

Conclusions, Review and Recommendation for Further Work

The conclusions drawn from the entire work in line with the aim, objectives and the research questions raised in Chapter One are presented here.

7.1 Conclusion

The following conclusions have been drawn from the work done in this thesis and have been highlighted below:

- Porosity, saturation and permeability were demonstrated from the literature and using Monte Carlo simulation runs on reserve estimates (STOIIP), Net Present Value (NPV) and Pressure Transient Analysis (PTA). These parameters and their variants appeared to be the most sensitive and also have strong influence on other parameters in different reservoir assets like sandstones, carbonates, fractured reservoirs.
- An extensive review of the existing models for porosity, fluid saturation and permeability indicates that there are lots of uncertainty inherent in the accurate prediction of these properties. They are not robust enough for all reservoirs hence they whip up confusion as to which should be used judging from the huge number of empirical models available in the literature.
- Default Raymer and Wyllie models are not suitable for porosity determination in unconsolidated formations as their accuracy seem to be related to consolidation. These inaccuracies in prediction appears to be more in terms of over-prediction than under-prediction in most of the cases examined. Thus the assumption that Wyllie gives effective porosity seems untrue.
- Application of the different sources for reserve estimation in fields of known reserve values showed that Schlumberger, Raymer and Wyllie Time Average could lead to errors of up to 17% in unconsolidated sandstones, about 5% in semi-consolidated sandstones and 3% in consolidated reservoir sandstones. The SVM model gave errors of less than 2% in all sandstones. This consistency across all sandstone types is essential to increase confidence and reduce uncertainty.

- There are over thirty models available for water saturation prediction. The stacked Ensemble model gave the best prediction of oil and water saturation from well logs than its base learners which are Gradient Boosting, Neural Networks and Random Forest. Also, the porosity-permeability relationship seems to be more sensitive to about 10-20% range of irreducible water.
- The existing relative permeability models gave different results for the same reservoir with some performing better at lower saturations but poorly at higher saturations and vice versa. The deep learning model gave reasonable predictions at all saturation values and therefore suitable for real time usage. This performance of the deep learning model is due to its scalability (performance always improving with more data), generalization and regularization ability.
- A 50% decrease in pay thickness appears to cause more destabilization on zero and constant flowrate pressure drawdown and its derivative than a 100% increase in pay thickness. Overall a high decrease in porosity, compressibility and pay thickness was more detrimental than a high increase. With this been said, it is very important to accurately define Effective Flow Interval.
- Well test permeability appeared to be the most accurate of the different sources. From using core, log and well test permeability to simulate oil flowrate, it was hard to see which source gave a better prediction for the cases considered as each did well at different wells but from simulating bottomhole flow pressure, well test permeability consistently performed better than core and log permeability probably due to its scale of investigation.
- Answering the different questions causing the bottlenecks of comparing core. Log and well test permeability such as scale, physics, environment etc of these sources makes it difficult to form a physical relationship between them and thus reduce the uncertainty across board. A machine learning approach has been adopted to solve this issue by learning the pattern of these different sources of data no matter how arbitrarily irregular they may be and therefore predict well test permeability from the more available core and log permeability. The work shows that using machine learning is a win-win situation thus it is a question of proper tuning.
- Just like the Schlumberger (1975) chart for porosity from the different sources, a nomograph has been developed enabling a quick guide for permeability determination.

More data still have to be gathered and then used to enhance the multiple nonlinear regression from which a numerical simulation was run to develop this model. This will no doubt improve the accuracy of the nomographic model.

- Machine learning removes the need for guessing certain properties such as saturation, porosity and permeability discussed in this work. This is critical as the different variables are related to each other hence error can easily propagate.

7.2 Review of Work

This section provides an explanation of the areas of the work that could be improved on as well as what should be taken forward for further research. No doubt a lot has been done in this work but no matter how good a work is, there is always room for improvement.

- There are myriads of machine learning models for regression analysis. Those not considered in this work such as ridge, gaussian process, lasso, stepwise regression etc might have done better than the ones used here. Again using these regression models in tandem with optimization techniques like genetic algorithms, ant colony can improve their performance if done effectively.
- More confidence on any model is proportional to the amount and variability of data used in building it. NorthSea sandstone reservoir data was used for this work. For the model performance to improve especially globally, data from different fields in Africa, North and South America, Asia and other parts of Europe has to be used for its development. The generalization ability would definitely have improved if more data from diverse fields were used. This will also make for a more comprehensive model validation using other sandstone reservoirs from outside the NorthSea not used for model development ofcourse in order to test model repeatability and reproducibility.
- Looking at the different sources across breadth, the most accurate source with respect to permeability was difficult to determine from oil flow rate forecast but well test permeability appears to be the most accurate source in terms of pressure prognosis. More still need to be done to really ascertain this.

- Various averages were used especially arithmetic averages due to study of the section of wells, tectonic history of the area to see if much tilting and folding has occurred and also bearing in mind the principle of original horizontality. Lack of well data cutting reservoirs, in some cases made it not clear which averaging technique to use. Again fields consisting of layers at some wells and more random permeability (where the geometric averaging is more appropriate) at other well locations causes problems.
- Machine learning and nomographic techniques were used to establish a relationship between K_{WT} , K_{CORE} and K_{LOG} necessary to determine the most accurate source from the other two when not available due to time and expense. Although very careful analysis of the averaging technique was done, a production logging tool log (PLT) would help to ascertain the direction of flow more accurately hence narrowing down on what averaging technique to use with more certainty.

7.3 Recommendations for further work

- More work should be done to understand how water/oil and gas/oil relative permeability vary significantly with different facies. These product of rock that have accumulated in a depositional environment could be studied in relation to their fluids. The database of this can then be built to improve predictions for the property. This understanding will also find application in Carbon Capture and Storage (CCS) projects necessary to combat climate change.
- Limited data were used in this work for training, validating and testing the developed models due to difficulty in sourcing for the right amount of data. Companies need to include such filtered database for different reservoir properties as part of joint venture agreement such that they could refer to it at all stages of the operation of the reservoir. This will help build more confidence and develop a more proactive approach/methodology towards reservoir characterization thus reducing Non-Productive Time (NPT) and cutting down uncertainty drastically.
- All the models, especially the analytical models should be further validated with data from different petroleum provinces like the Tertiary Niger Delta sandstone reservoirs which was the original plan for this work. The Agbada Formation in this basin contains

intercalations of sand and shale with the sandstone serving as the reservoirs and the shale in addition to the roll over anticlines serving as traps.

- Reservoir Saturation Tool (RST) well test data from reservoirs having logs and core to build a more robust, versatile and adaptable saturation model across the different sources would greatly improve the accuracy and radius of investigation of predicting fluid saturation away from the invaded zone.
- The work can be extended to include other reservoir types especially carbonate reservoirs as well as shale reservoirs. Would suggest one of the variables then to be facies type so the model can have an idea of the lithology play a part in the flow. This if properly done, will make the models more robust.

References

- Agarwal RG. A new method to account for producing time effects when drawdown type curves are used to analyze pressure buildup and other test data. A new method to account for producing time effects when drawdown type curves are used to analyze pressure buildup and other test data. *SPE Annual Technical Conference and Exhibition*: Society of Petroleum Engineers; 1980.
- Ahmadi MA, Chen Z. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs. *Petroleum*. 2019 Sep 1;5(3):271-84.
- Ahmed U, Crary S, Coates G. Permeability estimation: the various sources and their interrelationships. *Journal of Petroleum Technology*. 1991; 43(05):578-587.
- Al-Anazi AF, Gates ID. Support vector regression to predict porosity and permeability: effect of sample size. *Computers & Geosciences*. 2012 Feb 1;39:64-76.
- Al Salem SS. Environmental considerations for wastewater reuse in agriculture. *Water science and technology*. 1996; 33(10-11):345-353.
- Alavy K, Gaskell A, Leach S, Szymanski S. *On the edge of your seat: Demand for football on television and the uncertainty of outcome hypothesis*. 2006.
- Al-Bulushi N, King PR, Blunt MJ, Kraaijveld M. Development of artificial neural network models for predicting water saturation and fluid distribution. *Journal of Petroleum Science and Engineering*. 2009; 68(3-4):197-208.
- Al-Fattah SM, Al-Naim HA. Artificial-intelligence technology predicts relative permeability of giant carbonate reservoirs. *SPE Reservoir Evaluation & Engineering*. 2009; 12(01):96-103.
- Amaefule JO, Altunbay M, Tiab D, Kersey DG and Keelan DK. Enhanced reservoir description: using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. Enhanced reservoir description: using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. *SPE annual technical conference and exhibition*: Society of Petroleum Engineers; 1993.
- Aminzadeh F. Applications of AI and soft computing for challenging problems in the oil industry. *Journal of Petroleum Science and Engineering*. 2005; 47(1-2):5-14.
- Amyx JW, Bass DM, Whiting RL. *Petroleum Reservoir Engineering [by] James W. Amyx, Daniel M. Bass, Jr.[and] Robert L. Whiting*. : McGraw-Hill; 1960.
- Ani M, Oluyemi G, Petrovski A and Rezaei-Gomari S. Reservoir uncertainty analysis: The trends from probability to algorithms and machine learning. Reservoir uncertainty analysis: The trends from probability to algorithms and machine learning. *SPE Intelligent Energy International Conference and Exhibition*: Society of Petroleum Engineers; 2016.
- Archie GE. Introduction to petrophysics of reservoir rocks. *AAPG Bulletin*. 1950; 34(5):943-961.
- Arigbe, OD., Oyeneyin, M.B., Arana, I. and Ghazi, M.D., 2018. Real-time relative permeability prediction using deep learning. *Journal of Petroleum Exploration and Production Technology*, pp.1-14.
- Baker L. Three-phase relative permeability correlations. Three-phase relative permeability correlations. *SPE Enhanced Oil Recovery Symposium*: Society of Petroleum Engineers; 1988.

Barlai Z. Determination of Permeability and Specific Surface Area of the Pore Channels From Well Logs in Fine Grained Sandstone. Determination of Permeability and Specific Surface Area of the Pore Channels From Well Logs in Fine Grained Sandstone. *SPWLA 17th Annual Logging Symposium: Society of Petrophysicists and Well-Log Analysts*; 1976.

Baroni G, Tarantola S. A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. *Environmental Modelling & Software*. 2014; 51:26-34.

Bassiouni Z, Velić J. Prospecting for Bypassed Oil and Gas. *Geologia Croatica*. 1996; 49(2):197-202.

Bateman RM, Clavier CM. *Neutron well logging technique for gas detection*. 1977; .

Bear J, Braester C. On the flow of two immiscible fluids in fractured porous media. *Developments in Soil Science*. : Elsevier; 1972. p. 177-202.

Begg S, Bratvold R and Campbell J. Shrinks or Quants: Who will improve decision-making. Shrinks or Quants: Who will improve decision-making. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 2003.

Bethel F, Calhoun JC. Capillary desaturation in unconsolidated beads. *Journal of Petroleum Technology*. 1953; 5(08):197-202.

Bhatt A, Helle HB. Committee neural networks for porosity and permeability prediction from well logs. *Geophysical Prospecting*. 2002; 50(6):645-660.

Bjørlykke K, Jahren J. Sandstones and sandstone reservoirs. *Petroleum Geoscience*. : Springer; 2010. p. 113-140.

Bouchard A, Fox M. Incorporating core analysis data uncertainty in asset value assessment. *SCA-9905*. 1999

Bourdet D, Ayoub J, Pirard Y. Use of pressure derivative in well test interpretation. *SPE Formation Evaluation*. 1989; 4(02):293-302.

Bourdet D, Whittle T, Douglas A, Pirard Y. A new set of type curves simplifies well test analysis. *World Oil*. 1983; 196(6):95-106.

Box GE. Robustness in the strategy of scientific model building. *Robustness in statistics*. 1979; 1:201-236.

Brodley CE and Friedl MA. Identifying and eliminating mislabeled training instances. Identifying and eliminating mislabeled training instances. *Proceedings of the National Conference on Artificial Intelligence*; 1996. p. 799-805.

Buckley SE, Leverett M. Mechanism of fluid displacement in sands. *Transactions of the AIME*. 1942; 146(01):107-116.

Burke LI, Ignizio JP. Neural networks and operations research: an overview. *Computers & Operations Research*. 1992; 19(3-4):179-189.

Caers J. *Modeling uncertainty in the earth sciences*. : John Wiley & Sons; 2011.

Cantini S, Baldini D, Beretta E, Loi D and Mazzoni S. Reservoir permeability from wireline formation testers. Reservoir permeability from wireline formation testers. *EAGE Annual Conference & Exhibition incorporating SPE Europec: Society of Petroleum Engineers*; 2013.

- Carman P. Determination of the specific surface of powders I. Transactions. *J.Soc.Chemical Industries*. 1938; 57:225-234.
- Carman PC. *Flow of gases through porous media*. : Academic press; 1956.
- Carman PC. Fluid flow through granular beds. *Trans.Inst.Chem.Eng.* 1937; 15:150-166.
- Catuneanu O. *Principles of sequence stratigraphy*. : Elsevier; 2006.
- Chalkley HW, Cornfield J, Park H. A method for estimating volume-surface ratios. *Science*. 1949; 110(2856):295-298.
- Chaudhry A. *Oil well testing handbook*. : Elsevier; 2004.
- Clavier C, Hoyle W, Meunier D. Quantitative interpretation of thermal neutron decay time logs: part I. Fundamentals and techniques. *Journal of Petroleum Technology*. 1971; 23(06):743-755.
- Coates GR and Dumanoir JL. A new approach to improved log-derived permeability. A new approach to improved log-derived permeability. *SPWLA 14th Annual Logging Symposium: Society of Petrophysicists and Well-Log Analysts*; 1973.
- Corbett PW, Jensen J. Application of probe permeametry to the prediction of two-phase flow performance in laminated sandstones (lower Brent Group, North Sea). *Marine and Petroleum Geology*. 1993; 10(4):335-346.
- Corbett P, Jensen J. Quantification of variability in laminated sediments: A role for the probe permeameter in improved reservoir characterization. *Geological Society, London, Special Publications*. 1993; 73(1):433-442.
- Corbett P, Jensen J, Sorbie K. A review of up-scaling and cross-scaling issues in core and log data interpretation and prediction. *Geological Society, London, Special Publications*. 1998; 136(1):9-16.
- Cordiner, F. S., Gordon, D. T., & Jargon, J. R. (1972, January). Determination of residual oil saturation after waterflooding. In *SPE Improved Oil Recovery Symposium*. Society of Petroleum Engineers.
- Corey A, Rathjens C, Henderson J, Wyllie M. Three-phase relative permeability. *Journal of Petroleum Technology*. 1956; 8(11):63-65.
- Craig FF. *The reservoir engineering aspects of waterflooding*. : HL Doherty Memorial Fund of AIME New York; 1971.
- De Marsily G. *Quantitative hydrogeology*. 1986; .
- Delshad M, Pope GA. Comparison of the three-phase oil relative permeability models. *Transport in Porous Media*. 1989; 4(1):59-83.
- Deo P, Xue DC and Mendez F. Managing uncertainty of well log data in reservoir characterization. Managing uncertainty of well log data in reservoir characterization. *EUROPEC/EAGE Conference and Exhibition: Society of Petroleum Engineers*; 2009.
- Dietrich, J. K. (1975). A Method for Determining Reservoir Fluid Saturations Using Field Production Data (includes associated paper 6408). *Society of Petroleum Engineers Journal*, 15(06), 477-486.
- Du Yuqi OB, Dacun L. Literature review on methods to obtain relative permeability data. 2004; .

- Dubrule O, Haldorsen HH. *Geostatistics for permeability estimation*. : Academic Press, Orlando, Florida, USA; 1986.
- Dvorkin J, Nur A. Elasticity of high-porosity sandstones: Theory for two North Sea data sets. *Geophysics*. 1996; 61(5):1363-1370.
- Earlougher Jr R. Advances in well test analysis: SPE Monograph Vol. 5. *V, Society of Petroleum Engineers of AIME, Dallas*. 1977; .
- Eslinger E, Pevear DR. *Clay minerals for petroleum geologists and engineers*. ; 1988.
- Fayers F, Matthews J. Evaluation of normalized Stone's methods for estimating three-phase relative permeabilities. *Society of Petroleum Engineers Journal*. 1984; 24(02):224-232.
- Freeze RA, Cherry JA. *Groundwater: Englewood Cliffs. New Jersey*. 1979; .
- Friedman J, Hastie T, Tibshirani R. *The elements of statistical learning*. : Springer series in statistics New York; 2001.
- Garb FA. Assessing Risk in Estimating Hydrocarbon Reserves and in Evaluating Hydrocarbon-Producing Properties (includes associated papers 18606 and 18610). *Journal of Petroleum Technology*. 1988; 40(06):765-778.
- Gardner MW, Dorling S. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment*. 1998; 32(14):2627-2636.
- Gaymard R, Poupon A. Response of neutron and formation density logs in hydrocarbon bearing formations. *The Log Analyst*. 1968; 9(05).
- Geman S, Bienenstock E, Doursat R. Neural networks and the bias/variance dilemma. *Neural computation*. 1992; 4(1):1-58.
- Ginsburg RN, Pettijohn FJ. *Evolving concepts in Sedimentology*. 1973; .
- Glennie K, Underhill J. Origin, development and evolution of structural styles. *Petroleum geology of the North Sea: Basic concepts and recent advances*. 1998; :42-84.
- Grader AS, Horne RN. Interference testing: Detecting a circular impermeable or compressible subregion. *SPE Formation Evaluation*. 1988; 3(2):420-428.
- Gringarten A. Interpretation of transient well test data. *Developments in petroleum engineering*. 1985; 1:133-196.
- Gringarten AC. Well test analysis in practice. *The Way Ahead*. 2012; 8(02):10-14.
- Gringarten AC, Bourdet DP, Landel PA and Kniazeff VJ. A comparison between different skin and wellbore storage type-curves for early-time transient analysis. A comparison between different skin and wellbore storage type-curves for early-time transient analysis. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 1979.
- Gunter G, Finneran J, Hartmann D and Miller J. Early determination of reservoir flow units using an integrated petrophysical method. Early determination of reservoir flow units using an integrated petrophysical method. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 1997.
- Haldorsen HH. Simulator parameter assignment and the problem of scale in reservoir engineering. *Reservoir characterization*. 1986; 6.

Hamada GM, Elshafei MA. Neural network prediction of porosity and permeability of heterogeneous gas sand reservoirs using NMR and conventional logs. *Nafta*. 2010 Oct 29;61(10):451-60.

Heinemann ZE. Petroleum recovery. *Leoben: textbook series*. 2003; 16.

Helle HB, Bhatt A, Ursin B. Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. *Geophysical Prospecting*. 2001; 49(4):431-444.

Herrick D and Kennedy W. On The Quagmire Of Shaly Sand Saturation Equations. On The Quagmire Of Shaly Sand Saturation Equations. *SPWLA 50th Annual Logging Symposium: Society of Petrophysicists and Well-Log Analysts*; 2009.

Hewett T. Fractal distributions of reservoir heterogeneity and their influence on fluid transport. Fractal distributions of reservoir heterogeneity and their influence on fluid transport. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 1986.

Hinton GE, Sejnowski TJ, Poggio TA. *Unsupervised learning: foundations of neural computation*. : MIT press; 1999.

Holden L, Hauge R, Skare Ø, Skorstad A. Modeling of fluvial reservoirs with object models. *Mathematical Geology*. 1998; 30(5):473-496.

Honarpour M, Koederitz L, Harvey AH. Empirical equations for estimating two-phase relative permeability in consolidated rock. *Journal of Petroleum Technology*. 1982; 34(12):2,905-2,908.

Horner, S. C., & Sanyal, S. K. (1984). *Prediction of saturation using the carbon/oxygen log* (No. DOE/SF/11564-9; SUPRI-TR-40). Stanford Univ., CA (USA). Petroleum Research Inst..

Hossin A. Calcul des saturations en eau par la methode du ciment argileux (formule d'Archie generalisee). *Bull.Assoc.Francaise Tech.Petrol*. 1960; 140:31.

Huang Y, Gedeon TD, Wong PM. An integrated neural-fuzzy-genetic-algorithm using hyper-surface membership functions to predict permeability in petroleum reservoirs. *Engineering Applications of Artificial Intelligence*. 2001 Feb 1;14(1):15-21.

Hubick KT, Hubick KT. *Artificial Neural Networks in Australia: Executive summary 2. Introduction 2.1 Neuroscience 2.2. Development of computers 2.3 What is AI 3. Markets 4. Australian capabilities in neural networks 4.1 Industrial developments in Australia 4.2 Training in artificial neural network technology 4.3 Australian research 5. International developments in artificial neural networks 6. Future directions*. : Department of Industry, Technology & Commerce; 1992.

Huh C and Pope GA. Residual oil saturation from polymer floods: laboratory measurements and theoretical interpretation. Residual oil saturation from polymer floods: laboratory measurements and theoretical interpretation. *SPE Symposium on Improved Oil Recovery: Society of Petroleum Engineers*; 2008.

Hurst A, Rosvoll K, Lake L, Carroll H, Wesson T. Reservoir characterization II. 1991; .

James GM. Variance and bias for general loss functions. *Machine Learning*. 2003; 51(2):115-135.

Jensen JL, Lake LW. The influence of sample size and permeability distribution on heterogeneity measures. *SPE Reservoir Engineering*. 1988; 3(02):629-637.

