

FAST MODELLING OF GAS RESERVOIRS USING NON-INTRUSIVE REDUCED
ORDER MODELLING AND MACHINE LEARNING

JEMIMAH-SANDRA ADEBISI SAMUEL

**Imperial College
London**

Submitted for the degree of Doctor of Philosophy and the Diploma of Imperial College
London to the Department of Earth Science and Engineering Imperial College London
January 2022

ABSTRACT

This work focussed on developing approximate methods for rapidly estimating gas field production performance. Proper orthogonal decomposition (POD) - Radial basis function (RBF) and POD-Autoencoder (AE) Non Intrusive Reduced Order Models (NIROMs) were considered. The accuracy and speed of both NIROMs were evaluated for modelling different aspects of gas field modelling including reservoirs with time-varying and mixed production controls, reservoirs with and without aquifer pressure support, and for wells that were (or not) shut-in during production lifecycle. These NIROMs were applied to predicting the performance of four gas reservoir models: a homogeneous synthetic model; a heterogeneous gas field with 3 wells and structures similar to the Norne Field; a water coning model in radian grid; and a sector model of a real gas field provided by Woodside Petroleum.

The POD-RBF and POD-AE NIROMs were trained using the simulation solutions from a commercial reservoir simulator (ECLIPSE): grid distributions of pressure and saturations as well as time series production data such as production rates, cumulative productions and pressures. Different cases were run based on typical input parameters usually used in field performance studies. The simulation solutions were then standardised to zero mean and reduced into hyperspace using POD. In most cases, the optimum number of POD basis functions (99.9% energy criterion) of the solutions (training data) were used to reduce the training data into a lower-dimensional hyperspace space. The reduced training data and their corresponding parameter values were combined to form sample and response arrays based on a cause and effect pattern. RBF or AE was then used to interpolate the weighting coefficients that represented the dynamics of the gas reservoir as captured within the reduced training data. These weighting coefficients were used to propagate the prediction of new unseen simulation cases for the duration of predictions. The simulation results from either or both NIROMs was then compared against the simulation solution of the same cases in ECLIPSE.

It was found that the POD-RBF is a better predictive tool for gas field modelling. It is faster, more accurate and consistent than the POD-AE, giving satisfactory predictions with up to 99% accuracy and 2 orders of magnitude speed-up. No single POD-AE is sufficient for predicting different production scenarios, besides, the process of arriving at a suitable POD-AE involves finetuning several hyper-parameters by trial and error, which may be more burdensome for practising petroleum engineers.

The accuracy of NIROM's prediction of production variable is generally improved by using more than the optimal number of POD-basis functions, while predictions of grid distributed properties are satisfactorily predicted with the optimal number of POD-basis functions. NIROM's accuracy is dependent on whether the range of parameters of the prediction, their duration and specific production scenarios (such as having mixed production controls or aquifer pressure support) reflect

those of the training cases. Overall, the number of training runs, the size of the reservoir model as well as the number of time intervals at which simulation output data is required all affect the speed of training both NIROMs for prediction.

Other contributions of this work include showing that the linear RBF is the most suitable RBF for gas field modelling; developing a novel normalisation approach for time-varying parameters; and applying NIROMs to seasonally varying production scenarios with mixed production controls. This work is the first time that the POD-AE has been developed and evaluated for petroleum field development planning.

STATEMENT OF ORIGINALITY

I hereby certify that the work presented in this dissertation/ thesis is the result of my investigations conducted either wholly by myself or partially, along with collaborators during the PhD project. Text and results obtained from other sources are referenced and properly acknowledged.

London, January 2022

Jemimah-Sandra Adebisi Samuel

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

London, January 2022

Jemimah-Sandra Adebisi Samuel

This thesis is dedicated to God, my Creator and Father

... for in Him we live and move and have our being... Acts 17:18.

ACKNOWLEDGEMENTS

First, my immense gratitude goes to Professor Ann Muggeridge for considering me for this project and for the delightful supervisory experience. Thank you, Woodside Petroleum (Ian Sylvester, Lupo Guerrero, Ben Sellers and James), for sponsoring this project, your regular visits and sharing useful perspectives that helped to shape this project. I appreciate the advice, valuable insight and contributions of Professor Chris Pain, Professor Matt Jackson and Professor Peter King during my research milestones and seminar presentations. I also appreciate Professor Michael Christie and Professor Peter King for examining this thesis and their helpful input which has improved this thesis.

I am grateful to Larry Kostorz for the numerous discussions on MATLAB and NIROM, they gave me an invaluable head start in this project. Toby Philips, César Quilodran Casas, Navjot Kukreja, Rossella Arcucci and other members of the data science institute, I appreciate all the useful discussions on machine learning and modelling. I also thank Schlumberger and Mathworks for the use of their commercial software in this work, ECLIPSE and MATLAB respectively.

Thank you, Victoria Murphy, Emma Watson, Amanda Allotey and other support staff, at the Department of Earth Science and Engineering (ESE) for hosting activities and events that helped me socialise and contribute to the ESE community; Ruth Davies for being my PhD welcomer and setting the pace for my participation in ESE activities and events. Thank you, Jude Osamor, Maryam Abdulkarim, Lekan Ladipo, Zainab Titus, my Nigerian Imperial Postgraduate community and GradSoc committee members from 2018 to 2020.

Olalekan Keshinro and all my family members, I am obliged by all your love and support.

TABLE OF CONTENTS

Abstract.....	2
Statement of Originality.....	4
Copyright Declaration	5
Acknowledgements.....	7
Table of Contents.....	8
List of Figures	12
List of Tables	31
Chapter 1: Introduction	33
1.1 Motivation.....	33
1.2 Aim and Objectives	37
1.2.1 Aim	37
1.2.2 Objectives.....	37
1.3 Contributions	38
1.4 Overview of Thesis	38
Chapter 2: Literature Review	40
2.1 Modelling Gas Reservoir Performance	40
2.2 Fast Modelling Methods	46
2.3 Machine Learning Methods	54
Summary	58
Chapter 3: Theory and Implementation of NIROM	60
3.1 Overview of Approach	60
3.2 Training, Validation and Test Data	62
3.3 Processing of Training Data using POD	65
3.4 Parameterisation.....	68
3.5 Modelling Dynamics.....	71
3.5.1 Radial Basis Functions (RBFs)	72
3.5.2 Autoencoder	74
3.6 Simple Example of NIROM Workflow	76
3.7 Performance Measures.....	80
3.8 Summary	81
Chapter 4: Numerical Examples/ Test Cases.....	82
4.1 Synthetic Gas Reservoir Model.....	82
4.2 Heterogeneous Gas Reservoir	85
4.3 Water Coning Reservoir	88

4.4	Field X.....	90
Chapter 5: Training Model Results and Discussion.....		93
5.1	Synthetic Gas Reservoir	93
5.1.1	Production Controls	94
5.1.2	Aquifer models.....	99
5.1.3	Grid Refinement.....	108
5.2	Modified Norne Field	111
5.3	Interpolation Function Analysis and Error Measures	115
5.4	Time Analysis	117
Summary		118
Chapter 6: Constant Parameter Predictions - Results and Discussion		120
6.1	Synthetic Gas Model	120
6.1.1	Rate Controlled Production	121
6.1.2	BHP Controlled Production	124
6.1.3	Uncertain Constant Porosity Case	126
6.1.4	Initialization by Enumeration.....	128
6.2	Heterogeneous Gas Reservoir (Modified Norne Field).....	131
6.2.1	Rate Controlled Production	131
6.2.2	BHP Controlled Production	132
6.3	Water Coning Gas Reservoir	135
6.3.1	Rate Controlled Production	135
6.3.2	BHP Controlled Production	139
6.4	Field X.....	140
6.5	Error Analysis	145
6.6	Time Analysis	148
6.6.1	Synthetic Gas Reservoir	148
6.6.2	Modified Norne Gas Field and Water Coning Gas Reservoir	151
6.6.3	Field X.....	153
6.6.4.	Analysis of Simulation speedup	154
Summary		155
Chapter 7: Varying Parameter Predictions - Results and Discussion.....		157
7.1	Synthetic Gas Reservoir Model.....	157
7.1.1	Seasonally Varying Production Profile	157
7.1.2	Triennially Varying Production Profile	164

7.2	Modified Norne Gas Field	167
7.3	Field X.....	173
7.3.1	BHP Limit.....	173
7.3.2	THP Limit	179
7.4	Error Analysis	185
7.5	Time Analysis	188
7.5.1	Synthetic Gas Reservoir	188
7.5.2	Modified Norne Gas Field	190
7.5.3	Field X - Time Analysis.....	191
7.5.4.	Analysis of Simulation speedup	192
	Summary	193
	Chapter 8: POD-AE Results and Discussion.....	195
8.1	Cases with Shut-in Wells	196
8.2	Cases without Shut-in Wells	200
8.3	Error Analysis	204
8.4	Time Analysis	206
	Summary	207
	Chapter 9: Conclusions and Recommendations for Future Work	210
9.1	Specific Findings.....	211
9.2	General Findings	213
9.3	Non-Intrusive Reduced Order Model Workflow.....	214
9.4	Limitations and Recommendations for Future work.....	216
9.4.1	Limitations.....	216
9.4.2	Recommendations	217
	Nomenclature	219
	References	222
	Appendices.....	230
	Appendix A: Other Renormalisation Methods.....	230
A.1	Modified Normalization Method	230
A.2	Using a Proxy as Parameter	231
A.3	Constraining the Temporal Resolution with a Proxy Parameter	231
A.4	Results and Discussion	232
A.4.1	Modified Normalization Approach And Constrained Temporal Resolution	232
A.4.2	Boost and Sharpness of Turn Correlation Analysis	233

Appendix B: More POD-AE Results and Other Machine learning methods	234
B.1 POD-AE Results with 10 and 20 Training Runs.....	234
B.2 POD-AE with Dropout Layers	238
B.3 Other ML models considered in this work.....	239
B.3.1 Feed-Forward Neural Network (FFNN),	239
B.3.2 Long Short Term Memory (LSTM) Model	241
Appendix C: Reservoir Data files and NIROM Models	242
C.1 Reservoir Models	242
C.1.1 Synthetic gas Model with Fetkovich Aquifer	242
C.1.2 Modified Norne Gas Field Model For The Seasonally Varying Production Case	245
C.1.3 Water Coning Gas Reservoir Model.....	250
C.2 NIROM Models.....	253
C.2.1 POD-RBF Model with Seasonally Varying Production Profile	253
C.2.2 Autoencoder (AE) Model with 2HLs.....	260

LIST OF FIGURES

Figure 1.1: Global energy mix and forecast according to bp's rapid scenario ^[4]	33
Figure 1.2: Global CO ₂ emissions by fuel type per year since 1900 ^[2] . Gas emits less CO ₂ than Oil and Coal in billion tonnes.....	34
Figure 1.3: Gas consumption and production from 1999 to 2019. There is a consistent annual increase in the production and consumption of gas, except for 2009 ^[3, 6]	34
Figure 1.4: Comparison of gas consumption in the US since 2010 with the actual consumption during that period. In agreement with Harrison et al, the actual gas consumption according to EIA and bp exceeded EIA's 2003 forecast. But based on polynomial trend line projections, the actual current gas consumption is likely to exceed Harrison's generous forecast by 2020 ^[1, 3, 7, 8]	35
Figure 2.1: Havlena and Odeh's diagnostic material balance plot for determining the initial gas in place and the drive mechanism of a gas reservoir. The shape of the curve is an indication of the aquifer strength ^[11]	42
Figure 2.2. The linearity of p/Z plots affected by varying production rates a) in high and low production rate cases b) in a seasonally varying production case ^[11]	44
Figure 2.3: Forecasts of oil production rate in a test case of He and Durlofsky et al. (2014) using the POD-TPWL method. Their POD-TPWL is an intrusive ROM that relies on some modification of the reservoir simulator ^[32]	48
Figure 2.4: Jiang & Durlofsky (2019) ^[35] box plots comparing the errors obtained from POD-TPWL, POD-Only and POD-GNAT for 495 original test schedules (Δu) of their test case and cases with large perturbation schedules ($3\Delta u$), showing that GNAT gave better results than TPWL for predictions of oil and water rates.....	49
Figure 2.5: Finite difference stencil for Chen and Klie (2013) black-box stencil interpolation method ^[56] . Their BSIM exploits a similar discretisation approach as most simulators but was not tested for sufficient simulation scenarios (size and duration).	51
Figure 2.6: Xiao 2016 ^[38] POD-RBF estimation of the saturation in a two-material layers case showing solutions with 6, 12 and 18 POD bases functions.....	52
Figure 2.7: Kistorz et al.'s (2019) ^[37] breakthrough time (BT) PDFs generated from their simulator (Exact) and NIROM. Note that in both cases above, the NIROM's prediction of the peak BT PDFs are different, the case with fewer training runs is closer to the exact solution.	53
Figure 2.8: Kistorz et al.'s (2019) ^[37] NIROM approximations of a simple reservoir's P10 saturation distribution. Kistorz et al. showed that having larger numbers of training runs does not guarantee having better NIROM approximations.....	53
Figure 2.9: Effect of a hyper-parameter (number of hidden layers) on a NN's generalisation ^[111]	55

Figure 2.10: Temirchev et al.'s (2020) ^[13] comparison of different metamodels' prediction of the average values of the spatial pore pressure and oil saturation for a scenario of their test case. In this example, GRU means Gated Recurrent Unit, PCA means Principal Components Analysis, VAE means Variational Autoencoder, LIN means Linear Regression and LSTM means Long Short Term Memory.	57
Figure 3.1: Difference between (a) constant parameter distribution and (b) forced parameter distribution in the temporal resolution (x-axis). The x-axes (report steps) are in 300days interval. For constant parameters, production rates remain the same value in the temporal resolution while for forced parameters, the production rates vary in the temporal resolution.	63
Figure 3.2: A representation of training and prediction data stacked together according to their parameters in hyperspace. a) A set of training runs (in cyan, green and red colour) (b) a set of training runs and prediction runs (in magenta and black colour) which have been interpolated between the training runs ^[100]	71
Figure 3.3: Summary of POD-AE model workflow showing the hyperparameters applied to the model. The first and last steps are based on proper orthogonal decomposition projection and back-projection, while the middle steps are the autoencoder model with 5 layers for the most satisfactory cases.	76
Figure 3.4: Overview of the model workflow for POD-RBF and POD-AE NIROM presented in this work. The workflow components in blue colour are common to both models, while POD-RBF only related components are in orange colour and POD-AE only related components are in green colour.	81
Figure 4.1: Grid refinement cases for the synthetic gas reservoir. (a) Local grid refinement in the producer's grid cells and (b) Tartan grid in the horizontal layers of the reservoir.	84
Figure 4.2: Gas saturation distribution in the modified Norne Gas field with strong aquifer support and 3 producers.	85
Figure 4.3: An example of the intended versus actual production profiles for training the modified Norne gas field where production rates vary after at least 1000 days. The intended production profile stipulates the production rate profiles that are generated from MATLAB (a combination of unique rates and durations), and the conventional simulator (ECLIPSE) is expected to follow for two wells over their entire production lifecycle. However due to the well BHP limit, ECLIPSE is unable to maintain the intended production profile, and has switched to BHP production control during the wells' production life. BHP controlled production commences at about the 45th and 18th report steps for wells 1 & 2 respectively (yellow and purple circle lines). The production profile that ECLIPSE eventually used for the simulation is the actual production profile.	88
Figure 4.4: An example of the intended and actual production profiles for training the modified Norne gas field where production rates are varied every quarter of the year. Well controls switch from gas	

rate to BHP when ECLIPSE cannot maintain the intended production rate that was created in MATLAB. Eclipse actual production profile shows that well 3 was shut-in after the 230th report step (light blue circle-line).....	88
Figure 4.5: Gas saturation distribution in a gas reservoir model with significant water coning, where blue is low gas saturation and red is high gas saturation. The reservoir has one producer, completed at the end of the gas zone of the reservoir and just before the water (deep blue) zone of the reservoir.	89
Figure 4.6: The distribution of (a) porosity and (b) permeability in the horizontal direction of Field X. It is not clear why the relationship between the porosity and horizontal permeability in Field X is not directly proportional as most literature/ correlations show, i.e. the zones with the highest porosity have median permeability and the zones with the lowest porosity have the highest permeability values. Overall the porosity and permeability of Field X are generally high.	91
Figure 4.7: Gas saturation distribution in gas Field X produced with strong aquifer support.....	91
Figure 5.1: Pressure distribution across each layer of the synthetic gas reservoir at $t = 1800$ days. This reservoir has one well which is drained from the centre of the reservoir at a constant production rate of 6.77MMscf/day. The reservoir pressure is depleted the most at the well location and then radially outwards to other grid cells in the reservoir. Pressure depletion in the reservoir is also from top to bottom, such that lower layers are less depleted and have higher pressure than the top layers. The NIROM (using any of the 7 RBFs) can reproduce the radial and downward pressure depletion pattern seen in the synthetic gas reservoir with negligible errors in the NIROM results compared to ECLIPSE.	94
Figure 5.2: Pressure distribution in the first layer of the synthetic gas reservoir at $t = 1800$ days (in psi). NIROM adequately captures the radial and outward pressure depletion in the synthetic gas reservoir.	95
Figure 5.3: The average of the reservoir pressure distribution estimated by the NIROM at each report time in the synthetic gas reservoir at constant production rate control. The NIROM gives a satisfactory match to ECLIPSE with negligible errors. The same result is obtained with all 7 RBFs using 1 POD basis function (as suggested by the energy criterion).	95
Figure 5.4: Pressure distribution in all grid cells of the synthetic gas reservoir, normalised to zero mean and using 1 POD basis function (as suggested by the energy criterion). The normalised pressure in all 2394 grid cells of the reservoir are plotted as vectors for each of the 25 report steps. Every 399 cells represent a single layer of the reservoir, so the well is located in the middle of every 399 cells. Similar results were obtained from all 7 RBFs, showing that the reservoir pressure decreases during constant	

production rate control. The most errors were experienced in the reservoir layers hosting the production well, as seen by the darker streaks in the middle of each layer.	96
Figure 5.5: Bottom hole pressure and cumulative gas production profiles estimated by the POD-RBF NIROM in comparison to ECLIPSE using 1 POD basis function, as suggested by the 99.9% energy criterion (a & b). Errors in the estimation of the c) well bottom hole pressure were less than 80psi, while errors in the d) cumulative gas production were less than 6MMscf (0.01%) for all 7 RBFs.	96
Figure 5.6: Bottom hole pressure and cumulative gas production profiles estimated by the NIROM and predicted by ECLIPSE using 2 POD basis functions (a & b). Adding an extra basis function to the number of basis functions suggested by using the energy criterion gave negligible errors in c & d, for both the decline in bottom hole pressure and the increase in the cumulative gas production. Similar results were obtained with all 7 RBFs.....	97
Figure 5.7: The average of the synthetic gas reservoir pressure distribution estimated by the NIROM at each report time when the well was produced by BHP control. These results were obtained using 1 POD basis function (as predicted by the energy criterion). All RBFs (except RBF 6) gave consistent results with ECLIPSE, RBFs 1 and 5 gave the most accurate results.	98
Figure 5.8: Reservoir production data estimated using 1 POD basis function (as suggested by the energy criterion). The NIROM results from RBFs 1, 3, 5, 6 and 7 are consistent with those of ECLIPSE for the production duration of 3300 days (11 report steps) at constant production BHP limit while RBFs 2 and 4 gave satisfactory results for only the first 2400 days.	98
Figure 5.9: Reservoir production data estimated using 2 POD basis functions. By adding an extra basis function, NIROM results are slightly improved for all RBFs except RBF 6 which explodes, with overestimation. RBFs 2 and 4 still give satisfactory results up to 2400 days after which they over-predict the gas production rate and under predict the cumulative gas production.	99
Figure 5.10: The average of the reservoir pressure distribution estimated by the NIROM at each report time, obtained using 1 POD basis function (as estimated by the energy criterion) in a BHP controlled gas well, supported by a numerical aquifer. Aquifer support maintained the average reservoir pressure above the 1000psi BHP limit, resulting in continued gas production for 7500 days. All RBFs gave the same estimates of pressure distribution.	100
Figure 5.11: The variation over time of the average gas saturation distribution in the synthetic gas reservoir, predicted by NIROM using 3 POD basis functions (the predicted energy criterion). The reservoir is produced at a BHP limit of 1000psi with numerical aquifer support. All 7 RBFs give satisfactory match of the gas saturation decline.....	100

Figure 5.12: Gas Saturation distribution in the first layer of the synthetic gas reservoir at $t = 6300$ days. NIROM adequately captured the pattern of gas depletion in the synthetic gas reservoir with a numerical aquifer.....	101
Figure 5.13: Gas saturation distribution in all grid cells of the synthetic gas reservoir with a numerical aquifer. The error in NIROM estimates were less than 1% of the original solutions from ECLIPSE. Although most of the reservoir was already depleted by 6300 days, NIROM adequately captured the saturation distribution across the 6 layers of the synthetic gas reservoir.	101
Figure 5.14: Production profiles using 3 POD basis functions (the energy criterion) for the synthetic gas reservoir supported by a numerical aquifer. All RBFs except RBF 6 give satisfactory NIROM estimates of (a) gas production rate (b) cumulative gas production (c) water production rate (d) cumulative water production (f) aquifer influx rate (g) cumulative aquifer influx and (h) aquifer pressure. None of the RBFs gave an accurate estimate of (e) the water gas ratio in the reservoir.....	102
Figure 5.15: Error plot of production data from RBFs 1 - 5 and 7 using 3 POD basis functions. All RBFs give errors that are at least an order of magnitude lower than the actual data for all production profiles, except the water-gas ratio with almost 50% error.....	103
Figure 5.16: Error plot of production data from all RBFs using 4 POD basis functions (energy criterion = 3 POD basis functions). RBFs 1 - 5 and 7 give lower error values by adding an extra basis function to the NIROM simulation. Estimation of the water-gas ratio still shows the most significant error.....	104
Figure 5.17: NIROM's satisfactory estimation of water gas ratio using 7 POD basis functions (energy criterion = 3 POD basis functions). RBFs 1 - 5 and 7 give the same result.....	104
Figure 5.18: Average of the reservoir pressure distribution at each report step of the synthetic gas reservoir with a Carter Tracy aquifer for all RBFs when using 2 POD basis functions (the energy criterion). This analytical aquifer supports reservoir pressure with less than 100 psi pressure depletion throughout production.	105
Figure 5.19: Pressure distribution in all grid cells of the synthetic gas reservoir with a Carter Tracy aquifer. Reservoir pressure values were normalised to zero mean and the NIROM was implemented with 2 POD basis functions (as suggested by the energy criterion). The most visible errors in the NIROM predictions were experienced in the reservoir layers hosting the production well, as seen in the darker streak lines in the middle of each layer.	105
Figure 5.20: Average of the gas saturation distribution at each report time, obtained from the NIROM using 2 POD basis functions for the synthetic reservoir with a Carter Tracy aquifer. Only RBFs 1, 3 and 6 gave reasonable estimates of the gas depletion in the reservoir.....	106
Figure 5.21: Estimated reservoir production profiles using (a & b) 4 POD basis functions and (c) 5 POD basis functions (the energy criterion suggested 3 POD basis functions would be sufficient), for all 7	

RBFs. These production data required the most number of POD basis functions to obtain reasonable NIROM estimates. Two of these three production data are water-related.	106
Figure 5.22: Average reservoir pressure distribution in the synthetic gas reservoir with a Fetkovich aquifer using 1 POD basis function (99.9% energy criterion). All 7 RBFs gave similar satisfactory estimates with negligible error.	107
Figure 5.23: Pressure distribution in all grid cells of the synthetic gas reservoir with a Fetkovich aquifer. Reservoir pressure values were normalised to zero mean and the NIROM was implemented with 1 POD basis function (as suggested by the energy criterion).	107
Figure 5.24: Average gas saturation distribution in the synthetic gas reservoir with a Fetkovich aquifer using 3 POD basis functions (99.9% energy criterion). RBFs 1 - 6 give better results than RBF 7 especially after 22 report steps, where RBF 7 gives haphazard results.	108
Figure 5.25: NIROM estimate of the water gas ratio of the synthetic gas reservoir with a Fetkovich aquifer using 7 POD basis functions (energy criterion suggested that 3 POD basis functions were required).	108
Figure 5.26: Pressure distribution in the first layer of the synthetic gas reservoir at 1800 days, using 1 POD basis function. Pressure distribution in a) the global grid cells only, b) the local grid cells only and c) the global as well as local grid cells. The local grid cells emphasise the pressure distribution in the well's location.	109
Figure 5.27: Pressure distribution in the global and local grid cells of the synthetic gas reservoir at 1800 days, using 1 POD basis function and across all six layers. NIROM satisfactorily mimics ECLIPSE's solutions for local grid refinement.	110
Figure 5.28: Pressure distribution across the six layers in the synthetic gas reservoir with the tartan grid at 1800 days, using 1 POD basis function. NIROM satisfactorily mimics ECLIPSE's solutions from the tartan grid.	111
Figure 5.29: The average of the estimated pressure distribution at different report times in the Norne gas field as predicted by the POD-RBF NIROM and ECLIPSE simulation. The NIROM used 6 POD basis functions (as indicated by the energy criterion). RBFs 1, 2, 4, 5 and 7 mimic ECLIPSE with negligible error.	112
Figure 5.30: The average of the gas saturation distribution at different report times obtained by NIROM using 7 POD basis functions (as suggested by the energy criterion). Compared with ECLIPSE's results, the NIROM satisfactorily matches the gas saturation depletion pattern in the heterogeneous modified Norne field for RBFs 1, 2, 4, 5 and 7. RBFs 3 and 6 deviated from ECLIPSE's results after less than 1000 days.	112

Figure 5.31: Pressure distribution in all grid cells of the modified Norne gas field with pressure values normalised to zero mean. These results were obtained using 6 POD basis functions (the energy criterion) and the linear RBF (RBF 1). NIROM estimations have negligible error.....	113
Figure 5.32: Gas saturation distribution in all grid cells of the modified Norne gas field using 7 POD basis functions (the energy criterion) and the linear RBF (RBF 1). The error in NIROM estimates was mostly negligible.	113
Figure 5.33: NIROM estimates of water gas ratio in each of the 3 wells in the Norne gas field from RBFs 1, 3, 5, and 7. These were obtained using 14 POD basis functions. These water-related variables required the most number of POD basis functions for satisfactory NIROM results.	114
Figure 5.34: Error plot for NIROM estimations of water gas ratios from the 3 wells in the Norne gas field using RBFs 1, 3, 5, and 7. Fourteen (14) POD basis functions were used to obtain these results. Errors are negligible when using a large number of basis functions.	114
Figure 5.35: Total field gas production rate (far left), cumulative field gas production and gas production rates from wells 1 and 2 in the Norne gas field model as predicted by ECLIPSE and the NIROM. The NIROM used 4 POD basis functions (99.9% energy criterion). RBFs 1, 3, 5, & 7 gave identical satisfactory results.	114
Figure 5.36: Error plots of NIROM's estimation of the production data of the Norne gas field in Figure 5.35, predicted with 4 POD basis functions (99.9% energy criterion). Errors are mostly less than 1% for RBFs 1, 3, 5, & 7.	115
Figure 5.37: Error plots of NIROM's estimation of production data of the Norne gas field when estimated using 5 POD basis functions rather than 4 (99.9% energy criterion). Errors fall to less than $1.41e^{-3}\%$ when an additional POD basis function is included in the NIROM.	115
Figure 5.38: Average RMSE for pressure and saturation distribution estimations with NIROM for all 7 RBFs considered in this chapter. NIROM simulations with outliers have not been added to this figure.	117
Figure 6.1: NIROM predictions of the average pressure in the synthetic gas reservoir during constant rate production. NIROM predictions 2 to 5 give satisfactory results, while predictions 1 and 6 are worse. Predictions 1 and 6 are outside the range of parameters used in training the NIROM, although NIROM's prediction 6 has less error than prediction 1. These results were obtained using 5, 10, 15 and 20 equally spaced training runs and in most cases, NIROM errors are less than 10psi.	122
Figure 6.2: NIROM predictions of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir using 1 POD basis function (energy criterion). NIROM adequately predicts the pattern of pressure depletion and gas production from the reservoir, and better results can be obtained with an additional POD basis function. Similar results were obtained for predictions within	

the range of training runs, with less accurate results from predictions 1 and 6. NIROM predictions of the reservoir pressure are more accurate than those of the cumulative gas production.	123
Figure 6.3: NIROM predictions of the average reservoir pressure in the synthetic gas reservoir for BHP controlled production. All 6 predictions were reproduced by the NIROM with absolute errors of less than 6%, even for prediction cases outside the range of training parameters.	125
Figure 6.4: NIROM's prediction 3 of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir produced by BHP limit control. NIROM adequately predicts the pattern of pressure depletion and gas production from the reservoir and can give better results with an additional POD basis function. Similar results were obtained for predictions 1, 2, 4 to 6, although predictions 1 and 6 were the least accurate.....	126
Figure 6.5: NIROM predictions of the average reservoir pressure in the synthetic gas reservoir with different porosity values. Predictions 2 to 5 were satisfactorily reproduced by the NIROM while the predictions outside the range of training parameters (predictions 1 and 6) were less satisfactory. Despite the low porosity value in prediction 1 NIROM adequately predicted the pattern of pressure depletion in the reservoir.	127
Figure 6.6: NIROM's prediction 3 of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir with different constant porosity values. NIROM adequately predicts the pattern of pressure depletion and cumulative production from the reservoir.....	128
Figure 6.7: NIROM predictions of the average pressure in the synthetic gas reservoir during constant rate production that was initialised by enumeration. NIROM gives satisfactory predictions of cases within the range of training parameters using 5, 10, 15 and 20 equally spaced training runs. Predictions outside the range of training parameters are less accurate.....	130
Figure 6.8: NIROM predictions of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir when initialised by enumeration. NIROM adequately predicts the reservoir pressure and cumulative gas production from the reservoir. The best NIROM prediction (prediction 2) for 5, 10, 15, and 20 training run cases are shown in (i) to (iv) above. Similar results were obtained for other predictions within the range of training runs and less accurate results from predictions 1 and 6.	130
Figure 6.9: NIROM predictions of the average reservoir pressure across all the grid cells of the Norne gas reservoir. NIROM reproduces the average reservoir pressure in the reservoir with errors of less than 0.5bars, although predictions are less accurate in the last 4 report steps that NIROM was not trained on. Prediction 6 gives the highest error of all 6 prediction cases, as it is the farthest from the range of training runs.	131