Jensen J, Lake LW, Corbett PW, Goggin D. *Statistics for petroleum engineers and geoscientists*. : Gulf Professional Publishing; 2000.

Johnson H, Fisher M. North Sea plays: geological controls on hydrocarbon distribution. *Petroleum Geology of the North Sea: Basic Concepts and Recent Advances*. 1998; :463-547.

Jorden JR, Campbell FL. Well logging II-Electric and acoustic logging. 1986; .

Jouanna P, Fras G. Introduction A La Reconnaissance, Dans L'espace Des Frequences, De Milieux Fissures Par Essais D'eau Transitoire-Notion De Signatures Spectrales Et Application Au Cas D'horizons Fissures Reconnus Par Pompage Harmonique-Unicite Des Solutions. *CR HEBD ACAD SCI, AB*. 1979; 288(2).

Juanes R, Spiteri E, Orr F, Blunt M. Impact of relative permeability hysteresis on geological CO2 storage. *Water Resources Research*. 2006; 42(12).

Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: A survey. *Journal of artificial intelligence research*. 1996; 4:237-285.

Kamel MH, Mabrouk WM. An equation for estimating water saturation in clean formations utilizing resistivity and sonic logs: theory and application. *Journal of Petroleum Science and Engineering*. 2002; 36(3-4):159-168.

Katahara K. What is shale to a petrophysicist? *The Leading Edge*. 2008; 27(6):738-741.

Keijzer J and Kortekaas T. Comparison of deterministic and probabilistic simulation models of channel sands in the Statfjord reservoir, Brent Field. Comparison of deterministic and probabilistic simulation models of channel sands in the Statfjord reservoir, Brent Field. *European Petroleum Conference: Society of Petroleum Engineers*; 1990.

Kenari SAJ, Mashohor S. Robust committee machine for water saturation prediction. *Journal of Petroleum Science and Engineering*. 2013; 104:1-10.

Kozeny J. About capillaries conducting water in the earth. *Sitzber.Akd.Wiss.Wien, Math.Naurv.Kasse*. 1927; 136:271-306.

Kuchuk F, Biryukov D. Pressure-transient tests and flow regimes in fractured reservoirs. *SPE Reservoir Evaluation & Engineering*. 2015; 18(02):187-204.

Kumar B, Varghese J. Application of Monte Carlo Simulation to Quantify Uncertainties in Petrophysical Deliverables. 2005; .

Larionov V. Borehole radiometry. *Nedra, Moscow*. 1969; 127.

Lawal KA. Modelling Subsurface Uncertainties with Experimental Design: Some Arguments of Non-Conformists. Modelling Subsurface Uncertainties with Experimental Design: Some Arguments of Non-Conformists. *Nigeria Annual International Conference and Exhibition: Society of Petroleum Engineers*; 2009.

Lee LW, Krum GL, Yao T, Wattenbarger RC and Landis LH. Application of Model-Based Uncertainty Analysis. Application of Model-Based Uncertainty Analysis. *Abu Dhabi International Petroleum Exhibition and Conference: Society of Petroleum Engineers*; 2006.

Leveaux J, Poupon A. Evaluation of water saturation in shaly formations. *The Log Analyst*. 1971; 12(04).

Li K, Horne RN. Comparison of methods to calculate relative permeability from capillary pressure in consolidated water-wet porous media. *Water Resources Research*. 2006; 42(6).

- Lim J. Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea. *Journal of Petroleum Science and Engineering*. 2005; 49(3):182-192.
- Lomeland F, Ebeltoft E and Thomas WH. A new versatile relative permeability correlation. A new versatile relative permeability correlation. *International Symposium of the Society of Core Analysts, Toronto, Canada*; 2005. p. 1-12.
- Majdi A, Beiki M. Evolving neural network using a genetic algorithm for predicting the deformation modulus of rock masses. *International Journal of Rock Mechanics and Mining Sciences*. 2010 Feb 1;47(2):246-53.
- Manjnath A and Honarpour M. An investigation of three-phase relative permeability. An investigation of three-phase relative permeability. *SPE Rocky Mountain Regional Meeting: Society of Petroleum Engineers*; 1984.
- Mardi M, Nurozi H, Edalatkhah S. A water saturation prediction using artificial neural networks and an investigation on cementation factors and saturation exponent variations in an Iranian oil well. *Petroleum Science and Technology*. 2012; 30(4):425-434.
- Matheron G. *Eléments pour une théorie des milieux poreux*. 1967.
- Mazzullo S, Chilingarian GV. Diagenesis and origin of porosity. *Developments in petroleum science*. : Elsevier; 1992. p. 199-270.
- McVay DA, Dossary MN. The Value of Assessing Uncertainty. *SPE Economics & Management*. 2014; 6(02):100-110.
- McWhorter DB, Sunada DK. *Ground-water hydrology and hydraulics*. : Water Resources Publication; 1977.
- Mendrofa DM and Widarsono B. Analytical approach of seismic-derived resistivity. Analytical approach of seismic-derived resistivity. *GEO 2008*; 2008.
- Morrow EW. Oil industry megaprojects: our recent track record. Oil industry megaprojects: our recent track record. *Offshore Technology Conference: Offshore Technology Conference*; 2011.
- Nakayama MK. Multiple comparisons with the best using common random numbers for steady-state simulations. *Journal of Statistical Planning and Inference*. 2000; 85(1-2):37-48.
- Ng A, Ngiam J, Foo CY, Mai Y, Suen C, Coates A, et al. *Unsupervised feature learning and deep learning*. 2013.
- Nikraves M, Aminzadeh F. Past, present and future intelligent reservoir characterization trends. *Journal of petroleum science and engineering*. 2001 Nov 1;31(2-4):67-79.
- Noetinger B and Haas A. Permeability averaging for well tests in 3D stochastic reservoir models. Permeability averaging for well tests in 3D stochastic reservoir models. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 1996.
- Oraby E and Eubanks D. Determination of Irreducible Water Saturation Using Magnetic Resonance Imaging Logs (MRIL): A Case Study From East Texas, USAM. Determination Of Irreducible Water Saturation Using Magnetic Resonance Imaging Logs (MRIL): A Case Study From East Texas, USAM. *Middle East Oil Show and Conference: Society of Petroleum Engineers*; 1997.

- Pape H, Riepe L, Schopper JR. Interlayer conductivity of rocks—a fractal model of interface irregularities for calculating interlayer conductivity of natural porous mineral systems. *Colloids and Surfaces*. 1987; 27(4):97-122.
- Parker J, Lenhard R, Kuppasamy T. A parametric model for constitutive properties governing multiphase flow in porous media. *Water Resources Research*. 1987; 23(4):618-624.
- Patrick Corbett SZ, Pinisetti M, Mesmari A. The integration of geology and well testing for improved fluvial reservoir characterisation. 1998; .
- Pegrum R, Spencer A. Hydrocarbon plays in the northern North Sea. *Geological Society, London, Special Publications*. 1990; 50(1):441-470.
- Pickup G, Ringrose P, Corbett P, Jensen J, Sorbie K. Geology, geometry and effective flow. *Petroleum Geoscience*. 1995; 1(1):37-42.
- Poston S, Ysrael S, Hossain A, Montgomery III E. The effect of temperature on irreducible water saturation and relative permeability of unconsolidated sands. *Society of Petroleum Engineers Journal*. 1970; 10(02):171-180.
- Potocki D. Resolving Differences Between Core and Welltest Permeability in Basal Colorado Sandstones, Canada: The Role of Rock Heterogeneity, Relative Permeability and Formation Damage. Resolving Differences Between Core and Welltest Permeability in Basal Colorado Sandstones, Canada: The Role of Rock Heterogeneity, Relative Permeability and Formation Damage. *SPE Rocky Mountain Petroleum Technology Conference: Society of Petroleum Engineers*; 2001.
- Poupon A, Loy M, Tixier M. A contribution to electrical log interpretation in shaly sands. *Journal of Petroleum Technology*. 1954; 6(6):27-34.
- Quintard M, Whitaker S. Two-phase flow in heterogeneous porous media: The method of large-scale averaging. *Transport in Porous Media*. 1988; 3(4):357-413.
- Quintero L, Boyd A, Gyllensten A and El-Wazeer F. Comparison of permeability from NMR and production analysis in carbonate reservoirs. Comparison of permeability from NMR and production analysis in carbonate reservoirs. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 1999.
- Randall, R. R., Fertl, W. F., & Hopkinson, E. C. (1983). Time-derived sigma for pulsed neutron capture logging. *Journal of Petroleum Technology*, 35(06), 1-187.
- Ramey Jr H. Advances in Practical Well-Test Analysis (includes associated paper 26134). *Journal of Petroleum Technology*. 1992; 44(06):650-659.
- Rathmell J, Braun P, Perkins T. Reservoir waterflood residual oil saturation from laboratory tests. *Journal of Petroleum Technology*. 1973; 25(02):175-185.
- Ringen J, Halvorsen C, Lehne K, Rueslaatten H, Holand H. Reservoir water saturation measured on cores; case histories and recommendations. *SCA-9906*. 2001; .
- Ringrose P, Bentley M. Reservoir model types. *Reservoir Model Design*. : Springer; 2015. p. 173-231.
- Ringrose P, Corbett P. Controls on two-phase fluid flow in heterogeneous sandstones. *Geological Society, London, Special Publications*. 1994; 78(1):141-150.

Ringrose P, Sorbie K, Corbett P, Jensen J. Immiscible flow behaviour in laminated and cross-bedded sandstones. *Journal of Petroleum Science and Engineering*. 1993; 9(2):103-124.

Rosa A and Horne R. Automated type-curve matching in well test analysis using Laplace space determination of parameter gradients. Automated type-curve matching in well test analysis using Laplace space determination of parameter gradients. *SPE Annual Technical Conference and Exhibition*: Society of Petroleum Engineers; 1983.

Rose PR. Delivering on our E&P promises. *The Leading Edge*. 2004; 23(2):165-168.

Saner S, Kissami M, Al Nufaili S. Estimation of permeability from well logs using resistivity and saturation data. *SPE Formation Evaluation*. 1997; 12(01):27-31.

Sanyal SK, Sarmiento Z. Booking geothermal energy reserves. *GRC Transactions*. 2005; 29.

Saraf DN, Batycky JP, Jackson CH and Fisher DB. An experimental investigation of three-phase flow of water-oil-gas mixtures through water-wet sandstones. An experimental investigation of three-phase flow of water-oil-gas mixtures through water-wet sandstones. *SPE California Regional Meeting*: Society of Petroleum Engineers; 1982.

Saraf D, Fatt I. Three-phase relative permeability measurement using a nuclear magnetic resonance technique for estimating fluid saturation. *Society of Petroleum Engineers Journal*. 1967; 7(03):235-242.

Saraf D, McCaffery F. Relative permeabilities. *Developments in Petroleum Science*. 1985; 17:75-118.

Schlumberger. *Log Interpretation.-1: Principles*. : Schlumberger; 1972.

Schneider F, Owens W. Steady-state measurements of relative permeability for polymer/oil systems. *Society of Petroleum Engineers Journal*. 1982; 22(01):79-86.

Schneider F, Owens W. Sandstone and carbonate two-and three-phase relative permeability characteristics. *Society of Petroleum Engineers Journal*. 1970; 10(01):75-84.

Segesman F and Liu O. The excavation effect. The excavation effect. *SPWLA 12th Annual Logging Symposium*: Society of Petrophysicists and Well-Log Analysts; 1971.

Sen P, Straley C, Kenyon W, Whittingham M. Surface-to-volume ratio, charge density, nuclear magnetic relaxation, and permeability in clay-bearing sandstones. *Geophysics*. 1990; 55(1):61-69.

Shanley K, McCabe P, Flint S, Bryant I. Alluvial architecture in a sequence stratigraphic framework: a case history from the Upper Cretaceous of southern Utah, USA. *The geological modelling of hydrocarbon reservoirs and outcrop analogues*. : International Association of Sedimentologists; 1993. p. 21-56.

Shanmugam G. 50 years of the turbidite paradigm (1950s—1990s): deep-water processes and facies models—a critical perspective. *Marine and Petroleum Geology*. 2000; 17(2):285-342.

Sheldon J, Harris C and Bavly D. A method for general reservoir behavior simulation on digital computers. A method for general reservoir behavior simulation on digital computers. *Fall Meeting of the Society of Petroleum Engineers of AIME*: Society of Petroleum Engineers; 1960.

Sherborne J, Sarem A and Sandiford B. Flooding oil-containing formations with solutions of polymer in water. Flooding oil-containing formations with solutions of polymer in water. *7th World Petroleum Congress*: World Petroleum Congress; 1967.

- Shin D, Lee J, Lee J and Yoo H. 14.2 DNPU: An 8.1 TOPS/W reconfigurable CNN-RNN processor for general-purpose deep neural networks. 14.2 DNPU: An 8.1 TOPS/W reconfigurable CNN-RNN processor for general-purpose deep neural networks. *2017 IEEE International Solid-State Circuits Conference (ISSCC)*: IEEE; 2017. p. 240-241.
- Siddiqui S, Hicks PJ, Ertekin T. Two-Phase Relative Permeability Models in Reservoir Engineering Calculations. *Energy Sources*. 1999; 21(1-2):145-162.
- Simandoux P. Dielectric measurements on porous media, application to the measurements of water saturation: study of behavior of argillaceous formations. *Revue de l'Institut Francais du Petrol*. 1963; 18(supplementary issue):93-215.
- Skopec R. Proper coring and wellsite core handling procedures: The first step toward reliable core analysis. *Journal of Petroleum Technology*. 1994; 46(04):280-280.
- Smith MR and Martinez T. Improving classification accuracy by identifying and removing instances that should be misclassified. Improving classification accuracy by identifying and removing instances that should be misclassified. *The 2011 International Joint Conference on Neural Networks*: IEEE; 2011. p. 2690-2697.
- Spears RW. Lithofacies-based corrections to density-neutron porosity in a high-porosity gas-and oil-bearing turbidite sandstone reservoir, Erha field, OPL 209, Deepwater Nigeria. *Petrophysics*. 2006; 47(04).
- Stegemeier GL. Relationship of trapped oil saturation to petrophysical properties of porous media. Relationship of trapped oil saturation to petrophysical properties of porous media. *SPE Improved Oil Recovery Symposium*: Society of Petroleum Engineers; 1974.
- Stegemeier G. Improved oil recovery by surfactant and polymer flooding. *Vol. I, eds. Shah, DO and Schechter, RS, Academic Press, NY*. 1977; .
- Stieber S. Pulsed neutron capture log evaluation-louisiana gulf coast. Pulsed neutron capture log evaluation-louisiana gulf coast. *Fall Meeting of the Society of Petroleum Engineers of AIME*: Society of Petroleum Engineers; 1970.
- Stone H. Estimation of three-phase relative permeability and residual oil data. *J. Pet. Technol.; (United States)*. 1973; 12(4).
- Stone H. Probability model for estimating three-phase relative permeability. *Journal of Petroleum Technology*. 1970; 22(02):214-218.
- Suro-Perez V, Journel A. Indicator principal component kriging. *Mathematical Geology*. 1991; 23(5):759-788.
- Swanson B. A simple correlation between permeabilities and mercury capillary pressures. *Journal of Petroleum Technology*. 1981; 33(12):2,498-2,504.
- Teklu TW, Brown JS, Kazemi H, Graves RM and AlSumaiti AM. A critical literature review of laboratory and field scale determination of residual oil saturation. A critical literature review of laboratory and field scale determination of residual oil saturation. *SPE Production and Operations Symposium*: Society of Petroleum Engineers; 2013.
- Timur A. An investigation of permeability, porosity, & residual water saturation relationships for sandstone reservoirs. *The Log Analyst*. 1968; 9(04).
- Tixier M. Evaluation of permeability from electric-log resistivity gradients. *The Oil and Gas Journal*. 1949; 16:113-133.

Trevor H, Robert T, JH F. The elements of statistical learning: data mining, inference, and prediction. 2009; .

Usman Ahmed S, Crary S, Coates G. Permeability Estimations The Various Sources and Their Interrelationships. 1991; .

Uusitalo L, Lehtikainen A, Helle I, Myrberg K. An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*. 2015; 63:24-31.

von Schroeter T, Hollaender F and Gringarten AC. Deconvolution of well test data as a nonlinear total least squares problem. Deconvolution of well test data as a nonlinear total least squares problem. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 2001.

Washburne CW. The estimation of oil reserves. *Transactions of the AIME*. 1916; 52(01):645-648.

Waxman MH, Smits L. Electrical conductivities in oil-bearing shaly sands. *Society of Petroleum Engineers Journal*. 1968; 8(02):107-122.

Welge HJ. A simplified method for computing oil recovery by gas or water drive. *Journal of Petroleum Technology*. 1952; 4(04):91-98.

Wentworth CK. A scale of grade and class terms for clastic sediments. *The Journal of geology*. 1922; 30(5):377-392.

Wong P, Aminzadeh F, Nikravesh M. *Soft computing for reservoir characterization and modeling*. : Physica; 2013.

Worthington PF. The evolution of shaly-sand concepts in reservoir evaluation. *The Log Analyst*. 1985; 26(01).

Wreath DG. *A study of polymerflooding and residual oil saturation*. 1989; .

Wyllie M. A Note on the Interrelationship Between Wetting and Nonwetting Phase Relative Permeability. *Journal of Petroleum Technology*. 1951; 3(10):17-17.

Wyllie M, Gregory A, Gardner G. An experimental investigation of factors affecting elastic wave velocities in porous media. *Geophysics*. 1958; 23(3):459-493.

Yarus JM, Chambers RL. Practical geostatistics-an armchair overview for petroleum reservoir engineers. *Journal of Petroleum Technology*. 2006; 58(11):78-86.

Yeh N and Agarwal R. Pressure transient analysis of injection wells in reservoirs with multiple fluid banks. Pressure transient analysis of injection wells in reservoirs with multiple fluid banks. *SPE Annual Technical Conference and Exhibition: Society of Petroleum Engineers*; 1989.

Zemanek J. Low-resistivity hydrocarbon-bearing sand reservoirs. *SPE Formation Evaluation*. 1989; 4(04):515-521.

Zhang Q, Wei C, Wang Y, Du S, Zhou Y, Song H. Potential for Prediction of Water Saturation Distribution in Reservoirs Utilizing Machine Learning Methods. *Energies*. 2019 Jan;12(19):3597.

APPENDICES

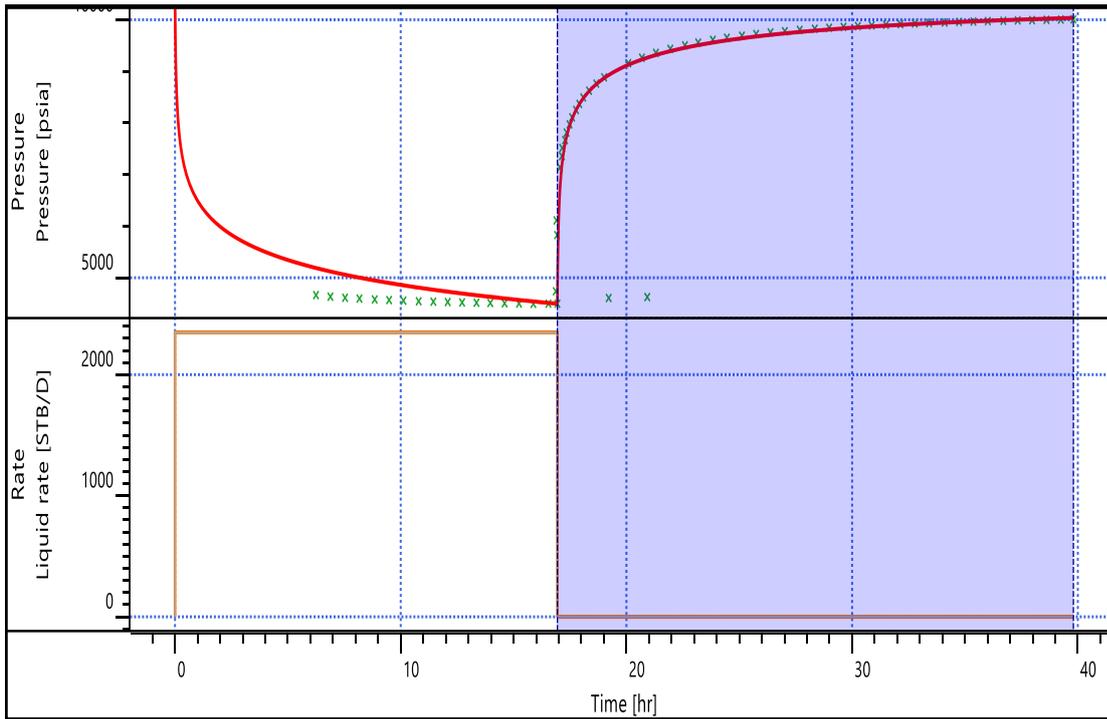
Appendix A- 1 Input parameters for well test analysis of well

Reservoir Properties	Values
Porosity (frac)	0.23
Pay thickness (ft)	100
Compressibility (1/psi)	3E-6
Fluid Viscosity (cp)	1.5
Formation Volume Factor (rb/stb)	1.0
Flowrate (stb/d)	1000
Well Radius (in)	3.6
Reservoir Area (acres)	2500
Initial Pressure (psia)	4988
Skin	2.1
Permeability (mD)	15.15

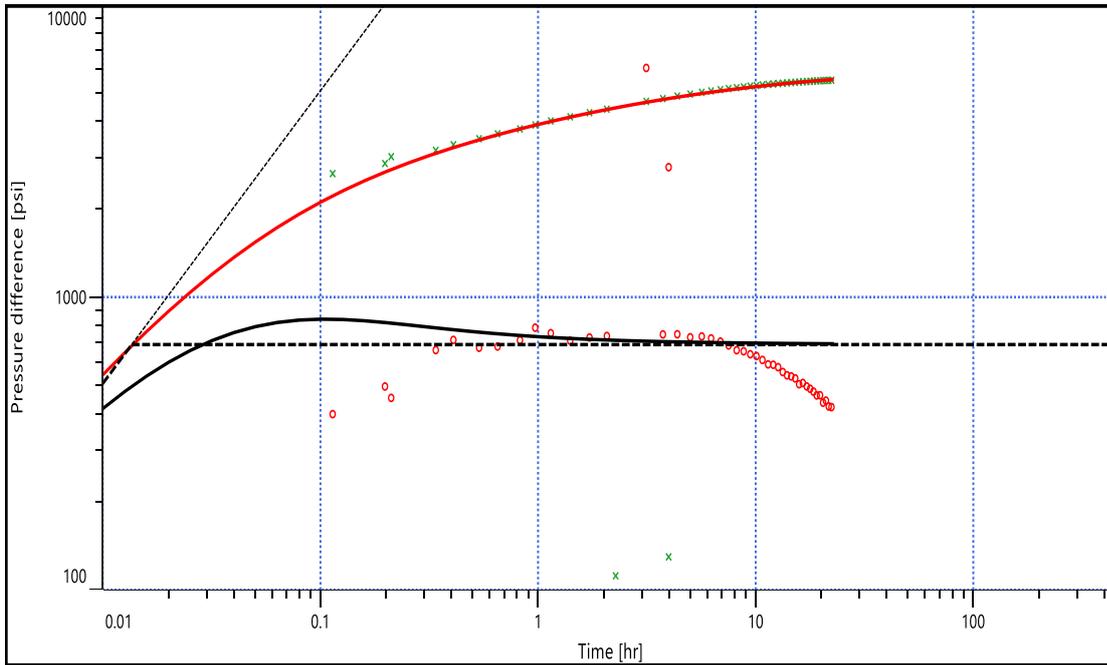
Appendix A- 2 Input parameters for well test analysis of the infinite conductive reservoir.