Figure 6.10: NIROM gives satisfactory predictions of the gas saturation distribution in the Norne gas field for predictions of similar production duration and range of training parameters. Outside the training runs' duration and range of parameters, NIROM predictions are less accurate. Prediction 6 has the most error in prediction and especially after the initial 21 report steps, because it is the most extreme prediction case.	132
Figure 6.11: Average reservoir pressure distribution at different report steps of the NIROM trained on 10 equally spaced training runs. Prediction 1, the farthest from the range of training runs, has the highest prediction error during the time interval for which the model was trained, and the least error deviation outside of the training duration of the NIROM. All the other errors are negligible, mostly less than 0.5%.	132
Figure 6.12: Average gas saturation in the Norne field model produced by constant BHP control with 10 training runs. Prediction errors are negligible, especially in the duration for which the NIROM has been trained.....	133
Figure 6.13: NIROM predictions of well and field gas and water production rates as well as BHP for prediction 5 of the modified Norne Field. Satisfactory results were obtained with NIROM trained on 10 training runs and 9 POD basis functions (energy criterion = 3 POD basis function). Similar results were obtained for predictions 2, 3 and 4, which are within the range of training runs and relatively less accurate results from predictions 1 and 6.	134
Figure 6.14: NIROM predictions of water gas ratios in all 3 wells and field gas production for prediction 5. These predictions were mostly for water-related production variables and do not adequately match those of ECLIPSE. This result is based on 9 POD basis functions, (energy criterion = 3 POD basis function).....	135
Figure 6.15: Average reservoir pressure predictions from NIROM match those from ECLIPSE, with negligible error for predictions 2 to 5 that are within the range of training parameters, while predictions outside the range of training parameters have larger errors. NIROM was trained with 10 training runs and 2 POD basis functions.	136
Figure 6.16: Average gas saturation distribution in the water coning model shows negligible error for predictions within the range of training parameters (predictions 2 to 5), and more errors outside the range of training parameters (predictions 1 and 6). NIROM errors are mostly less than 10% in comparison to ECLIPSE. NIROM adequately captured the non-linear decrease of gas saturation distribution in the reservoir with 88 POD basis functions.	136
Figure 6.17: Gas saturation distribution in all grid cells of the water coning gas reservoir when produced by constant rate control. The error in this NIROM prediction (prediction 3) was generally negligible.....	137

Figure 6.18: Pressure distribution in all grid cells of the water coning gas reservoir, normalised to zero mean and using 2 POD basis functions (as suggested by the energy criterion). The error in this NIROM prediction 3 was less than 1% of the original solutions from ECLIPSE.	137
Figure 6.19: NIROM's prediction 4 of different reservoir production data using 1 POD basis function (99.9% energy criterion) and 10 training runs. NIROM predictions of cumulative gas production and average reservoir pressure are the most accurate followed by those of gas production rate, water gas ratio, well BHP and water production rate, which were less satisfactory. Prediction of the cumulative water production was the least accurate. Adding 2 more POD functions gave a satisfactory prediction of all the production variables (Figure 6.20).	138
Figure 6.20: NIROM's prediction 4 of different reservoir production data with a satisfactory match. These results were obtained using 4 POD basis functions (99.9% energy criterion = 1 POD basis function) and 10 training runs. Adding 2 more POD functions gave accurate estimates of all the water and gas production variables.	139
Figure 6.21: NIROM predictions of average pressure distribution in a water coning model produced at constant BHP, using 2 POD basis functions. Predictions within the range of training parameters have lower error margins than those outside the range of training parameters. However, errors are mostly less than 1% from ECLIPSE's solution.	139
Figure 6.22: Average gas saturation predictions for the water coning gas reservoir model using 133 POD basis functions. Here the reservoir was produced by BHP control, and unlike for the rate controlled production example where both extrapolation predictions (1 and 6) were least accurate, prediction 6 was least accurate in this example. Prediction 6 had a much larger error margin than the other predictions cases that had less than 10% prediction error and were comparatively more accurate.	140
Figure 6.23: NIROM's prediction of the average pressure distribution in Field X when it was produced at different constant production rate values. Predictions 2 to 5 were always more accurate than predictions 1 and 6 (which were outside the range of training parameters). NIROM's predictions of cases 2 to 5 which are within the range of training parameters improved as the number of training runs increased, while NIROM's prediction of cases 1 and 6 got worse as the number of training runs increased from 5 to 20.	141
Figure 6.24: NIROM's predictions of the average gas saturation distribution in Field X at every report step. In these cases, Field X was produced at different constant production rate values. Here also, predictions 2 to 5 (the interpolation cases) are more accurate than predictions 1 and 6 (the extrapolation cases). However, the prediction cases that were within the range of training parameters	

get better as the number of training runs increased while the extrapolation cases got worse as the number of training runs increased.	142
Figure 6.25: NIROM's prediction 2 of Field X's production variables using 5 training runs. The water production and aquifer influx variables were the least accurately predicted variables. Otherwise, most of the other NIROM predictions were satisfactory.....	144
Figure 6.26: Average percentage RMSEs of all predictions in each constant parameter case of the synthetic gas reservoir model. NIROM predictions of grid distributed pressure and gas saturation give smaller errors than those of the production data that require more than the optimum number of POD basis functions to obtain satisfactory results.	146
Figure 6.27: Average of the percentage RMSE for NIROM predictions for the modified Norne field and water coning reservoir models. NIROM predictions of the pressure distribution in the reservoirs had lower errors than those for the gas saturation distribution. NIROM implemented with 5 training runs consistently gave the least RMSEs.....	147
Figure 6.28: Summary of the time taken to implement the constant production rate cases of the synthetic gas reservoir. NIROM speedups of between 8 and 13 times the speed of ECLIPSE are achievable.	149
Figure 6.29: Summary of the time taken to implement the constant BHP cases of the synthetic gas reservoir. NIROM speedups of between 10 and 17 times the speed of ECLIPSE are achievable.....	150
Figure 6.30: Summary of the time taken to implement the constant porosity cases of the synthetic gas reservoir. NIROM speedup of between 8 and 14 times ECLIPSE are achievable.	151
Figure 6.31: Summary of the time taken to implement initialisation by enumeration for constant production rate cases of the synthetic gas reservoir. NIROM speedups of between 9 and 18 times ECLIPSE are achievable.	151
Figure 6.32: A breakdown of the average time taken to implement NIROM simulations of the modified Norne field. Time spent on POD reduction is proportional to the number of training runs. NIROM is 15 to 30 times faster than ECLIPSE.	152
Figure 6.33: Breakdown of the average time taken to implement NIROM for the water coning reservoir. Time spent in actual NIROM prediction/ interpolation is negligible, leading to significant speedups of between 97 and 281 times ECLIPSE's run time.	153
Figure 6.34: The breakdown of the average time spent during NIROM predictions of constant rate cases of Field X. NIROM is between 160 and 290 times faster than ECLIPSE in simulating this real gas reservoir case.....	153
Figure 7.1: A seasonally varying production profile with mixed production control. In this example, predictions 1 and 2 switched to BHP limit control before the end of the production lifecycle. The other	

predictions are on varying rate control throughout the production duration. NIROM simulations of this case are shown in Figure 7.2 and Figure 7.7, below. The corresponding training runs for this prediction example are in Figure 7.4.....	158
Figure 7.2: Average pressure distribution in the synthetic gas reservoir when produced with the varying seasonally production profile in Figure 7.1. Prediction with 10 training runs gives the least prediction errors overall. Nevertheless, NIROM prediction errors are mostly less than 10%. The corresponding training runs for these predictions are in Figure 7.4. NIROM's predictions 1 and 2 are the least accurate because they have mixed production controls. The error in these two cases were especially worse during the interval when their production control changed from gas rate to BHP limit.	159
Figure 7.3: NIROM's predictions 2 and 4 of the pressure distribution in all the grid cells of the synthetic gas reservoir when produced by the seasonally varying production profile and mixed production controls of Figure 7.1. These pressure values have been normalised to zero mean. The error in these NIROM predictions is mostly negligible even for prediction 2 after its production control switched to BHP limit control.	160
Figure 7.4: Production profiles for 5, 10 and 20 training runs corresponding to the results in Figure 7.2. Each of these training runs has at least one case of mixed production control, hence their suitability for the prediction cases in Figure 7.1, as seen in the results in Figure 7.2 and Figure 7.3.	161
Figure 7.5: Production profiles of 5, 10 and 20 training runs corresponding to the results in Figure 7.6. These training runs are all varying production rates with no mixed production control, hence their erroneous predictions (Figure 7.6) of predictions 1 and 2 of Figure 7.1.	161
Figure 7.6: NIROM prediction of the average pressure distribution in the synthetic gas reservoir when NIROM was trained with non-mixed production controls (Figure 7.5), whereas the prediction cases have mixed production controls (Figure 7.1). Hence the NIROM has significantly over-predicted the average reservoir pressure during the BHP limit controlled durations of predictions 1 and 2. This emphasises the need for training the NIROM within similar conditions to the cases it is required to make predictions for.	162
Figure 7.7: NIROM prediction of the production data (cumulative production and reservoir pressure) for prediction 1. Using the optimum number of POD basis functions (1 POD basis function), NIROM satisfactorily identifies the intervals when the production control switched to the BHP limit. Increasing the number of training runs did not have a significant effect on the accuracy of NIROM's prediction.	163

Figure 7.8: Production profile for the triennially varying parameter case of the synthetic gas reservoir. The frequency at which the production rates vary is after 1000 – 1500 days (10 - 15 report steps) interval.	164
Figure 7.9: NIROM predictions of the average pressure distribution in the synthetic gas reservoir produced with a triennially varying production profile. NIROM predictions have less than $\pm 10\%$ error (mostly less than $\pm 5\%$) for all predictions except prediction 1.	165
Figure 7.10: Production profiles of the simulation runs used to train NIROM for the triennially varying production rate cases. The training runs are spread between 2000 and 8000Mscf/day.	166
Figure 7.11: NIROM prediction of the cumulative gas production and reservoir pressure for the triennially varying prediction 1. NIROM was able to mimic the pattern of the reservoir's performance and had the most errors when it was trained with 5 training runs.	167
Figure 7.12: Production profiles of the 3 wells in the modified Norne field that are used to test NIROM's prediction of varying production profiles with mixed boundary conditions. In all 3 wells, production control (boundary condition) was switched to BHP limit control after some period of being produced at varying production rates.	168
Figure 7.13: Production profiles for the training runs used to train NIROM for the results in Figure 7.14 to Figure 7.16. These training runs have mixed production profiles like the prediction runs in Figure 7.12.	168
Figure 7.14: NIROM predictions of the average pressure in the modified Norne gas field for the seasonally varying production rate and mixed production control scenarios of Figure 7.12.	169
Figure 7.15: Pressure distribution in all grid cells of the modified Norne gas field, normalised to zero mean (predictions 1 and 4). In both cases, the NIROM's predictions show the most errors between the 100th and 130th report step corresponding to the region with the highest error in Figure 7.14, i.e. the time interval when the production control changed. Nevertheless, NIROM's prediction of the pressure across the reservoir is satisfactory.	170
Figure 7.16: NIROM's prediction 1 and 4 of the gas saturation distribution in all the grid cells of the modified Norne gas field. NIROM satisfactorily captured the pattern of gas saturation distribution in the gas field, with negligible error. The interval when production control changed to BHP limit is not noticeable in these images.	171
Figure 7.17: NIROM prediction 4 of production variables of the modified Norne gas field trained with 20 runs and 9 POD basis functions (energy criterion = 1 POD basis function). NIROM adequately predicts most of the production variables and satisfactorily captured their performance pattern. .	173
Figure 7.18: Production profiles for the prediction runs of the seasonally varying parameter example of Field X, based on a BHP limit of 100bars. All 6 production profiles show that the field switched to	

BHP control during its production lifecycle, with instances of the producer shutting in after 230 report steps.....	173
Figure 7.19: NIROM's prediction of the average pressure distribution in Field X produced with a seasonally varying production profile and BHP limit. Predictions with 10 training runs were the most accurate with a maximum error of just over 4%. Predictions with 5 and 20 training runs have a maximum error of just over 6%.	174
Figure 7.20: NIROM's prediction of the average saturation distribution in Field X produced with a seasonally varying production profile and BHP limit of 100bars. Predictions with 10 training runs were the most accurate, but no case had errors greater than 0.001.....	175
Figure 7.21: NIROM's prediction 2 of the gas saturation distribution in all grid cells of Field X trained with 5 runs. NIROM satisfactorily captured the pattern of gas saturation distribution in the gas field, with negligible error.....	176
Figure 7.22: Production profiles used to train the NIROM for the prediction runs in Figure 7.18. The training runs span a similar range of parameters as the prediction runs, with some values being as high as prediction 1 of Figure 7.18. Given that these training runs cover the range of prediction runs, NIROM predictions in Figure 7.19 to Figure 7.21 and Figure 7.23 are satisfactory.	176
Figure 7.23: NIROM's prediction 4 of Field X's production variables using 5 training runs. With the optimum number of POD basis functions, NIROM did not accurately predict the time-varying production rates of Field X. However, NIROM satisfactorily predicts more than half of the production variables that were considered.	178
Figure 7.24: Production profile for the prediction runs of the seasonally varying parameter cases of Field X with THP limit. From the 150th report step, the production profiles for Field X were switching between THP limit control and episodes of shut-in due to a water gas ratio limit imposed on the reservoir.....	179
Figure 7.25: Production profiles of the training runs used for the NIROM predictions in Figure 7.24. These training runs have a similar shut-in and re-opening behaviour as the prediction runs in Figure 7.24, hence the satisfactory prediction results for these training cases as shown in Figure 7.26 to Figure 7.28.	179
Figure 7.26: NIROM predictions of the average pressure distribution in Field X for the seasonally varying production profiles with THP limit control in Figure 7.24. NIROM predictions with 5, 10 and 20 training runs are satisfactory with at most 4% error.....	180
Figure 7.27: NIROM's satisfactory predictions of the average gas saturation distribution in Field X when it was produced by a seasonally varying production profile with THP limit. Training NIROM with 20 runs gave the best results.....	181

Figure 7.28: NIROM's prediction 4 of the production variables of Field X when trained with 20 runs. NIROM satisfactorily predicted most of the production variables with the optimum number of training runs but found it challenging to capture the pattern of the varying production rates, and especially after the production control began to switch.	183
Figure 7.29: NIROM's prediction 4 of some production variables of Field X using the full POD basis functions. NIROM better predicts the varying production rates and BHP when more POD basis functions were added to its implementation.	184
Figure 7.30: Average percentage RMSEs of NIROM when implemented for predicting the performance of the synthetic gas reservoir and the Norne gas field, produced with varying production profiles and mixed boundary conditions. NIROM prediction of grid distributed pressure and saturations have lower RMSEs than its prediction of production variables, based on the optimal number of POD basis functions. There is also no consistent correlation between the number of training runs and the accuracy of NIROM predictions.	185
Figure 7.31: Average RMSEs for NIROM predictions of the pressure and saturation distribution in all grid cells of Field X for a) constant parameter cases b) seasonally varying parameter cases with BHP limit and c) seasonally varying production profile with THP limit.	187
Figure 7.32: The breakdown of time taken to train NIROM for the seasonally varying production profile of the synthetic gas reservoir, as well as the speed up from NIROM predictions. Significantly more time is taken during POD reduction for the cases with 10 and 20 training runs than with 5 training runs. The NIROM built with 10 or 20 training runs were slower than ECLIPSE.	189
Figure 7.33: Breakdown of the average time taken to implement NIROM for the triennially varying production profile of the synthetic gas reservoir. NIROM is 2 to 3.5 times faster than ECLIPSE for these cases. Training NIROM with 20 runs gives the least speed-up since it involves more data. The time spent in the actual NIROM interpolation/ prediction is negligible compared to the time spent in building the NIROM and if the NIROM is built in a compiled language it would be even faster.	189
Figure 7.34: Breakdown of the average time taken to implement NIROM predictions for the seasonally varying production parameter cases of the modified Norne field with mixed production controls. Less than 0.2% of the NIROM implementation time is spent on the actual NIROM predictions/ interpolation.	191
Figure 7.35: The breakdown of the average time taken for NIROM predictions of the seasonally varying cases of Field X. NIROM is between 2 and 4 times faster than ECLIPSE, although less than 1 second is spent on the actual NIROM interpolation step. This shows that NIROM has room for significantly more speedup if it is implemented in a compiled language like ECLIPSE.	191

Figure 8.1: Production profiles for the (a) seasonal and (b) triennially varying prediction cases of the modified Norne gas field with shut-in wells. These production profiles are for the prediction results in Figure 8.2 & Figure 8.4 respectively, with matching line colours.....	196
Figure 8.2: Predictions of the average pressure distribution in all grid cells of the modified Norne field when trained with seasonally varying production rates and (a) the POD-AE model with 2 HLs or (b) POD-RBF models with the full and optimal number of POD basis functions. Predictions from both POD-AE and POD-RBF are generally satisfactory with minute errors and especially for predictions 1 to 3. In (a), the POD-AE with 2HLs gave unstable predictions for the 4th and 5th predictions in the first 75 report steps encircled with a yellow rectangle. Also, predictions with both POD-RBF, i.e. (b), look identical, overlapping each other in most cases, and they give more consistent and stable predictions throughout the production lifecycle than the POD-AE prediction in (a).....	197
Figure 8.3: Error plot for the predictions of the pressure distribution in the modified Norne gas field using POD-AE model with 2 HLs as well as POD-RBF models with full and optimal (99.9% energy criterion) number of POD basis functions. The POD-AE model has the highest errors, especially for predictions 4 and 5, while the errors from both POD-RBF models have less variance and lower values. The full rank POD-RBF model has similar or lower errors than the optimal POD-RBF model.....	198
Figure 8.4: Predictions of average pressure distribution in all grid cells of the modified Norne gas field when trained with the triennially varying production profile. Predictions from (a) POD-AE and (b) POD-RBF models satisfactorily match those from ECLIPSE. The POD-RBF solutions, i.e. (b) are more consistent and accurate than the results in Figure 8.4a. Figure 8.4b also shows that the results from both POD-RBF models are identical.....	199
Figure 8.5: Percentage error in the predictions of the average pressure distribution in all grid cells of the Norne gas field when trained with a triennially varying production profile (Figure 8.1b). Both POD-RBF and POD-AE models give satisfactory predictions with less than 10% error. The POD-RBF solutions are more consistent and accurate (mostly less than 2% error). Results from the full and optimal rank POD-RBF models are mostly identical.	199
Figure 8.6: Production profiles for prediction runs of (a) seasonal and (b) triennially varying cases with no shut-in wells. These are the production profiles for the results in Figure 8.9 to Figure 8.12, with matching line colours. The seasonal profile is for production from 3 wells while the triennial profile is for production from 2 wells.	200
Figure 8.7: Predictions of the pressure distribution in all grid cells of the modified Norne gas field for prediction 3 of the seasonally varying production profile with no shut-in wells. All the models (POD-AE with 2 and 5HLs as well as POD-RBF with optimal and full POD basis functions) gave satisfactory predictions of the reservoir pressure distribution in over 44000 active grid cells. Predictions from the	

POD-RBF models (c & d) were better than those from the POD-AE models (a & b). The improvement in predictions with the full rank POD-RBF model is also obvious in these results.	201
Figure 8.8: Predictions of the average pressure distribution in all grid cells of the modified Norne gas field when trained with seasonally varying production rates and POD-AE models with 2 and 5 HLs. The POD-AE models gave predictions that satisfactory matched the pattern of pressure depletion in the reservoir, with some over- and under-predictions for 2 and 5 HLs respectively.	202
Figure 8.9: POD-RBF predictions of the average pressure distribution in all grid cells of the modified Norne gas field when trained with seasonally varying production rates. The results from the full and optimal rank POD-RBF models are similar and consistent throughout the prediction duration.	202
Figure 8.10: Percentage error in the predictions of the average pressure distribution in all grid cells of the Norne gas field when trained with seasonally varying production profiles for POD-AE models with 2 and 5 HLs as well as the full and optimal POD-RBF models. The POD-AE models with 2 and 5 HLs over and under predict the field's pressure distribution, respectively. The full and optimal rank POD models give similar results with lower prediction errors.	203
Figure 8.11: Predictions of average pressure distribution in the modified Norne gas field with production rates varying after at least 2.5 years. The POD-AE models are trained with 3 and 5 HLs and satisfactorily capture the pattern of pressure depletion in the gas field while predictions from the POD-RBF models are more accurate, with similar results from the full and optimal rank POD basis function POD-RBF NIROMs.....	203
Figure 8.12: Percentage errors in the predictions of the pressure distribution in the modified Norne gas field using POD-AE models with 3 and 5 HLs as well as POD-RBF models with full and optimal (99.9% energy criterion) number of POD basis functions. POD-AE models have the highest errors while the errors from both POD-RBF models are less than 2%, except for prediction 1. The full rank POD-RBF model has almost the same or slightly lower errors than the optimal POD-RBF model.....	204
Figure 8.13: Average absolute errors and RMSEs for predictions of the pressure distribution in the modified Norne field using POD-AE and POD-RBF NIROMs. Prediction cases with shut-in wells are less accurate than those without shut-in wells. POD-RBF predictions are more accurate than POD-AE predictions.	205
Figure 8.14: The breakdown of the time taken to train the POD-AE and POD RBF models that were implemented for predicting the pressure distribution in the modified Norne gas field, trained with 50 training runs. POD-AE is slower than ECLIPSE for both the seasonal and the triennially varying production profiles, while the full POD-RBF model is only faster than ECLIPSE for the triennially varying production profile cases.	207

Figure 9.1: Summary of the 3 stages that the non-intrusive reduced order model was evaluated for, a) estimating specific simulation cases, and predicting new simulation cases with b) constant parameters c) time-varying and mixed control parameters.	211
Figure 10.1: Summary of procedure for the modified normalization approach of forced parameter scenarios. Parameterization is executed in hyperspace to nullify the need for sorting parameters in any order of magnitude.	231
Figure 10.2: Examples of NIROM's attempt at predicting the average pressure distribution in 2 varying rate cases of the synthetic gas reservoir model. In both cases, a) modified normalization approach and b) temporal resolution constrained at equal intervals of cumulative production, NIROM's predictions were not representative. The corresponding production profiles for both cases are the (ii) plots...	232
Figure 10.3: Sharpness of turn and boost plots for the spatial pressure data of the synthetic gas reservoir. Both plots are compared for a) constant rate production and b) varying rate production, to evaluate the extent of instability inherent in the reservoir's pressure data and to evaluate if the NIROM can cope with the data for interpolation. The sharpen of turn plots in both cases are satisfactory, but the spikiness of the boost plot for the varying rate case suggests that it would be more challenging for the NIROM to accurately interpolate.	234
Figure 11.1: Results of modelling the modified Norne gas field with 10 training runs when at least one of the wells was shut-in during the production lifecycle. a) The production profiles for the prediction cases tested with POD-AE and POD-RBF b) The production profiles for the 10 training runs used in this example c) Predictions of the average pressure distribution in the field using POD-AE with 2 and 3 HLs d) Predictions of the average pressure distribution in the field using the optimal and full rank POD-RBF and e) the average error in the prediction of pressure distribution from the 4 models in c and d. ..	236
Figure 11.2: Results of modelling the modified Norne gas field with 20 training runs when at least one of the wells was shut-in during the production lifecycle. a) The production profiles for the prediction cases tested with POD-AE and POD-RBF b) The production profiles for the 20 training runs used in this example c) Predictions of the average pressure distribution in the field using POD-AE with 2, 4 and 5 HLs d) The average error in the prediction of pressure distribution from the 3 models in (c) as well as the optimal POD-RBF model.	238
Figure 11.3: Modelling a triennially varying production case of the modified Norne field case with shut-in well(s). This example is from the 4HLs model a) without dropout layers b) with 0.25 dropout rate and c) with 0.5 dropout rate. The best prediction was obtained with the fully connected model....	239
Figure 11.4: An example of a three-layer feedforward neural network ^[111]	240

Figure 11.5: Predictions of the average pressure distribution in the modified Norne gas field with shut-in wells. 5 FFNNs were tested and they all gave averaged values for all 5 predictions, being unable to distinguish the dynamics of the different prediction examples.	240
Figure 11.6: A schematic of the AE-LSTM model that was tested for this work. The model consists of an autoencoder whose innermost layer is an LSTM model.....	241
Figure 11.7: Results of using an AE-LSTM for modelling the average pressure distribution in a modified Norne field example with shut-in well(s). All the AE-LSTM models gave identical predictions for all the cases (HLs) that were tested.....	241
Figure 12.1: Relative permeability curves for the synthetic gas reservoir.	242

LIST OF TABLES

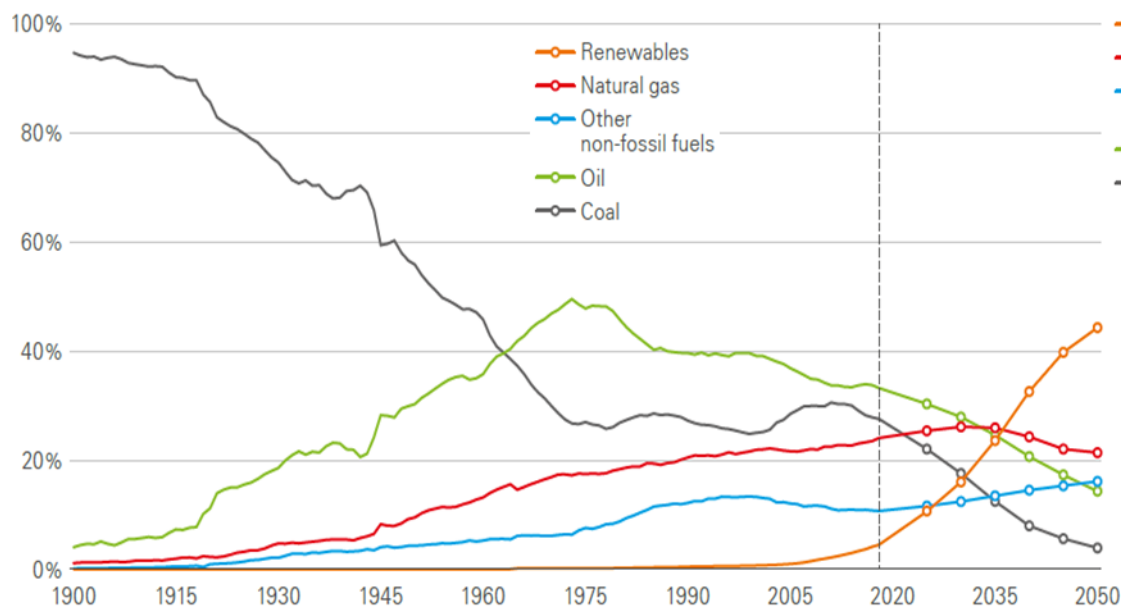
Table 3.1: Equations and schematics of ELU, ReLU and Leaky ReLU activation functions.	75
Table 3.2: Summary of the hyper-parameters used in the autoencoder model.	76
Table 4.1: Summary of reservoir rock and fluid properties of the synthetic dry gas reservoir.	83
Table 4.2: Summary of simulation parameters investigated in the constant parameter cases of the synthetic gas model. Training runs are spaced equally between the ranges of values specified.	85
Table 4.3: Summary of simulation parameters used to generate forced parameter production profiles for the synthetic gas reservoir. Parameters are selected from the list at random for the rate controlled production cases and these values are multiplied by distinct factors to generate varying profiles. ...	85
Table 4.4: Reservoir properties of the modified Norne gas field.	86
Table 4.5: Summary of constant parameters that were evaluated for the modified Norne gas field. Training runs are spaced equally between the range of training values while prediction runs are distributed within and outside the range of values for training the NIROM.	87
Table 4.6: Summary of production rates and BHP limit used to create the time-varying production profiles for the varying parameter cases of the modified Norne gas field. Production rates were multiplied by different unique factors to create distinct production profiles.	87
Table 4.7: Some reservoir and production properties of the water coning gas reservoir.	89
Table 4.8: Training and prediction simulation parameters for a gas reservoir model with significant water coning.	90
Table 4.9: Some reservoir properties of Field X.	90
Table 4.10: Summary of constant parameters that were evaluated for Field X. Training runs are spaced equally between the range of training values while prediction runs are distributed within and outside the range of values for training the NIROM.	92
Table 4.11: Summary of production rates and pressure limit used to create the varying production profiles for the varying parameter cases of Field X. Production rates were multiplied by different unique factors to create distinct production profiles.	92
Table 5.1: Summary of result on best-fit interpolation function(s) for various simulation cases.	116
Table 5.2: Average RMSE for pressure and saturation distribution estimations with NIROM for all 7 RBFs considered in this work. RBF 1 is the most consistently stable RBF with the least RMSE. The values in red colour are the outliers where NIROM overestimated its simulation or it could not give reasonable estimations.	116
Table 5.3: Simulation runtime (in seconds) and average speed ups from NIROM versus ECLIPSE for different gas production cases. Comparison of the cumulative computational time shows up to 60 times simulation speed up with NIROM.	118

Table 5.4: A summary of the cases that were considered in this chapter and the key findings.	119
Table 6.1: Number of POD basis functions required for NIROM predictions. The listed values represent the energy criterion except where it is mentioned otherwise.	147
Table 6.2: Summary of time taken to implement the constant production controlled (rate and BHP limit) cases of the synthetic gas reservoir. These values correspond to the plots in Figure 6.28 and Figure 6.29.	149
Table 6.3: Summary of the time taken to implement the constant porosity and initialisation by enumeration for constant production rate cases of the synthetic gas reservoir. These values correspond to Figure 6.30 and Figure 6.31.	150
Table 6.4: Breakdown of the average time taken to implement constant parameter simulations of the modified Norne field and the water coning reservoir model. POD reduction time increases as the number of training runs increase. NIROM achieves at least an order of magnitude speedup for both models.	152
Table 6.5: NIROM's simulation speed-up for the constant parameter test cases considered in this chapter.	154
Table 6.6: A summary of the prediction cases and key findings in this chapter.	156
Table 7.1: Breakdown of the time take to run the varying parameter cases of the synthetic gas reservoir model, in seconds. These results are plotted in Figure 7.32 and Figure 7.33.	188
Table 7.2: Average time taken to implement NIROM for the seasonally varying production profile of the modified Norne gas field. NIROM is a maximum of 2 times faster than ECLIPSE.	190
Table 7.3: NIROM's simulation speed-up for the varying parameter test cases considered in this chapter. NIROM's speed-up depends on the amount of data that was used to train the NIROM (model size) as well as the complexity of the reservoir model (homogenous or heterogeneous).	192
Table 7.4: A summary of the prediction cases and key findings in this chapter	194
Table 8.1: Average errors and RMSEs for pressure prediction cases (in bars). POD-RBF models show lower RMSEs than POD-AE models. And POD-RBF cases without shut-in wells have lower RMSEs (are more accurate) than cases with shut-in wells.	205
Table 8.2: Overview of the time taken to implement POD-RBF and POD-AE models for gas field simulations in comparison to ECLIPSE. POD-RBF models are faster than POD-AE models, but the optimal POD-RBF is the only model that is faster than ECLIPSE for the seasonally varying production profile cases.	206
Table 8.3: A summary of the prediction cases and key findings in this chapter.	209

CHAPTER 1: INTRODUCTION

1.1 MOTIVATION

Global energy consumption has been on the rise to sustain the growth in technological advancement and development, and fossil fuel supplies over 80% of this energy ^[1]. For several more years, fossil fuels, comprising coal, oil and natural gas ^[1-4], will continue to be relevant in the global energy mix, especially natural gas ^[3, 4]. Natural gas is used in electric power generation, for heating residential and commercial buildings, for transportation, and industrially as a raw material for hydrogen production as well as in fertilizer and petrochemical industries ^[3-5]. According to bp ^[4] and the EIA in 2020 ^[3], natural gas currently constitutes 24.2% ^[6] of the global energy mix and it is expected to be used more in the next decade, despite the current transition away from fossil fuels, Figure 1.1.