Reservoir Properties	Values
Porosity (frac)	0.20
Pay thickness (ft)	30.1
Compressibility (bbl/psi)	0.0011582
Fluid Viscosity (cp)	1.0
Formation Volume Factor (rb/stb)	1.2
Flowrate (stb/d)	2300
Well Radius (in)	3.6
Area of Reservoir (acres)	2295.68
Initial Pressure (psia)	14618.6
Skin	3.15
Permeability (mD)	8.034

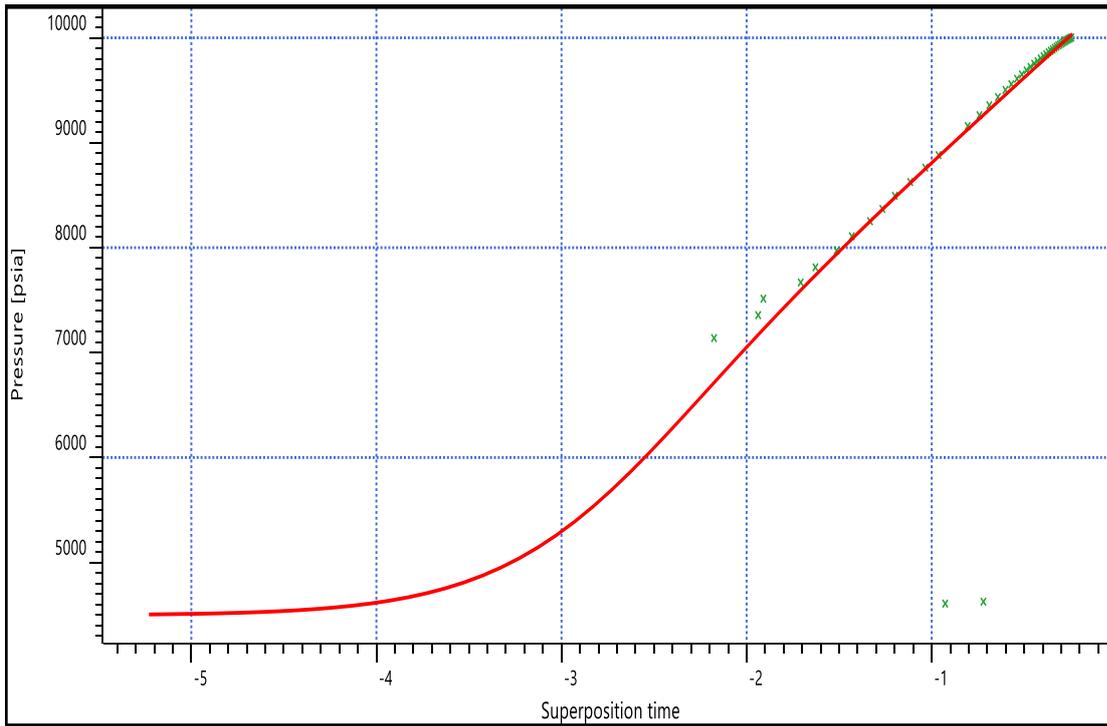
Appendix A- 3 History plot indicating pressure buildup analysis



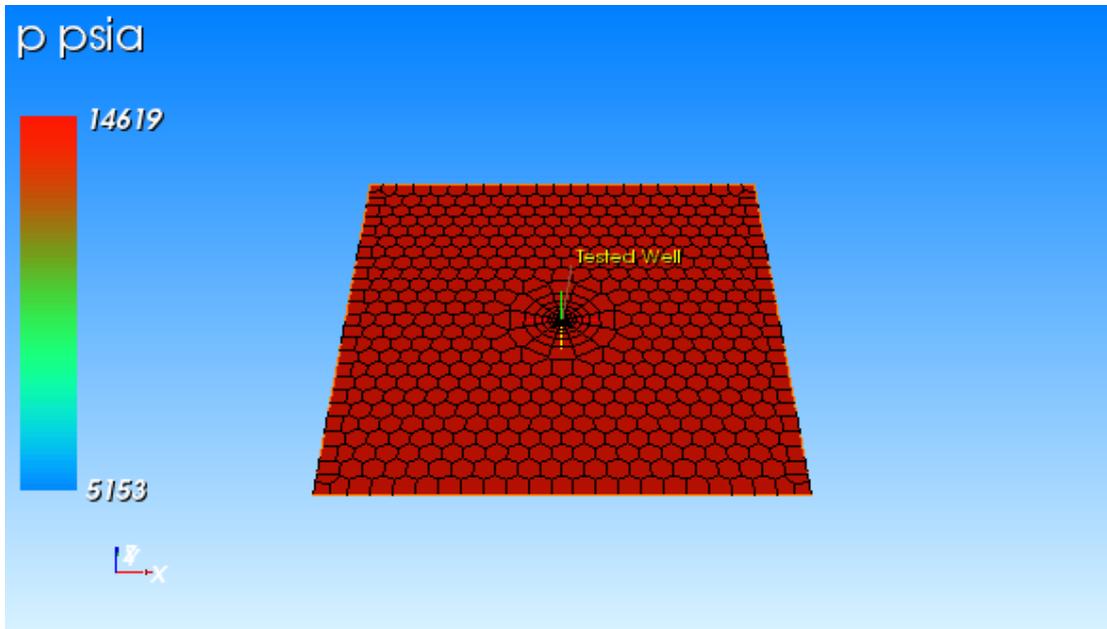
Appendix A- 4 Log-Log diagnostic plot showing pressure and derivative



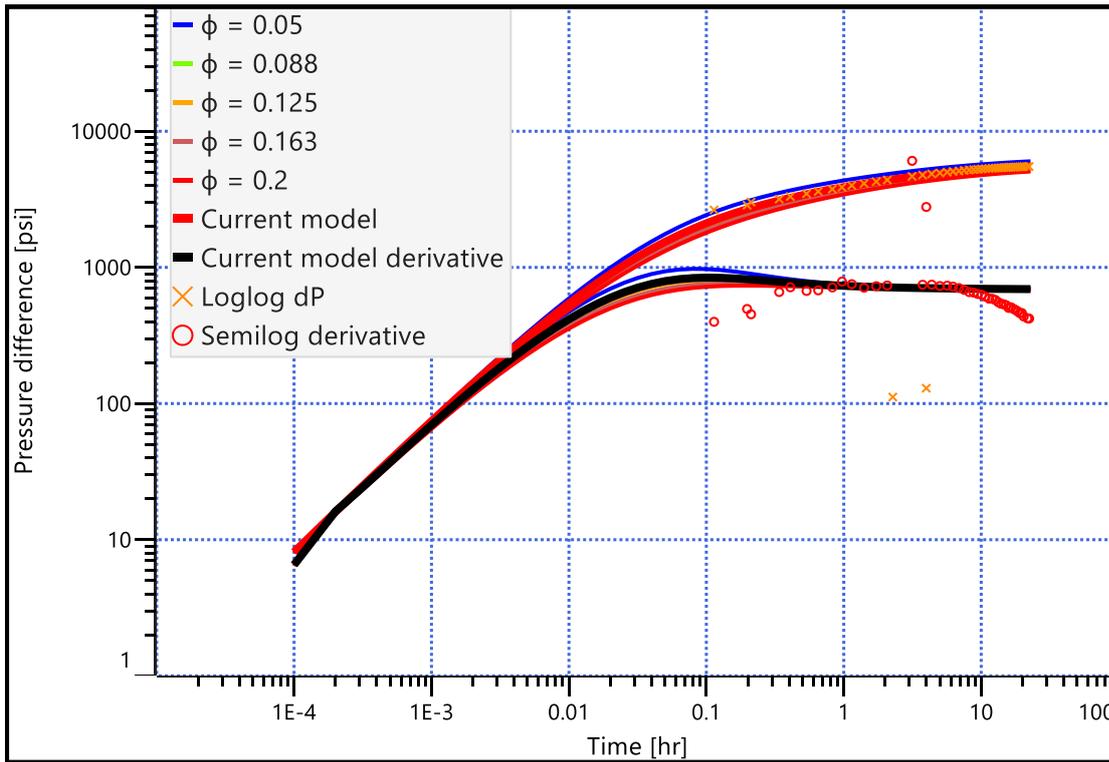
Appendix A- 5 Superposition plot of actual and simulated plot for well



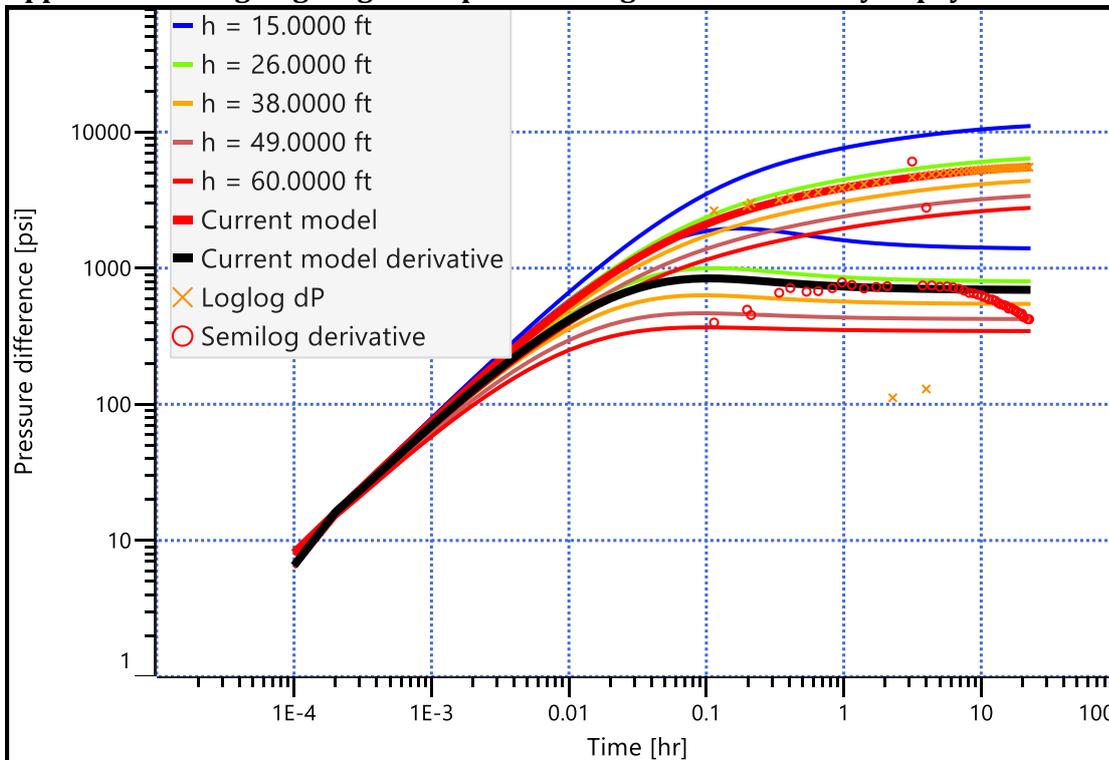
Appendix A- 6 Numerical simulation of the tested well



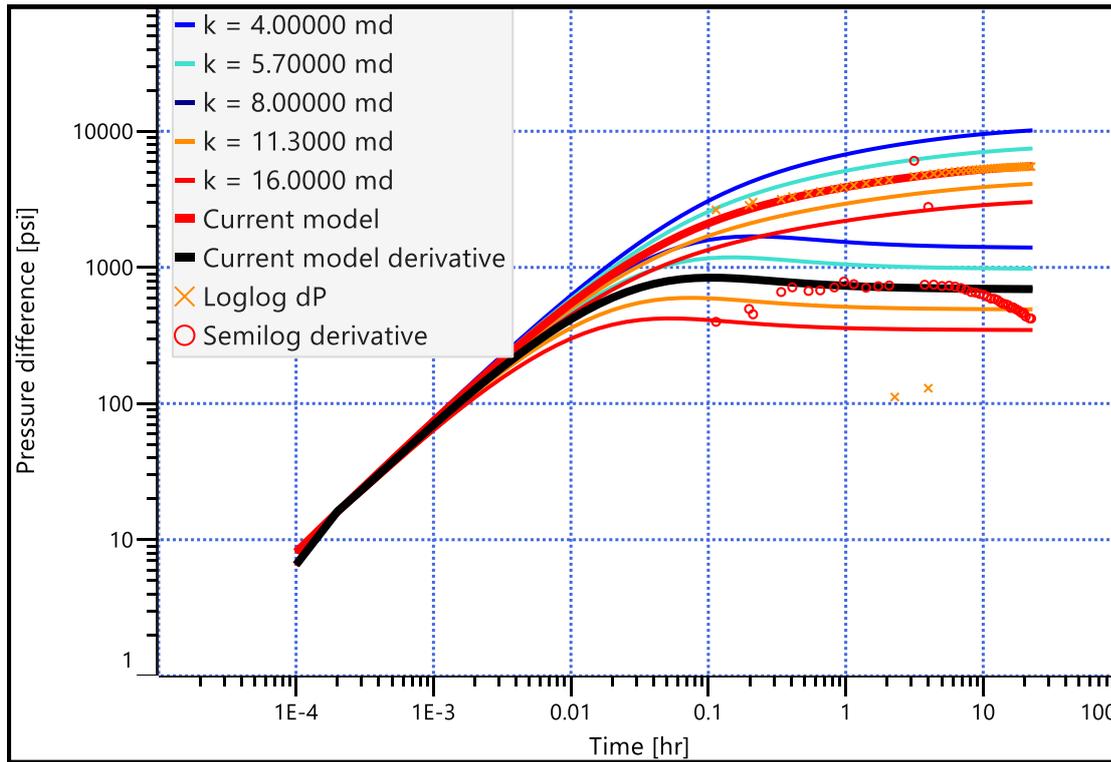
Appendix A- 7 Log-Log diagnostic plot showing model sensitivity to porosity



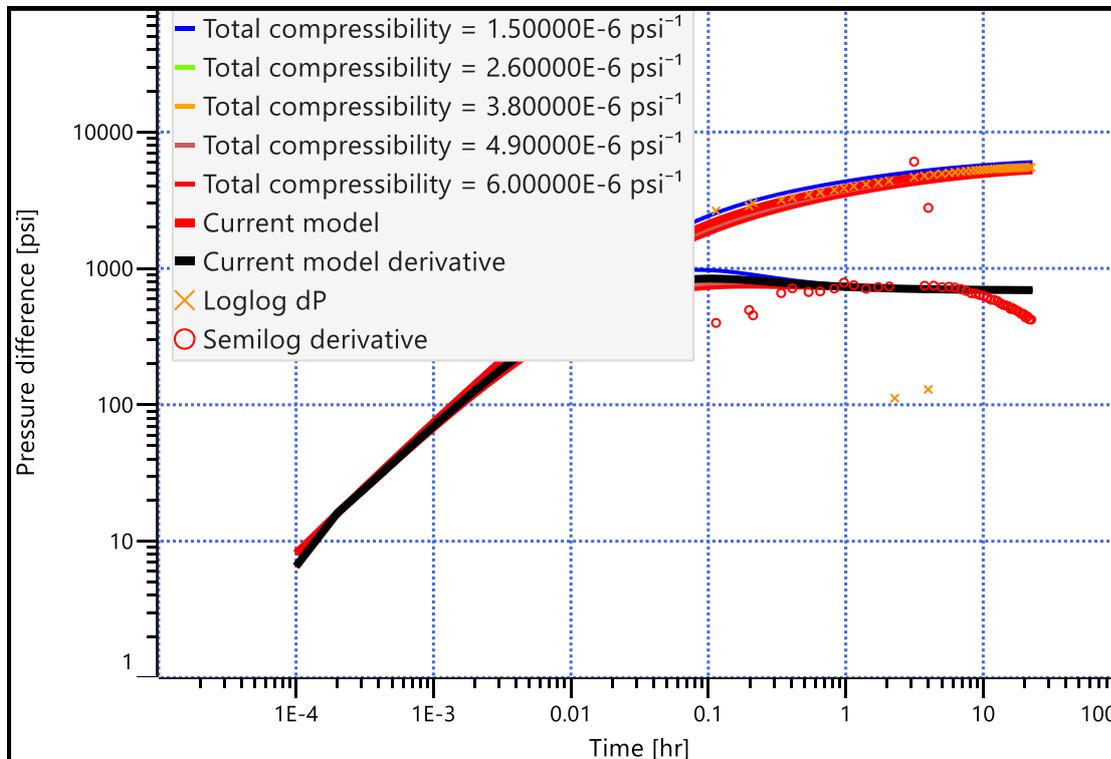
Appendix A- 8 Log-Log diagnostic plot showing model sensitivity to pay thickness



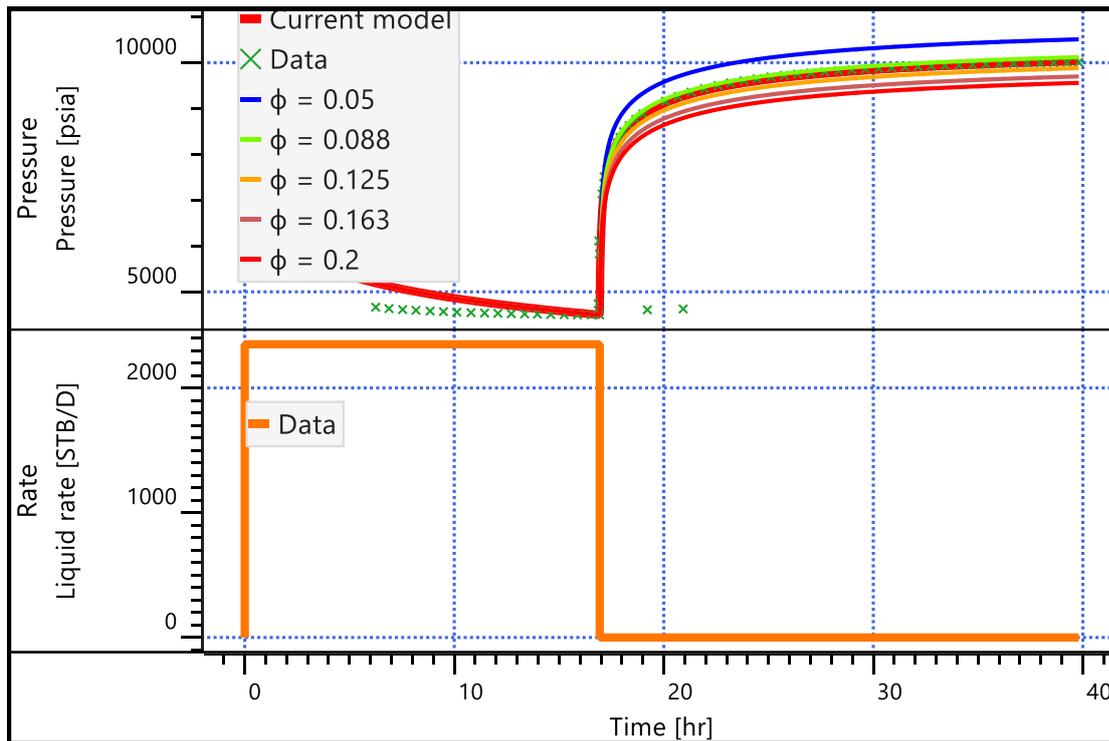
Appendix A- 9 Log-Log diagnostic plot showing model sensitivity to permeability