The energy transition towards low carbon sources is important to combat the rise in global temperatures and climate change. Natural gas is a significant part of this transition ^[3, 4] since it has lower carbon emissions than coal and oil, Figure 1.2 ^[2]. According to bp ^[4], even if there is a “rapid” transition towards lowering the carbon emissions from fossil fuels, such as through targeted policies and carbon price measures, the place of gas in the global energy mix should still experience some increase for up to 15 years before it starts to decline, Figure 1.1.

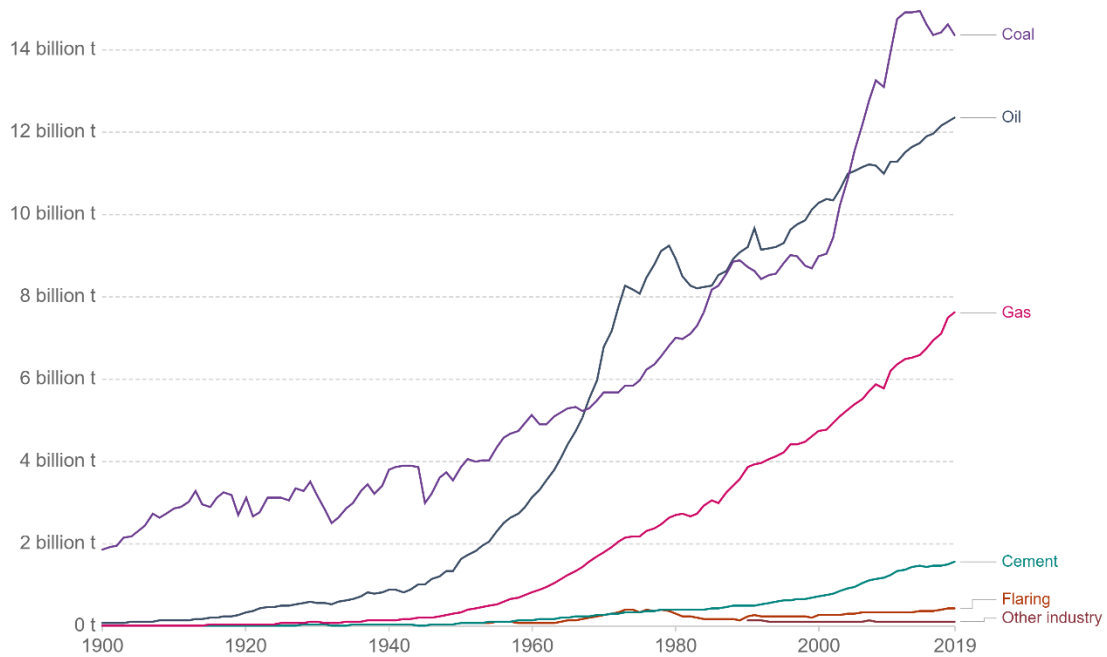


Figure 1.2: Global CO₂ emissions by fuel type per year since 1900 ^[2]. Gas emits less CO₂ than Oil and Coal in billion tonnes.

It is also interesting to notice that except for 2008/2009, global gas production and consumption have been increasing continuously every year for the past twenty years, Figure 1.3 ^[3, 4]. In the US (the largest consumer of natural gas), gas consumption has even defied prior forecasts by Harrison et al. (2006) and EIA (Figure 1.4) ^[1, 3, 7, 8]. Therefore, it is worthwhile to pay attention to studies that facilitate the improvement of gas exploration and production, for the next decade.

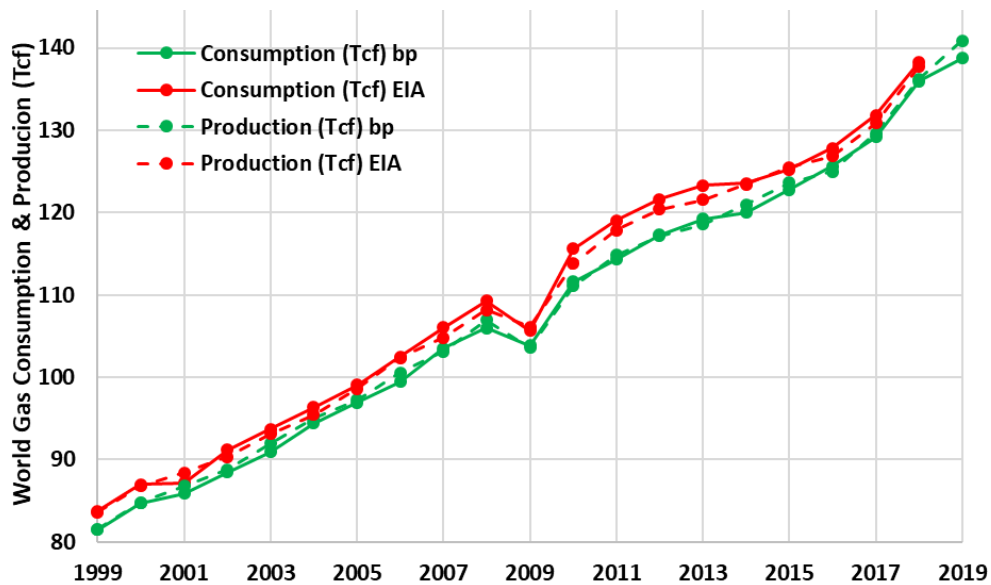


Figure 1.3: Gas consumption and production from 1999 to 2019. There is a consistent annual increase in the production and consumption of gas, except for 2009 ^[3, 6].

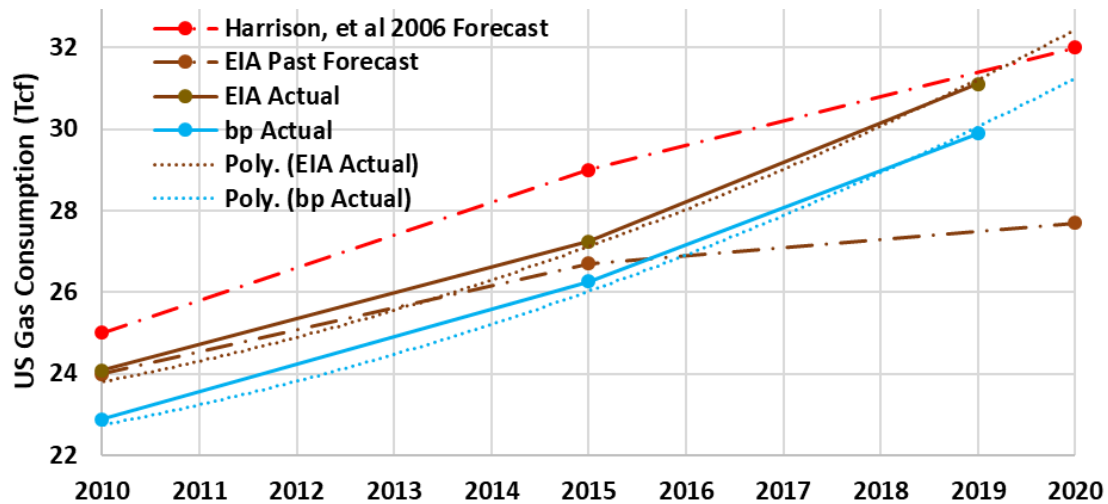


Figure 1.4: Comparison of gas consumption in the US since 2010 with the actual consumption during that period. In agreement with Harrison et al, the actual gas consumption according to EIA and bp exceeded EIA's 2003 forecast. But based on polynomial trend line projections, the actual current gas consumption is likely to exceed Harrison's generous forecast by 2020 ^[1, 3, 7, 8].

Similar to exploration for oil, gas exploration and development starts with appraising the presence and amount of gas that can be produced from an asset, based on proved technical and commercial considerations, the gas reserves ^[9, 10]. This information is critical to deciding if the reservoir or field is economically worth exploring, in addition to planning its development and production strategies ^[10]. In most cases, gas reservoirs are only produced when there is a contractual agreement between sellers and buyer(s), where there is the prospect of selling the gas with minimal storage time and cost ^[5]. Gas contracts typically cover the lifecycle of gas reservoirs, 15 to 25 years. Therefore information on gas reserves as well as potential gas supply rates is required in advance of drawing up the terms of gas contracts or commencing actual production. This prior information on the potential gas supply is typically anticipated based on field development planning which involves understanding the impact of geological uncertainty and other risks on the different range of production outcomes. After all, not adhering to the agreed amounts of gas supply in a contract can attract penalties. In addition, production scenarios from field development plans should be indicative of the economic recovery factors, with possibilities for improved recovery to meet up with daily and annual contractual quantities, despite unforeseen technological, commercial or policy changes that could occur while the contract is in place, or indeed weather ^[5, 11].

The dynamics of gas reservoirs are quite different from those of oil, although like oil, its production behaviour depends on factors like drive mechanism, reservoir pressure and heterogeneities. Gas is highly compressible so gas reservoirs are usually produced by pressure depletion and water influx as a consequence of the pressure depletion can reduce production ^[11, 12]. When gas reservoirs have no associated aquifers, gas production is less complicated and straightforward and there is a likelihood of high gas recovery from pressure depletion drive. Some gas properties are also pressure-dependent

and change substantially as the reservoir's pressure varies. Predicting gas production from water drive reservoirs is more complicated as it is more sensitive to geological heterogeneity and requires additional data for simulation such as relative permeabilities and the size and properties of the aquifer. In heterogeneous reservoirs, non-uniform gas drainage can be further aggravated by the water influx, causing gas to be trapped in the parts of the reservoir that are invaded by water, and thereby decreasing gas recovery ^[13]. In some cases, gas recovery can be optimised by strategically producing the gas at varying rates to suit the prevailing reservoir conditions. For example, a gas reservoir with an associated aquifer can be initially produced at high rates before the less mobile water begins to encroach into the reservoir, after which the gas is produced at lower rates to manage the rate of water influx and gas bypassing ^[11, 14].

Traditionally gas reservoir performance was predicted using material balance methods, but material balance methods do not provide information on rates versus time, neither do they provide spatial information on water and pressure distributions in the reservoir or allow optimisation of field performance by changing well rates. Numerical simulation can provide this information which is necessary to interpret the flow mechanisms within a reservoir and for informing further development decisions such as the location of drilling development or infill wells to improve recovery ^[10, 15, 16]. Reservoir simulation is core to integrated petroleum field modelling and involves evaluating and planning development and operating strategies in terms of their impact on a project's robustness and economics. Field development planning entails running several simulations, first to build a dynamic model that is consistent with the observed field's geology and (production) history and later to evaluate the sensitivity of the field's performance to alternative project decisions such as the number and location of wells. In addition field development planning involves evaluating the impact of different uncertainties and other risks that could accompany the field's future performance ^[17].

Conventional simulation (based on discretising and solving the PDEs) to describe the flow of fluids in actual reservoirs may take a few hours to several days to complete ^[10, 17, 18]. Reservoir simulation can be very time-intensive for reservoirs with complex structures and behaviours, such as oil and gas reservoirs with dual porosity and/or permeability and unconventional reservoirs as well as during certain investigations such as for improved recovery, sensitivity and uncertainty analysis ^[10, 13, 18-20]. As timing is key to the profitability of a petroleum asset and a large chunk of project time is spent on reservoir simulation ^[10, 17], the need for fast approximate simulations cannot be over-emphasized.

Computational speed challenges exist in other areas/ fields of study e.g. physical and biological sciences ^[21-25], as well as engineering ^[26-29] and some techniques from unrelated fields have found applications in petroleum engineering e.g. reduced order modelling, machine learning, parallel and cloud computing ^[13, 29-31]. Some of these methods need to be linked to the source code of the

simulators that generate their data (i.e. intrusive models), ^[21, 32-35] while others are independent techniques ^[36-38] and are preferable in petroleum engineering since commercial reservoir simulation companies do not allow users access to their source codes.

In petroleum engineering, most of the fast computational approaches have been tested for predicting production performance, they provide little application to the distribution of pressure and fluid saturations within the reservoir since this spatial information is more extensive and time-consuming to train ^[13, 32, 33, 36-41]. Likewise, the focus of optimising the speed of computational solutions in petroleum engineering research has been inclined towards oil-water displacements such as in waterflooding ^[36, 39, 42, 43] and EOR ^[16, 20, 44], while also neglecting to predict time-varying production or mixed boundary conditions (BHP or rate production control). Given the increasing importance of global gas production and consumption, there is a need to develop simulation methods that are suitable for practical gas reservoir production.

1.2 AIM AND OBJECTIVES

1.2.1 AIM

This research aimed to develop a fast, yet straightforward and accurate non-intrusive reduced order model (NIROM) for gas reservoir modelling based on the output of a conventional black oil simulator and to evaluate the NIROM's usefulness for gas field modelling. A non-intrusive model is more relevant to practising petroleum engineers and straightforward to use since it does not require access to the simulator's proprietary codes.

1.2.2 OBJECTIVES

- To review the literature on Non-Intrusive Reduced Order Modelling (NIROM) and evaluate its application to integrated gas field development, with emphasis on the estimation of the spatial distribution of saturation and pressure over time.
- To develop NIROM for modelling gas reservoir performance/ behaviour based on a full physics commercial reservoir simulation software (Schlumberger ECLIPSE).
- To build a conventional model of a typical gas field to be used to test the applicability of NIROM.
- To validate NIROM and identify decisions that it can support.
- To develop theory and workflow to enable NIROM to be created by non-specialist engineers, with minimum training runs and simple to use.
- To evaluate NIROM on real field cases in comparison to the conventional approach.
- To explore machine learning methods for gas field modelling in comparison to NIROM.

1.3 CONTRIBUTIONS

The main contribution of this work is that it presents a method of implementing the NIROM that suits the prediction of time varying parameter systems and demonstrates this for estimating the performance of gas reservoirs with time varying production. Time-varying scenarios are an important consideration for seasonally varying gas production. The following are the novel contributions of this work:

1. Developed two (2) frameworks/ workflows that suffice for the fast modelling of pressure, saturation and production performance in gas reservoirs:
 - a. Proper Orthogonal Decomposition (POD) - Radial Basis Function (RBF) NIROM,
 - b. POD - Autoencoder (AE) machine learning framework.
2. Showed the application of both methods for modelling different cases of gas field production developments with constant and time-varying production scenarios:
 - a. training and prediction (unseen) cases,
 - b. production rate and pressure limit controls,
 - c. homogeneous, heterogeneous and real gas fields/ reservoirs,
 - d. reservoirs with and without aquifer pressure support.
3. Demonstrated the extent of accuracy, simplicity and speed-up that is obtainable from both POD based methods for gas reservoir modelling.
4. Demonstrated different methods of normalising reduced data in hyperspace and their practicality for gas field modelling.
5. Demonstrated for the first time an application of machine learning for modelling gas reservoir performance as a regression problem.
6. Demonstrated that although POD-AE is a suitable machine learning method for gas field modelling, the best implementation of autoencoders for gas field modelling is case dependent.
7. Comparative analysis of NIROM and machine learning for accuracy and speed of forecasting gas reservoir simulation results from a commercial reservoir simulator.

1.4 OVERVIEW OF THESIS

The objectives of this thesis are addressed as follows:

Chapter 2 reviews the current knowledge and gaps in conventional methods for modelling gas reservoirs in addition to fast modelling methods, with emphasis on reduced order models and machine learning. In chapter 3, the theory and workflow for the 2 POD based fast modelling frameworks that have been adopted for this work are described. In chapter 4, four gas reservoir models are described, comprising their geometry, geological and production properties, as well as the different production controls and scenarios for which they are evaluated in the result chapters of this

work. Chapters 5 to 8 contain results and discussions of the results from the two proposed POD-models as well as how they have been evaluated and validated for gas reservoir modelling. These chapters also showcase the range of decisions that NIROM can support in gas field development and production.

Chapter 5 shows the initial implementation of POD-RBF NIROM and the choice of hyper-parameters that are applicable for gas reservoir modelling. The choice of RBF that best suits gas reservoir modelling is validated for a synthetic gas reservoir and a heterogeneous gas field, in addition to the number of POD basis functions to be used for gas reservoir modelling. NIROM was evaluated for reproducing its training runs.

In chapter 6, POD-RBF NIROM is fine-tuned for predicting cases that it has not been trained on, hence the introduction of parameterisation. Prediction cases require several training runs and parameterisation is a means of identifying individual training and prediction runs based on their specific production conditions. In this chapter, the POD-RBF NIROM is evaluated for predicting unseen cases of four reservoirs that are evaluated for constant production controls: synthetic, heterogeneous, water coning and real gas reservoirs.

The POD-RBF NIROM is then explored and evaluated for modelling the performance of 3 reservoirs with time-varying production controls: synthetic, heterogeneous and a real gas field. In the test cases in chapter 7, the NIROM is evaluated for modelling unseen cases with more practical production scenarios such as mixed production controls and well shut-in conditions, alongside the time-varying production controls. In chapter 8 the POD-AE is evaluated for modelling the heterogeneous gas field with varying production controls and boundary (shut-in) conditions in unseen/ prediction cases. The performance of the POD-AE is also compared to that of POD-RBF on their suitability for gas reservoir modelling.

The conclusions from this work and some recommendations for research and industry are discussed in chapter 9.

CHAPTER 2: LITERATURE REVIEW

This chapter provides some background into the conventional approaches for modelling gas reservoirs. The applications and limitations of these approaches are reviewed, leading up to reservoir simulation. Thereafter, some fast methods for modelling petroleum reservoir performance are reviewed, based on reservoir simulation as well as their research and knowledge gaps.

2.1 MODELLING GAS RESERVOIR PERFORMANCE

Several methods are used for estimating and forecasting the initial gas in place, recoverable reserves, and gas reservoir performance, including analogue, experimental and mathematical methods. Each method has its limitations and scope of application ^[18]. Analogue methods involve comparing a reservoir to other reservoirs with similar characteristics or properties while experimental methods are mostly applied for reservoir characterisation ^[18, 19, 45, 46]. Mathematical methods include the volumetric, material balance and decline curve analysis approaches ^[18]. The volumetric method relies on the static geological and petrophysical reservoir data to estimate the original hydrocarbon/ gas in place, GIIP ^[18, 46, 47]. The volumetric method does not provide any dynamic information about the reservoir's performance and is mainly used during the asset's appraisal stage, because information about GIIP becomes increasingly certain as production advances and more data is available ^[11, 12, 46-48].

Gas is highly compressible but has a low viscosity such that a product of its compressibility and viscosity is typically of the order of magnitude of 10^{-6}cp/psi (in field units). The low viscosity means that gas reservoirs have a high hydraulic diffusivity constant ($k/\phi\mu c$), where k is permeability, ϕ is porosity, μ is viscosity and c is compressibility. For this reason, gas reservoirs achieve pressure equilibrium quite rapidly and can often be treated as a tank model, if their permeability is greater than 1mD. This is a condition for the material balance approach and means that material balance methods are particularly applicable to gas reservoir modelling ^[10, 11, 49].

Material balance methods confirm estimates of the GIIP and help to confirm the drive mechanism of the reservoir by assuming that the reservoir has a uniform pressure and is homogenous. It is independent of the reservoir's size, geometry, varying rock and fluid properties such as fluid saturations and relative permeability ^[10, 11, 18]. There are 2 main material balance methods: the Havlena-Odeh approach and the p/Z plots ^[11]. In terms of reservoir volumes, the production, expansion and influx of a reservoir are captured in the material balance equation as:

$$\begin{aligned}
\text{Gas \& water Production} &= \text{Gas expansion} + \text{Water expansion/rock compaction} + \text{Water influx} \\
G_p B_g + W_p B_w &= G(B_g - B_{gi}) + GB_{gi} \frac{c_w S_{wc} + c_f}{1 - S_{wc}} \Delta p + W_e B_w
\end{aligned} \tag{2.1}$$

where the subscripts p, g, i , and w refer to produced, gas, initial and water respectively; G is the volume of Gas, W is the volume of water, B is the formation volume factor, c is compressibility, S_{wc} is the connate water saturation, W_e is water influx and Δp is the total pressure drop in the reservoir.

By adopting Havlena and Odeh's nomenclature where:

$$F = G_p B_g + W_p B_w \quad ; \quad E_g = B_g - B_{gi} \quad \text{and} \quad E_{fw} = B_{gi} \frac{c_w S_{wc} + c_f}{1 - S_{wc}} \Delta p$$

2.1 can be rewritten as:

$$F = G(E_g + E_{fw}) + W_e B_w \tag{2.2}$$

Practically, gas expansion is much greater than the expansion of water or the reservoir formation, i.e. $E_g \gg E_{fw}$, so 2.2 becomes:

$$F = GE_g + W_e B_w \tag{2.3}$$

Rearranging 2.3 further gives,

$$\frac{F}{E_g} = G + \frac{W_e B_w}{E_g} \tag{2.4}$$

Havlena and Odeh's approach involves plotting $\frac{F}{E_g}$ versus G_p from 2.4 to obtain a curve that looks like one of the curves in Figure 2.1. Figure 2.1 is a diagnostic plot of the Havlena and Odeh approach for determining the GIIP and drive mechanism of the reservoir ^[11]. The Havlena and Odeh approach is sensitive to the presence of water influx in gas reservoirs and is a better tool for establishing a reservoir's drive mechanism.

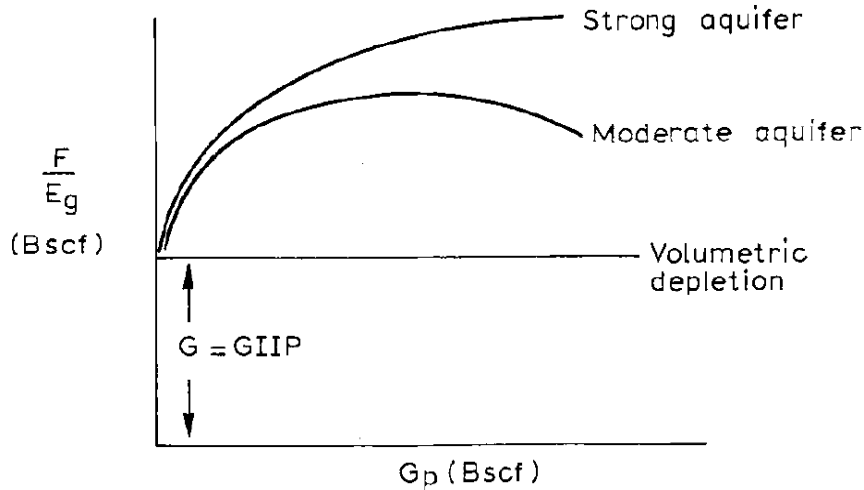


Figure 2.1: Havlena and Odeh's diagnostic material balance plot for determining the initial gas in place and the drive mechanism of a gas reservoir. The shape of the curve is an indication of the aquifer strength ^[11].

The other way of expressing the material balance approach is the p/Z plot. The p/Z plot approach relies on the equation of state (EOS) below:

$$pV = ZnRT \quad \text{which can be rearranged to} \quad V = \frac{ZnRT}{p} \quad 2.5$$

where P is pressure; V is the gas volume; Z is the dimensionless gas compressibility factor (Z factor); n is the amount of gas in moles; R is the gas constant and T is temperature.

For hydrocarbons, the gas expansion factor (E) relates the reservoir volume to the surface volume, and it can be expressed in terms of the EOS in 2.5:

$$E = \frac{\text{Surface Volume}}{\text{Reservoir Volume}} = \frac{V_{\text{surface}}}{V_{\text{reservoir}}} \quad 2.6$$

$$E = \frac{\left(\frac{ZnRT}{p}\right)_{\text{surface}}}{\left(\frac{ZnRT}{p}\right)_{\text{reservoir}}} = \frac{n\left(\frac{ZRT}{p}\right)_{\text{surface}}}{n\left(\frac{ZRT}{p}\right)_{\text{reservoir}}} \quad 2.7$$

$$\text{In field units (scf/rcf),} \quad E = 35.37 \left(\frac{p}{ZRT}\right)_{\text{reservoir}} \quad 2.8$$

The material balance equation for gas in a gas reservoir with no water influx (volumetric depletion drive), can thus be written as:

$$\begin{aligned} \text{Gas produced} &= \text{Gas initially in place} - \text{Gas remaining in the reservoir} \\ G_p &= G - \frac{G}{E_i} E \end{aligned} \quad 2.9$$

2.9 can be rewritten as:

$$E = Ei \left(1 - \frac{G_p}{G} \right) \quad 2.10$$

where E is the gas expansion factor and $\frac{G}{Ei}$ is the original or initial gas in the reservoir. At a constant reservoir temperature, 2.10 can be written in terms of the definition of the gas expansion factor in 2.8 to give:

$$\frac{p}{Z} = \left(\frac{p}{Z} \right)_i \left(1 - \frac{G_p}{G} \right) \quad 2.11$$

So if a gas reservoir is in fact undergoing volumetric depletion, then a plot of p/Z versus the cumulative gas production (G_p) should be linear such that when $p/Z = 0$, then the cumulative gas production (G_p) should be equivalent to the original volume of gas that was in the reservoir, GIIP^[11, 12]. If the plot of p/Z versus G_p is not linear, then the reservoir is undergoing water influx or some other mechanism such as compaction, extra gas being removed by a competitor or leakage from an adjacent compartment. The extent of deviation of the plot from linearity is an indication of the strength of the water influx/ aquifer, if other mechanisms are eliminated. In some cases, the p/Z plot needs to be extrapolated to determine the GIIP. The challenge with this approach is that if the plot is erroneously extrapolated due to a deviation from linearity in the plot that is not noticeable by inspection, then the reservoir could be assumed to be behaving as if it were volumetric and/or the GIIP can be overestimated^[11, 48]. Furthermore, Dake 2001^[11] indicates that if a volumetric depletion reservoir is produced at varying production rates and conditions to the extent that the linearity of its p/Z plot is altered, the gas reservoir could be incorrectly interpreted to be a water drive reservoir. Figure 2.2 shows examples where the linearity of the p/Z plot is altered by high or low production rates or seasonal production.

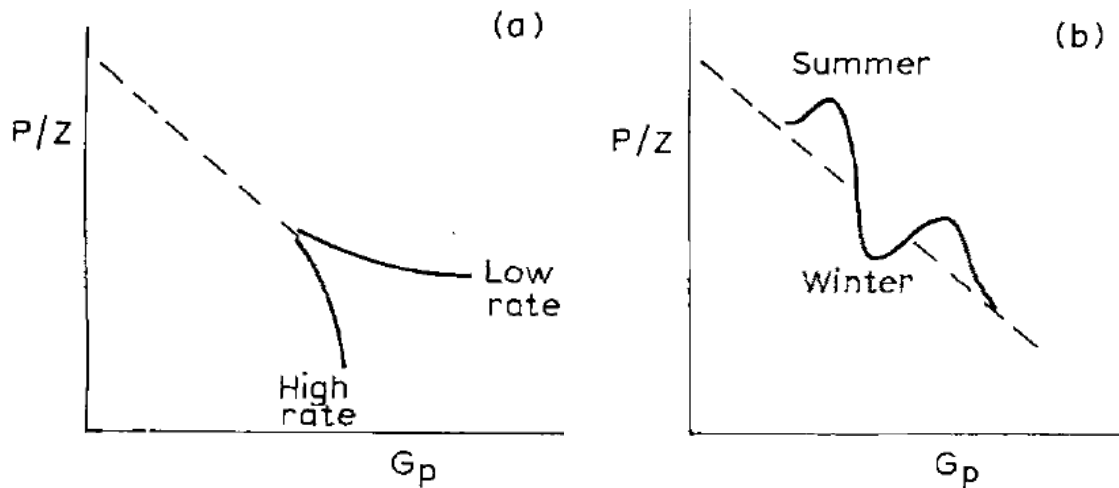


Figure 2.2. The linearity of p/Z plots affected by varying production rates a) in high and low production rate cases b) in a seasonally varying production case ^[11].