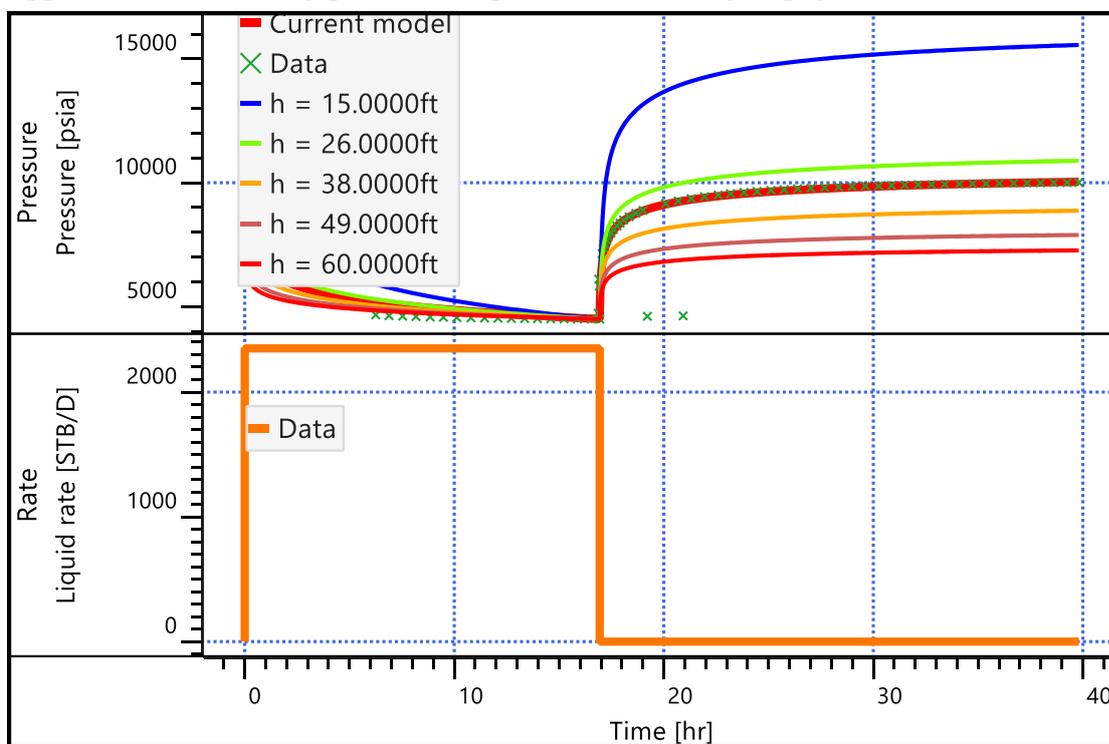
Appendix A- 10 Log-Log diagnostic plot showing model sensitivity to compressibility



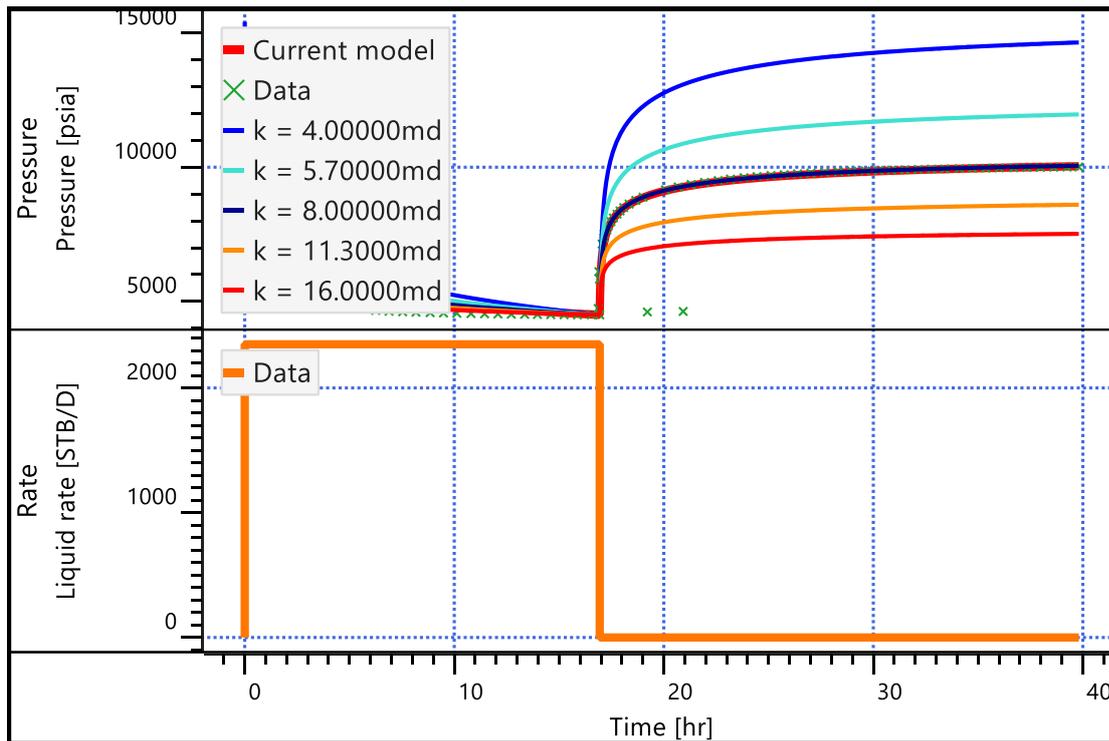
Appendix A- 11 History plot showing model sensitivity to porosity



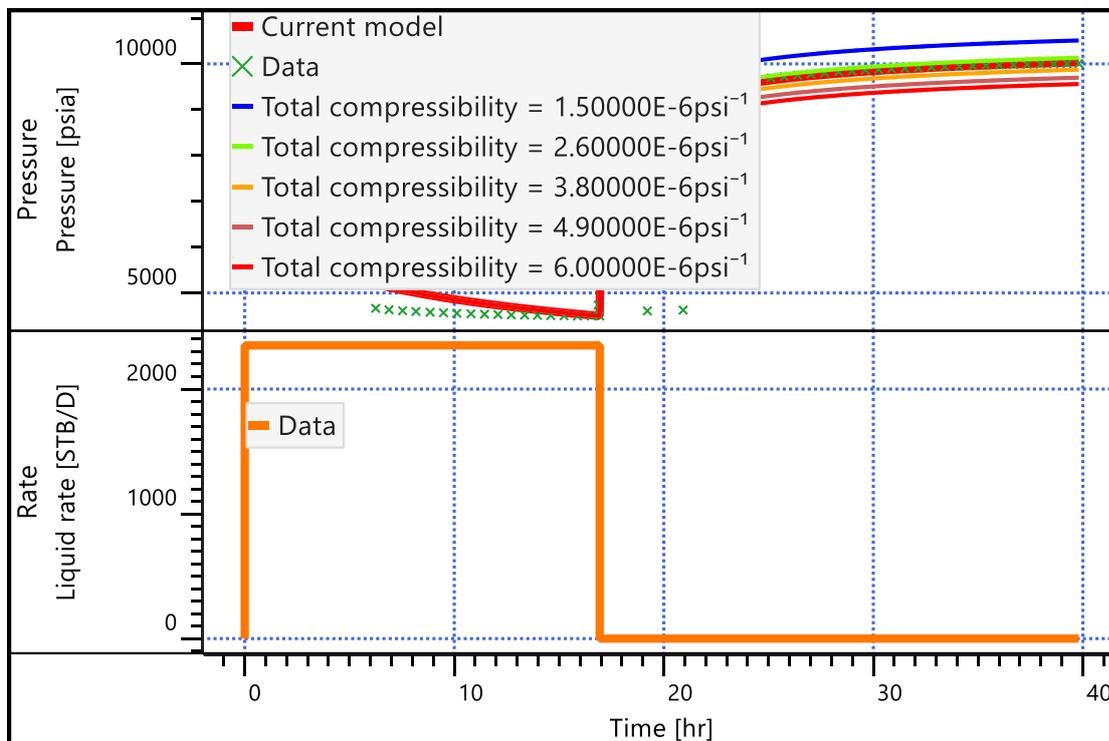
Appendix A- 12 History plot showing model sensitivity to pay thickness



Appendix A- 13 History plot showing model sensitivity to permeability



Appendix A- 14 History plot showing model sensitivity to compressibility



Appendix B- 1 Codes for Support Vector Porosity Model

Install Packages

```
install.packages("neuralnet")
install.packages("caret")
install.packages("e1071")
```

Load Packages

```
library(neuralnet)
library(caret)
library(e1071)
```

Load data

```
svm_phi<-read.table("C:/Users/ovkoe/Documents/gaussianProcess/svm_phi.csv",header=T,
sep=",")

svm_phi <- as.data.frame(svm_phi)
plot(svm_phi)
str(svm_phi)
```

#Plot data

```
plot(svm_phi$nphi,svm_phi$rhob,col=svm_phi$core)
plot(svm_phi$nphi, svm_phi$son, col=svm_phi$core)
```

Create training and testing datasets randomly

```
Indo <- sample(1:nrow(svm_phi), 435)
trainPhi <- svm_phi[Indo,]
testPhi <- svm_phi[-Indo,]
```

Organise the formular names

```
allVars <- colnames(svm_phi)
predictorsvars <- allVars[!allVars%in%"core"]
predictorsvars <- paste(predictorsvars,collapse = "+")
form <- as.formula(paste("core~", predictorsvars, collapse = "+"))
```

model fitting using neuralnet function

```
svm_phi <- svm(core~.,data=trainPhi, cost=128, gamma=0.01, epsilon=0.15)
```

predict for test data set

```
predictPhi <- predict(svm_phi,testPhi[,2:4])
a=cbind(predictPhi,testPhi$core)
View(a)
```

```

error <- testPhi$core-predictPhi
svrPredRMSE <- sqrt(mean(predictPhi-testPhi$core)^2)
svm_phi_tune <- tune(svm, core~., data = trainPhi, ranges = list(epsilon = seq(0,1,0.1), cost =
1:100)
)
plot(svm_phi_tune,contour=t)
print(svm_phi_tune)

```

retune

```

tuneResult <- tune(svm, core~., data=trainPhi, ranges = list(epsilon =
seq(svm_phi_tune$best.model$epsilon-.15,
      svm_phi_tune$best.model$epsilon+.15,
      0.1),
      cost = seq(2^(log2(svm_phi_tune$best.model$cost)-1),
      2^(log2(svm_phi_tune$best.model$cost)+1),
      length = 6)
))

```

```

best_mod <- svm_phi_tune$best.model
best_mod_pred <- predict(best_mod,testPhi)
error_best_mod <- testPhi$core - best_mod_pred
best_mod_rmse <- sqrt(mean(error_best_mod^2))
points(testPhi$core, best_mod_pred,col="blue", pch=4)

```

Appendix B- 2 Codes for Ensemble Total Water Saturation Model

Install the packages

```

install.packages("h2o")
install.packages("caretEnsemble")

```

Load package

```

library(caretEnsemble)
library(h2o)

```

Initialize h2o

```

h2o.init()

```

sessionInfo()

```

plot(ensData)

```

Import data from document

```
ensData<-read.csv("c:/Users/ovkoe/Documents/gaussianProcess/sw_ensemble.csv", header =
T, sep = ",")
```

```
#Make it an h2o dataset
```

```
ensData <- as.h2o(ensData)
dim(ensData)
```

Separate data into ensemble, blending and testing sets

```
splits <- h2o.splitFrame(ensData,
                        ratios = c(0.9, 0.051),
                        seed = 148)
```

```
train <- splits[[1]]
valid <- splits[[2]]
test <- splits[[3]]
```

Organise formular names

```
y <- "Sw"
x <- names(ensData)[names(ensData)!=y]
nfolds=5
```

Random grid search

```
search_Criteria <- list(strategy="RandomDiscrete", max_runtime_secs=600)
hidden_opt <- list(c(200,200), c(100,300,100), c(500,500))
l1_opt <- c(1e-5,1e-7)
hyper_params <- list(hidden = hidden_opt, l1 = l1_opt)
grid <- h2o.grid(algorithm = "deeplearning", hyper_params = hyper_params, search_criteria =
search_Criteria, x=predictors,
y=labelName, training_frame = ensembleSwi_h2o, validation_frame= blenderSwi_h2o)
```

Instruct how the train function provides the best parameters for the model

```
modelControl <- trainControl(method = 'repeatedcv', number = 10, repeats = 3, savePredictions
= TRUE, classProbs = TRUE)
```

Train & Cross-validate a GBM

```
my_gbm <- h2o.gbm(x = x,
                 y = y,
                 training_frame = train2,
                 distribution = "gaussian",
                 model_id = "gbm_def",
                 max_depth = 7,
                 learn_rate = 0.2,
                 nfolds = nfolds,
                 fold_assignment = "Modulo",
```

```
keep_cross_validation_predictions = TRUE,  
seed = 1)
```

Eval perf

```
perf_gbm_train <- h2o.performance(my_gbm)  
perf_gbm_test <- h2o.performance(my_gbm, newdata = ensData1)  
pred_gbm <- h2o.predict(my_gbm, newdata = test[,1:5])  
h2o.mae(pred_gbm)  
plot(testSwi_h2o$Swi, pred$Swi, col = 'black', pch=1, cex=1, type = "p", xlab = "Actual", ylab =  
"Predicted")  
cbind(pred_gbm, test$Sw)  
View(test$Swi)  
View(pred_gbm)  
h2o.varimp_plot(my_gbm)
```

Train & Cross-validate a RF

```
my_rf <- h2o.randomForest(x = x,  
  y = y,  
  training_frame = train,  
  ntrees = 30,  
  model_id = "rf_def",  
  nfolds = nfolds,  
  fold_assignment = "Modulo",  
  keep_cross_validation_predictions = TRUE,  
  seed = 1)
```

Eval perf

```
perf_rf_trainSw <- h2o.performance(rf)  
perf_rf_testSw <- h2o.performance(rf, newdata = testSw)  
pred1_rfSw <- h2o.predict(rf, newdata = testSw[,1:5])  
pred2_rfSw <- h2o.predict(rf, newdata = testSw[,1:5])  
h2o.varimp_plot(rf)  
plot(rf)
```

Train and cross validate a GLM

```
my_glm <- h2o.glm(x = x,  
  y = y,  
  training_frame = train,  
  family = "gaussian",  
  model_id = "glm_def",  
  nfolds = nfolds,  
  standardize = TRUE,  
  fold_assignment = "Modulo",  
  keep_cross_validation_predictions = TRUE,  
  seed = 1)
```

Eval perf

```

perf_glm_train <- h2o.performance(my_glm)
perf_glm_test <- h2o.performance(my_glm, newdata = test)
pred_glm <- h2o.predict(my_glm, newdata = test[,1:5])
pred1_glm <- h2o.predict(my_glm, newdata = ensData1[,1:5])
plot(testSw_h2o$Sw,pred$Swi,col = 'black',pch=1,cex=1,type = "p",xlab = "Actual",ylab =
"Predicted")
cbind(pred_glm,test$Sw)
View(pred_glm)
h2o.varimp_plot(my_glm)

```

Train and cross validate a NN

```

my_nn <- h2o.deeplearning(x = x, y = y, distribution = "gaussian",
      training_frame = train,
      nfolds = nfolds,
      model_id = "dl_def",
      fold_assignment = "Modulo",
      hidden = 20,
      epsilon = 1e-08,
      rate = 0.005,
      l2=1e-05,
      keep_cross_validation_predictions = TRUE,
      seed = 1)

```

Eval perf

```

perf_dl_train <- h2o.performance(my_dl)
perf_dl_test <- h2o.performance(my_dl, newdata = test)
pred_dl <- h2o.predict(my_dl, newdata = test[,1:5])
View(pred_dl)
View(pred2_dl)
h2o.varimp_plot(my_dl)

```

Train a stacked ensemble using the GBM and NN above

```

models_ids <- list(my_gbm@model_id, my_nn@model_id)
SE <- h2o.stackedEnsemble(x = x,
      y = y,
      training_frame = train,
      validation_frame = blenderSwi_h2o,
      model_id = "SE_gbm_glm_rf_xrf_dl",
      base_models = models_ids,
      metalearner_algorithm = "deeplearning",
      metalearner_nfolds = nfolds,
      seed = 1
)

```

Predict Using the Stacked Model

```

pred_SE <- h2o.predict(SE,newdata = test[,1:5])
plot(pred_stack,testSw_h2o$Sw)

```

```

perf_SE_train <- h2o.performance(SE)
perf_SE_test <- h2o.performance(SE, newdata = ensData1)
plot(h2o.performance(SE))
baselearner_best_rmse_train <- min(h2o.rmse(perf_gbm_train), h2o.rmse(perf_dl_train))
baselearner_best_rmse_test <- min(h2o.rmse(perf_gbm_test), h2o.rmse(perf_dl_test))
SE_rmse_train <- h2o.rmse(perf_SE_train)
SE_rmse_test <- h2o.rmse(perf_SE_test)

h2o.shutdown()

```

Appendix B- 3 Codes for Relative permeability Deep Learning Model

Install the packages

```

install.packages("h2o")
install.packages("caretEnsemble")

```

Load package

```

library(caretEnsemble)
library(h2o)

```

Data import for Krw anmd Kro

```

RelPerm_orig=read.table("C:/Users/ovkoe/Documents/WorkingDIR/relpermpredkrw.csv",hea
der=T, sep=",")
RelPerm_Krw=as.data.frame(RelPerm_orig)
plot(relperm_krw)

```

```

relperm_orig=read.table("C:/Users/ovkoe/Documents/relPermeability/relpermpredkro.csv",h
eader=T, sep=",")
relperm_kro=as.data.frame(relperm_orig)
plot(relperm_kro)

```

h2o initialisation

```

h2o.init(ip = "localhost",port = 54321)

```

Data Normalization for both Krw and Kro

```

Traindlw = RelPerm_Krw[1:106,]
Testdlw = RelPerm_Krw[107:118,]
valdlw = RelPerm_Krw[119:137,]

```

```

Traindlo = RelPerm_Kro[1:106,]
Testdlo = RelPerm_Kro[107:118,]
valdlo = RelPerm_Kro[119:137,]

```

```

MaxValue <- apply(relperm_krw,2,max)
MinValue <- apply(relperm_krw,2,min)
traindlw_df <- as.data.frame(scale(traindlw,center = MinValue,scale = MaxValue-MinValue))
traindlw_h2o <- as.h2o(traindlw_df,destination_frame = "traindlw_h2o")

```

```

Maxvalue <- apply(relperm_kro,2,max)
Minvalue <- apply(relperm_kro,2,min)
traindlo_df <- as.data.frame(scale(traindlo,center = Minvalue,scale = Maxvalue-Minvalue))
traindlo_h2o <- as.h2o(traindlo_df,destination_frame = "traindlo_h2o")

```

Defining x and y

```

yw = "krw"
xw = setdiff(colnames(traindlw_h2o),yw)

```

```

yo = "kro"
xo = setdiff(colnames(traindlo_h2o),yo)

```

Hyper_parameter tuning with grid search

```

hyper_params <- list(
  activation=c("Rectifier","Tanh","Maxout","RectifierWithDropout"),
  hidden=list(c(20,20,20,20,20),c(50,50,50,50,50),c(30,30,30,30),c(200,200,200,200,200)),
  input_dropout_ratio=c(0,0.05),
  l1=seq(0,1e-4,1e-6),
  l2=seq(0,1e-4,1e-6)
)

```

```

# Stop once the top 5 models are within 1% of each other (i.e., the windowed average varied less than 1%)

```

```

help("h2o.grid")

```

```

search_criteria = list(strategy="RandomDiscrete", stopping_rounds=10, seed=1234567,
stopping_metric="AUTO", stopping_tolerance=1e-3)

```

```

dl_random_grid <- h2o.grid(
  algorithm = "deeplearning",
  grid_id = "dl_grid",
  training_frame=traindlw_h2o,
  validation_frame=valDlw_h2o,
  x=xw,
  y=yw,
  epochs=10,
  stopping_tolerance=1e-2, ## stop when logloss does not improve by >=1% for two scoring events

```

```

# score_validation_samples=10000, ## downsample validation set for faster scoring

```

```

score_duty_cycle=0.025,
max_w2=10, ## can help improve stability for rectifier
hyper_params = hyper_params,
search_criteria = search_criteria)
summary(dl_random_grid)

library(jsonlite)

grid <- h2o.getGrid("dl_random_grid",sort_by = "err",decreasing = FALSE)

```

Model fitting

Regularization with l1 and l2 to further solve the problem of overfitting

```

modeldlw = h2o.deeplearning(x=xw,
  y=yw,
  seed = 1234,
  training_frame = as.h2o(traindlw_df),
  nfolds = 5,
  standardize = FALSE, # since it has already been normalized
  stopping_rounds = 5,
  epochs = 400,
  overwrite_with_best_model = TRUE,
  ignore_const_cols = FALSE,
  activation = "Rectifier",
  hidden = c(100,100),
  l2=6e-5,
  diagnostics = TRUE,
  variable_importances = TRUE,
  loss = "Automatic",
  distribution = "AUTO",
  stopping_metric = "RMSE")