From the above review, it can be seen that the material balance equations make simplifying assumptions to satisfy volumetric conditions that may not always be valid. The material balance approach is sensitive to changes in the reservoir conditions such as water influx, gas expansion, multiphase flow in the reservoir, or varying production performance. This means that it can be used for quick assessments to see if there are other important mechanisms than just gas expansion in a gas reservoir. However the equations for the material balance approaches have also been modified to make them better adapted to different practical gas reservoir conditions ^[10, 12, 51]. These material balance methods also do not have consideration or reference to time.

Meanwhile, the decline curve analysis approach relies on the production data from a reservoir and does not take into account the reservoir's properties or geometry, neither does it give or provide information on the distribution of pressure and saturation in a reservoir. Decline curve analysis is based on the assumption that the trend of decline in a reservoir's production rate is constant, i.e. the reservoir's production rate would continue to decline in the future as it was in the past. However some reservoir operations such as well workovers and changes in production controls or drive mechanisms are sometimes unavoidable. Such changes can alter a reservoir's production decline pattern and make decline curve analysis impractical for managing the dynamic strategies of field development planning ^[10, 15, 18, 46, 52].

What is common to all the above traditional analytical methods of evaluating reservoir performance is that they are limited in the range of reservoir performance that they explore and they are not useful throughout a reservoir's lifecycle. These models do not account for the heterogeneity in real reservoirs and do not allow for understanding a gas reservoir's performance from changes in the distribution of pressure or gas across different parts of a reservoir ^[10, 18, 46]. All of these methods have underlying assumptions that simplify the complexity of their underlying equations, allowing quick solutions with

minimum computational requirements. They individually do not provide all the details that can be obtained from reservoir simulation, or the capacity to use different sources of information to solve and build dynamic reservoir models ^[10, 11, 18, 46].

Reservoir simulation is more computationally intensive than all the conventional methods for reservoir modelling. It involves solving the material balance equation in addition to the partial differential equations of Darcy's law that describe flow in porous media whilst including various physical behaviour such as phase behaviours ^[18, 46, 53]. Reservoir simulation requires engineers to combine reservoir data from different sources into a dynamic model that represents an actual reservoir. These dynamic models are made up of many grid cells that together represent the physical description of the actual reservoir. In reservoir simulation, each grid cell of a dynamic reservoir model is representative of a specific location in the actual reservoir, with its characteristic rock and fluid properties such as thickness, porosity, permeability, transmissibility, pressure and saturations. It can be difficult for engineers to get sufficient information about every location of a reservoir, besides the scale of measurement of some of the available reservoir data may not match the scale of the dynamic model. In such cases and based on the available data, different realizations of a reservoir's properties are acquired from geostatistical software. To deal with the uncertainty in these processes and techniques, dynamic models are calibrated by performing multiple simulations and fine-tuning the dynamic model until its performance matches those of the actual reservoir's production history. This is particularly important since the resultant dynamic simulation models are expected to be able to predict the actual reservoir's future performance, although dynamic models may not always give correct predictions even if they were well history matched ^[11, 18, 46, 53, 54].

Reservoir simulation is a versatile method for reservoir performance prediction and a realistic tool for field development planning. It can be updated as more data becomes available from the reservoir's operation and can be used to monitor a reservoir's performance throughout its lifecycle. Reservoir simulation models can provide information on the spatial distribution of reservoir properties at different time intervals, as well as other information that can help an engineer infer a reservoir's drive mechanisms, factors limiting recovery rates, the efficiencies of different production scenarios, in addition to sensitivities to uncertainty and operational risks ^[10, 11, 46, 53, 54]. During the life of a reservoir, reservoir simulation serves as a predictive tool for understanding the possible consequence of different development strategies on the reservoir, including when and where to perform well workover, drill infill wells, commence EOR and plan for abandonment. Information from reservoir simulation is also helpful to support decisions on gas deliverability, gas sweep, surface facilities design and the net present value of an asset ^[10, 11, 15, 18, 46, 54]. It gives engineers an idea of what could be

expected from a field during its production life and hence the opportunity to prepare and plan in advance ^[15, 17].

Nevertheless, reservoir simulation is a time-consuming technique to set up and run ^[10, 13, 17-20, 55]. Reservoir simulation takes into account the many grid cells of a reservoir's dynamic model in addition to the wells and field-wide performance, so depending on a reservoir's complexity and the type of information that a model is built to study, reservoir simulation can be time-consuming. This is particularly significant when many simulation runs are required during history matching, sensitivity and uncertainty analysis as well as for production optimization ^[32, 33, 35, 36, 43, 53, 56, 57]. For instance, models that represent reservoirs with multiphase flow, enhanced oil recovery, retrograde gas, natural fractures and unconventional reservoirs typically require more computational resources (time and memory) ^[13, 16, 18, 21, 46, 55, 58].

Technological advancements as well as new exploration frontiers such as CO₂ sequestration and emerging technologies for cleaner hydrocarbon extraction continue to promote the role of reservoir simulation. In fact the technological advancements in computing capacity mean that more sophisticated reservoir models can be considered ^[53]. Given the benefits of reservoir simulation in executing elaborate field development plans along with its accompanying computational challenges, it is worthwhile to investigate complementary methods for speeding up its runtime. These complementary methods should ideally improve the time spent in building and running dynamic models. This way, engineers can spend more time reviewing the practicality of the dynamic models as well as the development strategies that are considered based on the dynamic model's solutions of the reservoir's performance ^[17].

2.2 FAST MODELLING METHODS

Reservoir simulation solves the equations of flow in porous media in the many grid cells that represent a reservoir ^[49, 53]. The number of grid cells that are used to build a reservoir model and describe the reservoir in space is determined by the reservoir complexity and heterogeneities in addition to the size of computer and time required for simulation. Since reservoir simulation solves the equations of flow in porous media in each grid cell, then the more cells there are in a reservoir model, the more the simulation will be computationally demanding ^[32, 33, 35, 36, 43, 49, 56, 57]. Hence in the petroleum industry, as in many other fields of science and engineering, various methods are applied to creating approximate models that represent complex systems. These approximate models are subsequently used to forecast the system's performance in less time and with less computational capacity.

Some of the methods that have been applied to improve the computational speed of solving complex problems include using coarser simulation grids/ upscaling ^[55, 59-61], developing alternative methods for solving the equations of flow, such as streamline simulation ^[61-63], manipulating the iterative solver methods employed in reservoir simulators ^[64, 65], incorporating GPU technologies ^[57, 66, 67], projection-based reduction techniques ^[33, 40, 68-70], neural networks or proxy models ^[21, 46, 71, 72] and cloud sourcing/ computing ^[31, 73, 74].

The use of coarse grids without upscaling can lead to significant numerical dispersion and optimistic predictions of sweep due to homogenization of the geological heterogeneity ^[17]. Upscaling approaches can overcome these difficulties but rely on further fine-scale (and hence time-consuming) simulations, besides, they are often complex to apply and solutions are specific to particular initial and boundary conditions ^[40, 59, 61, 75]. Streamline simulations are only accurate for applications with limited pressure depletion and so are not so suitable for gas reservoir simulation. GPUs perform better than CPUs in terms of computational time, but they require associated algorithms that are better implemented in high-level languages and data-parallel structures in addition to in-depth technical expertise to harness their implementation to the fullest ^[57, 66].

Projection-based reduced order models (ROMs) project large scale problems into lower-dimensional subspaces that approximate conventional simulators ^[39-41, 76, 77]. ROMs have been applied in aerodynamics and aerospace ^[26, 28, 29, 78, 79], physical and computational sciences ^[22, 23, 25, 27, 80-87], engineering ^[88, 89], medicine and the health sector ^[24, 90] for fast, flexible and less expensive computations. In the last couple of years, ROM has been explored in the petroleum industry. Most applications of ROM in petroleum engineering have been for water flooding in oil reservoirs ^[33, 43], multiphase flow in porous media ^[91, 92], history matching ^[36, 41, 56, 77] and production optimization ^[33, 76, 93].

Recently projection based reduced order models have been developed based on proper orthogonal decomposition (POD) which transform high dimensional data into a lower-dimensional space. In 2011, Chaturantabut and Sorensen ^[69] applied a POD based discrete empirical interpolation method (DEIM) for modelling viscous fingering in a 2D system with 15000 full order dimensions. They used POD to reduce the dimension of their system to 40 and applied DEIM for approximating the non-linear function of their system. Chaturantabut and Sorensen (2011) found that POD-DEIM was efficient in capturing the dynamics of viscous fingering with a computational speedup of up to 3 orders of magnitude. He and Durlofsky (2014) ^[32] applied a POD based trajectory piecewise linearization (TPWL) technique for compositional simulation of a 2 phase oil and gas system i.e. gas injection into oil and gas reservoirs. To minimise the amount of storage space required to implement their POD-TPWL, He and Durlofsky (2014) applied an intrusive approach that requires modifying the conventional

simulator from which they obtained their training simulations/ snapshots. They also highlighted the complexity of specifying practical well controls, such as time varying BHPs, in reduced order modelling and propose that the range of controls applied to training cases should be similar to those for the prediction runs. Their intrusive POD-TPWL method gave speedups of over 500 times compared with the Stanford General Purpose Research simulator (GPRS) (Figure 2.3) and was also applied for a production optimization problem.

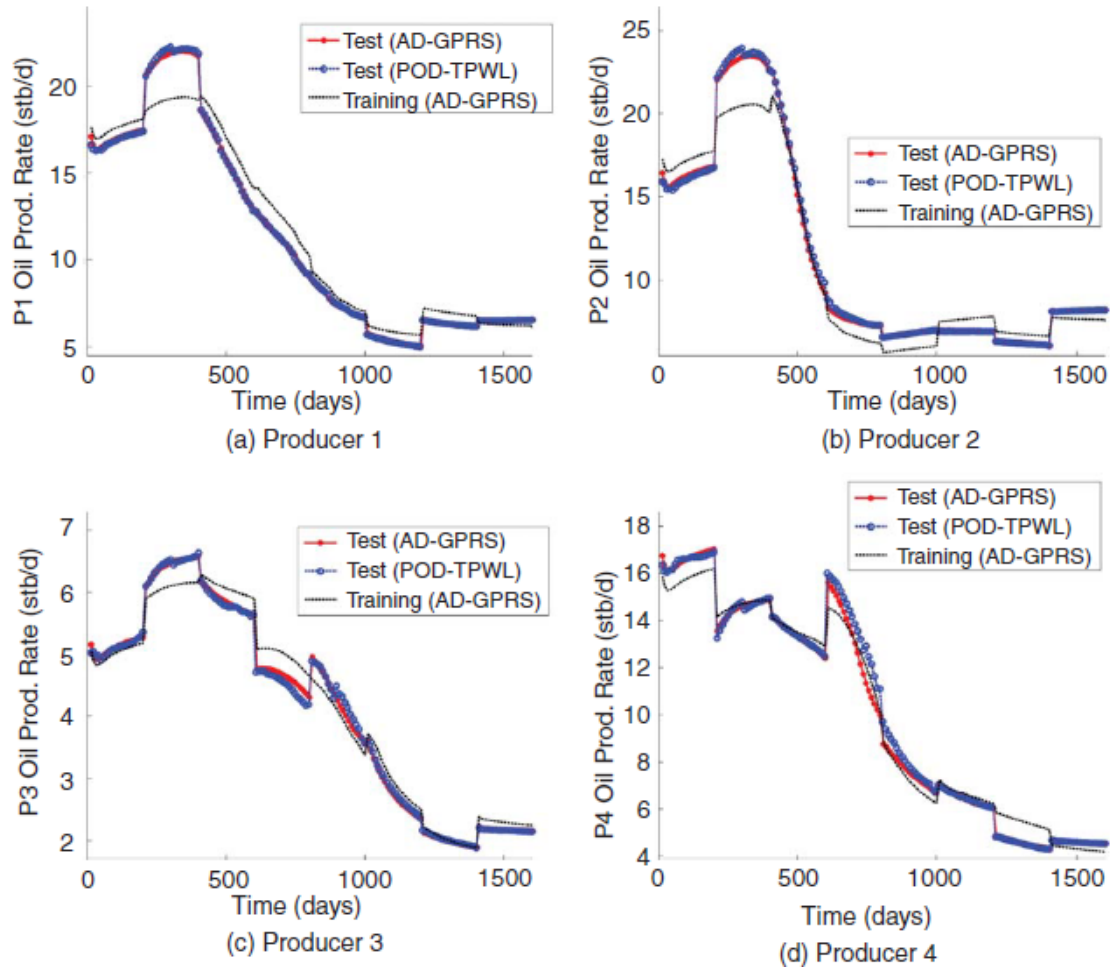


Figure 2.3: Forecasts of oil production rate in a test case of He and Durlofsky et al. (2014) using the POD-TPWL method. Their POD-TPWL is an intrusive ROM that relies on some modification of the reservoir simulator ^[32].

In 2016, Ghasemi and Gildin ^[94] noted that POD models suffer from computationally expensive back-projection especially when evaluating nonlinear systems with large data. They also mentioned that POD-DEIM may require a large number of basis functions with no guarantee of a stable system, while TPWL is computationally expensive in terms of the space required to store its data. So Ghasemi and Gildin (2016) proposed a quadratic bilinear formulation for modelling the saturation equation of a 2-phase flow oil-water model undergoing waterflooding. Their approach was applied to nonlinear systems without any approximations like in DEIM, it initially increases the dimension of the problem

before reducing the system with POD. They believed that this technique improved the POD basis function selection and the prospects of capturing the dynamics of nonlinear systems while preserving some properties of the system like its stability. However, their approach depends on the governing equations of a system, which they transform into a new quadratic bilinear form. Jiang and Durlofsky (2019) ^[35] have also modelled oil-water subsurface flow with a POD based Gauss-Newton with approximated tensors (GNAT) method. GNAT approximates the non-linear terms of a system like DEIM, but GNAT allows for the number of interpolation points to be different from the number of basis functions used for its approximation. Jiang and Durlofsky (2019) ^[35] found that POD-GNAT was more accurate than POD-TPWL in predicting the production rates of oil and water (Figure 2.4), but it gives larger errors as the non-linearity in the system increases and it is not as fast as POD-TPWL. They also suggested that the GNAT stage of their method can inherit the inaccuracy of the POD stage.

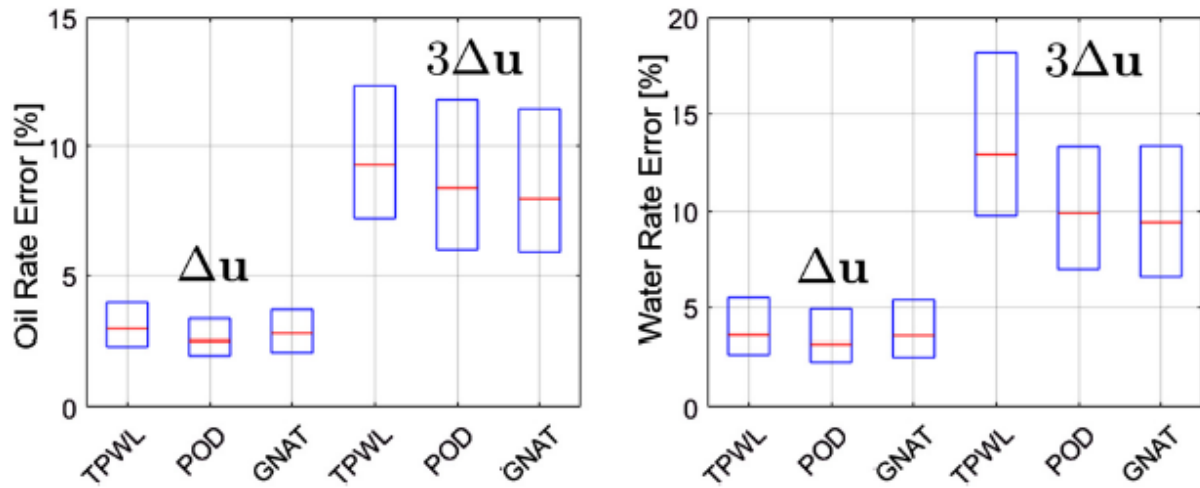


Figure 2.4: Jiang & Durlofsky (2019) ^[35] box plots comparing the errors obtained from POD-TPWL, POD-Only and POD-GNAT for 495 original test schedules (Δu) of their test case and cases with large perturbation schedules ($3\Delta u$), showing that GNAT gave better results than TPWL for predictions of oil and water rates.

In all the above cases, except for Chaturantabut and Sorensen's method ^[69], the ROMs either relied on the governing equations of the reservoir or the conventional simulator had to be modified to accommodate the ROM formulation, these are called intrusive ROMs. However most commercial reservoir simulators that are widely used in the industry do not allow access to the source code or the algorithms for their simulators, for proprietary reasons. Some other simulators have models that might not be easily changed or altered to accommodate intrusive ROMs ^[21, 36]. Such restrictions in commercial simulation software limit the use of intrusive ROM techniques for simulation speedup. Besides intrusive ROMs are not always stable or computationally efficient when producing solutions for nonlinear problems involving the inversion of a large matrix such as the spatial distribution of reservoir properties.

There are however some non-intrusive (data-driven) reduced order models (NIROMs) that do not need the governing equations of a system and do not require any changes to the commercial simulator that produces their training runs. NIROMs are standalone programs that are “trained” using physical and numerical data including output solutions of commercial simulators ^[36-38]. They can be used in conjunction with any commercial simulators, making them readily available to practising engineers irrespective of their preferred simulation software.

In 2013, Klie ^[36] highlighted that since NIROMs do not depend on the proprietary codes of commercial simulators, their predictability is dependent on the amount of training data that is available for them to learn the system’s dynamics, i.e. NIROM needs more training data than intrusive ROMs. Klie (2013) proposed a NIROM based on POD, DEIM and radial basis functions (RBFs), where POD or DEIM was used to reduce the size of the training data (for the linear and nonlinear terms of the system respectively) allowing for a compact RBF approximation. He emphasised that RBFs exploit data locality when they approximate the relationship between data input and output, and RBFs have lower sensitivity to noisy data as well as lower training speed compared to artificial neural networks (ANN). Using the Gaussian basis function, Klie (2013) applied his NIROM for predicting the gas and oil production rates from two “simple” multiphase reservoirs as well as their pressure and oil saturation field with a maximum of 3% approximation error. In the same year, Chen et al. ^[56] also proposed a black-box stencil interpolation method (BSIM). Similar to Klie 2013 ^[36], their training data was reduced with POD or DEIM but their interpolation was done with an ANN that had 50 nodes in its hidden layer. Their BSIM takes advantage of the discretisation approach applied in most simulators where the flow dynamics in any given grid cell has local support from its surrounding grid cells in all directions, stencil locality (Figure 2.5). The local relationship between grid cells is by far less complex than the flow dynamics in the entire reservoir, therefore Chen et al. (2013) ^[56] used stencil samples from their POD or DEIM reduced dataset as input data for their ANN. They tested their model on oil-water and oil-gas cases with errors of around 10%, which they likened to what is obtainable from intrusive ROMs. Their test cases were oil reservoirs with few grid cells or short production durations, i.e. reservoirs with 10000, 300 and 13200 grid cells, evaluated for 500days, 12years and 50days, respectively.

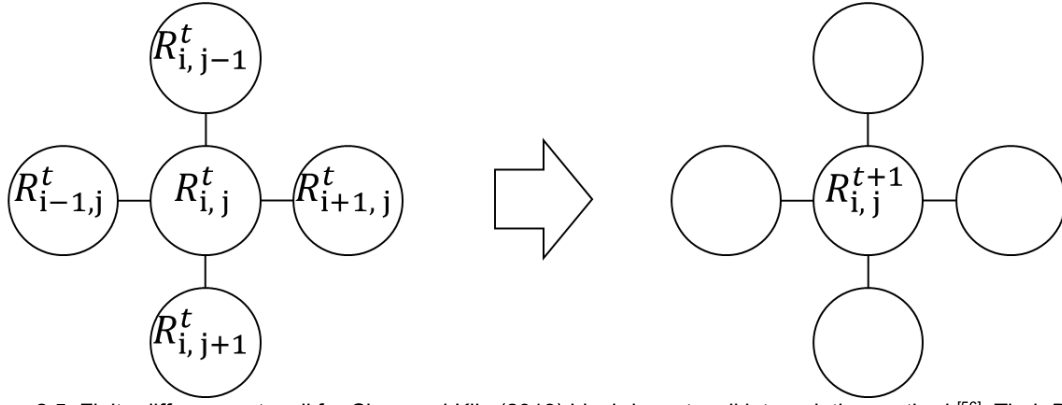


Figure 2.5: Finite difference stencil for Chen and Klie (2013) black-box stencil interpolation method ^[56]. Their BSIM exploits a similar discretisation approach as most simulators but was not tested for sufficient simulation scenarios (size and duration).

From 2015 to 2017, Xiao et al. ^[38, 91, 92, 95, 96] proposed three NIROM methods, based on POD. These methods are the POD-second order Taylor series expansion approach, POD-Smolyak sparse grid interpolation method, and POD-Radial Basis Function (POD-RBF) method. They proposed that the POD-RBF technique was a method of choice in speeding up simulations without sacrificing the accuracy of modelling results. Their POD-RBF NIROM models a time-dependent solution and they showed its application to solving problems in unstructured grids, i.e. the Navier-Stokes equations in a 3D non-hydrostatic system and for multiphase flow in porous media with the Imperial College Finite Element Reservoir Simulator ^[97]. They considered multiphase (oil-water) flow simulations undergoing water flooding, which were also simplified to have homogeneous porosity, two permeability regions and with no gravity considered. Xiao et al.'s test cases had 984, 1386 and 31776 dimensionless data points (grid cells) and they showed that using more than 9 POD basis functions (18 or 36 at the most) improved the accuracy of their solutions by more than 50% while adequately capturing the flooding saturation shock front (Figure 2.6). Although they had speed ups of 3 orders of magnitude, their model was built using output from a research simulator that is not commonly used by practising engineers. Also, their test cases did not consider practical aspects of field development planning such as production from heterogeneous reservoirs with real PVT properties or the application of NIROM for modelling different production scenarios with different production controls ^[38, 91, 92, 95, 96].

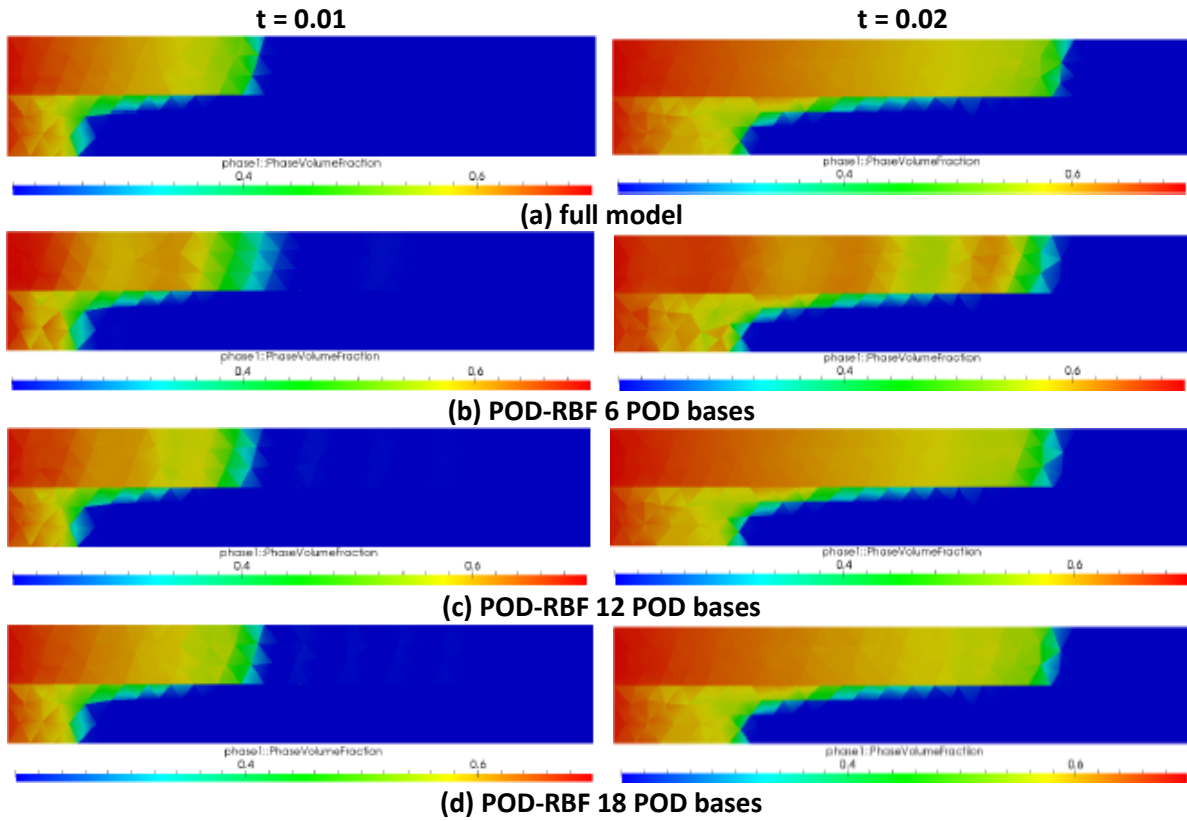


Figure 2.6: Xiao 2016 ^[38] POD-RBF estimation of the saturation in a two-material layers case showing solutions with 6, 12 and 18 POD bases functions.

In 2019, Kostorz et al. ^[37] extended the application of POD-RBF to modelling new or unseen simulation cases, in addition to designing the POD-RBF NIROM to accommodate being trained with multiple simulation parameters (parameterisation) simultaneously. Their NIROM was applied to a 2D oil reservoir undergoing gas injection and they proposed the use of RBF and polyharmonic splines as good interpolation methods. They further highlighted that other alternative interpolation methods like machine learning and kriging require more computational time and process optimisation to give similar results as RBF or polyharmonic spline methods. On the relevance of parameterization, they noted the importance of “renormalizing” the training data so that even though the multiple parameters being considered for the same system have different magnitudes, renormalisation enables all the parameters to be brought to a similar regularly spaced structure. Kostorz et al.’s parameterisation approach however caters for only constant, time-invariant parameters i.e. multiple simulations with different constant parameters like oil-gas density ratios, directional permeability, reservoir dip, or a single production rate throughout the production lifecycle. Their parameterisation design did not consider simulation cases with multiple time-varying parameters (forced or varying parameters) and how to appropriately renormalize these for NIROM interpolation. This is particularly relevant in the petroleum industry where for instance a reservoir can be produced at several different production rates during its production lifecycle. For swift experimentation and development, Kostorz

et al.'s NIROM was implemented for a simple reservoir with 1000 grid cells, and with an in-house finite difference implicit pressure explicit saturation porous media flow simulator^[98]. They highlighted that the NIROM does not adequately model complex physical behaviours like displacement problems with a shock front. The NIROM's accuracy did not improve when the number of training runs was doubled (Figure 2.7 and Figure 2.8).

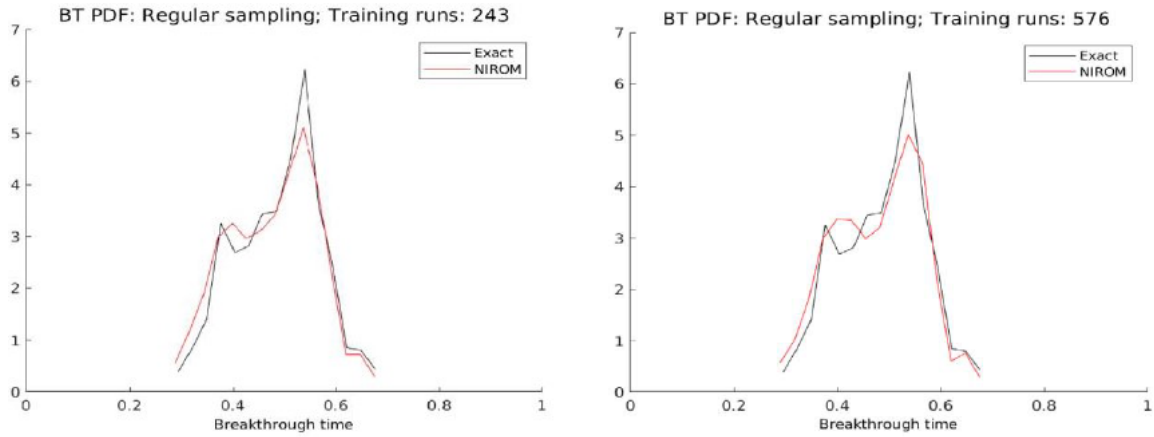


Figure 2.7: Kostorz et al.'s (2019)^[37] breakthrough time (BT) PDFs generated from their simulator (Exact) and NIROM. Note that in both cases above, the NIROM's prediction of the peak BT PDFs are different, the case with fewer training runs is closer to the exact solution.

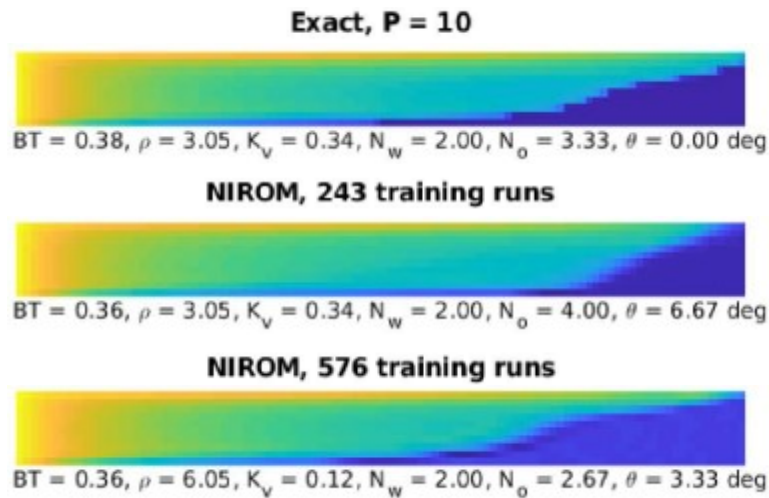


Figure 2.8: Kostorz et al.'s (2019)^[37] NIROM approximations of a simple reservoir's P10 saturation distribution. Kostorz et al. showed that having larger numbers of training runs does not guarantee having better NIROM approximations.

All the above applications of NIROM did not consider gas-water (that is with aquifer support) or gas-only (that is with very weak or no aquifer support) reservoirs, and they all applied their NIROMs to simple (homogeneous and impractical) or small (less than 11×10^3 acre-ft) reservoirs. In some cases, the proposed NIROMs have been used to replicate their training runs^[38] and not new unseen prediction cases. And in other cases^[37, 99, 100], the test reservoir simulators are not commercially available to practising engineers. None of the above authors provided a NIROM framework that is easy

to implement, neither did they consider realistic well controls for their test cases. They did not provide insight into how they came about their choice of RBF or compare the speed-up of their NIROM to their simulators (in relatable time). Nevertheless from the above works, the POD-RBF has repeatedly been shown to have potential for fast approximation of petroleum reservoir simulations, POD-RBF is the main NIROM that is considered in this work.

2.3 MACHINE LEARNING METHODS

Machine learning (ML) is an application of artificial intelligence ^[101, 102]. Like most statistical methods, machine learning techniques identify the underlying patterns in data and approximate these patterns in new instances and on a larger scale ^[19, 103]. ML approaches work well with data that exhibit the non-linearity and complex behaviours that are characteristic of petroleum reservoirs, for instance time-varying well controls ^[13, 104]. They also work well for problems with noisy and incomplete data such as in image reconstruction ^[13, 20, 42, 54, 104, 105]. Artificial neural networks (ANN) are a type of algorithm in deep learning, which is a subdivision of machine learning. Precisely, any ANN with more than 3 layers (an input, output and hidden layer) is a deep neural network ^[101, 102, 106]. The idea of neural networks is similar to RBFs since they both involve a network of layer(s) interconnected with interpolation or activation function(s). In supervised learning schemes, neural networks can learn the behaviour of the initially given data in a cause and effect manner and can eventually be used as a prediction tool, for making forecasts ^[20, 42, 54, 105]. Other artificial intelligence models include capacitance-resistance models ^[42], fuzzy logic systems ^[42, 105, 107], genetic algorithms ^[42, 105, 107] and support vector models ^[42, 105, 107].

There are two common types of problems that NNs are applied to, classification and regression problems. In classification problems, prediction outputs are grouped into predefined categories based on whether or not they meet or pass certain quotas e.g. binary or multiclass classification problems where final prediction outputs are often probabilities that sum up to unity or percentages. Regression problems have individual or subjective outputs ^[105, 106]. Since the problem in this work requires outputs for individual grid cells and/or at different time instances, it is a regression problem. It is important to be aware of the problem type in order to make informed decisions when designing a neural network (NN). For instance, the problem type influences the choice of activation functions, loss functions and error metrics ^[105, 106, 108]. Sigmoid, softmax and hyperbolic tangent activation functions are often used for classification problems while rectified linear unit (ReLU) and its associated functions such as Leaky ReLU and exponential linear unit (ELU) are applied to regression problems. The sigmoid function always gives outputs between 0 and 1, softmax outputs always add up to unity and the hyperbolic tangent function gives outputs between -1 and 1 ^[21, 106, 109].

Hyper-parameters are also an important part of neural network models. Hyper-parameters are design components/ configurations that make up the NN, they are predefined before training the model. These include the optimiser algorithm, batch size, epoch, dropout, activation function(s), number of hidden layers, and number of nodes in each hidden layer. The epoch is the total number of times that the full training data is passed through during training, the batch size refers to the size of a subset of the full data that is considered at every instance during training, while the optimiser is the method by which the NN's weights are improved/ updated during iterations ^[109-111]. The topology or architecture of the NN (number of hidden layers and number of nodes or units per hidden layer) also needs to be set up in advance of training the model. The effect of the number of hidden layers on a NN's generalisation is shown in Figure 2.9 ^[111]. The most suitable choice of hyper-parameters for NNs are often arrived at by trial and error to minimise their prediction error ^[20, 105, 106, 111].

Whilst the hyper-parameters of NN models are tuned to obtain the most suitable model, the standard ML practice is to also provide the model with validation data. This validation dataset is different from the data that is used to train the model as well as the “test” data that the NN model would later be used to forecast. During the NN's training process, the validation data is used to gauge the NN's predictive capacity and for continuously adapting the NN to be able to generalise on “new” data until a robust predictive model with minimised estimation error is obtained ^[19, 58, 105, 111].

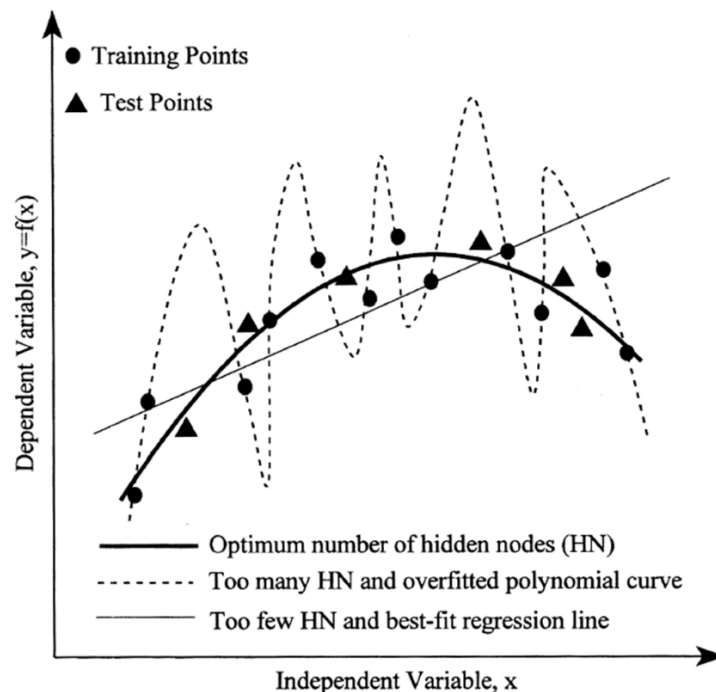


Figure 2.9: Effect of a hyper-parameter (number of hidden layers) on a NN's generalisation ^[111].

Recently ML techniques are increasingly being applied to solving computational tasks in the petroleum industry. While significant achievements have been made in applying ML in the petroleum industry

since the 1990s, these applications have been limited to certain aspects of petroleum engineering ^[45, 105, 112, 113]. For instance, ML has been applied to modelling enhanced oil recovery (such as water, gas or steam flooding) ^[13, 42, 44, 46], estimating the flow rate of gas ^[114] and identification of flow types and pressure gradients in multiphase flow situations ^[109]. When coupled with optimisation techniques, ML has also been used to model the optimum net present value of petroleum exploration and production projects ^[16, 54, 73].

Significantly, because there is a large amount of output data in the petroleum industry, this has fostered the application of machine learning in the industry. But despite several works that show that ML is useful in the petroleum industry, more still needs to be done to optimize and integrate these techniques into mainstream field development operations ^[103]. Also, only a handful of these applications have considered the prediction of reservoir dynamics in terms of pressure and fluid saturation distributions in entire reservoirs or fields. Similar to reduced order models, there have been even fewer works that have considered production dynamics in gas reservoirs, to help engineers reduce the time it takes to arrive at quality decisions for developing hydrocarbon reservoir assets, well placements or sensitivity of field development strategies to production rates and wellhead pressures ^[36, 45, 115]. What is common to most of these ML applications in petroleum engineering is that they have been tested on problems for oil-water or multiphase reservoirs. Some notable ML applications are discussed below, especially applications to gas reservoirs and some applications that provide more insight into some challenges of ML.

In 2014, Al-Dhamen ^[116] used ANN to predict gas rates in a 2 phase condensate reservoir. Their focus was on creating a tool that can better estimate gas rates for a wide range of condensate gas ratio values. Their model was trained with data from well tests such as choke sizes, upstream and downstream pressures as well as condensate gas ratios. Al-Dhamen's model had a 10.2% error and was better than analogous correlations, but their work did not delve into the 3D dynamics of the gas reservoirs or forecasting reservoir property distribution. Ersahin et al. (2019) ^[20] applied ANNs to forecasting the oil production profiles and oil viscosity of cyclic steam injection operations in a dual-porosity reservoir. They highlighted that in building the architecture for neural networks there is no rule of thumb for determining the number of neurons per hidden layer (HL) and the heuristic process is relied on for fine-tuning models to give better results. They evaluated their models' performance on a reservoir with a single well and radial grid geometry, after training the model with 1428 training cases.

Gas rate is typically not measured for each well in field settings, it is often evaluated from surface multiphase flowmeters and other less direct methods such as well test analysis. So in 2019, Kalam et al. ^[114] developed and evaluated 3 artificial intelligence methods for confirming gas rate calculations

from correlations and well tests. Kalam et al. (2019) ^[114] used the tan-sigmoid activation function for the HL of their feedforward neural network (FFNN) and linear function for the output layer. Their artificial neural network model had 17097 data points for training and 4846 data points for testing as well as 19 neurons in the only hidden layer. Although they achieved above 90% accuracy from the model and worked with over 20000 data points, their methods' evaluation was for estimating gas rates and does not convey other engineering implications for decision making or development strategies that accompany the dynamics of the dry gas field. Wood (2020) ^[117] investigated the transparent open box data matching algorithm to predict the effective permeability, effective porosity and water saturation in a giant Algerian gas reservoir, from the well log data. Their model gave useful insight into the lithology of a 100m reservoir and was considered to be fit-for-purpose in the petrophysical metric prediction. They addressed a reservoir characterisation problem that does not have direct implication for the reservoir's future performance.

Probably the most relevant application of NNs to this work is the work of Temirchev et al. (2020) ^[13]. Temirchev et al. (2020) developed a deep learning technique called a metamodel, which is based on a convolutional variational autoencoder and a recurrent neural network (RNN) with gated recurrent units (GRU). The metamodel was applied to single-phase oil and gas reservoirs as well as a multiphase production case with and without water injection. Their reservoir cases had a single well and they were horizontally homogenous with vertically varying permeability. Most interesting is that they went beyond predicting the reservoirs' cumulative hydrocarbon production for 30 years, to applying the metamodel to predicting the 3D pressure and saturation dynamics within these reservoir cases. They also did a comparative analysis of 2 main dimension reduction techniques, proper orthogonal decomposition (also known as principal components analysis - PCA) with 120 and 150 basis functions and autoencoders (Figure 2.10).

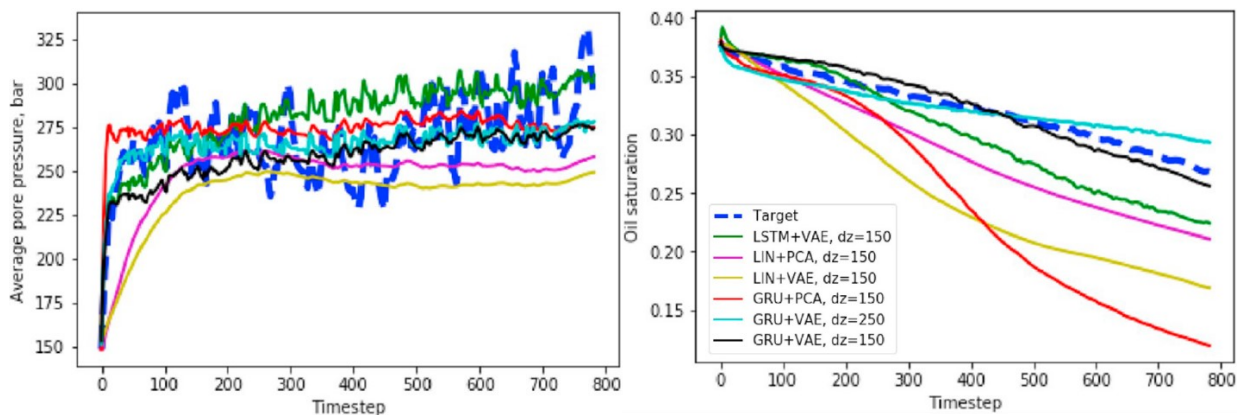


Figure 2.10: Temirchev et al.'s (2020) ^[13] comparison of different metamodels' prediction of the average values of the spatial pore pressure and oil saturation for a scenario of their test case. In this example, GRU means Gated Recurrent Unit, PCA means Principal Components Analysis, VAE means Variational Autoencoder, LIN means Linear Regression and LSTM means Long Short Term Memory.

The convolutional neural network part of the metamodel was used to compress their input data into a lower-dimensional vector structure while the RNN part was used to implement the actual interpolation of the reservoir's dynamics through time. The metamodel's input data consisted of grid distributions of pressure and fluid saturations within the reservoir as well as production bottom hole pressure and water injection rate. To mitigate the high computational time spent in inverting latent space predictions to full space, they only recovered the full-order metamodel predictions at time steps of interest. They had promising results from their work and showed that using GPUs made their predictions 12 times faster as opposed to 7 times speedup with an ordinary CPU. They did not specify if this speed up was from their technique or solely due to their hardware. In addition, they indicated the need to scale up and adapt the metamodel for real field cases that are fully heterogeneous, with more complex fluid systems and PVT representations, as well as for injection of other fluid types in addition to low permeability reservoirs.

From the ML applications seen so far, it is quite obvious that some of the downsides of machine learning include understanding the hyper-parameters (pre-defined model configuration parameters) as well as how to effectively fine-tune them and striking a balance between having abundant data for training a model without overfitting the model. There is also the issue of dealing with high dimensionality data (the curse of dimensionality) and getting stuck at a local minima as opposed to the global minima when training the model. The most significant challenge with ML is the computational capacity required to run these models since standard RAM and CPUs do not always suffice for running deep networks. Some of these ML issues are mitigated by creating hybrid models of machine learning with optimisation or feature selection techniques, applying stochastic gradient descent methods, introducing model reduction techniques, adding regularization parameters to models and increasing computational capacity via GPUs, parallel and/ or cloud computing ^[19, 55, 105].

SUMMARY

This chapter has shown that natural gas production and exploration constitutes an important/ integral part of the global energy mix and reservoir simulation is the most widely used and general tool for solving the dynamics of gas reservoir's nonlinear flow equations, even though it is time-consuming. For industrial applications, the POD-RBF NIROM and machine learning interpolation techniques constitute a scope of practical solutions for creating fast approximations of reservoir simulation solutions. The POD-RBF NIROM can be used by practising engineers irrespective of their preferred simulation software and it can be implemented for multiple parameter scenarios. Machine learning interpolations can also be applied to the time-varying production scenarios that are practised in the petroleum industry. Specifically, the rest of this work focuses on applying POD-RBF NIROM and POD

coupled with Autoencoder to different practical aspects of gas reservoir simulation and management in real gas reservoirs, especially the spatial distribution of pressure and saturation in the reservoirs.

CHAPTER 3: THEORY AND IMPLEMENTATION OF NIROM

This chapter describes the theory and workflow for implementing the NIROMs considered in this work, the proper orthogonal decomposition radial basis function (POD-RBF) and the proper orthogonal decomposition autoencoder (POD-AE) methods. In both frameworks, POD is used to project the reservoir simulation data into a lower-dimensional hyperspace and RBF or AE is used to interpolate or approximate the dynamics of the data in hyperspace. Subsequently, POD is applied to transform the data back to the full space where it can be compared with the commercial simulator.

The NIROM workflow that is applied in this work is based on the approach that was developed by Kistorz et al. (2021) ^[100] but with a “renormalisation” that this work finds to be suitable for time-varying (forced) parameters. Kistorz et al.’s (2021) ^[100] “renormalisation” method is better suited to constant parameter cases. The POD part of the workflow is the same for both NIROMs and is fully implemented in MATLAB ^[118] while the RBF interpolation is implemented in MATLAB and the AE interpolation is implemented in Python, Jupyter Notebook ^[119]. Hence, the code that is used for this work is written in MATLAB and Python, which are interpreted programming languages ^[118, 120]. These codes are available in Appendix C.

Different production scenarios have been considered in this work, such as constant and time-varying (forced) production parameters in addition to different well controls (boundary conditions). The impact of these parameters on the implementation of the NIROMs is shown in the following sections. Two types of simulation solution data have also been considered, the spatial distribution of pressure and saturation in reservoirs as well as the production performance data for the reservoirs. The NIROM workflow is explained for both data types.

3.1 OVERVIEW OF APPROACH

For any given test case or gas field example, the first step is to define the key parameters that will be varied during the study and the ranges over which these parameters will be varied, in addition to the output variables that will be estimated by the proxy models. Input parameters may be associated with the reservoir characteristics such as porosity, permeability and fault distributions or with the development strategy and operation of the field such as well placement(s) and well controls. In this work the following parameters are considered: gas production rate, well bottom hole pressure (BHP), tubing head pressure (THP) and reservoir porosity. The output variables (outputs) will typically be pressure and saturation distribution over time and the production data (gas rate, water rate, gas-water ratio and/or BHP). The pressure and saturation distributions will be estimated in each grid cell that constitutes the gas reservoir’s dynamic model. The engineer also needs to specify the time

intervals at which output variables are written out by the conventional simulator (time steps or report steps), these can equally spaced in time.

The next step is to build the base reservoir model as required for the conventional reservoir simulator of choice. Solutions to flow in porous media from the conventional reservoir simulator are taken as the true/ original representation of the actual reservoirs. In this work, a finite difference based commercial simulation software, Schlumberger's ECLIPSE black oil simulator (ECLIPSE) ^[121], is used to build the dynamic models of the gas reservoirs. The proxy model (NIROM) is expected to learn the base reservoir model's dynamics from scenarios of ECLIPSE simulation output and subsequently be used to make predictions.

In this work, the NIROM is implemented as a coupled model (NIROM and conventional simulator together), so the approximate model/ NIROM is set-up to call and run ECLIPSE from within the programming language of choice, thereby generating the training data. Depending on the user's preference, the approximate model can be configured to write INCLUDE files (ECLIPSE text files) with the adjustable parameters (such as production profiles) and report step intervals for ECLIPSE, before each ECLIPSE simulation run. Other fixed simulation parameters/ information are built into the base reservoir model in advance. In this work, the entire POD-RBF NIROM is built and implemented in MATLAB while most of the pre- and post-processed data for the autoencoder (AE) model is done in MATLAB. This was to facilitate rapid development and testing. If NIROM were to be used for real applications, then it would be more efficient to write the code in a compiled language such as C++. The autoencoder model itself is built, trained and used for predictions in Python's Jupyter Notebook ^[118-120].

In the implementation of the NIROM in this work, the next step is for the approximate model to select the value(s) of the adjustable parameters from the range of values that the user has provided, then it calls the conventional simulator (Schlumberger ECLIPSE) to run the reservoir simulation case with the chosen parameter. The approximate model in turn reads the simulator's output solutions and saves them within its memory space and moves on to the next parameter to repeat the same process. When all parameter cases have been run for the total number of training runs specified in the approximate model, the model couples the output data from all the training cases together to form what is hereafter referred to as the approximate model's training dataset.

The next step is to train the fast approximate model (NIROM) with the training runs data that was extracted and saved in its memory space. The extracted training data is normalised by mean extraction and projected into a lower-dimensional hyperspace with proper orthogonal decomposition (POD). Proper orthogonal decomposition projects the data into a lower dimensional space but it is expected

that the reservoir's dynamics is retained within the reduced training data in the new hyperspace. This reduced data is then interpolated in hyperspace using radial basis functions (RBFs) or autoencoders (AE), to learn the reservoir's dynamics as interpolation weights. These weights are used for the prediction of new unseen simulation cases. For the new unseen cases, only the initial condition data (data at the first time instance) of the new unseen simulation cases as well as the parameters for the new simulation case serve as the input data for RBF or AE predictions. All these steps are executed within the coupled NIROM-ECLIPSE model.

The next step is to run predictions. In this work, the prediction steps are implemented as a continuous step after training the NIROM, so within the coupled set up new sets of parameters are chosen and ECLIPSE is called to run simulations for the new parameters. The difference with the training cases is that only the ECLIPSE's output data of the first report step (initial condition) would be saved and used for prediction. This initial condition data is the starting data for NIROM prediction and it is propagated through time with NIROM for the expected production duration with RBFs or AE. After the RBF or AE prediction in hyperspace, the NIROM predicted simulation data is projected back into the full space to be compared with the full ECLIPSE simulation of the new unseen case. The accuracy and the speed of the fast models/ NIROMs are then analysed. The above workflow is discussed in details in the following sections.

3.2 TRAINING, VALIDATION AND TEST DATA

The training data is used to learn the dynamics of the gas reservoir from its simulation solutions, while the validation and test data are new, unique and different simulation data. The validation data is only provided for the AE interpolation, to facilitate the process of fine-tuning the AE to generalise well on new unseen data. This is best practise in machine learning. The test data represents the simulation cases that are later predicted with the trained NIROM and compared with ECLIPSE's solutions, i.e. accuracy and speed analysis. What differentiates each data set is the blend of parameter(s) that are used to create its simulation case. This work has considered constant properties (such as permeability or porosity) or time-varying sink properties (such as production rates or pressure limits) as simulation parameters. Since NIROM is data-driven, it has a similar implementation for constant or varying (forced) parameters. The original theory for the NIROM implementation that is applied in this work was designed by Kostorz et al. (2019-2021) ^[37, 99, 100] for constant parameter cases. However, their implementation is not suitable for time-varying parameter. Therefore, this work provides a modification of this implementation that makes it applicable to varying parameter cases.

To generate the data for the constant parameter scenarios, in this work, the maximum and minimum allowable values are chosen for any of the listed parameters and 5 to 50 training runs are generated,

equally spaced between the extreme parameter values. The parameters used to generate the training runs are the training parameters. The validation and test data are then selected at random intervals between the training parameter values, ensuring that no validation or test parameter is the same value as any of the training parameter values. For the test data, it is also ensured that two of the parameters fall outside the maximum and minimum training parameter values, to test the NIROM's capacity to make predictions of scenarios outside its range of training parameters (extrapolation). For example, consider a gas development project that requires that a reservoir is produced at a production rate between 50MMscf/day and 60MMscf/day. In this illustration, the adjustable parameter is the gas production rate. So in the constant parameter case, the NIROM can be trained with 5 training parameters i.e. 50, 52.5, 55, 57.5, and 60MMscf/day (Figure 3.1a). The rationale here is that after the NIROM has been trained with ECLIPSE solutions at the 5 production rates above, the NIROM will be expected to subsequently predict the reservoir's production performance at other production rates (test/ prediction parameters). These prediction parameters are different from the training parameter values listed above, for instance at 56.5 or 59MMscf/day. This constant parameter case is common in different fields of science, hence Kostorz et al. (2019-2021) ^[37, 99, 100] originally implemented the NIROM for such scenarios.

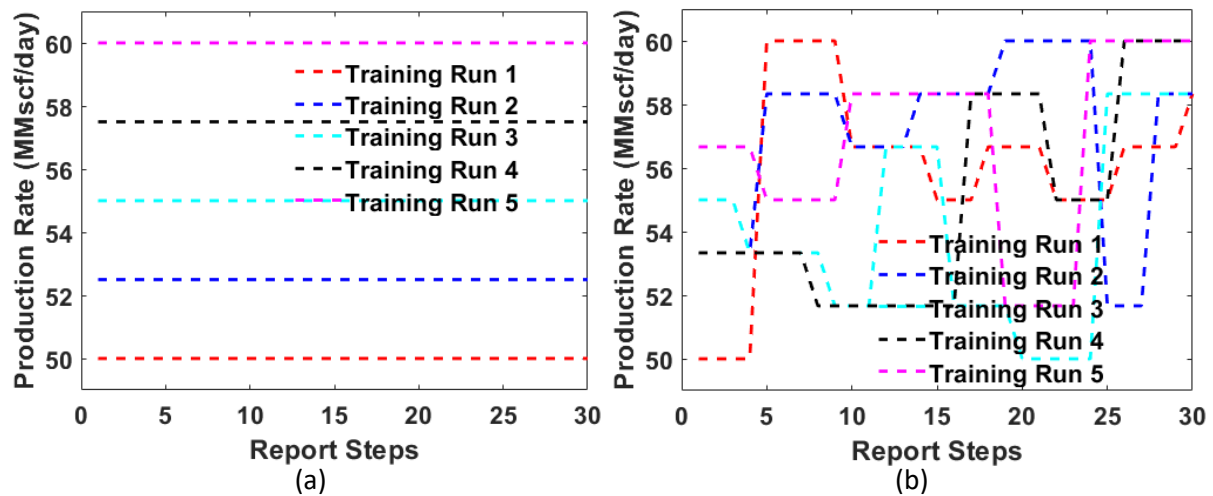


Figure 3.1: Difference between (a) constant parameter distribution and (b) forced parameter distribution in the temporal resolution (x-axis). The x-axes (report steps) are in 300days interval. For constant parameters, production rates remain the same value in the temporal resolution while for forced parameters, the production rates vary in the temporal resolution.

However, for hydrocarbon production in the petroleum industry, it is more common for production rates and/ or pressures to vary at different production intervals due to increasing or decreasing demand. For the same gas reservoir example considered above, instead of producing the reservoir at a constant gas rate between 50 to 60MMscf/day, gas production might commence at 50MMscf/day and 2.5 years later the production rate is raised to 57MMscf/day for 3.5 years. Next, the demand for gas

might be lower and the production is cut back to 40MMscf/day for 1.5 years and finally, the reservoir may be produced at 55Mscf/day for its remaining lifecycle. In this later production scenario, the simulation parameters have been adjusted on 4 different occasions, even though the changes in production rate could have occurred more or less frequently throughout the production regime of the reservoir. This is an example of a time-varying parameter case (Figure 3.1b), sometimes referred to as forced parameter case ^[99, 100]. In this work, the varying parameter cases have been generated as a combination of production parameter values (such as rates), within an acceptable range of values, that change or take on different values after short and random time intervals.

To adapt the varying production rate cases to the concept of gas supply contracts in which production rates vary seasonally, this work has developed and considered production profiles in which production rates change up to four times every year. These seasonally varying profiles have a higher gas production during winter, a lower production during autumn and spring and the least production during summer. Another practical production scenario that has been considered in this work is when wells that fail to meet minimum THP or BHP constraints are shut-in even before the end of the reservoir's production lifecycle or when there are well control switches during the production life of a reservoir, to accommodate predetermined pressure limits.

These varying or forced parameter scenarios are important in the petroleum industry as they are designed to allow for changing boundary conditions from rate to BHP control. Varying or forced parameter cases are mathematically different boundary conditions from the constant parameter cases and hence could prove difficult for the fast models to capture correctly. In 2020, Samuel and Muggeridge showed that Kostorz et al.'s (2019)^[37] normalisation techniques for constant parameter cases do not suffice for forced parameter cases, so different normalisation techniques are required to implement the POD-RBF NIROM for varying or forced parameter cases. Although they presented 3 methods for normalising these cases, they showed that those methods were not satisfactory for the varying parameter cases. Their 3 normalisation methods involved normalisation in a second hyperspace, normalisation based on a proxy parameter such as cumulative production and changing the temporal axis of the NIROM to a proxy parameter. These methods are discussed in Appendix A. Hence, this work presents a NIROM implementation that accommodates varying parameter cases for both POD-RBF and POD-AE (see section the later paragraphs of section 3.3 and section 3.4). The outcome of this different NIROM implementation is illustrated in the result sections of this work (chapters 7 and 8).

Two types of simulation solutions are considered in this work, grid distributed reservoir properties and production profiles. Pressure and saturation distribution solutions from each simulation run are saved within the approximate model interface (MATLAB) in addition to their corresponding time series

production profiles/ data. In this work, two ECLIPSE output data files have been read to obtain these simulation results, the “.UNRST” and the “.RSM” files. These unformatted ECLIPSE output files are read into MATLAB with the aid of existing computer programs ^[122, 123], which were modified to suit the grid refinement test examples that are considered in this work. Other measures that are taken when generating training and prediction data include ensuring that the report time intervals for ECLIPSE output results are equally spaced, for instance 300days intervals between report steps. This measure simplified the NIROM code/ set up since a simulation’s time instance can be easily inferred by multiplying a specific report step by 300days, for example. Furthermore, Kostorz et al. (2020) ^[99] suggested that having shorter/ smaller time intervals between training runs can lead to more accurate interpolation of a system’s underlying dynamics. Their suggestion was employed in this work, as well as ensuring that the range of input parameters for training runs spanned the expected range of prediction parameters to be investigated.

As mentioned earlier, the NIROM implementation in this work was trained with 5 to 50 training runs. This was because the most suitable number of training runs for each test case is not known in advance of implementing and evaluating the NIROM. Hence this provided the opportunity for investigating the ideal number of training runs for the different production scenarios that are considered in this work. The specific number of training runs that were used to implement the NIROM for each test case are stated in the results chapters (chapters 5 to 8). The result chapters also provide some insight on the relationship between NIROM’s performance and the number of runs it was trained on.