```

```
help("h2o.deeplearning")
```

```
plot(as.data.frame(h2o.varimp(modeldlw))) # variable importance of the model
```

modeling oil relative permeability

```

modeldlo = h2o.deeplearning(x=xo,
  y=yo,
  seed = 1234,
  training_frame = as.h2o(traindlo_df),
  nfolds = 5,
  # standardize =FALSE
  stopping_rounds = 5,
  epochs = 400,
  overwrite_with_best_model = TRUE,

```

```

ignore_const_cols=FALSE,
activation = "Rectifier",
diagnostics = TRUE,
variable_importances = TRUE,
hidden = c(100,100,100,100),
l2=6e-5,
loss = "Automatic",
distribution = "AUTO",
stopping_metric = "RMSE")

```

```
as.data.frame(h2o.varimp(modeldlo))
```

predictions

```

predictiondlw = as.data.frame(predict(modeldlw,as.h2o(testdlw_df)))
predictiondlw_v = as.data.frame(predict(modeldlw,as.h2o(valdlw_df)))
g=cbind(predictiondlw,testdlw_df$krw)
View(g)
h=cbind(predictiondlw_v,valdlw_df$krw)
View(h)
h2o.varimp_plot(modeldlo)

```

```

predictiondlo = as.data.frame(predict(modeldlo,as.h2o(testdlo_df)))
predictiondlo_v = as.data.frame(predict(modeldlo,as.h2o(valdlo_df)))
i=cbind(predictiondlo,testdlo_df$kro)
View(i)
j=cbind(predictiondlo_v,valdlo_df$kro)
View(j)
predictiondlw$predict
h2o.sensitivity(modeldlw)

```

plotting predicted values vs actual values

```

par(mfrow=c(2,2))
plot(testdlw_df$krw,predictiondlw$predict,col = 'black',main = 'dnn validation krw',
     pch=1,cex=1,type = "p",xlab = "Actual",ylab = "Predicted")
plot(valdlw_dfi$Krw,predictiondlw_v$predict,col = 'black',main = 'dnn test Krw',
     pch=1,cex=1,type = "p",xlab = "Actual",ylab = "Predicted")
plot(Testdlo_dfi$Kro,predictiondlo$predict,col = 'black',main = 'dnn validation Kro',
     pch=1,cex=1,type = "p",xlab = "actual",ylab = "predicted")
plot(valdlo_dfi$Kro,predictiondlo_v$predict,col = 'black',main = 'dnn test Kro',
     pch=1,cex=1,type = "p",xlab = "actual",ylab = "predicted")

```

MSE determination

```

MSEdlw <- sum((predictiondlw$predict-Testdlw_dfi$Krw)^2)/nrow(Testdlw_dfi)
MSEdlw_v<- sum((predictiondlw_v$predict-valdlw_dfi$Krw)^2)/nrow(valdlw_dfi)

```

```
MSEdlo <- sum((predictiondlo$predict-Testdlo_dfi$Kro)^2)/nrow(Testdlo_dfi)
MSEdlo_v<- sum((predictiondlo_v$predict-valdlo_dfi$Kro)^2)/nrow(valdlo_dfi)
```

```
h2o.shutdown()
```

Appendix B- 4 Log, Core and Test Support Vector Regression Model Alongside other tested Models

Read Original data table

```
set.seed(123)
```

```
LCTun=read.table("C:/Users/ovkoe/Documents/WorkingDIR/LCTun_ML.csv",header=T,
sep=",")
```

Convert to data frame

```
LCTun=as.data.frame(LCTun)
```

Install/load Packages

```
install.packages("caret")
```

```
install.packages("h2o")
```

```
install.packages("e1071")
```

Load Packages

```
library(caret)
```

```
library(e1071)
```

```
library(shiny)
```

Create training and testing datasets

```
Indw <- sample(1:nrow(LCTun), 120)
```

```
Traink <- LCTun[Indw,]
```

```
Testk <- LCTun[-Indw,]
```

```
dim(Traink)
```

Organise the formular names

```
allVars <- colnames(LCTun)
```

```
predictorsvars <- allVars[!allVars%in%"wt"]
```

```

predictorsvars <- paste(predictorsvars,collapse = "+")
form <- as.formula(paste("wt~", predictorsvars, collapse = "+"))

y <- "wt"
x <- names(LCTun)[names(LCTun)!=y]

# model fitting using neuralnet function
neuralnetk <- nnet(wt~,data = Traink,size = 4, decay=5e-04)
predictionk <- predict.nnet(neuralnetk,Testk[,-7])
View(predictionk)

# Support Vector Machine Regression
svmk <- svm(wt~,data=Traink2, cost=1000, gamma=0.0001)

predictionk1 <- predict(svmk,TestK[,-7])
predictionk1 <- predict(svmk, TestKval[,-7])

cbind(predictionk,TestK$wt)
cbind(predictionk,TestKval$wt)

# Tuning svm model parameter
svmk_tune <- tune(svm, wt~, data = Traink3, ranges = list(epsilon = seq(0,1,0.1), cost = 1:100)
)
plot(svmk_tune)
print(svmk1_tune)

# retune if need be
tuneResult <- tune(svm, wt~, data=Traink, ranges = list(epsilon =
seq(svmk1_tune$best.model$epsilon-.15,
      svmk_tune$best.model$epsilon+.15,
      0.1),
cost = seq(2^(log2(svmk1_tune$best.model$cost)-1),
      2^(log2(svmk1_tune$best.model$cost)+1),

```

```
length = 6)))
```

Make plot

```
plot(Testk$wt, predictionk,col="dark green", pch=4, ylim=c(0,1000), xlim=c(0,1000),  
xlab="actual perm (mD)", ylab="svm perm (mD)")
```

```
abline(0,1, col="brown")
```

```
sqrt(mean(prediction-Testk$wt)^2)
```

```
error <- testPhi$score-predictPhi
```

```
svrPredRMSE <- sqrt(mean(predictPhi-testPhi$score)^2)
```

Make SVM interface using the Shiny Package

```
# Define UI for application that draws a histogram
```

```
ui <- fluidPage(  
  # Application title
```

```
  # Application title
```

```
  titlePanel("Well Test SVM Model"),
```

```
  # Sidebar with a slider input for number of bins
```

```
  sidebarLayout(  
    sidebarPanel(  
      sliderInput("kcore",  
        "Core Permeability (mD):",  
        min = 1,  
        max = 1000,  
        value = 50, step = 1),  
      sliderInput("klog",  
        "Log Permeability (mD):",  
        min = 1,  
        max = 1000,  
        value = 50, step = 1),  
      sliderInput("kcore^1/2",  
        "Square Root Core Permeability (mD):",  
        min = 1,
```

```
      sliderInput("kcore",
```

```
        "Core Permeability (mD):",
```

```
        min = 1,
```

```
        max = 1000,
```

```
        value = 50, step = 1),
```

```
      sliderInput("klog",
```

```
        "Log Permeability (mD):",
```

```
        min = 1,
```

```
        max = 1000,
```

```
        value = 50, step = 1),
```

```
      sliderInput("kcore^1/2",
```

```
        "Square Root Core Permeability (mD):",
```

```
        min = 1,
```

```

    max = 60,
    value = 20, step = 1),
sliderInput("klog^1/2",
  "Square Root Log Permeability (mD):",
  min = 1,
  max = 60,
  value = 2, step = 1),
sliderInput("kcore^2",
  "Core Permeability Squared (mD):",
  min = 1,
  max = 1000000,
  value = 1000, step = 100),
sliderInput("klog^2",
  "Log Permeability Squared (mD):",
  min = 1,
  max = 1000000,
  value = 1000, step = 1000)),
# Show a plot of the generated distribution
mainPanel(
  tableOutput("Plot")
)
)
)

# Define server logic required to draw a histogram
server <- function(input, output) {

  output$distPlot <- renderTable({

    LCTun3=read.table("C:/Users/ovkoe/Documents/WorkingDIR/LCTun_ML3.csv",header=T,
sep=",")

    LCTun3=as.data.frame(LCTun3)

    Indw3<- sample(1:nrow(LCTun3), 85)

```

```

Traink3 <- LCTun3[Indw3,]
TestK3 <- LCTun3[-Indw3,]
TestK3val <- Traink3[1:10,]

y <- "wt"
x3 <- names(LCTun3)[names(LCTun3)!=y]

svm3 <- svm(wt~.,data=Traink3, cost=28, gamma=0.01, epsilon=0.6)
New_kcore <- data.frame(kcore=input$kcore)
New_klog <- data.frame(klog=input$klog)
New_kcore^1/2 <- data.frame(kcore^1/2=input$kcore^1/2)
New_klog^1/2 <- data.frame(klog^1/2=input$klog^1/2)
predictionk3 <- predict(svm3, TestK3[,-7])
predictionk3
})
}

# Run the application
shinyApp(ui = ui, server = server)

# Other models also tried
# Initialize H2O Package
h2o.init()

# Train & Cross-validate a GBM
k_gbm <- h2o.gbm(x = x,
  y = y,
  training_frame = as.h2o(Traink),
  distribution = "gaussian",
  model_id = "gbm_def",
  max_depth = 7,

```

```

learn_rate = 1,
nfolds = 10,
fold_assignment = "Modulo",
keep_cross_validation_predictions = TRUE,
seed = 1)

```

Eval perf

```

perf_gbm_traink <- h2o.performance(k_gbm)
perf_gbm_testk <- h2o.performance(k_gbm, newdata = as.h2o(Testk))
pred_gbm <- h2o.predict(k_gbm, newdata = as.h2o(Testk[,-7]))
h2o.rmsle(perf_gbm_Testk)
cbind(pred_gbm, Testk$wt)
plot(Testk$wt, pred_gbm, col = 'black', pch=1, cex=1, type = "p", xlab = "Actual", ylab = "Predicted")
h2o.varimp_plot(k_gbm)

```

Train & Cross-validate a RF

```

k_rf <- h2o.randomForest(x = x,
  y = y,
  training_frame = as.h2o(Traink),
  ntrees = 50,
  model_id = "rf_def",
  nfolds = 5,
  fold_assignment = "Modulo",
  keep_cross_validation_predictions = TRUE,
  seed = 1)

```

Eval perf

```

perf_rf_traink <- h2o.performance(k_rf)
perf_rf_testk <- h2o.performance(k_rf, newdata = as.h2o(Testk))
pred_rfk <- h2o.predict(k_rf, newdata = as.h2o(Testk[,-7]))
h2o.varimp_plot(k_rf)

```

Train & Cross-validate a extremely-randomized RF

```

k_xrf <- h2o.randomForest(x = x,
  y = y,
  training_frame = as.h2o(Traink),
  model_id = "xrf_def",
  ntrees = 50,
  histogram_type = "Random",
  nfolds = 5,
  fold_assignment = "Modulo",
  keep_cross_validation_predictions = TRUE,
  seed = 1)

# Eval perf

perf_xrf_trainsw1 <- h2o.performance(k_xrf)
perf_xrf_testsw1 <- h2o.performance(k_xrf, newdata = as.h2o(Testk))
pred_xrfs1 <- h2o.predict(k_xrf, newdata = as.h2o(Testk[,-7]))

```

Train and cross validate a GLM

```

k_glm <- h2o.glm(x = x,
  y = y,
  training_frame = as.h2o(Traink),
  family = "gaussian",
  model_id = "glm_def",
  nfolds = 5,
  standardize = TRUE,
  fold_assignment = "Modulo",
  keep_cross_validation_predictions = TRUE,
  seed = 1)

# Eval perf

perf_glm_traink <- h2o.performance(k_glm)
perf_glm_testk <- h2o.performance(k_glm, newdata = as.h2o(Testk))
pred_glmk <- h2o.predict(k_glm, newdata = as.h2o(Testk[,-7]))
View(pred_glmk)

```

```
h2o.varimp_plot(k_glm)
```

Train and cross validate a deep learning

```
k_dl <- h2o.deeplearning(x = x, y = y,  
  distribution = "gaussian",  
  training_frame = as.h2o(Traink),  
  nfolds = 10,  
  model_id = "dl_def",  
  fold_assignment = "Modulo",  
  hidden = 10,  
  epsilon = 1e-05,  
  rate = 0.005,  
  l2=1e-08,  
  keep_cross_validation_predictions = TRUE,  
  seed = 1)
```

Evaluate Perf

```
perf_dl_traink <- h2o.performance(k_dl)  
perf_dl_testk <- h2o.performance(k_dl, newdata = as.h2o(Testk))  
pred_dlk <- h2o.predict(k_dl, newdata = as.h2o(Testk[,-7]))  
View(pred_dlk)  
cbind(pred_dlk, TestK1$wt)  
h2o.varimp_plot(k_dl)
```

Train a stacked ensemble

```
models_ids <- list(k_gbm@model_id, k_dl@model_id, k_rf@model_id, k_glm@model_id)  
k_SE <- h2o.stackedEnsemble(x = x,  
  y = y,  
  training_frame = as.h2o(Traink),  
  model_id = "SE_gbm_rf_dl",  
  base_models = models_ids,  
  metalearner_algorithm = "deeplearning",  
  metalearner_nfolds = nfolds,
```

seed = 1

)

h2o.shutdown()

Appendix C- 1 Log, Core and Test permeability data

core	log	core^2	Log^2	sqrt(core)	sqrt(log)	wt
2.00	19.86	4.00	394.59	1.41	4.46	4.70
45.00	272.50	2025.00	74258.16	6.71	16.51	6.40
81.00	89.86	6561.00	8074.82	9.00	9.48	31.00
0.16	1.02	0.03	1.03	0.40	1.01	104.00
16.00	82.30	256.00	6773.29	4.00	9.07	0.23
52.00	48.86	2704.00	2387.30	7.21	6.99	94.31
4.80	32.78	23.04	1074.53	2.19	5.73	211.10
947.00	56.71	896809.00	3216.02	30.77	7.53	100.02
0.21	11.90	0.04	141.61	0.46	3.45	10.00
0.22	4.09	0.05	16.73	0.47	2.02	6.20
0.69	2.75	0.48	7.56	0.83	1.66	6.90
121.00	44.09	14641.00	1943.93	11.00	6.64	35.50
2.34	100.04	5.48	10007.52	1.53	10.00	61.70
1.60	52.68	2.56	2774.85	1.26	7.26	51.50
47.00	47.63	2209.28	2268.97	6.86	6.90	26.30
0.03	0.02	0.00	0.00	0.17	0.14	14.60
1.39	6.94	1.92	48.17	1.18	2.63	33.50
0.03	1.40	0.00	1.96	0.17	1.18	114.00
14.34	12.78	205.72	163.40	3.79	3.58	2.00
6.09	1.99	37.08	3.97	2.47	1.41	3.30
34.41	55.34	1184.19	3062.90	5.87	7.44	5.60
70.12	48.86	4916.25	2387.20	8.37	6.99	8.40
24.66	9.53	608.26	90.81	4.97	3.09	56.30
7.06	16.26	49.79	264.29	2.66	4.03	30.70

96.74	58.11	9358.43	3376.79	9.84	7.62	78.20
1.10	8.82	1.20	77.86	1.05	2.97	8.40
1.06	5.38	1.12	28.90	1.03	2.32	185.10
20.09	95.87	403.77	9191.05	4.48	9.79	78.80
101.98	48.86	10400.12	2387.20	10.10	6.99	5.10
27.65	51.38	764.41	2639.93	5.26	7.17	19.60
25.11	54.00	630.26	2915.76	5.01	7.35	94.80
3.74	16.81	14.02	282.43	1.93	4.10	1474.00
251.08	70.23	63040.16	4931.75	15.85	8.38	14.60
259.37	106.54	67272.28	11351.20	16.10	10.32	115.80
368.44	51.38	135747.30	2639.93	19.19	7.17	37.60
1.55	5.86	2.41	34.40	1.25	2.42	4.40
0.81	7.84	0.65	61.47	0.90	2.80	11.10
0.81	8.16	0.65	66.56	0.90	2.86	13.80
0.52	8.49	0.27	72.02	0.72	2.91	13.80
0.06	8.82	0.00	77.86	0.24	2.97	13.10
20.41	33.71	416.73	1136.40	4.52	5.81	9.90
22.61	39.69	511.39	1575.30	4.76	6.30	3.50
18.84	51.38	354.87	2639.93	4.34	7.17	111.60
65.66	65.48	4310.84	4287.69	8.10	8.09	383.70
1114.20	152.47	1241432.73	23248.05	33.38	12.35	511.30
1758.00	200.93	3090564.00	40373.11	41.93	14.17	124.12
154.00	190.43	23716.00	36262.37	12.41	13.80	60.94
0.12	6.66	0.01	44.34	0.35	2.58	15.48
12.00	133.14	144.00	17725.57	3.46	11.54	44.69
9678.00	380.83	93663684.00	145031.87	98.38	19.51	142.17
664.00	219.40	440896.00	48134.30	25.77	14.81	158.29
45.00	135.78	2025.00	18436.29	6.71	11.65	214.24
21.00	24.53	441.00	601.79	4.58	4.95	381.00
12.00	19.76	144.00	390.48	3.46	4.45	7.51

2.90	12.78	8.41	163.40	1.70	3.58	12.99
1.90	10.67	3.61	113.77	1.38	3.27	14.05
0.21	7.23	0.04	52.29	0.46	2.69	15.48
0.01	0.06	0.00	0.00	0.10	0.24	11.90
0.19	3.05	0.04	9.32	0.44	1.75	17.24
0.08	2.76	0.01	7.60	0.28	1.66	18.84
0.21	0.49	0.04	0.24	0.46	0.70	12.34
0.23	4.92	0.05	24.20	0.48	2.22	19.85
0.31	6.12	0.10	37.47	0.56	2.47	14.32
7.00	7.84	49.00	61.47	2.65	2.80	38.85
0.18	2.62	0.03	6.84	0.42	1.62	31.32
0.17	2.48	0.03	6.15	0.41	1.58	40.87
1.50	7.23	2.25	52.29	1.22	2.69	37.85
3.10	7.84	9.61	61.47	1.76	2.80	40.97
4.70	7.53	22.09	56.72	2.17	2.74	35.78
5.30	5.62	28.09	31.55	2.30	2.37	46.75
4.00	5.62	16.00	31.55	2.00	2.37	39.92
0.23	5.14	0.05	26.46	0.48	2.27	57.21
1.20	6.66	1.44	44.34	1.10	2.58	48.31
0.28	3.05	0.08	9.32	0.53	1.75	6.93
0.19	2.48	0.04	6.15	0.44	1.58	4.00
0.30	2.35	0.09	5.53	0.55	1.53	2.98
2.80	9.53	7.84	90.81	1.67	3.09	7.24
3.70	8.49	13.69	72.02	1.92	2.91	8.84
6.90	9.17	47.61	84.12	2.63	3.03	5.74
5.90	10.67	34.81	113.77	2.43	3.27	9.86
16.00	16.81	256.00	282.43	4.00	4.10	42.22
7.90	7.53	62.41	56.72	2.81	2.74	37.31
3.90	11.07	15.21	122.47	1.97	3.33	42.71
1.90	9.90	3.61	97.97	1.38	3.15	38.26

8.10	8.49	65.61	72.02	2.85	2.91	33.28
12.00	8.49	144.00	72.02	3.46	2.91	36.17
4.20	9.90	17.64	97.97	2.05	3.15	38.60
4.50	12.78	20.25	163.40	2.12	3.58	34.42
1065.00	208.17	1134225.00	43335.00	32.63	14.43	78.97
133.00	70.23	17689.00	4931.75	11.53	8.38	107.02
268.00	62.45	71824.00	3900.37	16.37	7.90	0.81
15.00	30.99	225.00	960.25	3.87	5.57	126.78
4.62	18.54	21.34	343.56	2.15	4.31	0.38
5.80	41.85	33.64	1751.38	2.41	6.47	20.37
	0.00	0.00	0.00	0.00	0.00	0.73
3.27	50.11	10.69	2510.78	1.81	7.08	0.06
120.00	87.93	14400.00	7731.09	10.95	9.38	15.65
1.33	2.23	1.77	4.96	1.15	1.49	31.60
0.02	2.62	0.00	6.84	0.14	1.62	18.20

Real-Time Relative Permeability Prediction Using Deep Learning

O.D.Arigbe, M.B.Oyeneyin, I.Arana, M.D.Ghazi

School of Engineering, Sir Ian Wood Building, Robert Gordon University, Aberdeen, United Kingdom.

Abstract

A review of the existing two and three phase relative permeability correlations shows a lot of pitfalls and restrictions imposed by (a) their assumptions (b) generalization ability and (c) difficulty with updating in real time for different reservoir systems. These increase the uncertainty in its prediction which is crucial owing to the fact that relative permeability is useful for predicting future reservoir performance, effective mobility, ultimate recovery, injectivity among others. Laboratory experiments can be time consuming, complex, expensive and done with core samples which in some circumstances may be difficult or impossible to obtain.

Deep Neural Networks (DNNs) with their special capability to regularize, generalize and update easily with new data has been used to predict oil-water relative permeability. The details have been presented in this paper. In addition to common parameters influencing relative permeability, Baker and Wyllie parameter combinations were used as input to the network after comparing with other models such as Stones, Corey, Parker, Honapour using Corey and Leverett-Lewis experimental data. The DNN automatically used the best cross validation result (in a 5-fold cross validation) for its training until convergence by means of Nesterov accelerated gradient descent which also minimizes the cost function.