3.3 PROCESSING OF TRAINING DATA USING POD

When the training data is obtained from ECLIPSE, they are pre-processed as follows. For each training run, the output solution from ECLIPSE is extracted into a 2-dimensional ($nx \times nt$) matrix array U_s , with nx corresponding to the spatial dimension and nt corresponding to the frequency at which data is output/ requested from ECLIPSE (total number of report steps or snapshots). nx is the total number of active grid cells in the reservoir or the total number of production data variables being monitored in the simulation (nx should be the same value in every training run), while the subscript s refers to a particular simulation run/ snapshot. For instance, if we want to train a proxy model to predict the pressure distribution in a reservoir of $NX \times NY \times NZ$ active grid cells, then $nx = NX \times NY \times NZ$ or if we intend to monitor the production data (for example well BHP, gas rate and water-gas ratio from 2 wells in the reservoir), then $nx = 6$. Also, suppose that production from the 2 wells is at different rates which change every month for 20 years with a fixed THP of 120bars, then in this example the number of parameters, $np = 3$ (two production rates and one THP limit).

After all the ns training runs have been completed, and their simulation solutions have been extracted and saved, the entire training data ($\mathbf{U}_1$ to $\mathbf{U}_{ns}$), i.e. all the individual matrices containing each simulation solution are concatenated into a single dataset, $\mathbf{U}$.

$$\mathbf{U} = [\mathbf{U}_1; \mathbf{U}_2; \mathbf{U}_3; \dots; \mathbf{U}_{ns}] \quad 3.1$$

Then the temporal mean, $\bar{\mathbf{U}}$, of the full training dataset is computed (Equation 3.2) and subtracted from individual training data, $\mathbf{U}_s$ (Equation 3.3). The temporal mean of the reservoir property of interest, for instance pressure, is the average pressure in each grid cell of the reservoir throughout the production lifecycle and for all training snapshots that are considered.

$$\bar{\mathbf{U}} = \frac{1}{ns \times nt} \sum_{t=1}^{ns \times nt} \mathbf{U} \quad 3.2$$

$$\tilde{\mathbf{U}}_s = \mathbf{U}_s - \bar{\mathbf{U}} \quad 3.3$$

This step standardises the training data to have zero mean and ensures that only the changes in the reservoir performance are being modelled in the NIROM. After this, the training dataset is ready to be reduced or projected into a lower-dimensional space with proper orthogonal decomposition (POD).

POD is mostly used as a dimensionality reduction technique to improve the efficiency of working with high dimensional data. POD involves the singular value decomposition of high dimensional data. Here POD is used to project the training data into a lower-dimensional space as shown in Equations 3.4 to 3.7.

$$\tilde{\mathbf{U}} = \mathbf{V} \boldsymbol{\lambda} \mathbf{W} \quad 3.4$$

where $\mathbf{V}$ is the left singular vectors' matrix,
 $\mathbf{W}$ is the right singular vectors' matrix,
 $\boldsymbol{\lambda}$ contains the singular values of the matrix $\tilde{\mathbf{U}}$, and
 $\boldsymbol{\lambda}^2$ corresponds to the eigenvalues ($\boldsymbol{\Sigma}$) of $\tilde{\mathbf{U}}^T \tilde{\mathbf{U}}$ or $\tilde{\mathbf{U}} \tilde{\mathbf{U}}^T$.

After decomposing the training dataset into its corresponding singular values and singular vectors, the largest singular values, which are usually in an ordered sequence, and their corresponding singular vectors are filtered from the full left singular matrix, $\mathbf{V}$. These largest singular values and their corresponding vectors represent the optimal number of POD basis functions that can describe the full dataset, $\tilde{\mathbf{U}}$, in its linear subspace. The criteria for determining the cut-off for the optimal singular values is referred to as the energy criteria. Depending on the threshold or energy criteria for truncating the singular values, most of the information in the high dimensional data is preserved with minimal truncation error when the projected data is transformed back to its original dimension [21, 32, 36, 55, 94, 99,

^{100, 124]}. In this work the POD energy threshold/ criterion of at least 99.9% of the full training dataset has been applied as shown in Equation 3.5. Using an optimal amount of POD basis functions ensures that the proxy model approximations are faster and retains the most significant dynamics of the gas reservoir's performance.

$$\frac{\sum_{t=1}^{t=d} \lambda_t^2}{\sum_{t=1}^{t=ns \times nt} \lambda_t^2} \geq 99.9\% \approx 100\% \quad 3.5$$

where λ_t is the t th singular value, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$, $d \leq ns \times nt$, and λ_t^2 is an element-wise product of each singular value with itself.

Based on the optimal number of POD basis functions evaluated from Equation 3.5 (an equivalent of 99.9% of the singular values of the zero-mean representation of the training data ($\tilde{U}$) from Equation 3.3), the optimal singular values or the POD transform matrix, V_d , is obtained from Equation 3.4 and Equation 3.5 as:

$$V_d = [v_1, v_2, v_3, \dots v_d] \quad 3.6$$

where v are column vectors of V , which is the left singular vectors' matrix from Equation 3.4 and d is evaluated from the energy criterion in Equation 3.5

Next, the transpose of the optimal POD transform matrix from Equation 3.6 is multiplied with individual training snapshot data, $\tilde{U}_s$, to obtain the individual reduced training snapshot data, α_s .

$$\alpha_s = V_d^T \tilde{U}_s \quad 3.7$$

Therefore each training snapshots' matrix, U_s , is projected into a reduced space, α_s , defined by the transpose of the left singular vectors of the full training data.

$$R^N \in U_s \rightarrow \alpha_s = V_d^T (U_s - \bar{U}) \in R^d \quad 3.8$$

where $N = nx \times nt \times ns$ and $d \leq ns \times nt$

After the training snapshot data have been reduced with the optimal POD transform matrix, they are normalised within a uniform range before they are interpolated with RBF or AE. Normalisation helps to bring the entire training data into a consistent range of values where smaller values do not lose their significance in the data set. There are different possible methods for normalising data during the pre-interpolation stage. Some of these methods include scaling the dataset into a range of values from 0 to 1, -1 to 1 or from a small value to 1. The standardization approach involves mean extraction and division by the data's standard deviation, and large values are sometimes pre-processed into a logarithmic scale ^[58, 111].

In simulations with more than one training run, Kotorz et al. (2020) ^[99] emphasised the importance of linking individual training runs with their corresponding parameters and proposed an efficient normalisation method that accounts for the communication between the training data and parameters. Their normalisation approach (referred to here as renormalisation) involves scaling the data based on a division by the average Euclidean (L2 norm) or pairwise distance of the training data or parameter, to obtain equal spacing of unity between training runs. Their approach works well for constant parameter (time-invariant) cases but it was not designed for varying parameter cases where simulation output are based on non-uniform time-varying controls ^[125]. This is further discussed in the parameterisation section below. Kotorz et al.'s (2020) ^[99] renormalisation approach for constant parameter cases proposes that each training run data, α_s , be normalised to α_s' as follows:

$$\alpha_s' = \frac{\alpha_s}{\Delta_\alpha} \quad 3.9$$

where

$$\Delta_\alpha = \frac{\sum_{s=1}^{n_s} (\sum_{t=1}^{n_t-1} \|\alpha_s^t - \alpha_s^{t-1}\|)}{\sum_{s=1}^{n_s} (n_t - 1)}$$

For forced parameter cases this work proposes that the dataset should be normalised based on the maximum value of the data (spatial property or parameter), to obtain unit values between -1 and 1. This is achieved by dividing each of the reduced training snapshot data, α_s , by the maximum absolute value of the full reduced training data set, $|\alpha_{max}|$, to obtain $\widetilde{\alpha}_s$.

$$\widetilde{\alpha}_s = \frac{\alpha_s}{|\alpha_{max}|} \quad 3.10$$

where $\alpha_{max} = \text{maximum} ([\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_{ns}]^{1:nt})$

The NIROM workflow described so far is similar for both POD-RBF and POD-AE, except that for the POD-AE, the full rank of the POD transform matrix is used to project the training data into hyperspace, i.e. energy criterion =100%.

3.4 PARAMETERISATION

In the previous section, the process of normalising the training data was highlighted. Likewise, when NIROM is used for making predictions, it is important to include simulation parameters into the training dataset before interpolation. The parameters serve as unique identifiers for individual simulation runs. The output simulation data and input parameter(s) may have different units as well

as be in different orders of magnitude. For example when the pressure distribution (in thousands of psi or hundreds of bars) in a reservoir is being modelled and the production rates (in hundreds of thousands of Mscf or m³) and THP (in thousands of psi or hundreds of bars) are the input parameters, then there are at least 2 different units to work with and the data spans at least 2 different orders of magnitude. So each simulation run needs to be identifiable by its unique set of parameter(s) and each set of simulation runs and parameter(s) needs to be normalised into a consistent range of values that captures their corresponding significance. Parameterisation is the process of normalising the simulation parameters to be consistent with the simulation data and to represent the combined data in a form that ensures that the significance of the parameters is reflected in the simulation dataset [36, 58, 99, 100].

According to Kostorz et al. (2020 & 2021) [99, 100], parametrization inculcates the dependence and communication between individual simulation runs and their parameters into the simulation data in preparation for the interpolation step. The parameterisation approach of Kostorz et al. (2020 & 2021) works well for constant parameter cases (e.g. Figure 3.1a). Their approach is achieved by renormalizing the training data as well as their parameters with the mean magnitude of their pairwise distances apart. This method leads to a spacing of unity between individual training runs or their parameters. So if the corresponding training parameters for the training dataset α_1 to α_{n_s} are

$$\mathbf{p} = [\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots, \mathbf{p}_{n_s}] \quad 3.11$$

where $\mathbf{p}_s$ is the constant parameter for the s th training run,

then the renormalisation factor for simulation parameters will be:

$$\Delta_p = \frac{\sum_{s=1}^{n_s} \|\mathbf{p}_{s+1} - \mathbf{p}_s\|}{n_s} \quad 3.12$$

And the parameterisation is achieved by dividing each training parameter by the parameter scaling factor:

$$\mathbf{p}_s' = \frac{\mathbf{p}_s}{\Delta_p} \quad 3.13$$

In this work, Kostorz et al.'s (2020 & 2021) approach has been used for the constant parameter prediction results that are discussed later in Chapter 6. However in the varying parameter case, for example in the varying production profile example in Figure 3.1b, each simulation snapshot has several parameter values at different report times (time-varying). So it would be difficult to distinguish between different training runs and assign them with unique parameters. In theory, $\|\mathbf{p}_{s+1} - \mathbf{p}_s\|$

cannot be computed since there is no single representative value for each varying parameter run. An attempt to compute the parameter scaling factor at the report step level so that $\|\mathbf{p}_{s+1} - \mathbf{p}_s\|$ in Equation 3.12 is replaced with $\|\mathbf{p}_s^t - \mathbf{p}_s^{t-1}\|$ will yield a scaling factor that does not produce unit (or uniform) spacing between report steps. The continuous random rise and fall in varying parameter values during each run will lead to a resultant Euclidean distance that is impacted by exaggerated effects of the non-uniform rises and falls of the varying parameters and hence lead to unrepresentative dependence of training runs on their corresponding varying parameter.

Therefore figuring out how to appropriately normalise the varying parameter scenarios was a major challenge in this work. This chapter is focused on the methods that were most suitable for the NIROM predictions described in this work. So the most suitable method for normalizing simulation parameters is described here and other approaches that were evaluated are described in Appendix A. This work proposes an approach where each varying parameter data is normalized to unit values by dividing through each parameter occurrence with the maximum parameter value encountered during the entire training run (Equation 3.14). So $\mathbf{p}_s$ is normalised to $\widetilde{\mathbf{p}}_s$ for each simulation run and $\widetilde{\mathbf{p}}_s$ holds the combination of varying parameters for the s th simulation run.

$$\widetilde{\mathbf{p}}_s = \frac{p_s}{|\mathbf{p}_{max}|} \quad 3.14$$

where $\mathbf{p}_{max} = \text{maximum}([\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots, \mathbf{p}_{ns}]^{1:nt})$

The next normalization step, before simulation data is fed into RBF, AE or any other interpolation approach, is to concatenate the normalised training data with their corresponding normalized parameters to get $\mathbf{A}_s$ in Equation 3.15.

$$\mathbf{A}_s = [\widetilde{\boldsymbol{\alpha}}_s; \widetilde{\mathbf{p}}_s] \quad 3.15$$

The final step in pre-processing the training data involves separating the data into sample and response arrays. This is a requirement for supervised learning algorithms where data is initially given to the approximate model in a cause and effect manner, to learn the dynamics of the data and eventually serve as a prediction tool. This concept is actually in line with the principle/ process by which conventional reservoir simulators predict reservoir performance, in that a reservoir's performance or property distribution at any given time interval is a function of its performance or description in the preceding time interval. So the normalised training data, $\mathbf{A}_s$, is separated into 2 matrix arrays where each column of training data in the sample array corresponds to the preceding training data in the same column of the response array. For example if the training data matrix is

$A_s = [A_s^1, A_s^2, A_s^3, A_s^4, \dots, A_s^{nt}]$, then the sample matrix array will be $A_s^t = [A_s^1, A_s^2, A_s^3, A_s^4, \dots, A_s^{nt-1}]$ and the response matrix array will be $A_s^{t+1} = [A_s^2, A_s^3, A_s^4, A_s^5, \dots, A_s^{nt}]$.

3.5 MODELLING DYNAMICS

Recall from chapter 2 that the accuracy and predictive capacity of POD based models are improved by introducing interpolation or approximation techniques such as DEIM, TPWL, GNAT, RBF and machine learning methods. In this work, the RBF and/ or Autoencoder (a machine learning technique) have been chosen for modelling the dynamics (interpolation) of the NIROM.

In 2021, Kotorz et al. ^[100] described the relationship between training and prediction runs in a parameterised hyperspace with the schematic in Figure 3.2. In Figure 3.2a, three training runs are stacked within the same hyperspace, covering a range of constant parameter(s), and the intermediate regions between the training runs are interpolated in Figure 3.2b, as prediction cases with the same model dynamics as the training runs in Figure 3.2a. This work applies a similar approach for modelling the dynamics of both POD-RBF and POD-AE NIROMs.

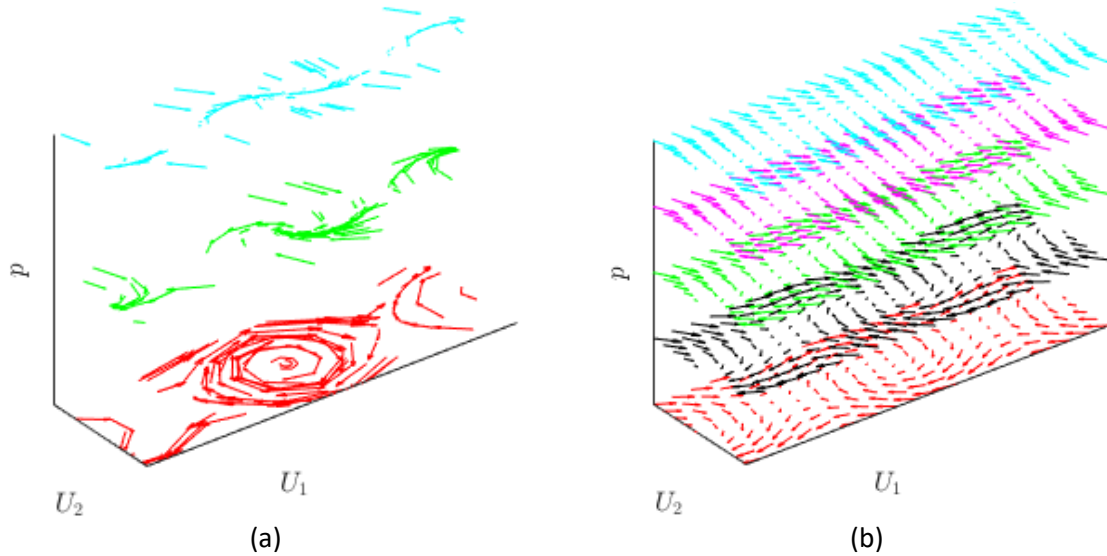


Figure 3.2: A representation of training and prediction data stacked together according to their parameters in hyperspace. a) A set of training runs (in cyan, green and red colour) (b) a set of training runs and prediction runs (in magenta and black colour) which have been interpolated between the training runs ^[100].

As in Kotorz et al. 2020 ^[99], a gas reservoir's production dynamics in hyperspace can be represented as:

$$A_s^{t+1} = A_s^t + f(A_s^t) \quad 3.16$$

Where the matrix $A_s = f(\tilde{a}_s; \tilde{p}_s)$ from Equation 3.15, A_s^t is the matrix describing the projection of the training solution in hyperspace at time instance t and f is an unknown function that will be reconstructed by interpolation of the reduced training data in hyperspace, using RBF or AE. After the

system's dynamics have been interpolated with RBF or AE, the NIROM can be used to propagate the prediction of the initial condition of new unseen cases, one time step at a time (as represented in Equation 3.16) until a full prediction is obtained like in Figure 3.2b.

For the prediction step, initial condition prediction data (the first report step of the prediction data) is reduced and normalised with the same values ($\bar{U}$, V^T , $|\alpha_{max}|$ and $|p_{max}|$) that were used to reduced and normalise the training data to zero mean, lower-dimensional hyperspace and unit value form. This is achieved by following the same processes as in Equations 3.3, 3.7, 3.10, 3.13 and 3.15. The AE or RBF model is then used to make forecasts of the full prediction data from the normalised initial condition prediction data.

Once a suitable interpolation method has been used to obtain satisfactory outputs for the full prediction duration or production lifecycle of the reservoir, the forecasted output, $\boxed{\bar{\alpha}_s}$ is un-normalised and projected back to the original high dimensional space from before POD reduction, and the mean that was initially extracted is added back to the forecasted results, as shown in Equation 3.17. These are the prediction results that are analysed for accuracy and speed in comparison to the commercial simulator's (ECLIPSE) solutions. Note that the POD transform matrix is used for back projection as opposed to its transpose matrix.

$$U_s \cong \boxed{U_s} = \{V (\boxed{\bar{\alpha}_s} \times |\alpha_{max}|)\} + \bar{U} \quad 3.17$$

3.5.1 RADIAL BASIS FUNCTIONS (RBFs)

Radial basis functions are universal approximation functions that facilitate accurate interpolation of un-sampled parameters between locations of known data, the kind of scenarios that reservoir engineers are faced with in estimating reservoir properties and behaviour at un-sampled locations. RBFs are a simple type of neural network (NN) that is robust for modelling noisy data without overfitting. RBFs are characterised by a continuous spread, decreasing or increasing from or about a centre, and determining the coefficients needed to perform their interpolation for a given system is relatively fast. In theory, they are a summation of the linear multiplication of weighting coefficients and their corresponding interpolation function coefficient matrix ^[36, 38, 71, 105, 126]. Therefore, a function, $y = f(x)$ can be written as:

$$f(x) = \sum_{i=1}^N \omega_i \times \varphi(\|x - \hat{x}_i\|) \quad 3.18$$

where $f(x)$ is the response from the interpolation function,
 N is the number of column vectors in the dataset being interpolated
 ω_i is a weighting coefficient matrix of interpolation weights,

φ is a radial basis interpolation function,
 $\hat{\mathbf{x}}_i$ represents the known data point or centre point for the RBF,
 $(\|\mathbf{x} - \hat{\mathbf{x}}_i\|)$ is the Euclidean distance between individual data points and the centre points, and
 $\varphi(\|\mathbf{x} - \hat{\mathbf{x}}_i\|)$ is the only hidden layer in the interpolation and is a function of the input sample data, $\mathbf{x}$.

The accuracy of NIROM depends on factors such as the choice of interpolation function, φ , and the smoothness of the interpolation function [36, 37, 72, 92, 96, 107, 126]. Klie (2013) [36] used Gaussian radial basis functions, Xiao et al. (2016) [38] used Gaussian and multi-quadratic radial basis functions, and Kostorz et al. (2020) [99] used a linear basis function. Neither Klie (2013) nor Xiao et al. (2016) provided any justification for their choice of RBF. Kostorz et al. (2020) stated that they found that linear basis functions worked best for the problem they investigated (waterflooding of oil reservoirs) but did not present any results demonstrating this. In this work, we have evaluated seven different RBFs:

- 1) Linear function: $\varphi(r) = r^1$
- 2) Inverse multi-quadratic function: $\varphi(r) = \frac{1}{\sqrt{1+r^2}}$
- 3) Modified thin-plate spline function: $\varphi(r) = r^2 \ln(r + 1)$
- 4) Inverse quadratic function: $\varphi(r) = \frac{1}{1+r^2}$
- 5) Gaussian function: $\varphi(r) = \exp(-r^2)$
- 6) Cubic function: $\varphi(r) = r^3$
- 7) Multi-quadratic function: $\varphi(r) = \sqrt{1 + r^2}$ where $r = \|\mathbf{x} - \hat{\mathbf{x}}_i\|$.

During RBF construction, Equations 3.16 and 3.18 can be combined and written as:

$$\mathbf{A}_s^{t+1} = f(\mathbf{A}_s^t) = \sum_{t=1}^{nt-1} \omega_t \varphi(\|\mathbf{A}_s^t\|) \quad 3.19$$

where φ denotes one of the seven radial basis functions listed above,
 $\mathbf{A}_s^{t+1}$ and $\mathbf{A}_s^t$ are the response and sample matrix arrays from section 3.4

In the RBF construction in Equation 3.19, the weighting coefficient matrix (interpolation weights), ω_t is the only unknown when training the NIROM, since $\mathbf{A}_s^{t+1}$ is the response array and $\mathbf{A}_s^t$ is the sample array. So by substituting the respective interpolation function, in addition to the sample and response training data arrays into a rearranged Equation 3.19, ω_t is evaluated.

For NIROM prediction simulations, the computed weighting coefficients, ω_t , and the initial condition prediction data (the new $\mathbf{A}_s^1$) are now substituted into the right-hand side of Equation 3.19 to interpolate the reservoir's behaviour for prediction cases, one report step at a time. The idea is that

with the first column of prediction data (initial condition data) and the weighting coefficients, NIROM makes a forecast of the prediction case at the second report step. Following NIROM's prediction of what the reservoir's performance should be at the second step, NIROM makes a forecast for the third time interval and so on for the number of time intervals that NIROM has been set to make forecasts.

Kostorz et al. (2021) ^[100] have however noted that with this NIROM interpolation/ approximation, there is the likelihood that any prediction errors in one time interval are propagated and accumulate from that time interval to the following time intervals. This type of problem will be more severe with certain interpolation functions, like the Gaussian and polynomial type of functions.

3.5.2 AUTOENCODER

For this work, the best ML technique was found to be the autoencoder (AE), other machine learning techniques including Feedforward NN (FF), Long Short Term Memory (LSTM) and convolutional NNs were evaluated and are discussed in Appendix B. Nevertheless, the general idea behind ML techniques is similar to that of RBFs described earlier, except that ML techniques can have several layers where interpolation weights are being computed and optimised with different sophisticated algorithms.

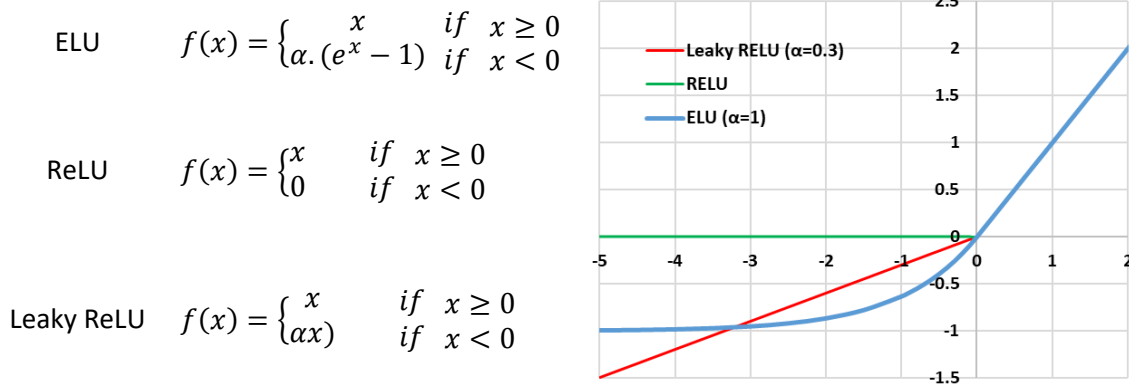
An autoencoder is a supervised machine learning algorithm for non-linear data reduction. AEs can automatically learn from training data, by compressing input data into a lower-dimensional latent space with encoder layers and subsequently decompressing the data with decoder layers. Training data is reduced by minimising the loss functions, after which they are projected back to their initial dimension before they were reduced by the encoder layers. Decoder layers are usually similar in architecture to their encoder layers but in reversed order ^[106, 110].

In this work, autoencoders with 1 to 5 fully connected encoder layers are considered. Fully connected (dense) layers mean that all the nodes from a previous layer are applied to generating the weights for the next layer. Fully connected (dense) NNs have longer training times than NNs without dense layers (models with dropout layers), but AEs with less dense layers were found to be unsuitable for this work (as shown in Appendix B.2). Also, the full hyperspace representation of the training data, an energy criterion of 100%, was used for AE modelling to minimise the training data loss, i.e. from Section 3.2, it ensured that the full-rank (un-truncated) POD training data was used for training the autoencoder. The innermost layer of the encoder was set to have 10 nodes, irrespective of the number of hidden layers in the model, this is the starting dimension for the decoder layer which projects the output of the encoder layers back to the same dimension as the original full-rank POD reduced data, $\mathbf{A}_5$. The number of nodes in the innermost hidden layer of the autoencoder was determined based on the work of Teoh et al. (2006) and Cai et al. (2019), who showed that the optimal rank of POD can serve

as a pruning criterion for determining the number of nodes in the hidden layer of a neural network. The 10 nodes in the innermost layer of the autoencoder models used in this work represent an average number of POD basis functions that make up the optimal rank of POD basis functions.

The other hyper-parameters applied to this work are the Exponential Linear Unit (ELU) activation function, the Adam optimiser (a stochastic gradient descent method) and 1000 epochs with 200 batch sizes ^[110, 127]. ELU was chosen because it gave the best results in comparison to ReLU and Leaky ReLU functions. The Adam optimiser was suggested in Kanin et al. (2019) and Temirchev et al. (2020), and was found to be suitable for the test cases in this work. Sensitivity analysis on the learning rate of the Adam optimiser showed that 0.001, the default value, was best suited to this work. The network's learning rate is an indication of the speed at which the optimizer's weights are updated during interpolation. Activation functions in ML are synonymous with RBF interpolation functions, but in ML the different network layers can be interpolated with different functions. In this work, the ELU was chosen after comparative analysis with ReLU and Leaky ReLU, these are all functions that penalise the negative values of a function, however ELU penalises negative values as a smooth exponential function as shown in Table 3.1 below.

Table 3.1: Equations and schematics of ELU, ReLU and Leaky ReLU activation functions.



The choice of epoch and batch size means that the POD-AE model processes 200 data samples per batch until it has passed through the entire training dataset and this process is repeated 1000 times. Since this work is based on a regression-type gas reservoir dataset, the mean square error loss function was used to compare the sample versus response array data after every batch's observation, in addition to the mean absolute error metrics. Besides the sample and response training data arrays that are used to train the AE, another dataset is provided solely for validation of the AE model during its training process. Table 3.2 shows a summary of the hyper-parameters used in the autoencoder model. Figure 3.3 is a schematic summary of the AE workflow. In this work, the AE model was built with NN modules from the Keras TensorFlow Python library ^[110].

Table 3.2: Summary of the hyper-parameters used in the autoencoder model.

Number of encoder layers	1 to 5
Number of decoder layers	1 to 5
Number of nodes in the innermost hidden layers	10
Number of epochs	1000
Batch size	200
Optimiser Algorithm	Adam optimiser
Activation Function	Exponential Linear Unit
Loss Function	Mean Square Error
Error Metrics	Mean Absolute Error

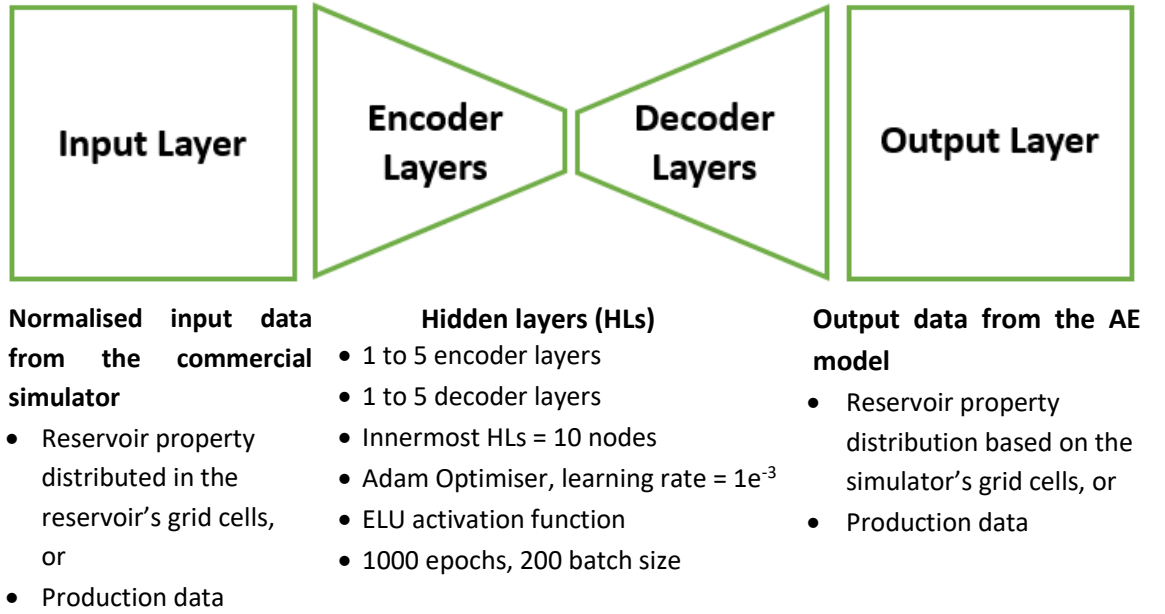


Figure 3.3: Summary of POD-AE model workflow showing the hyperparameters applied to the model. The first and last steps are based on proper orthogonal decomposition projection and back-projection, while the middle steps are the autoencoder model with 5 layers for the most satisfactory cases.

3.6 SIMPLE EXAMPLE OF NIROM WORKFLOW

In this section, the implementation of the POD-RBF NIROM is presented from 2 training runs/ snapshots of a 1D reservoir with 5 grid cells. Let the first snapshot be U_1 and the second snapshot is U_2 . In this example, the pressure distribution in the 5 grid cells of the 1D reservoir is modelled in 4 report steps i.e. $nt = 4$, $nx = 5$ and $ns = 2$. The 1D reservoir is also assumed to have one production well producing at different rates during the 4 report steps, these production rates are saved in row vectors as P_1 and P_2 , for the two training runs. Therefore if

$$U_1 = \begin{bmatrix} 251 & 242 & 209 & 123 \\ 251 & 200 & 151 & 127 \\ 292 & 201 & 155 & 130 \\ 279 & 268 & 231 & 217 \\ 292 & 251 & 244 & 211 \end{bmatrix} \quad \text{and} \quad U_2 = \begin{bmatrix} 284 & 263 & 223 & 115 \\ 286 & 252 & 206 & 170 \\ 266 & 256 & 251 & 170 \\ 287 & 217 & 176 & 139 \\ 274 & 250 & 233 & 162 \end{bmatrix}$$

then according to Equation 3.1, $U = [U_1; U_2; U_3; \dots; U_{ns}]$ as shown below and $\bar{U}$ (from Equation 3.2) is obtained as the mean pressure value in each grid cell across all time steps and snapshots.

$$U = \begin{bmatrix} 251 & 242 & 209 & 123 & 284 & 263 & 223 & 115 \\ 251 & 200 & 151 & 127 & 286 & 252 & 206 & 170 \\ 292 & 201 & 155 & 130 & 266 & 256 & 251 & 170 \\ 279 & 268 & 231 & 217 & 287 & 217 & 176 & 139 \\ 292 & 251 & 244 & 211 & 274 & 250 & 233 & 162 \end{bmatrix} \quad \text{And} \quad \bar{U} = \begin{bmatrix} 213.75 \\ 205.38 \\ 215.13 \\ 226.75 \\ 239.63 \end{bmatrix}$$

Next $\tilde{U}_1$ and $\tilde{U}_2$ are computed from $\tilde{U}_s = U_s - \bar{U}$ (Equation 3.3) as shown below:

$$\tilde{U}_1 = \begin{bmatrix} 37.25 & 28.25 & -4.75 & -90.75 \\ 45.63 & -5.38 & -54.38 & -78.38 \\ 76.88 & -14.13 & -60.13 & -85.13 \\ 52.25 & 41.25 & 4.25 & -9.75 \\ 52.38 & 11.38 & 4.38 & -28.63 \end{bmatrix} \quad \text{and} \quad \tilde{U}_2 = \begin{bmatrix} 70.25 & 49.25 & 9.25 & 98.75 \\ 80.63 & 46.63 & 0.63 & -35.38 \\ 50.88 & 40.88 & 35.88 & -45.13 \\ 60.25 & -9.75 & -50.75 & -87.75 \\ 34.38 & 10.38 & -6.63 & -77.63 \end{bmatrix}$$

$\tilde{U}_1$ and $\tilde{U}_2$ are then concatenated to obtain $\tilde{U}$, which is decomposed into its singular values and vectors as shown in Equation 3.4, i.e. $\tilde{U} = V\lambda W$. This decomposition was implemented with the $\text{svd}(\tilde{U})$ MATLAB library as shown in the POD-RBF model in Appendix C.2.

Therefore,

$$V = \begin{bmatrix} -0.56 & 0.10 & -0.79 & 0.13 & 0.20 \\ -0.47 & -0.35 & 0.29 & 0.64 & -0.41 \\ -0.49 & -0.50 & 0.29 & -0.52 & 0.39 \\ -0.35 & 0.72 & 0.46 & 0.16 & 0.36 \\ -0.33 & 0.31 & 0.00 & -0.53 & -0.72 \end{bmatrix}, \quad \text{an } 5 \times 5 \text{ matrix}$$

$\lambda =$ Contains the singular values of $\tilde{U}$, i.e 286.80, 123.40, 56.24, 46.59, and 11.32,

$$W = \begin{bmatrix} -0.40 & 0.03 & 0.54 & -0.54 & -0.02 & 0.07 & 0.02 & 0.50 \\ -0.08 & 0.37 & -0.16 & 0.18 & 0.79 & 0.32 & 0.09 & 0.26 \\ 0.19 & 0.43 & -0.49 & -0.12 & -0.34 & -0.27 & 0.19 & 0.54 \\ 0.49 & 0.36 & 0.34 & -0.09 & -0.22 & 0.52 & 0.38 & -0.18 \\ -0.47 & 0.06 & 0.19 & 0.56 & -0.18 & -0.15 & 0.62 & 0.01 \\ -0.24 & -0.29 & -0.32 & 0.17 & -0.36 & 0.72 & -0.15 & 0.25 \\ -0.01 & -0.45 & -0.35 & -0.47 & 0.20 & 0.05 & 0.63 & -0.13 \\ 0.52 & -0.51 & 0.25 & 0.31 & 0.13 & -0.09 & 0.10 & 0.52 \end{bmatrix}, \quad \text{a } 8 \times 8 \text{ matrix}$$

Next the optimal singular values are computed based on Equation 3.5, with the energy criterion of $\geq 99.9\%$. λ_t^2 is an element-wise product of each singular value with itself, as shown below:

λ_t^2	%	Cumulative %
82256.83	79.90	79.90
15226.88	14.79	94.69
3163.11	3.07	97.77
2170.72	2.11	99.88
128.08	0.12	100.00
Cum: 102945.62	100	

For the 99.9% energy criterion, $d = 5$, i.e. the fifth singular value and hence the full set of left singular vectors, $V_d = [v_1, v_2, v_3, v_4, v_5] = V$.

Each training run is then reduced into hyperspace with the transpose of the full singular vectors' matrix, V^T to obtain α_1 and α_2 as shown in Equation 3.7 (i.e. $\alpha_s = V_d^T \tilde{U}_s$). Hence,

$$\alpha_1 = \begin{bmatrix} -115.28 & -24.32 & 54.78 & 141.90 \\ 3.46 & 45.10 & 53.04 & 44.76 \\ 30.29 & -9.17 & -27.75 & 19.36 \\ -25.15 & 8.27 & -5.68 & -4.31 \\ -0.19 & 8.93 & -3.88 & -2.47 \end{bmatrix} \text{ and } \alpha_2 = \begin{bmatrix} -134.14 & -69.39 & -3.27 & 149.73 \\ 7.66 & -35.52 & -55.84 & -62.66 \\ 10.42 & -17.75 & -19.71 & 14.31 \\ 25.90 & 8.05 & -21.68 & 14.60 \\ -2.01 & -4.03 & 2.21 & 1.44 \end{bmatrix}.$$

α_1 and α_2 are then concatenated as shown below to obtain α and α_{max} .

$$\alpha = \begin{bmatrix} -115.28 & -24.32 & 54.78 & 141.90 & | & -134.14 & -69.39 & -3.27 & 149.73 \\ 3.46 & 45.10 & 53.04 & 44.76 & | & 7.66 & -35.52 & -55.84 & -62.66 \\ 30.29 & -9.17 & -27.75 & 19.36 & | & 10.42 & -17.75 & -19.71 & 14.31 \\ -25.15 & 8.27 & -5.68 & -4.31 & | & 25.90 & 8.05 & -21.68 & 14.60 \\ -0.19 & 8.93 & -3.88 & -2.47 & | & -2.01 & -4.03 & 2.21 & 1.44 \end{bmatrix} \text{ and } |\alpha_{max}| = 149.73$$

α_1 α_2

According to Equation 3.10, $\tilde{\alpha}_s = \frac{\alpha_s}{|\alpha_{max}|}$, hence α is normalised to $\tilde{\alpha}$, values between -1 and 1.

$$\tilde{\alpha} = \begin{bmatrix} -0.77 & -0.16 & 0.37 & 0.95 & | & -0.90 & -0.46 & -0.02 & 1.00 \\ 0.02 & 0.30 & 0.35 & 0.30 & | & 0.05 & -0.24 & -0.37 & -0.42 \\ 0.20 & -0.06 & -0.19 & 0.13 & | & 0.07 & -0.12 & -0.13 & 0.10 \\ -0.17 & 0.06 & -0.04 & -0.03 & | & 0.17 & 0.05 & -0.14 & 0.10 \\ 0.001 & 0.06 & -0.03 & -0.02 & | & -0.01 & -0.03 & 0.01 & 0.01 \end{bmatrix}$$

$\tilde{\alpha}_1$ $\tilde{\alpha}_2$

These are the normalised pressure values for each grid cell of the simple 1D gas reservoir with 5 grid cells, across 4 report steps. Next, the production rates (parameters) from the well are normalised for the 2 training runs to values between 1 and -1. The production parameters for the well for both training runs are:

$$P_1 = 4208822 \quad 4104387 \quad 3808826 \quad 2868819 \quad \text{and} \quad P_2 = 4470860 \quad 4041551 \quad 3714933 \quad 3068818$$

The maximum parameter in both training runs is $|p_{max}| = 4470860$, hence as shown in Equation 3.14, dividing each parameter by the maximum parameter value yields the normalised parameter values:

$$\tilde{p}_1 = 0.94 \quad 0.92 \quad 0.85 \quad 0.64 \quad \text{and} \quad \tilde{p}_2 = 1.00 \quad 0.90 \quad 0.83 \quad 0.69$$

According to Equation 3.15, the normalised pressure values and the normalised parameter values are now combined into a single matrix, A_s , as shown below:

$$\begin{bmatrix} \tilde{\alpha} \\ \tilde{p} \end{bmatrix} = \begin{bmatrix} -0.77 & -0.16 & 0.37 & 0.95 & | & -0.90 & -0.46 & -0.02 & 1.00 \\ 0.02 & 0.30 & 0.35 & 0.30 & | & 0.05 & -0.24 & -0.37 & -0.42 \\ 0.20 & -0.06 & -0.19 & 0.13 & | & 0.07 & -0.12 & -0.13 & 0.10 \\ -0.17 & 0.06 & -0.04 & -0.03 & | & 0.17 & 0.05 & -0.14 & 0.10 \\ 0.001 & 0.06 & -0.03 & -0.02 & | & -0.01 & -0.03 & 0.01 & 0.01 \\ 0.94 & 0.92 & 0.85 & 0.64 & | & 1.00 & 0.90 & 0.83 & 0.69 \end{bmatrix}$$

A_1 A_2

These values are then separated into a sample and a response array (A_s^t and A_s^{t+1}), in a cause and effect manner described in the last paragraph of section 3.4. These arrays are used during the interpolation stage.

$$\begin{array}{cc}
 \begin{bmatrix} -0.77 & -0.16 & 0.37 & -0.90 & -0.46 & -0.02 \\ 0.02 & 0.30 & 0.35 & 0.05 & -0.24 & -0.37 \\ 0.20 & -0.06 & -0.19 & 0.07 & -0.12 & -0.13 \\ -0.17 & 0.06 & -0.04 & 0.17 & 0.05 & -0.14 \\ 0.001 & 0.06 & -0.03 & -0.01 & -0.03 & 0.01 \\ 0.94 & 0.92 & 0.85 & 1.00 & 0.90 & 0.83 \end{bmatrix} & \begin{bmatrix} -0.16 & 0.37 & 0.95 & -0.46 & -0.02 & 1.00 \\ 0.30 & 0.35 & 0.30 & -0.24 & -0.37 & -0.42 \\ -0.06 & -0.19 & 0.13 & -0.12 & -0.13 & 0.10 \\ 0.06 & -0.04 & -0.03 & 0.05 & -0.14 & 0.10 \\ 0.06 & -0.03 & -0.02 & -0.03 & 0.01 & 0.01 \\ 0.92 & 0.85 & 0.64 & 0.90 & 0.83 & 0.69 \end{bmatrix} \\
 A^t & A^{t+1}
 \end{array}$$

In the RBF interpolation stage, the first step is to compute the interpolation weight in Equation 3.19 by rearranging Equation 3.19 i.e. $\omega_t = \frac{A_s^{t+1}}{\varphi(\|A_s^t\|)}$. For this simple example with 2 training runs and with the linear interpolation function (φ),

$$\omega_t = \begin{bmatrix} 0.10 & -0.84 & -0.08 & -0.07 & -0.12 \\ 0.02 & -0.50 & 0.33 & 0.02 & 0.04 \\ -0.06 & -0.17 & -0.20 & 0.07 & -0.003 \\ 0.61 & 0.61 & 0.003 & -0.12 & 0.10 \\ 0.60 & 0.41 & 0.18 & 0.40 & -0.02 \end{bmatrix}$$

These are the interpolation weights that are used to predict future pressure values and propagate the pressure distribution in the 5 grid cells of the 1D reservoir through the remaining 3 report steps/ time.

In the prediction phases, for example, with a new initial condition data (U_{new} at report step = 1):

$$U_{new}^1 = \begin{bmatrix} 293 \\ 271 \\ 260 \\ 274 \\ 283 \end{bmatrix}, \quad \alpha_{new}^1 = \begin{bmatrix} -127.6 \\ 10.29 \\ -8.72 \\ 13.83 \\ -7.54 \end{bmatrix}, \quad \tilde{\alpha}_{new}^1 = \begin{bmatrix} -0.85 \\ 0.07 \\ -0.06 \\ 0.09 \\ -0.05 \end{bmatrix} \quad \begin{array}{l} P_{new} = 4237710 \ 3508168 \ 3259739 \ 2828780 \\ \text{and} \\ \tilde{p}_{new} = 0.95 \ 0.78 \ 0.73 \ 0.63 \end{array}$$

$$\text{i.e. } \tilde{U}_{new}^1 = U_{new}^1 - \bar{U}; \quad \alpha_{new}^1 = V^T \tilde{U}_{new}^1; \quad \tilde{\alpha}_{new}^1 = \frac{\alpha_{new}^1}{|\alpha_{max}|} \quad \text{and} \quad \tilde{p}_{new} = \frac{P_{new}}{|P_{max}|}.$$

$$\text{Therefore, } A_{new}^1 = \begin{bmatrix} \tilde{\alpha}_{new}^1 \\ \tilde{p}_{new}^1 \end{bmatrix} = \begin{bmatrix} -0.85 \\ 0.07 \\ -0.06 \\ 0.09 \\ -0.05 \\ 0.95 \end{bmatrix}.$$

Next, A_{new}^2 is computed using Equation 3.19 and ω_t , which was computed earlier. A_{new}^3 is computed using A_{new}^2 and ω_t , then A_{new}^4 is computed using A_{new}^3 and ω_t to obtain the full A_{new} (i.e. propagated across 3 more report steps).

For the autoencoder interpolation, A^t and A^{t+1} are fed into the autoencoder as the training data along with an extra set of validation data, A_{s+1}^t and A_{s+1}^{t+1} . The autoencoder parameters (weights and prediction algorithm) are computed using the hyperparameters that are specified in section 3.5.2. These weights and prediction model are then used to propagate new cases such as A_{new}^1 through time (3 report steps).

The prediction values from both models are then back projected to full space using Equation 3.17 and they can be compared to the solutions from the conventional simulator.

3.7 PERFORMANCE MEASURES

To evaluate the performance of NIROM to ECLIPSE (the commercial reservoir simulator), identical prediction cases were compared. Simulation output from identical POD-RBF, POD-AE and ECLIPSE cases were compared for the spatial distribution of reservoir properties (pressure and gas saturation) as well as the reservoir production performance (production rates and ratios).

Comparative analysis of the difference between ECLIPSE and the two NIROM estimations have been carried out throughout this work, this difference is simply called “Error”. In cases where NIROM prediction errors have negative values, such as for production output data like gas and water rates, it is not visually appealing for these plots to have a negative axis, so the absolute errors between NIROM and ECLIPSE are computed. In the cases where both NIROMs were applied to the same production scenario, other measures of analysis have been adopted including Mean Absolute Error, Root Mean Square Error (RMSE), visual inspection and engineering judgement. These are the main measures that have been used to analyse the accuracy of predictions from the two proposed NIROMs (POD-RBF and POD-AE).

In this work, absolute error refers to the absolute value of the difference between the ECLIPSE and NIROM solutions, while mean absolute error is the average of the absolute error for all the prediction runs in a particular test case or scenario. For instance, the average RMSE or mean absolute error can be computed as a single value for 6 different prediction cases of the water coning reservoir produced at constant BHP. RMSE is better illustrated with Equation 3.20 below. For all these error measures, smaller values are preferable ^[58].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Eclipse\ Solution - NIROM\ Solution)^2} \quad 3.20$$

where n is the number of data samples over which the RMSE is being evaluated

Other performance measures that have been applied in this work include a framework developed by Kostorz et al. 2021 ^[100] for estimating the potential of using NIROM for gas reservoir modelling as well as the generalisation criteria used to evaluate when a NN starts to exhibit overfitting tendencies i.e. when the prediction error starts to build up after each epoch ^[111]. Kostorz et al.’s (2021) ^[100] framework analyses the behaviour of simulation data in hyperspace based on 2 criteria: the extent to which the simulation data experiences changes in direction between successive temporal resolutions (sharpness of turn correlation) and the amount of steadiness or explosion that the data experiences at every time interval when a new parameter is introduced into the system (boost). The detailed theory for the

sharpness of turn correlation and boost approaches can be found in Kostorz et al. (2021) ^[100], while a comparison of the performance of constant and varying parameter examples are considered and discussed in Appendix A.

3.8 SUMMARY

In this chapter, the workflow for implementing 2 NIROMs are discussed alongside the theory behind their application. The process begins with deciding on the boundary conditions and production parameters of interest, followed by generating training data from the conventional simulator, ECLIPSE. The conventional simulation output data of the gas field's dynamics is projected into a lower-dimensional space based on proper orthogonal decomposition and normalised to zero mean and unit values for consistency. This data is then used to train the predictive models which are 1) Autoencoder - a supervised machine learning technique for non-linear problems or 2) RBF - an interpolation technique. These trained models are subsequently used as tools for forecasting the reservoir's performance in new unseen prediction cases, which are compared to the conventional simulator's solution for those same prediction cases. Figure 3.4 is a summary of this workflow.

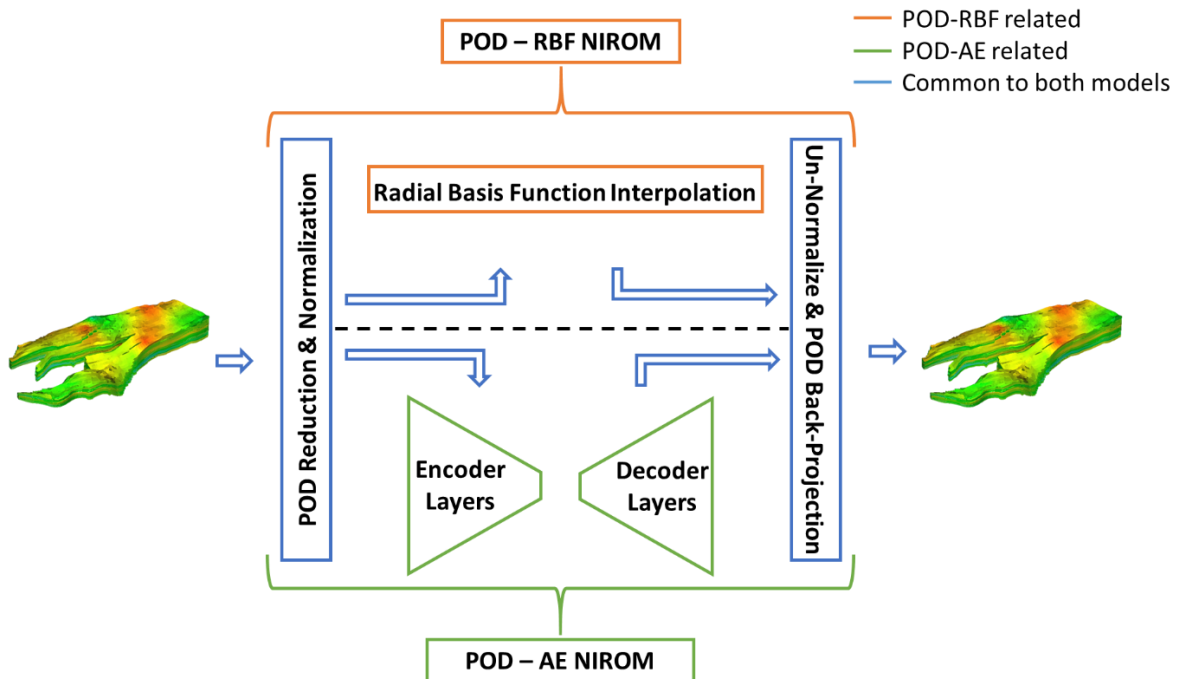


Figure 3.4: Overview of the model workflow for POD-RBF and POD-AE NIROM presented in this work. The workflow components in blue colour are common to both models, while POD-RBF only related components are in orange colour and POD-AE only related components are in green colour.

CHAPTER 4: NUMERICAL EXAMPLES/ TEST CASES

The base gas reservoir models that have been investigated for the application of POD-RBF and POD-AE to modelling gas reservoirs are discussed in this chapter. The reservoir geometry, adjustable parameters as well as aquifer descriptions, where applicable, are provided in the respective sections. The following models are considered in this chapter: a synthetic gas reservoir with homogenous geological properties, a heterogeneous model based on the real Norne oil field model (adapted to gas as the reservoir fluid) ^[128, 129], a water coning gas reservoir model ^[130] and a sector model of an actual gas field (Field X). Except for the water coning model which was built using a radial grid, all the other models were built with Cartesian grid. These dynamic gas reservoir models were all built for Schlumberger's ECLIPSE simulator (ECLIPSE) ^[121]. The approximate models of choice (POD-RBF and POD-AE) were initially trained with simple gas reservoirs and boundary conditions (homogenous reservoirs with constant parameters). Subsequently, more complex cases were considered i.e. the heterogeneous reservoir models with varying parameters. The base case reservoir models used in this work are available in Appendix C, with more details about the gas reservoirs such as their relative permeability values/ curves.

Some of the adjustable parameters considered in this work include porosity, production rates, bottom hole pressures (BHPs) and tubing head pressures (THPs). Specific information on these adjustable parameters is provided in the respective sections for each reservoir. Some of the parameter values that were used for comparing the prediction cases against ECLIPSE's solutions, were chosen to be outside the range of training parameters. This is to evaluate the accuracy of the NIROM when predicting reservoir performance outside of the range of parameters for which it is trained. There is also a case where the NIROM was made to make predictions for longer durations than it was trained for. The range of training and prediction parameters for each case is provided in the subsequent sections.

4.1 SYNTHETIC GAS RESERVOIR MODEL

The synthetic dry gas reservoir represents a sector of a homogenous dry gas reservoir with uniform properties throughout. It has similar dimensions and properties to an average North Sea dry gas reservoir ^[131-137], with a GIIP of 62BCF and reservoir thickness of 330ft. The model is represented by a uniform Cartesian grid of 19×21×6 grid blocks each with dimensions 105ft × 105ft × 55ft, making a total of 2394 active grid cells or data points, i.e. when modelling any grid property distribution in the reservoir $nx = 2394$. The reservoir's rock physics and fluid models were built with Schlumberger's PETREL pre-set models for a dry gas reservoir with sandstone petrophysics ^[138]. Other properties of

the reservoir are given in Table 4.1. This gas reservoir is produced from one well in its centre, grid location (10, 11).

Table 4.1: Summary of reservoir rock and fluid properties of the synthetic dry gas reservoir.

Property	Value
Porosity	0.2
Horizontal permeability (mD)	50
Vertical permeability (mD)	5
Initial pressure (psi)	4200
Reservoir depth/ top (ft)	8095
Gas-water contact (ft)	325 below the top of the reservoir
Gas density (lb/ft ³)	0.047
Gas viscosity (cP)	0.2
Initial gas saturation	0.88
Water density (lb/ft ³)	71.4
Water formation volume factor	1.02
Water viscosity (cP)	0.56
Rock compressibility (psi ⁻¹)	3×10^{-5}
Well diameter (ft)	0.5
Production lifecycle (days)	7500

For this homogeneous gas reservoir, POD-RBF NIROM was used to reproduce the same ECLIPSE simulation runs that were used to train the NIROM. Two types of production control parameters were considered, gas production rate and well BHP. The NIROM was evaluated at a constant production rate of 6.77MMscf/day or a constant BHP limit of 1000psi. For the BHP controlled production, there was an initial surge in gas production that is neither practical nor long-lasting, but this was allowed to investigate the NIROM's ability to estimate the reservoir's behaviour for different simple well controls.

Cases with water influx from an aquifer were also investigated. With the aquifer providing pressure support to the reservoir, the reservoir was produced for the full production lifecycle of 7500 days with BHP limit control. Three different ECLIPSE aquifer models were considered; a numerical aquifer and 2 analytic aquifer models (Carter Tracy and Fetkovich) ^[121]. In each case, the aquifer is modelled as a bottom water drive aquifer in the lowest layer of the reservoir, with a porosity of 0.2 and a permeability of 200mD. The aquifer dimensions were as follows, respectively: cross-sectional area of $9 \times 10^9 \text{ft}^2$ and length of 200ft; a thickness of 500ft and inner radius of 5000ft; as well as a productivity index of 500STB/day/psi with an initial volume of $1 \times 10^{10} \text{STB}$.

The NIROM was also evaluated for its rigour in reproducing ECLIPSE's simulation of the synthetic gas model with grid refinement. Reservoir models are often evaluated with finer grid cells near locations of interest such as at well locations and these simulations are usually time-consuming. This was because it was important to ascertain if NIROM was capable of reproducing such simulations and also

to determine the impact of grid refinement on the time taken to build the NIROM and the speed with which the NIROM ran. This was also done to investigate the ability of the NIROM to capture the rapidly changing pressures in the near well region. In the local grid refinement simulation case, finer local grids were introduced into the grid blocks where the producer was located, (location 10, 11, 1:6). Each of the 6 grid cells was refined to have local dimensions of $5 \times 5 \times 12$, a total of 300 additional data points. A “tartan” grid refinement was also investigated. Here the dynamic model maintained the same number of grid cells as the base case, but the size of the horizontal grid cells was increased logarithmically away from the production well. In both instances, the reservoir was produced with gas rate control.

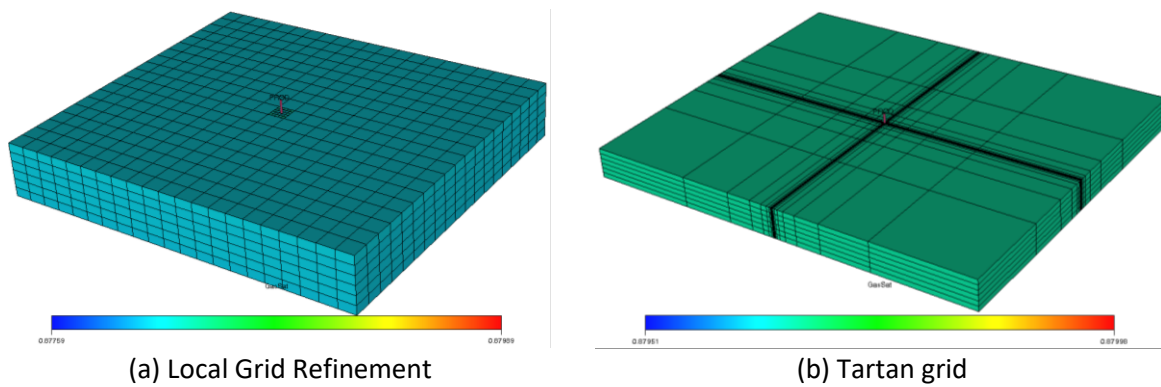


Figure 4.1: Grid refinement cases for the synthetic gas reservoir. (a) Local grid refinement in the producer's grid cells and (b) Tartan grid in the horizontal layers of the reservoir.

The next instance where the synthetic gas reservoir was evaluated was for forecasting the reservoir's behaviour in cases that it was not trained on, unseen or prediction cases. Here the homogeneous gas reservoir was evaluated for different unseen porosity, gas production rate and well BHP values. Simulation parameters for the constant parameter cases are summarised in Table 4.2, including a constant parameter case where the reservoir was initialised by enumeration (from a history restart file) rather than equilibration. For the varying (forced) parameter case, gas production rate and well BHP were investigated as parameters. Different unique production rates from Table 4.3 with unique durations (in multiples of 30 or 100 days) were combined to form the varying rate production profiles, each profile being distinct. Simulation parameters for the varying parameter cases are summarised in Table 4.3. In the unseen or prediction cases, NIROM is usually trained with between 5 and 20 training runs, to show the effect of the number of training runs on NIROM's performance and to determine the ideal number of training runs that were most suitable for each production scenario. As the speed of POD reduction is dependent on the size of the training data, the more training runs, the more time is taken to decompose the data into hyperspace ^[125]. All simulations are run for 7500 days with

simulation outputs being written every 300days (25 report steps) for constant parameter cases and every 30 or 100 days for varying parameter cases, except where it is mentioned otherwise.

Table 4.2: Summary of simulation parameters investigated in the constant parameter cases of the synthetic gas model. Training runs are spaced equally between the ranges of values specified.