Predictions of non-wetting and wetting phase relative permeability gave good match with field data obtained for both validation and test sets. This technique could be integrated into reservoir simulation studies, save cost, optimize the number of laboratory experiments and further demonstrates machine learning as a promising technique for real time reservoir parameters prediction.

Keywords

Deep Neural Networks Relative Permeability Training Validation Testing

Introduction

Relative permeability is the most important property of porous media to carry out reservoir prognosis in a multiphase situation (Delshad and Pope 1989; Yuqi and Dacun 2004) and therefore needs to be as accurate and readily accessible as possible. Theoretically, it is the ratio of effective and absolute permeability. It is useful for the determination of reservoir productivity, effective mobility, wettability, fluid injection for EOR, late-life depressurization, gas condensate depletion with aquifer influx, injectivity, gas trapping, free water surface, residual fluid saturations, temporary gas storage amongst others (Fig. 1). It is well known that a significant variation in relative permeability data can have a huge impact on a macroscopic scale.

The oil and gas industry have a need for easily available and reliable relative permeability data, expense reduction on experiments and a more general model for the parameter judging by the pitfalls pointed out by several researchers after testing the existing two and three phase relative permeability models. Such workers like Fayers-Matthews (1984) and Juanes et al. (2006) after testing non-wetting relative permeability interpolation models such as Baker and Stone's I and II, against Saraf et al. (1982), Schneider and Owens (1970), Saraf and Fatt (1967) and Corey et al. (1956) experimental data, presented the same conclusion that they give similar results for high oil saturations but are different as it tends towards residual oil saturation. Manj Nath and Honarpour (1984) concluded that Corey gives higher values for non-wetting phase relative permeability after comparing against Donaldson and Dean data.

Based on the assumption that water and gas relative permeability depends only on their saturation and not on that of other phases, Delshad and Pope (1989) concluded after a comparative study of 7 relative permeability models that Baker and Pope performed better but also stated the need for better models. Siddiqui et al. (1999) found Wyllie-Gardner and Honarpour to yield consistently better results at experimental condition after testing 10 relative permeability models. Al-Fattah (2009) found Honarpour regression model to be the best after comparing with 5 other models and also developed his own regression model. Since the coefficients of these regression models are not generalized, they are not suitable for real time applications.

Furthermore, for wetting phase relative permeability in consolidated media, Li and Horne (2006) showed that the Purcell model best fits the experimental data in the cases studied by them provided the measured capillary pressure curve had the same residual saturation as the relative permeability curve which is sometimes not the case. Saraf and McCaffery (1985) could not recommend a best model due to scarcity of three phase relative permeability data. The different relative permeability correlations have limitations and assumptions which no doubt have implications thus increasing the uncertainty in reservoir simulation studies hence the need for a more generalized model.

Therefore, the purpose of this study is to implement a Deep Neural Networks model for the prediction of relative permeability accounting for reservoir depletion, saturation and phase changes with time. Guler et al. (1999) developed several neural network models for relative permeability considering different parameters that affects the property and selected the best model to make predictions for the test set while Al-Fattah (2010) also used a Generalized Regression Neural Network to predict relative permeability. Issues of better prediction for out of sample datasets (better generalization) and them requiring far more neurons (and hence an increased computational time) to achieve better results as Deep Learning models. Again most of the reviewed empirical models are static but Deep Neural Networks (with its advanced features) if appropriately tuned can capture the transients faster and more accurately throughout the reservoir life while also getting better as more data becomes available with time. Training can be done offline and the trained networks are suitable for on-board generation of descent relative permeability profiles as their computation requires a modest CPU effort hence not a concern to real time application.

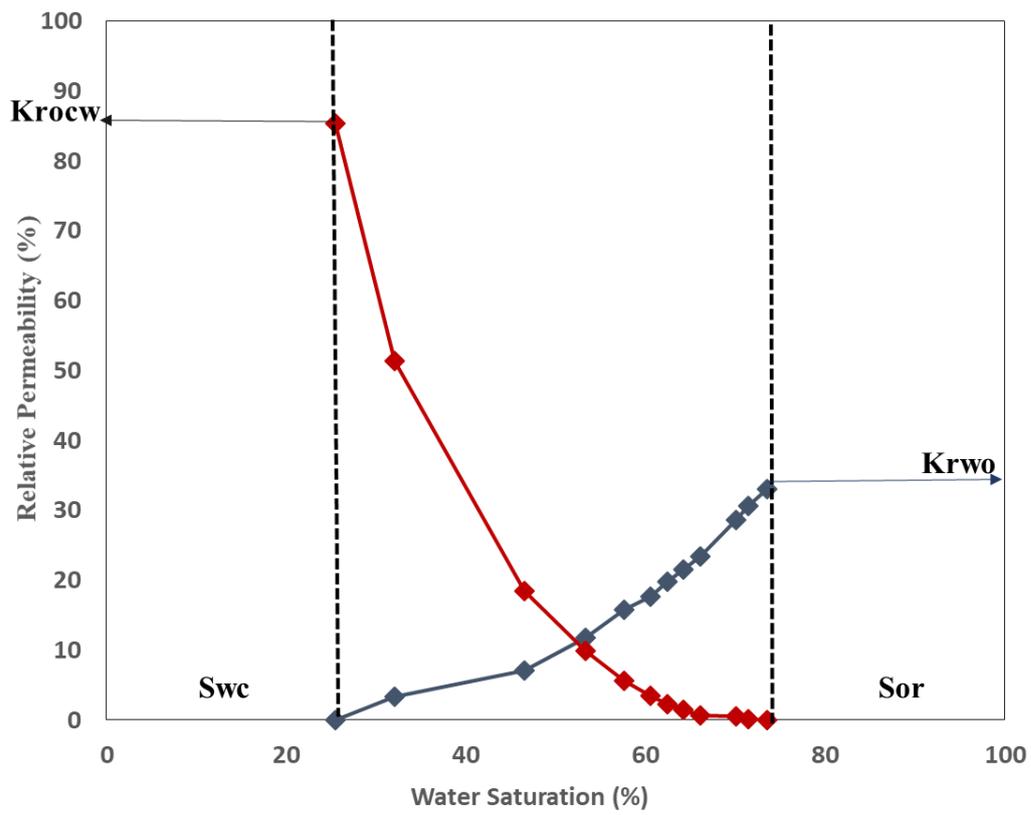


Fig. 1 Schematic of oil-water relative permeability curve

Table 1 Assumptions and application of the commonly used two and three phase relative permeability correlations

MODEL	CORRELATION	PHYSICS	ASSUMPTIONS	APPLICATION WINDOW
Corey et al. (1956)	$K_{ro} = \left(\frac{S_o - S_{or}}{1 - S_{or}} \right)^{\frac{2+3\lambda}{\lambda}}$ $K_{rg} = \left(\frac{1 - S_o}{1 - S_{or}} \right)^2 \left[1 - \left(\frac{S_o - S_{or}}{1 - S_{or}} \right) \right]^{\frac{2+\lambda}{\lambda}}$	An extension of Purcell (1941) and Burdine (1953) which is based on the mean hydraulic radius concept of Kozeny-Carman (bundle of capillaries model) for each pore size in a rock with large variety of pores and tortuosity expressed in terms of fluid saturation.	<p>$K_{ro} \propto$ oil pore area and saturation of water and gas phases</p> <p>Relative permeability of wetting and non-wetting phase independent of saturation of other phases.</p>	<p>Requires a single suite of K_{rg}/K_{ro} data at constant S_w to calculate K_{rg} and K_{ro} for all saturations</p> <p>Not flexible to force end points of isoperms to match measured data.</p> <p>Applies only to well-sorted homogenous rocks.</p>
Wyllie (1951)	$K_{rw} = \left(\frac{S_w - S_{wc}}{1 - S_{wc}} \right)^4$ $K_{rg} = \frac{S_g^2 [(1 - S_{wc})^2 - (S_w + S_o - S_{wc})^2]}{(1 - S_{wc})^4}$ $K_{ro} = \frac{S_o^3 (2S_w + S_o - 2S_{wc})}{(1 - S_{wc})^4}$	Based on bundle of capillaries cut and rejoined along their axis with related entrapment of the wetting phase.	Considers irreducible water as part of the rock matrix.	Applied when water saturation is at irreducible level.
Honarpour et al. (1982)	<p>Water wet</p> $K_{rw} = 0.035388 \left(\frac{S_w - S_{wc}}{1 - S_{wc} - S_{orw}} \right) - 0.0108074 \left(\frac{S_w - S_{orw}}{1 - S_{wc} - S_{orw}} \right)^{2.9} + 0.56556 (S_w)^{3.6} (S_w - S_{wc})$ <p>Any wettability</p> $K_{ro} = 0.76067 \left[\frac{(S_o/1 - S_{wc}) - S_{or}}{1 - S_{orw}} \right]^{1.8} \left[\frac{S_o - S_{orw}}{1 - S_{wc} - S_{orw}} \right]^{2.0} + 2.6318 \emptyset (1 - S_{orw}) (S_o - S_{orw})$	Based on proposed empirical relationships describing experimentally determined permeabilities.	Assumes normally distributed variables	New constant will have to be developed for other areas to have a good fit.

	$K_{rg} = 1.1072 \left(\frac{S_g - S_{gr}}{1 - S_{wc}} \right)^2 K_{rgo} + 2.7794 S_{org} \left(\frac{S_g - S_{gr}}{1 - S_{wc}} \right) K_{rgro}$			
Parker et al. (1987)	$K_{ro} = (\bar{S}_t - \bar{S}_w)^{1/2} \left[\left(1 - \bar{S}_w^{1/m} \right)^m - \left(1 - \bar{S}_t^{1/m} \right)^m \right]^2$	Based on relative permeability, saturation-fluid pressure functional relationships with a flow channel distribution model in two or three phase flow subject to monotonic saturation path and to estimate effective mean fluid conducting pore dimensions.	Wettability takes the water > oil > gas sequence. Irreducible fluid saturation is independent of fluid properties or saturation history No Gas/water contact occurs in the three phase region until the level where oil exists as discontinuous bolbs or pendular rings	Limited to cases where a satisfactory fit to the two-phase data is provided by the fitting equations $m = 1 - 1/n$
Baker (1988)	$K_{ro} = \frac{(S_w - S_{wc})K_{row} + (S_g - S_{gr})K_{rog}}{(S_w - S_{wc}) + (S_g - S_{gr})}$ $K_{rw} = \frac{(S_o - S_{or})K_{rwo} + (S_g - S_{gr})K_{rwg}}{(S_o - S_{or}) + (S_g - S_{gr})}$ $K_{rg} = \frac{(S_o - S_{or})K_{rgo} + (S_w - S_{wc})K_{rgw}}{(S_o - S_{or}) + (S_w - S_{wc})}$	As the saturation of a phase tends to zero, that of the other two-phase will dominate.	The end points of the three phase relative permeability isoperms coincide with the two-phase relative permeability data.	Weighting factors $(S_w - S_{wc})$ and $(S_g - S_{gr})$ must be both positive
Frode (2005)	$K_{row} = K_{ro}^x \frac{(1 - S_{wn})^{L_w^o}}{(1 - S_{wn})^{L_w^o} + E_o^w S_{wn}^{T_w^o}}$ $K_{rw} = K_{rw}^o \frac{S_{wn}^{L_w^o}}{S_{wn}^{L_w^o} + E_w^o S_{wn}^{T_w^o}}$ $S_{wn} = \frac{S_w - S_{wi}}{1 - S_{wi} - S_{orw}}$	Based on the mean hydraulic radius concept of Kozeny-Carman (bundle of capillaries model)	Assumes that the whole spectrum of the relative permeability curve can be captured with the L, E, T parameters.	It exhibits enough flexibility to reconcile the entire spectrum of experimental data.

Methodology

The most commonly available factors influencing relative permeability such as porosity, ϕ , viscosity, μ , permeability, k , saturation, s , together with Baker and Wyllie parameter combinations were used as inputs for the network. Baker gave correlation coefficients of 0.96 and 0.86 while Wyllie has correlation coefficients of 0.91 and 0.89 for Corey and Leverett-Lewis datasets respectively (Table 2). There were a total of 12 input parameters fed into the network as shown in Table 3 after testing the sensitivity of several parameter combinations.

Ten (10) sets of water-oil relative permeability data with 132 data points from a North Sea field with four-fifths used as training set and one-fifth as validation set. Another set of water-oil relative permeability data from a separate field were used as the testing set after data wrangling and normalization. A seed value was set to ensure the repeatability of the model. An optimised number of hidden layers was used to reduce the need for feature engineering. The best cross validation result in a 5-fold arrangement was automatically used to train the DNN models until convergence using Nesterov accelerated gradient descent (which minimize their cost function). The Rectifier Linear Units (ReLU) were used in the DNN modelling to increase the nonlinearity of the model, significantly reduce the difficulty in learning, improve accuracy and can accept noise (Equation 1). This allows for effective training of the network on large and complex datasets making it helpful for real time applications compared to the commonly used sigmoid function which is difficult to train at some point.

$$f(x) = \max(0, x + Y) \quad (1)$$

Where $Y \sim \mathcal{N}(0, \sigma(x))$ is the Gaussian noise applied to the ReLUs.

Separate models was constructed for wetting and non-wetting phases as have also been found to improve predictions (Guler et al. 1999). They were then validated and tested to check the generalization and stability of the models for out of training sample applications.

Table 2 Comparison of relative permeability models (vertical) with different datasets (horizontal) using correlation coefficient (Modified after Baker 1988)

DATA	COREY	LEVERETT AND LEWIS	REID	SNELL	SARAF ET AL	HOSAIN	GUCK ERT
STONE I	0.97	0.76	0.90	0.57	0.82	0.85	0.48
STONE 11	0.77	0.75	0.87	0.75	0.68	0.33	0.50
AZIZ AND SETARRI	0.8	0.75	0.95	0.75	0.74	0.9	0.48
COREY	0.88	0.83	0.89	0.48	0.50	0.74	0.6
BAKER	0.96	0.86	0.88	0.58	0.9	0.84	0.57
NAAR AND WYGAL	0.74	0.67	0.78	0.50	0.55	0.54	0.50
PARKER	0.85	0.73	0.88	0.56	0.87	0.93	0.52
LAND	0.93	0.8	0.89	0.50	0.66	0.74	0.55
WYLLIE	0.91	0.89	-	-	-	-	-

The developed Deep Neural Networks model could further be applied to predict other experimental data carried out based on Buckley-Leverett (1942) frontal advance theory (Fig. 2) and Welge (1952) method for average water saturation behind the water front using the saturation history to make predictions of relative permeability as a function of time.

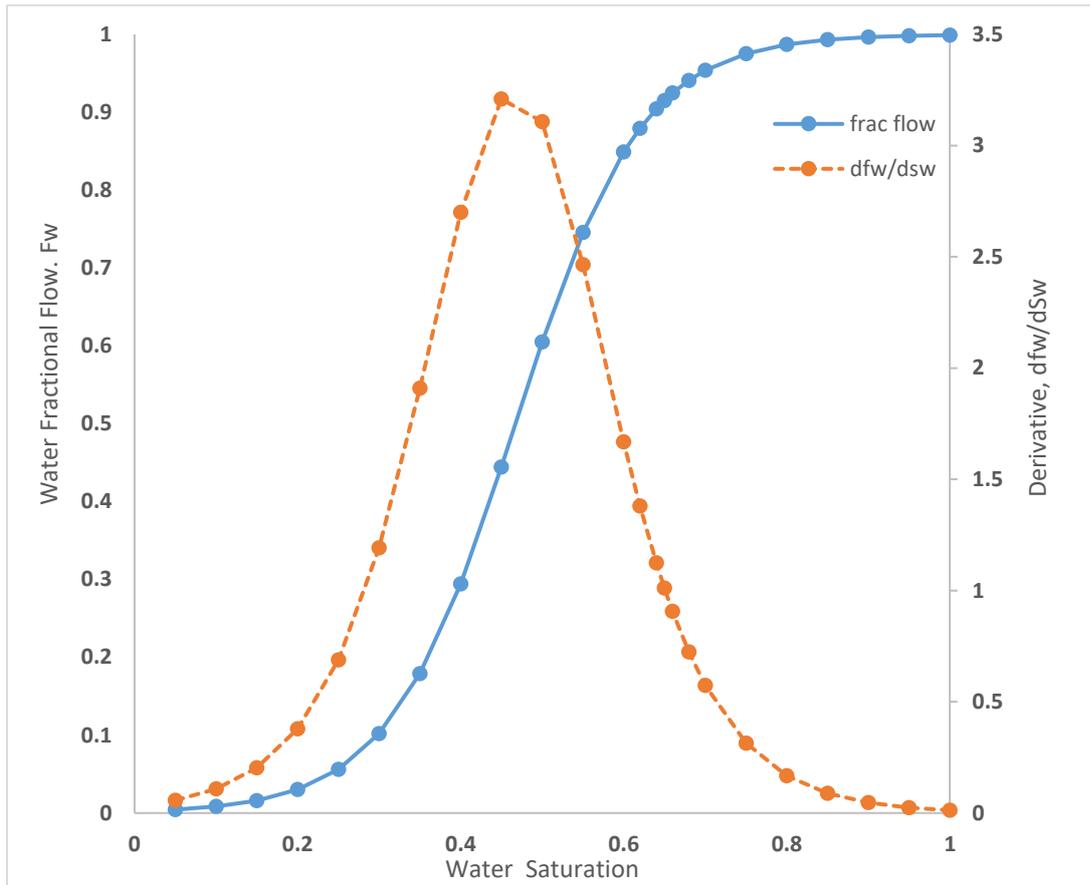


Fig. 2 Water fractional flow curve with its derivative for the field considered.

Deep Neural Networks

Deep Neural Networks (sometimes referred to as stacked neural network) is a feed-forward, artificial neural network with several layers of hidden units between its inputs and outputs. One hundred hidden layers with twelve neurons each (100, 12) were used in this work. The ability of the model to transfer to a new context and not over-fit to a specific context (generalization) was addressed using cross validation which is described in detail below. All networks were trained until convergence with Nesterov accelerated gradient descent which also minimizes the cost function. In addition, both λ_1 and λ_2 regularization (Equation 2) were used to add stability and improve the generalization of the model. This regularization ability was further improved by implementing dropout. A copy of the global models parameters on its local data is trained at each computed node with multi-threading asynchronously and periodically contributes to the global model through averaging across the network.

Mathematically,

$$J(\theta) = \frac{1}{2} \sum_{i=1}^n (\theta^T x^{(i)} - y^{(i)})^2 + \lambda \sum_{j=1}^p \theta_j^2 \quad (2)$$

Where x are inputs, θ are parameters, λ is a measure of complexity by introducing a penalty for complicated and large parameters represented as λ_1 or λ_2 (preferred to λ_0 for convexity reasons). They are well suited for modelling systems with complex relationships between input and output (Burke, 1992; Hubick, 1992) which is what is obtainable in natural earth systems. In such cases

with no prior knowledge of the nature of non-linearity, traditional regression analysis is not adequate (Gardner and Dorling, 1998). It has been successfully applied to real time speech recognition, computer vision, optimal space craft landing etc.

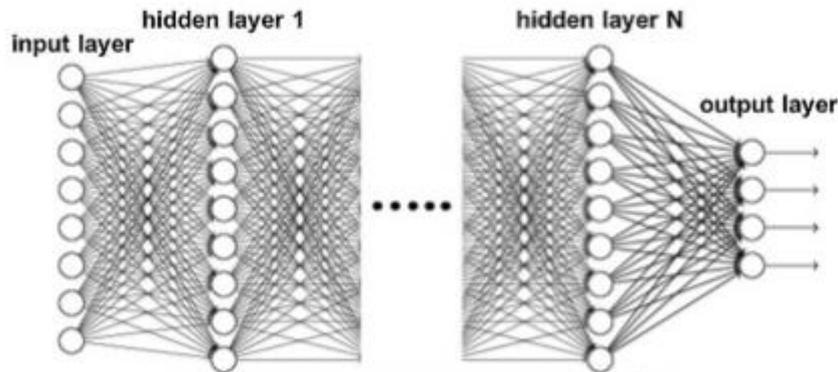


Fig. 3 Deep Neural Network model architecture showing input, hidden and output layers (Lee et al. 2017)

Cross Validation

Overfitting which is the single major problem of prediction when independent datasets is used was reduced through cross validation by estimating out of sample error rate for the predictive functions built to ensure generalisation. Other issues like variable selection, choice of prediction function and parameters and comparison of different predictors were also addressed. A 5-fold cross validation technique was used to split the data set into training and test set, build a model on the training set, evaluate on the test set and then repeat and average the errors estimated. A weight decay was chosen to improve the generalization of the model by suppressing any irrelevant component of the weight vector while solving the learning problem with the smallest vector. This also suppresses some of the effects of static noise on the target if chosen correctly and increase the level of confidence in the prediction.

Results and Discussion

Deep neural networks model have been validated using separate out of sample datasets not used for the training. The good agreement between experimental data and DNN's model predictions indicates that the complex, transient, non-linear behaviour of reservoir fluids can be effectively modelled as their saturation and phase changes with time.

Figs. 4, 5 and 6 give a comparison between actual experimental values and model predictions using neural networks without cross validation, neural networks with cross validation and the deep neural networks. The objective here was to see how Deep Learning out performs ordinary networks on new data. These cross plots show the extent of agreement between the laboratory and predicted values. For the testing set drawn from a different field from the training set, the Deep Neural Networks for both the wetting and non-wetting phase relative permeability (Fig. 6 b&d) gives very close values to the perfect correlation line in all data points compared to the other models. Fig. 4 a&c representing Neural Networks without cross validation, gave an RMS value of 0.2484 and 0.0767 while Neural net with cross validation gave an RMS of 0.0624 and 0.0765 (Fig. 5 a&c). The Deep Neural Net gave an RMS value of 0.2517 and 0.065 (Fig. 6 a&c) for

both wetting and non-wetting relative permeability. It is clear that all the models did well for the validation set although the deep neural networks performed better than the other two models. The different models were then shown new data from a separate field to see how they performed. For the test set (which is an out of sample dataset) obtained from a different field, the RMS for neural network without cross validation is 0.9996 and 0.8483 (Fig. 4 b&d), 0.2295 and 0.8022 with cross validation (Fig. 5 b&d) while DNNs gave 0.0759 and 0.15 (Fig. 6 b&d) for wetting and non- wetting relative permeability respectively.

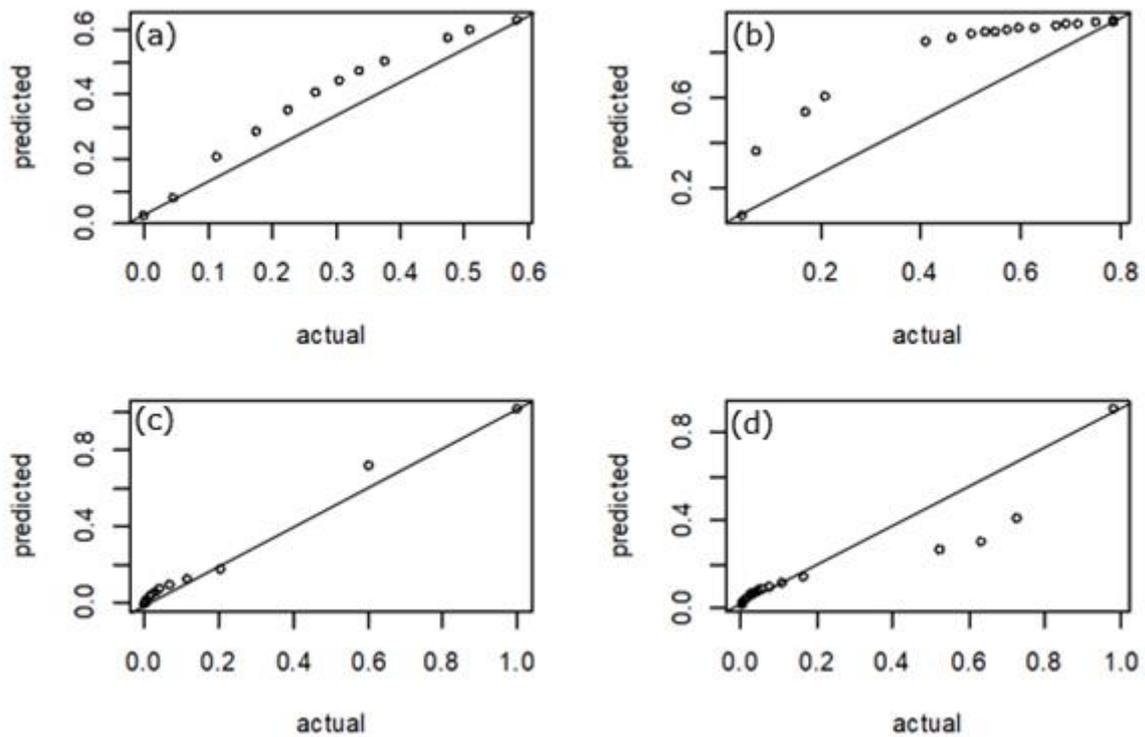


Fig. 4. Actual vs predicted value for neural networks without cross validation (cross validation not considered as part of the model formulation) with (a) wetting phase relative permeability for validation set (b) wetting phase relative permeability for test set (c) non-wetting relative permeability for validation set (d) non-wetting relative permeability for the test set.

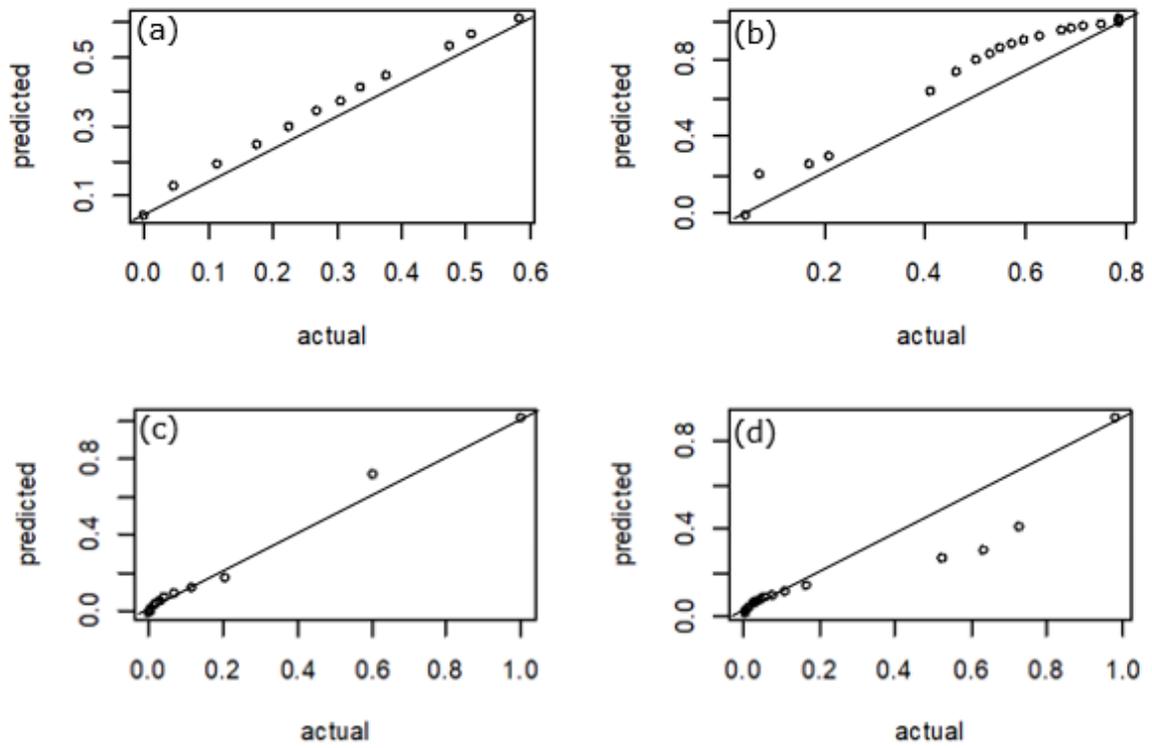


Fig. 5 Actual vs predicted value for neural networks with cross validation technique used for its model formulation and it improved prediction ability of the network with (a) wetting phase relative permeability for validation set (b) wetting phase relative permeability for test set (c) non wetting relative permeability for validation set (d) non-wetting relative permeability for the test set.

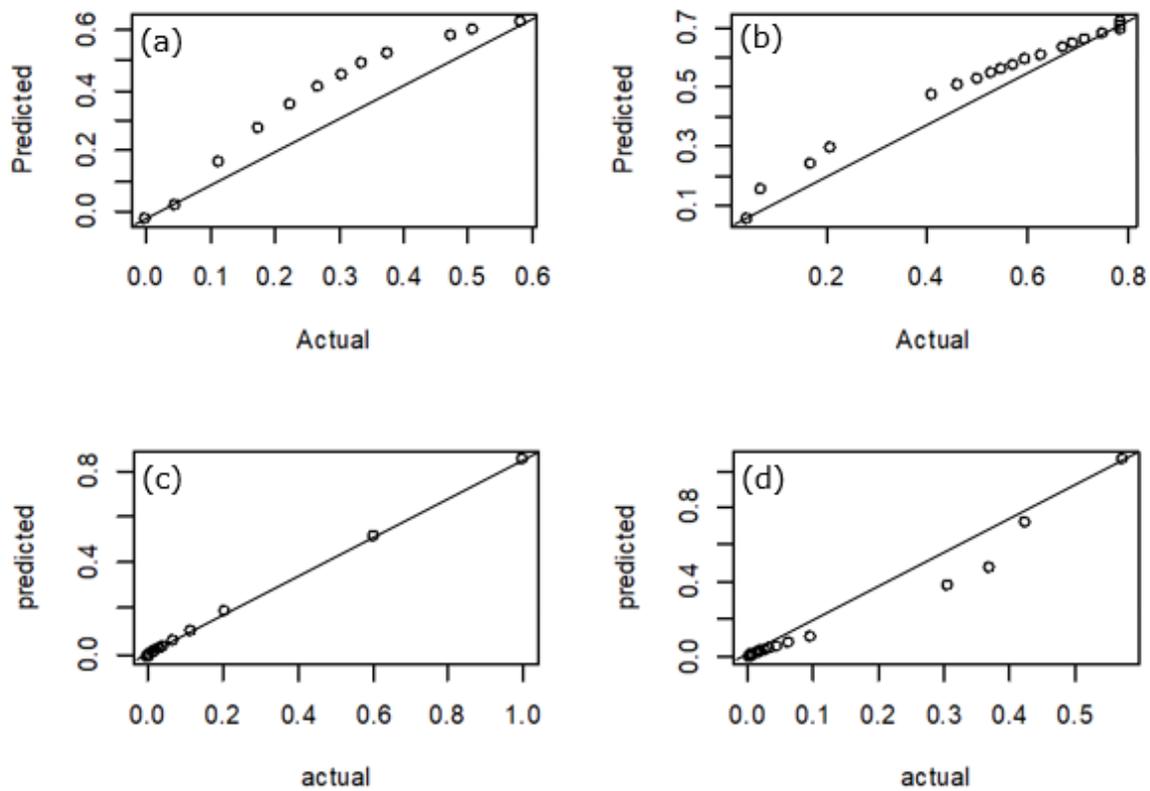


Fig. 6 Actual vs predicted value for Deep Neural Networks model with (a) wetting phase relative permeability for validation set (b) wetting phase relative permeability for test set (c) non wetting relative permeability for validation set (d) non-wetting relative permeability for the test set.

The deep learning model used the fourth cross validation model which happen to be the best for the wetting phase with a correlation coefficient of about 97% (Table 3) and the lowest training error of 0.0014 while the second cross validation model was used for the non-wetting phase relative permeability having 96% correlation coefficient and the lowest training error value of 0.030 (Table 4).

Table 3 Accuracy of the Deep Learning model for the wetting phase with cross validation for the five folds.

	mean	sd	5-Fold Cross Validation Results				
			1	2	3	4	5
mae	0.0489	0.0068	0.0558	0.0477	0.0612	0.0330	0.0468
mrdr	0.0052	0.0022	0.0053	0.0047	0.0108	0.0014	0.0038
mse	0.0052	0.0022	0.0053	0.0047	0.0108	0.0014	0.0038
r2	0.9259	0.0186	0.9121	0.9086	0.9018	0.9745	0.9325
rd	0.0052	0.0022	0.0053	0.0047	0.0108	0.0014	0.0038
rmse	0.0689	0.0150	0.0728	0.0684	0.1037	0.0380	0.0615
rmsle	0.0541	0.0130	0.0509	0.0558	0.0854	0.0277	0.0509

Table 4 Accuracy of the Deep Learning model for the non-wetting phase with cross validation for the five folds.

	mean	sd	5-Fold Cross Validation Results				
			1	2	3	4	5
mae	0.0470	0.0109	0.0633	0.0395	0.0593	0.0583	0.0521
mrd	0.0052	0.0019	0.0065	0.0038	0.0089	0.0075	0.0060
mse	0.0052	0.0019	0.0065	0.0038	0.0089	0.0079	0.0060
r2	0.9214	0.0217	0.8800	0.9636	0.9099	0.9043	0.9492
rd	0.0052	0.0019	0.0065	0.0038	0.0089	0.0065	0.0060
rmse	0.0690	0.0153	0.0805	0.0619	0.0941	0.0705	0.0774
rmsele	0.0489	0.0090	0.0641	0.0466	0.0578	0.0541	0.0492

Figs. 7 and 8 display the trend comparing the different models using the standard relationship between saturation and relative permeability. The Deep Learning model clearly out performs the other models giving better predictions for both the wetting and non-wetting phases. Measurement error which causes input values to differ if the same example is presented to the network more than once is evident in the data. This limits the accuracy of generalization irrespective of the volume of the training set. The Deep Neural Networks model deeply understands the fundamental pattern of the data thus able to give reasonable predictions than ordinary networks and empirical models (Figs. 9 and 10). The curves show that significant changes in the saturation of other phases has large effect on the wetting phase ability to flow as observed from the less flattening of the water relative permeability curve and vice versa for the flattened curve. Although this flattening behaviour is usual in the secondary drainage and imbibition cycles but mainly in the wetting phase when flow is mainly through small pore networks. Again, the curve flattening of the oil relative permeability curve could from experience be from brine sensitivity and high rates causing particle movements resulting to formation damage.

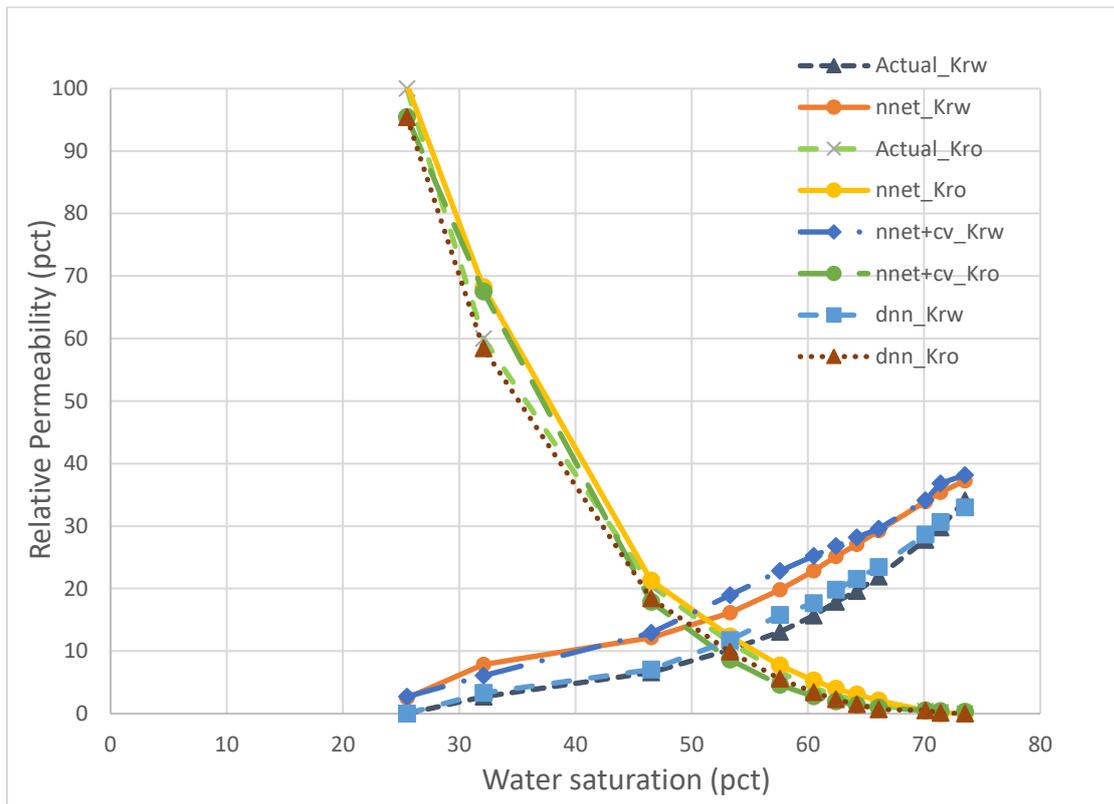


Fig. 7 Experimental and predicted relative permeability models using neural network with and without cross validation and deep neural networks on the validation set. The neural network model with cross validation (cv) partitioned the dataset into 5-fold and then trained and tested the model using the different folds.

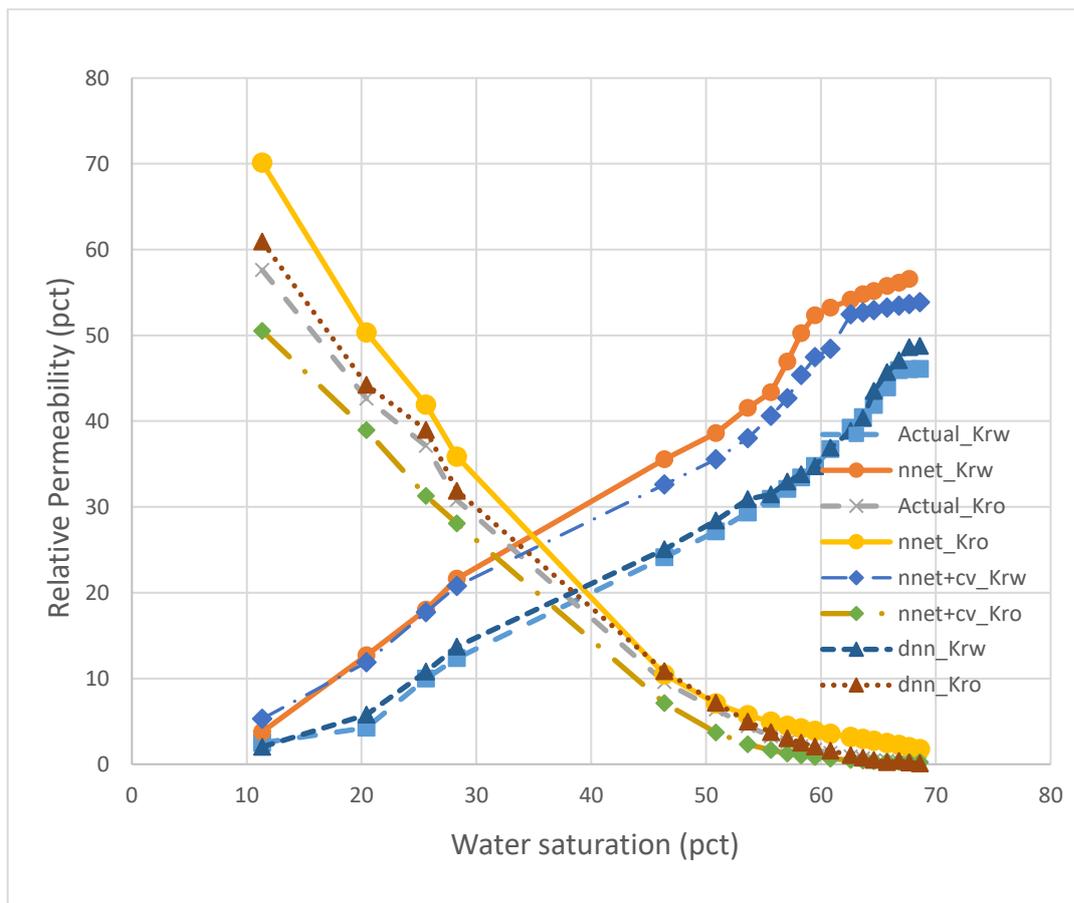


Fig. 8 Experimental (actual) and predicted relative permeability models using neural network (both with and without cross validation) and Deep Neural Networks on the out of sample test set (Stafford reservoir). Cross validation (cv) involved in the network helped to improve its accuracy for out of sample datasets.

Figs. 9 and 10 compares the Deep Neural Network model with commonly used empirical relative permeability models like Baker, Wyllie, Honarpour, Stones, Corey, Parker. Despite the fact that some of these models were developed using lots of datasets way more than the amount used for training the Deep Neural Networks, it still outperformed them showing that it is more able to capture the transients and eddies in real time scenarios due to its ability to regularize and generalize using its robust parameters as discussed earlier.

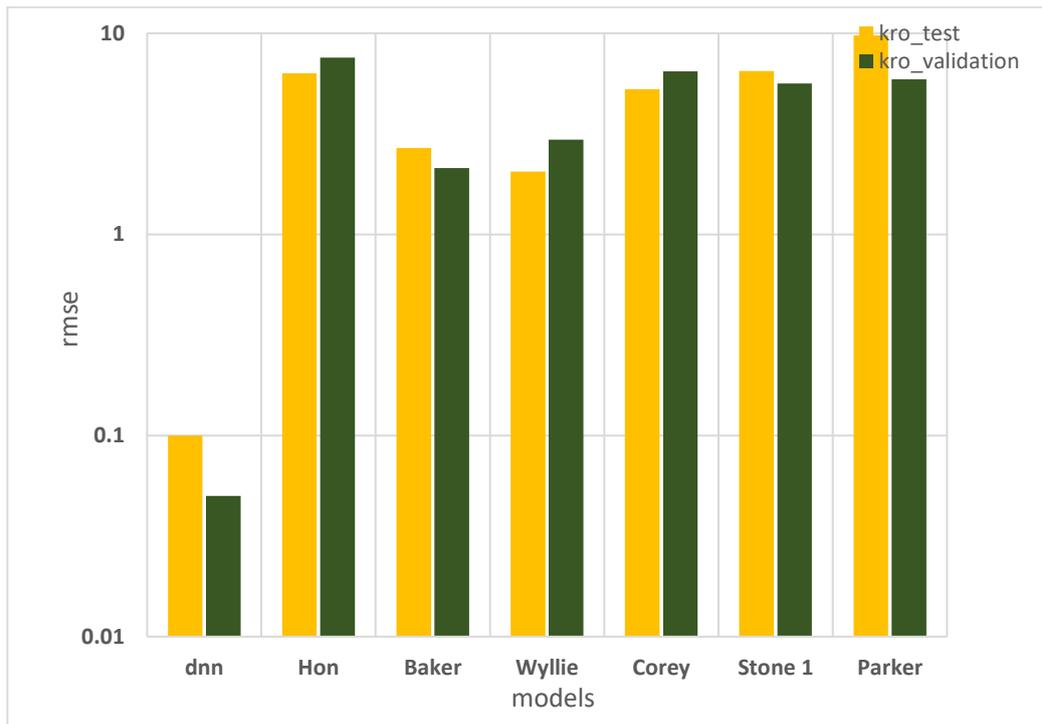


Fig. 9 Comparison of Wyllie, Corey, Parker, Stone, Baker, Honarpour, Deep Neural Networks for the Brent reservoir, North Sea. The DNN gave better prediction than the existing models for this validation set. Corey's λ , taken to be 2 and Parker's n parameter

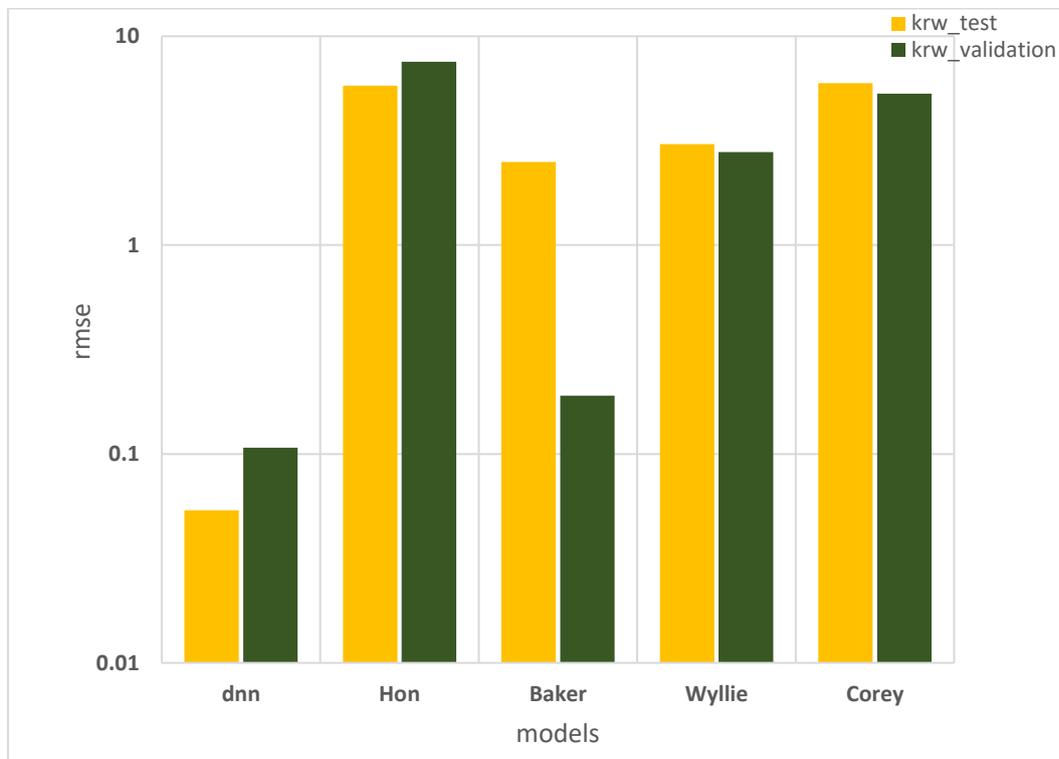


Fig. 10 Comparison of Wyllie, Corey, Baker, Honarpour, Deep Neural Networks models for the Stratford reservoir, NorthSea.

Figs. 11 and 12 corroborates the earlier observation that the Deep Learning model predicts better compared to most of the relative permeability models used in reservoir modelling software. It is important to note here that the empirical models (Figs. 9&10) have a problem of generalization especially as every reservoir is unique. Again, the assumptions associated with their formulation might not be practically true in all cases but this reservoir uniqueness or generalization is captured by the Deep Learning model bearing in mind that it will perform even better as more real time data is added to the training set.

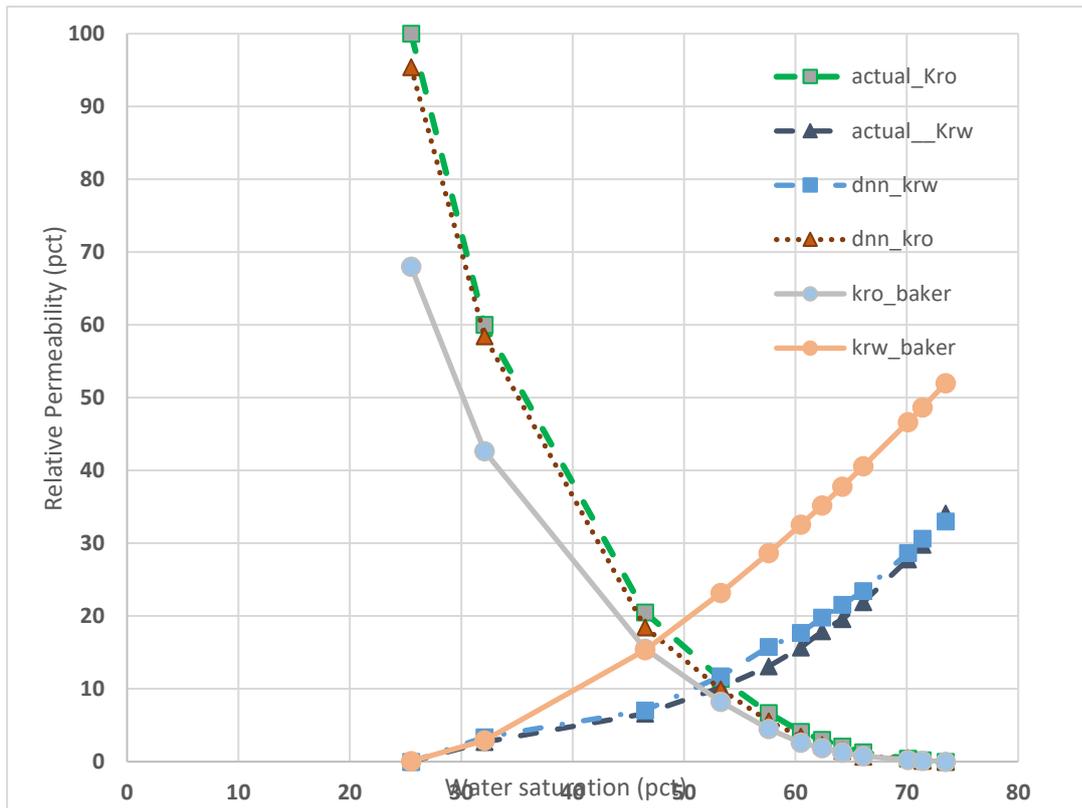


Fig. 11 Comparison of Deep Neural Networks and Baker with the measured wetting and non-wetting relative permeability models for the validation set (Brent reservoir).

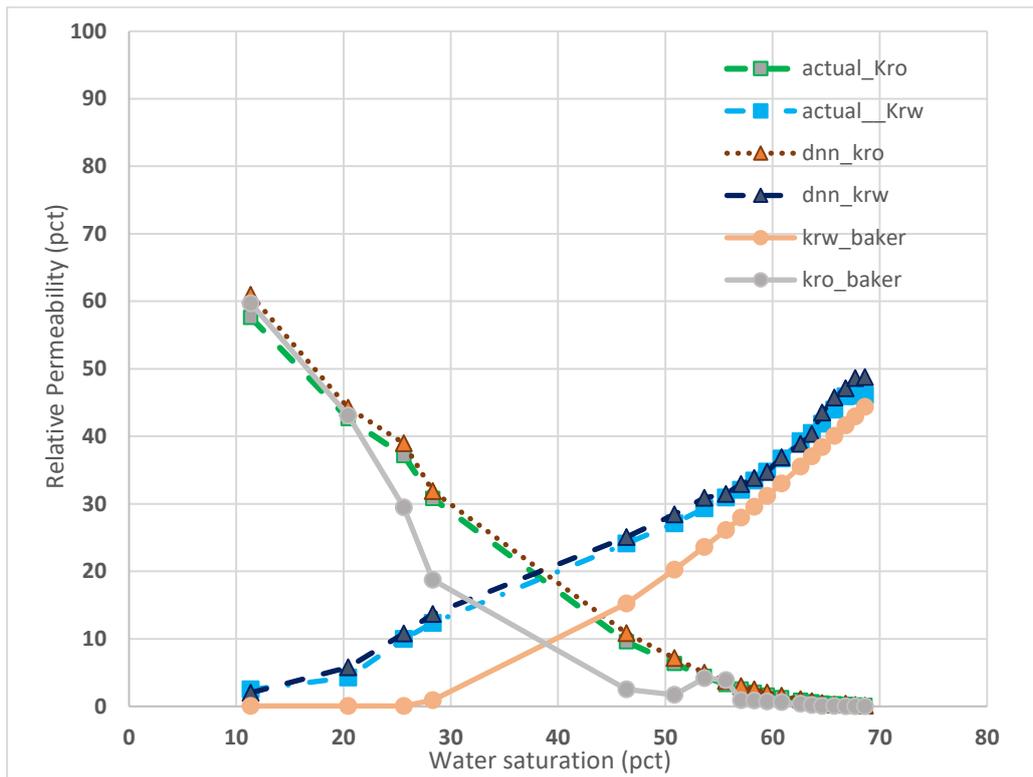


Fig. 12 Comparison of Deep Neural Networks and Baker with the wetting and non-wetting phase relative permeability models with for the test sets (Stratjford reservoir). Baker was used since it performed best among the models compared.

Fig. 13 and 14 describe the relative importance (sensitivity) of the variables used for the wetting and non-wetting Deep Learning relative permeability models. The wetting phase model was more sensitive to its saturation and relatively less sensitive to that of the non-wetting phase while the non-wetting phase model was very sensitive to both its saturation and that of the wetting phase. Both models were also more sensitive to their own viscosities than the other. These models seem to obey the basic physics underlying relative permeability modelling. The least important variable still contributed above the median mark although in general, all variables show greater sensitivity in the non-wetting model than in the wetting relative permeability model. Table 5 shows the performance of the different variables combinations for both the wetting and non-wetting phase model.

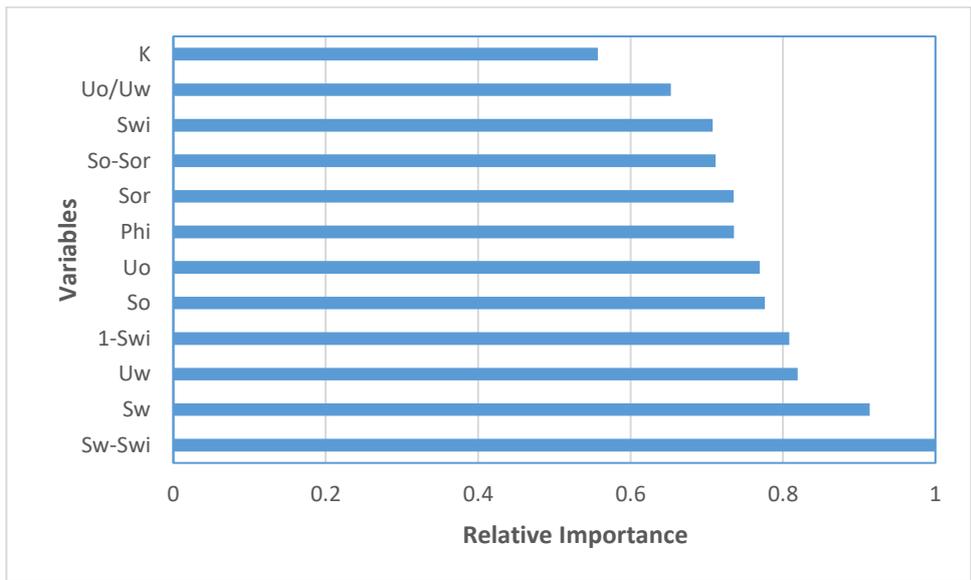


Fig. 13. Sensitivity analysis of individual variables used for building the wetting phase Deep Learning relative permeability model.

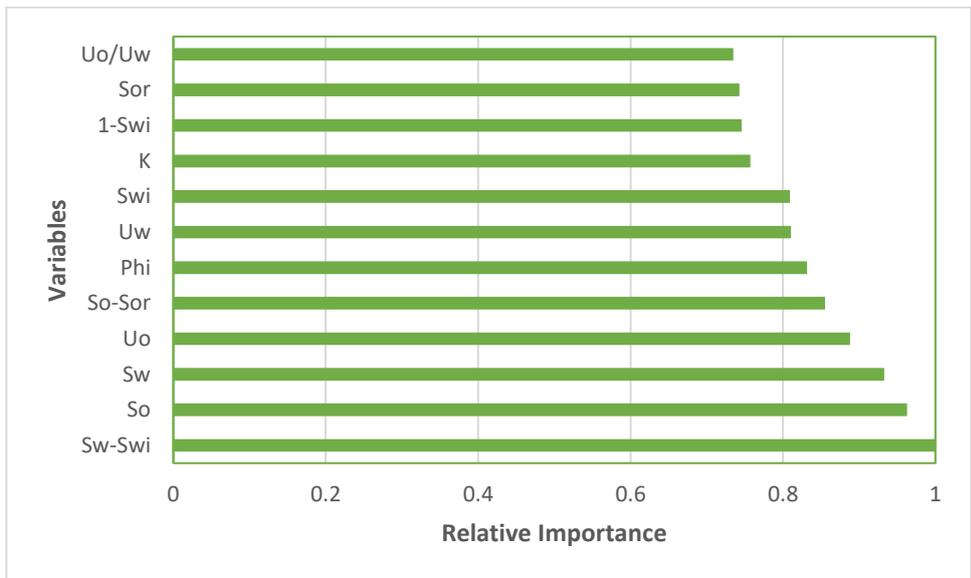


Fig. 14. Sensitivity analysis of individual variable used for building the non-wetting phase Deep Learning relative permeability model.

Table 5 Sensitivity analysis showing the importance of the different to both water and oil relative permeabilities

Cases	Input Parameters	Functional Links (From Baker and Wyllie)	Model Metric (RMSE, fraction)	
			K_{rw}	K_{ro}
1	S_w, S_o		0.1204	0.1532
2	S_w, S_o, S_{wi}		0.1201	0.1057
3	S_w, S_o, S_{wi}, S_{or}		0.1153	0.0712
4	$S_w, S_o, S_{wi}, S_{or}, k$		0.0906	0.0698
5	$S_w, S_o, S_{wi}, S_{or}, k, phi$		0.0705	0.0671
6	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o$		0.0616	0.0691
7	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$		0.0481	0.0681
8	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc})$	0.0463	0.0667
9	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc}), (S_o - S_{or})$	0.0449	0.0652
10	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc}), (S_o - S_{or}), (1 - S_{wc})$	0.0508	0.0732
11	$S_w, S_o, S_{wi}, S_{or}, k, phi, \mu_o, \mu_w$	$(S_w - S_{wc}), (S_o - S_{or}), (1 - S_{wc}), (\mu_o/\mu_w)$	0.0380	0.0619

Conclusion

A Deep Neural Network methodology has been formulated for wetting and non-wetting phase relative permeability predictions taking into account phase and saturation changes hence its capability for real time applications. This work has the following conclusions:

1. Deep Neural Network has shown to be a good predictive and prescriptive tool for relative permeability than ordinary networks. Its ability to generalize and regularize helped to stabilize and reduce the main problem of all predictive tools which is over fitting.
2. Different results were obtained from different relative permeability models for the same reservoir with some of the models giving better predictions at lower saturations but performs poorly at higher saturations and vice versa hence lots of uncertainty. Therefore, it is needful for practitioners to know the limitations of any correlation used for the prediction of wetting and non-wetting phase relative permeability.
3. In an industry where big data is now available, Deep Learning can provide the platform to systematically forecast reservoir fluid and rock properties in order to drastically optimize the cost and time needed for laboratory experiments. Even with the amount of data used, the power

of the Deep Neural Networks is evident in that it gave reasonable predictions which will dramatically improve if more data were available.

Nomenclature

K_{ro} = oil relative permeability

K_{rw} = water relative permeability

S_w = water saturation

S_{wc} = irreducible water saturation

S_o = oil saturation

cv = cross validation

val = validation dataset

dnn = deep neural network

σ = standard deviation

n = number of samples

References

AHMAD, T. 2001. *Reservoir engineering handbook*, Houston, TX: Gulf Professional.

AL-FATTAH, S.M. and AL-NAIM, H.A., 2009. Artificial-intelligence technology predicts relative permeability of giant carbonate reservoirs. *SPE Reservoir Evaluation & Engineering*, 12(01), pp. 96-103.

BAKER, L., 1988. Three-phase relative permeability correlations. *SPE Enhanced Oil Recovery Symposium*. Society of Petroleum Engineers.

BURKE, L.I. and IGNIZIO, J.P., 1992. Neural networks and operations research: an overview. *Computers & Operations Research*, 19(3-4), pp. 179-189.

COREY, A. et al., 1956. Three-phase relative permeability. *Journal of Petroleum Technology*, 8(11), pp. 63-65.

DAKE, L.P. 1978. *Fundamentals of Reservoir Engineering*, No. 8. Amsterdam: Developments in Petroleum Science, Elsevier Science BV.

DELSHAD, M. and POPE, G.A., 1989. Comparison of the three-phase oil relative permeability models. *Transport in Porous Media*, 4(1), pp. 59-83.

DONALDSON, E.C. and DEAN, G.W., 1965. Two-and three-phase relative permeability studies,

DURY, O., FISCHER, U. and SCHULIN, R., 1999. A comparison of relative non wetting-phase permeability models. *Water Resources Research*, 35(5), pp. 1481-1493.

DU YUQI, O.B. and DACUN, L., 2004. Literature review on methods to obtain relative permeability data.

FAYERS, F. and MATTHEWS, J., 1984. Evaluation of normalized Stone's methods for estimating three-phase relative permeabilities. *Society of Petroleum Engineers Journal*, 24(02), pp. 224-232

GARDNER, M.W. and DORLING, S., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14), pp. 2627-2636.

GULER, B., ERTEKIN, T., and GRADER, A. S., 1999. An Artificial Neural Network Based Relative Permeability Predictor. *Petroleum Society of Canada*. doi:10.2118/99-91.

HONARPOUR, M., KOEDERITZ, L. and HARVEY, A.H., 1982. Empirical equations for estimating two-phase relative permeability in consolidated rock. *Journal of Petroleum Technology*, 34(12), pp. 2,905-2,908.

HUBICK, K.T. and HUBICK, K.T., 1992. *Artificial Neural Networks in Australia*: Department of Industry, Technology & Commerce, pp. 71-120.

JUANES, R. et al., 2006. Impact of relative permeability hysteresis on geological CO₂ storage. *Water Resources Research*, 42(12), pp.

LI, K. and HORNE, R.N., 2006. Comparison of methods to calculate relative permeability from capillary pressure in consolidated water-wet porous media. *Water Resources Research*, 42(6), pp. 39-48.

LOMELAND, F., EBELTOFT, E. and THOMAS, W.H., 2005. A new versatile relative permeability correlation. *International Symposium of the Society of Core Analysts*, Toronto, Canada. pp. 1-12

MANJNATH, A. and HONARPOUR, M., 1984. An investigation of three-phase relative permeability. *SPE Rocky Mountain Regional Meeting*. Society of Petroleum Engineers.

PARKER, J. C., LENHARD, R. J., and KUPPUSAMY, T. 1987. A parametric model for constitutive properties governing multiphase flow in porous media. *Water Resources Research*, 23(4), pp. 618-624.

SARAF, D.N., BATYCKY, J.P., JACKSON, C.H. and FISHER, D.B., 1982. An experimental investigation of three-phase flow of water-oil-gas mixtures through water-wet sandstones. *SPE California Regional Meeting*. Society of Petroleum Engineers.

SARAF, D. and FATT, I., 1967. Three-phase relative permeability measurement using a nuclear magnetic resonance technique for estimating fluid saturation. *Society of Petroleum Engineers Journal*, 7(03), pp. 235-242

SARAF, D. and MCCAFFERY, F., 1985. Relative Permeabilities. *Developments in Petroleum Science*, 17, pp. 75-118

SCHNEIDER, F. and OWENS, W., 1970. Sandstone and carbonate two-and three-phase relative permeability characteristics. *Society of Petroleum Engineers Journal*, 10(01), pp. 75-84

SIDDIQUI, S., HICKS, P.J. and ERTEKIN, T., 1999. Two-Phase Relative Permeability Models in Reservoir Engineering Calculations. *Energy Sources*, 21(1-2), pp. 145-162

STONE, H., 1970. Probability model for estimating three-phase relative permeability. *Journal of Petroleum Technology*, 22(02), pp. 214-218

STONE, H., 1973. Estimation of three-phase relative permeability and residual oil data. *J.Pet.Technol.:(United States)*, 12(4)

WYLLIE, M., 1951. A Note on the Interrelationship between Wetting and Non-wetting Phase Relative Permeability. *Journal of Petroleum Technology*, 3(10), pp. 17-17