Parameter	Training Values	Prediction Values
Production Rate (MMscf/day)	3.8 – 7.4	3.54, 4.34, 5.14, 5.94, 6.74, 7.54
Production Bottom Hole Pressure (psi)	450 – 1250	400, 586, 773, 959, 1145, 1330
Porosity	0.135 – 0.3	0.12, 0.163, 0.207, 0.25, 0.293, 0.336
Enumeration Case	Same values as for production rate control	

Table 4.3: Summary of simulation parameters used to generate forced parameter production profiles for the synthetic gas reservoir. Parameters are selected from the list at random for the rate controlled production cases and these values are multiplied by distinct factors to generate varying profiles.

Parameters	Values
Production Rate (MMscf/day)	3.54, 4.34, 5.14, 5.94, 6.74, 7.54
Fixed BHP Limit (psi)	500

4.2 HETEROGENEOUS GAS RESERVOIR

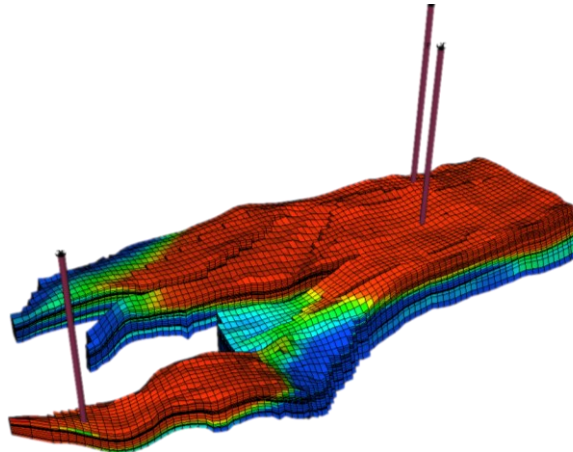


Figure 4.2: Gas saturation distribution in the modified Norne Gas field with strong aquifer support and 3 producers.

This is a geologically and structurally realistic model of a gas field based on the heterogeneous Norne oil field (Figure 4.2) ^[128, 129]. The gas reservoir is located at a depth of 2580m with a gas-water contact (GWC) at 2618m. The reservoir has 3 producers and no injector. The reservoir's dynamic model consists of 46 cells in the x-direction, 112 cells in the y-direction and 22 cells in the z-direction, although only 44,927 of these grid cells are active. Only the active grid cells are considered for NIROM evaluation. This means that the spatial resolution of the NIROM model was $nx = 44927$ for each snapshot of pressure and gas saturation distribution or $nx = 18$ for each production data snapshot (including gas and water production rates and cumulative rates, BHPs and gas-water ratios). A summary of the reservoir's properties is provided in Table 4.4 below.

Table 4.4: Reservoir properties of the modified Norne gas field.

Property	Values
Porosity	0.09 – 0.35
Horizontal permeability (mD)	0.08 – 999
Vertical permeability (mD)	0.004 – 1
Initial pressure (bars)	271.96
Initial Gas in Place (sm ³)	2.714×10^{10}
Gas density (kg/m ³)	0.991
Water density (kg/m ³)	1022
Rock compressibility (bar ⁻¹)	5.3×10^{-5}
Water formation volume factor	1.03
Production lifecycle (days)	7500

The first case where this modified Norne gas field was evaluated was for reproducing its training runs during BHP limit controlled production. In this case, there were no injection wells and the 3 production wells in the field were produced at 200, 220, and 200bars respectively (similar to the original Norne field model). Due to the initial surge in production and subsequent abrupt shut-in of a well, aquifer pressure support was introduced into the reservoir model to maintain its production for 7500 days. Schlumberger's ECLIPSE was set to give simulation output data at the start of each year, so a total of 21 report steps in the field's 7500 days lifecycle.

The reservoir was also examined for unseen predictions with NIROM. In the constant parameter cases, the modified Norne gas reservoir was again evaluated for predictions with aquifer pressure support. The strong bottom drive aquifer has an area of $1 \times 10^{14} \text{m}^2$, length of 200m and 200mD permeability. Another highlight of evaluating the NIROM with constant parameters is that NIROM was tested on its ability to make predictions beyond the production duration that it was trained for. The training data had a duration of 7500days lifecycle but field performance was predicted for 9000days, an extrapolation of 1500days or four report steps. A summary of the production parameters for the training and prediction runs in the constant parameter cases are given in Table 4.5. To ensure that there was gas production during the extra 1500 days that the NIROM was evaluated for, the gas production rate from the 3 producers were set to a ratio of 1.1: 1.5: 0.3, and during the BHP limit controlled production, wells 1 and 3 were produced at the production BHP specified in Table 4.5. Well 2 was produced at a BHP of 10bars higher than the other two wells, i.e. 10bars lower than in the first case (reproducing training runs) to accommodate the extra 1500 days of production that the NIROM was evaluated for.

Table 4.5: Summary of constant parameters that were evaluated for the modified Norne gas field. Training runs are spaced equally between the range of training values while prediction runs are distributed within and outside the range of values for training the NIROM.

Parameters	Training Values	Prediction Values
Production Rate (Sm^3/day) $\times 10^5$	2.0 – 5.2	1.85, 2.6, 3.3, 4.1, 4.8, 5.5
Production BHP (bars)	140 – 250	255, 231, 207, 183, 159, 135

For the varying parameter cases, the modified Norne gas field was investigated for 7500 days of production, and two types of production profiles were considered. When building the varying production profiles, more than 25 report steps were introduced into the profile to mimic a reasonable occurrence of gas rate changes during the production lifecycle. One of the production profiles had 75 report steps, where the gas production rate changed every two to three years (Figure 4.3) and there was another production profile that varied seasonally (summer, autumn, winter, spring), with at least 240 report steps (Figure 4.4). Moreover, in these cases, all the wells were initially set to rate controlled production with a BHP limit. In other words, the wells were produced by gas rate control that switches to BHP control when ECLIPSE estimates that a well is unable to meet up with the intended production rate. The production rates in Table 4.6 were provided for NIROM to create the production profiles for the wells, while the BHP limit was fixed at 60bars. The varying production profiles were formed by combining different sequences of rates, chosen from the values in Table 4.6, and durations for the rates. Another highlight of these varying parameter cases is that production wells can be shut-in during the production lifecycle of the field, and NIROM was evaluated for predicting their performance. Training and prediction runs should ideally have similar production profile scenarios, but the combination of rates and durations for each run was unique. The seasonally varying parameter cases of this heterogeneous gas reservoir have been evaluated with both POD-RBF and POD-AE, both results are compared in Chapter 8. These prediction cases were investigated for gas recovery factors of over 70%, and an average field pressure drop of 200bars during the 20years of production.

Table 4.6: Summary of production rates and BHP limit used to create the time-varying production profiles for the varying parameter cases of the modified Norne gas field. Production rates were multiplied by different unique factors to create distinct production profiles.

Parameters	Values
Production Rate (sm^3/day) $\times 10^5$	18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8
Fixed Production BHP limit (bars)	60

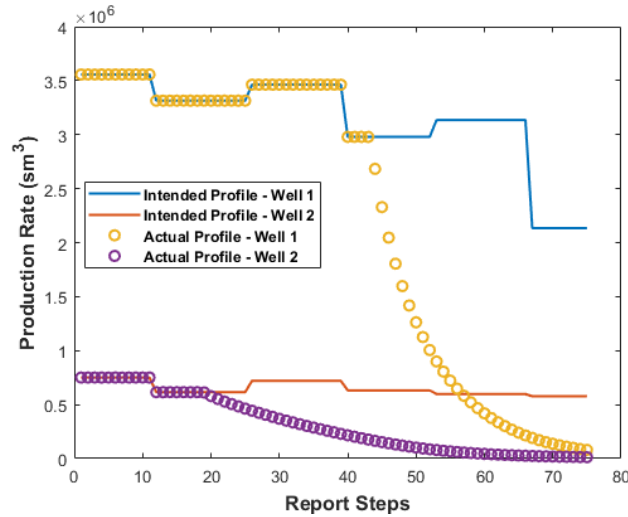


Figure 4.3: An example of the intended versus actual production profiles for training the modified Norne gas field where production rates vary after at least 1000 days. The intended production profile stipulates the production rate profiles that are generated from MATLAB (a combination of unique rates and durations), and the conventional simulator (ECLIPSE) is expected to follow for two wells over their entire production lifecycle. However due to the well BHP limit, ECLIPSE is unable to maintain the intended production profile, and has switched to BHP production control during the wells' production life. BHP controlled production commences at about the 45th and 18th report steps for wells 1 & 2 respectively (yellow and purple circle lines). The production profile that ECLIPSE eventually used for the simulation is the actual production profile.

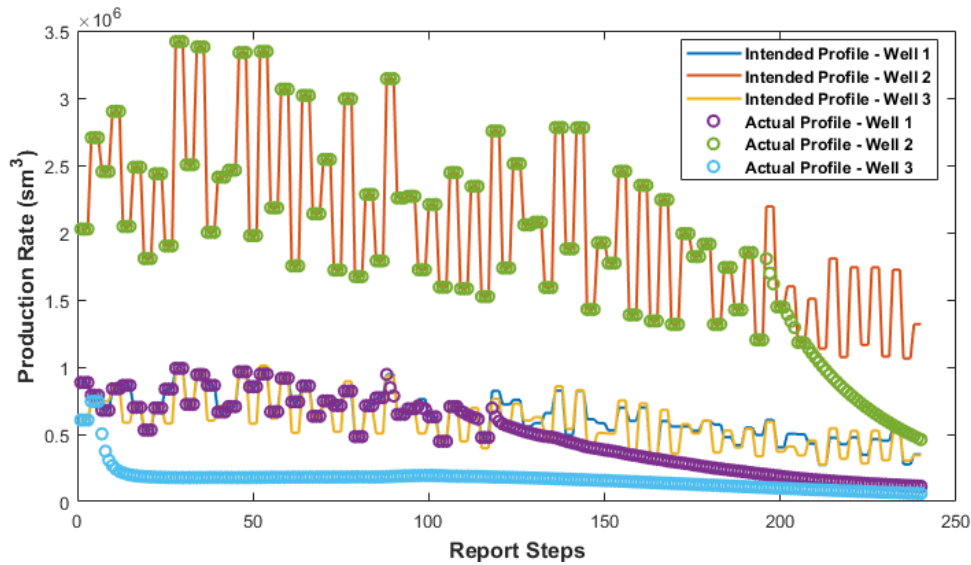


Figure 4.4: An example of the intended and actual production profiles for training the modified Norne gas field where production rates are varied every quarter of the year. Well controls switch from gas rate to BHP when ECLIPSE cannot maintain the intended production rate that was created in MATLAB. Eclipse actual production profile shows that well 3 was shut-in after the 230th report step (light blue circle-line).

4.3 WATER CONING RESERVOIR

Water influx in gas reservoirs may impact negatively on their production performance. When water breaks through a gas reservoir, it can lead to gas being trapped in the reservoir thereby partially or totally retarding gas production [11, 13]. To investigate NIROM's capacity for modelling this practical gas reservoir problem, a gas reservoir with significant water coning is considered. This model is based on the water coning model of Armenta (2003) [130] where a radial grid pattern is used to illustrate the

water coning mechanism in the gas reservoir. This water coning gas reservoir is located at a depth of 5000ft, with a thickness of 640ft and GWC at 5050ft. The reservoir has a single producer well that is completed just above the GWC. The reservoir's dynamic model is made up of $26 \times 4 \times 128$ grid cells in the r -, θ - and z -directions respectively. Figure 4.5 shows the gas saturation in the reservoir and the extent of water encroachment into the gas zone (cone height) at the end of the production life of the well. Table 4.7 provides a summary of some other properties of the water coning gas reservoir. Other information on the reservoir model are available in Appendix C.

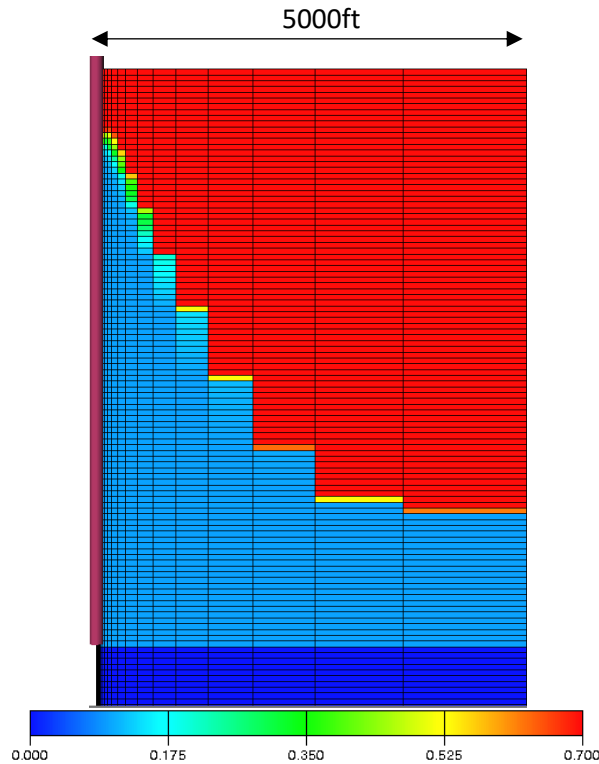


Figure 4.5: Gas saturation distribution in a gas reservoir model with significant water coning, where blue is low gas saturation and red is high gas saturation. The reservoir has one producer, completed at the end of the gas zone of the reservoir and just before the water (deep blue) zone of the reservoir.

Table 4.7: Some reservoir and production properties of the water coning gas reservoir.

Property	Value
Porosity	0.25
Permeability (mD)	$k_r = 10$, $k_\theta = 100$, $k_z = 5$
Initial pressure (psi)	1501
Number of active grid cells	12,192
Initial Gas in Place (BCF)	58.6
Gas density (lb/ft ³)	0.046
Water viscosity (cP)	0.68
Water density (lb/ft ³)	64
Rock compressibility (psi ⁻¹)	10×10^{-6}
Well inner radius (ft)	0.333
Production lifecycle (days)	7296 (30.4days $\times$ 240)

This water coning gas reservoir's performance has been explored for unseen constant parameter predictions and Table 4.8 shows the range of parameters that are investigated. The 20 years production performance data of the reservoir is outputted in 20 report steps (annually) and NIROM was trained with a maximum of 20 training runs.

Table 4.8: Training and prediction simulation parameters for a gas reservoir model with significant water coning.

Case	Parameters	Training Values	Prediction Values
1	Production Rate (MMscf/day)	1.00 – 1.45	0.9, 1.02, 1.14, 1.26, 1.38, 1.5
2	Production Bottom Hole Pressure (kpsi)	0.65 – 1.40	0.6, 0.77, 0.94, 1.11, 1.28, 1.45

4.4 FIELD X

The Field X model is a sector model of a real gas field that was provided by the project sponsor, Woodside Petroleum. It is located between a depth of 2916.2m and 3194.4m. The model was built using a 116×202×41 Cartesian grid with 127,945 active grid cells. The gas production is from a single well that is supported by a Carter Tracy aquifer model with an inner radius of 10000m and 5m thickness. Other properties of Field X are summarised in Table 4.9 and shown in Figure 4.6 and Figure 4.7. In this work, Field X has been produced with the seasonal production rate profiles with a fixed pressure limit of 120bars for BHP or 100bars for THP. In some instances, the well is shut in some months to the end of its production life. This field has a production lifecycle of 7560days and its performance data is outputted from ECLIPSE in 252 report steps (30 days interval). POD-RBF and POD-AE were both evaluated for modelling the performance of Field X, for 5 to 50 training runs, and their performances are compared in Chapter 8.

Table 4.9: Some reservoir properties of Field X

Property	Value (metric unit)
Porosity	0.27 – 0.3
Permeability (mD)	27.5 - 2700
Initial pressure (bars)	326.9
Gas density (kg/m ³)	0.80081
Water density (kg/m ³)	1014.2
Rock compressibility (bar ⁻¹)	2.3206×10 ⁻⁵
Water formation volume factor	1.0252
Well diameter (m)	0.2159
Initial Gas in Place (sm ³)	14.57×10 ⁹
Production lifecycle (days)	7560

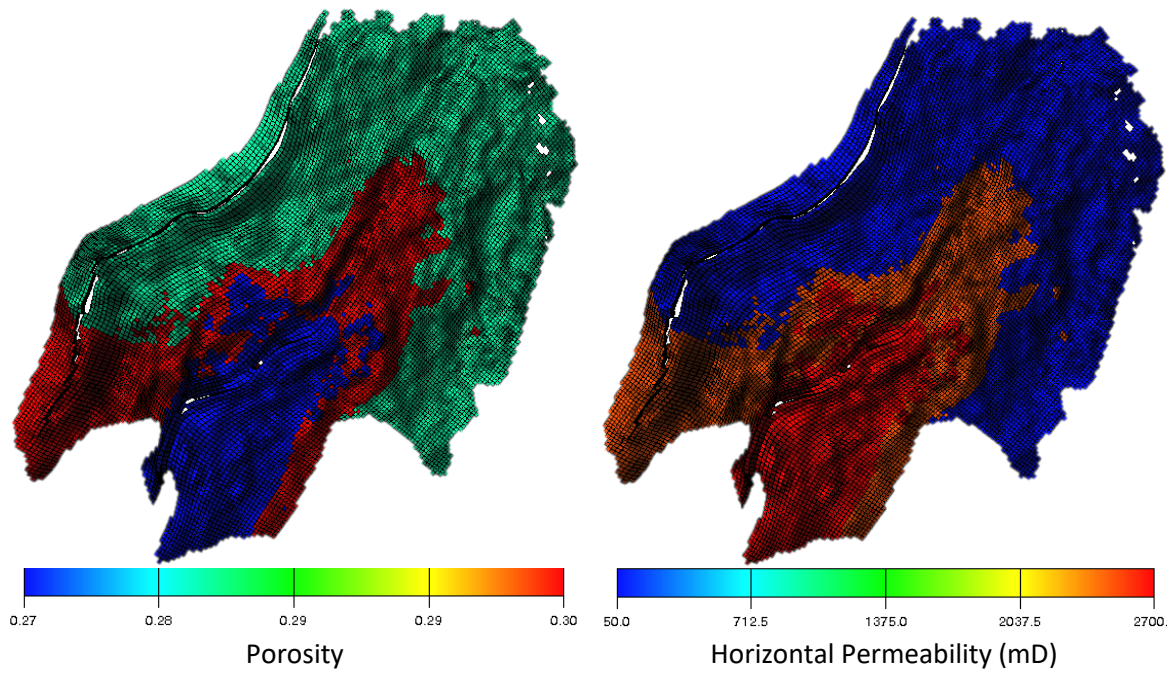


Figure 4.6: The distribution of (a) porosity and (b) permeability in the horizontal direction of Field X. It is not clear why the relationship between the porosity and horizontal permeability in Field X is not directly proportional as most literature/ correlations show, i.e. the zones with the highest porosity have median permeability and the zones with the lowest porosity have the highest permeability values. Overall the porosity and permeability of Field X are generally high.

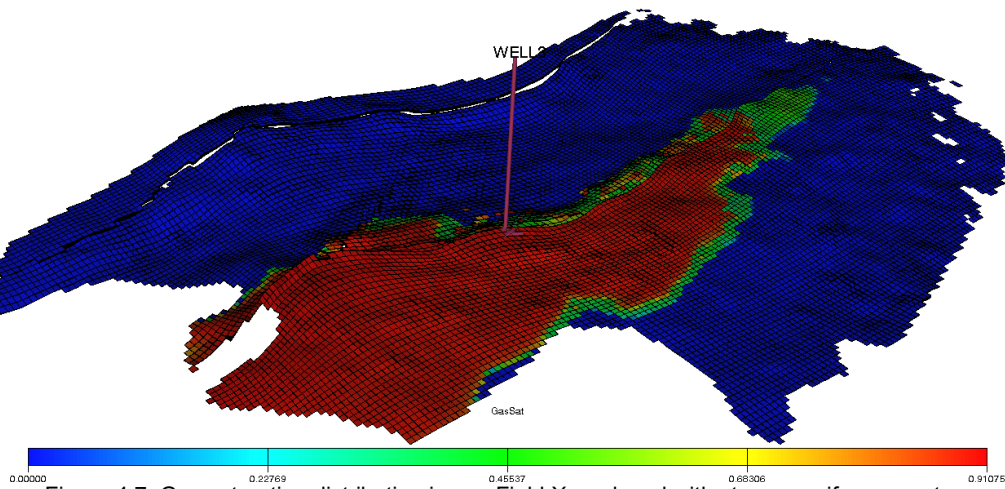


Figure 4.7: Gas saturation distribution in gas Field X produced with strong aquifer support.

In chapter 7, NIROM is analysed for modelling the performance of Field X for both constant and varying parameter cases. The constant parameter prediction cases that are evaluated with NIROM are shown in Table 4.10 while the parameters used for evaluating the varying parameter cases are in Table 4.11. The constant parameter cases were analysed for rate controlled production while the varying parameter cases were evaluated for seasonally varying production rates with BHP or THP limit production controls. The seasonally varying parameters were multiplied with random factors to generate the intended varying production profiles. Field X was evaluated with NIROM for 5, 10 and 20

training runs for 7500 days of production that were reported at the start of every year (21 years or report steps).

Table 4.10: Summary of constant parameters that were evaluated for Field X. Training runs are spaced equally between the range of training values while prediction runs are distributed within and outside the range of values for training the NIROM.

Parameters	Training Values	Prediction Values
Production Rate (sm^3/day) $\times 10^5$	9 – 20	8, 10.6, 13.2, 15.8, 18.4, 21

Table 4.11: Summary of production rates and pressure limit used to create the varying production profiles for the varying parameter cases of Field X. Production rates were multiplied by different unique factors to create distinct production profiles.

Parameters	Values
Production Rate (sm^3/day) $\times 10^5$	27, 25.5, 24, 22.5, 21, 19.5, 18, 16.5, 15, 13.5, 12
Fixed Production BHP limit (bars)	100
Fixed Production THP limit (bars)	120

CHAPTER 5: TRAINING MODEL RESULTS AND DISCUSSION

In this work, the process of developing a NIROM that can model gas reservoir performance in unseen scenarios started with using very simple synthetic models and then adding complexity and realism to the models. The NIROM results discussed in this chapter are the first attempt at coupling a NIROM built in MATLAB to the ECLIPSE simulator. This chapter is considered to be the proof of concept stage of this research, to ascertain that NIROM can be implemented for fast gas reservoir performance estimation from a commercial simulator. The initial emphasis here was to find out which hyper-parameters were most suitable for modelling the performance of gas reservoirs. The first objective was to determine which RBF(s) are most suitable for modelling different gas production scenarios and also to determine if the energy criterion of 99.9% from 3.5 (suggested by Xiao et al. ^[38]) is sufficient for gas reservoir modelling. To determine the most suitable RBF(s), the NIROM was applied to reproducing the training runs for which it was trained, with 7 different radial basis interpolation functions, i.e. linear, inverse multi-quadratic, modified thin-plate spline, inverse quadratic, Gaussian, cubic and multi-quadratic functions. NIROM estimates were obtained by modelling the dynamics of the reservoir based on the initial condition training data at $t = 0$ for each test case. These estimates were then compared to the actual solutions from ECLIPSE.

The variables that were estimated include the pressure distribution in the reservoir, and if there was aquifer influx, the gas saturation distribution. For production data, either well BHP or gas production rate (depending on the chosen well control), cumulative gas production and, if there was an aquifer, the producing water-gas ratios were estimated. Variables were estimated at time intervals of 300 days and all simulations ran for 7500 days (about 20 years), with one exception which is discussed later. Initially, an energy criterion of 99.9% (3.5) was used to determine the number of POD basis functions needed to project the simulations into hyperspace, although in some cases more POD basis functions were needed to obtain an acceptable NIROM solution. To better visualise the pressure or saturation distribution in the reservoirs, especially for the reservoirs with more than 2400 grid cells, the average value of the pressure or saturation across the reservoir is plotted at every report step instead of their values in each of the grid cells. These results are discussed in the following sections.

5.1 SYNTHETIC GAS RESERVOIR

This is a homogeneous gas reservoir with $19 \times 21 \times 6 = 2394$ active grid cells. It has been set to produce and generate pressure and saturation distribution output data every 300 days for 25 report steps, a total of 7500 days. All pressure units are in psi.

5.1.1 PRODUCTION CONTROLS

5.1.1.1 GAS RATE CONTROL

In this case, the average reservoir pressure declined continuously throughout the 20 years of production and there was a pressure drop of over 20psi across the reservoir at each report time. Figure 5.1 shows the pressure distribution predicted by ECLIPSE and the POD-RBF NIROM in each layer of the synthetic model at the 6th report step (1800days), together with the difference between the two solutions. Figure 5.2 further highlights that the NIROM correctly captures both the simulated pressure distribution in each of the six layers and the overall pressure depletion from the well location to the boundaries of the reservoir.

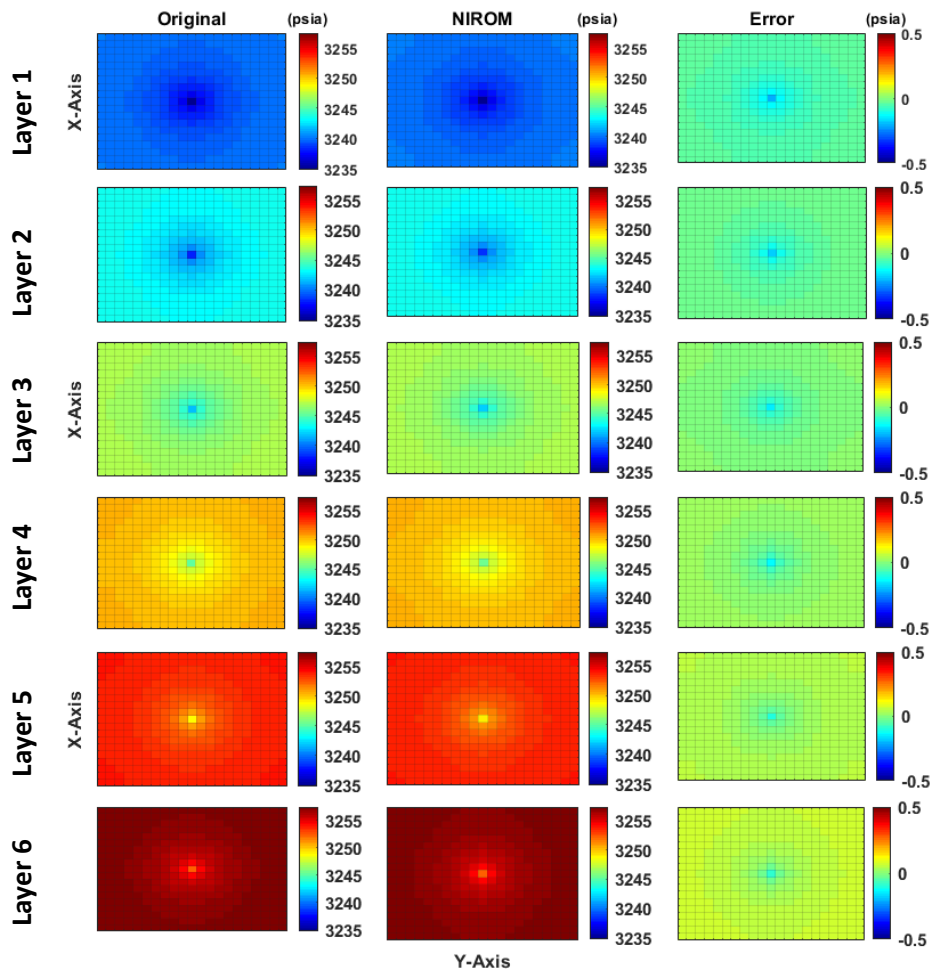


Figure 5.1: Pressure distribution across each layer of the synthetic gas reservoir at $t = 1800$ days. This reservoir has one well which is drained from the centre of the reservoir at a constant production rate of 6.77MMscf/day. The reservoir pressure is depleted the most at the well location and then radially outwards to other grid cells in the reservoir. Pressure depletion in the reservoir is also from top to bottom, such that lower layers are less depleted and have higher pressure than the top layers. The NIROM (using any of the 7 RBFs) can reproduce the radial and downward pressure depletion pattern seen in the synthetic gas reservoir with negligible errors in the NIROM results compared to ECLIPSE.

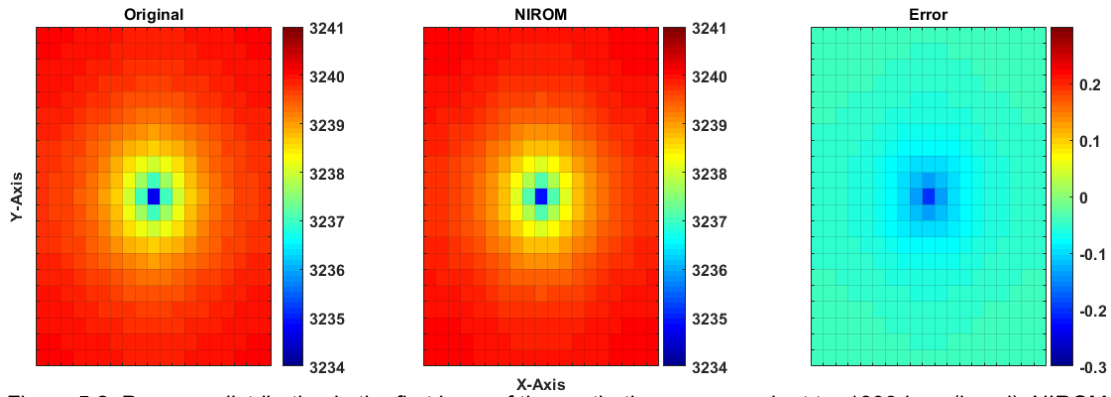


Figure 5.2: Pressure distribution in the first layer of the synthetic gas reservoir at $t = 1800$ days (in psi). NIROM adequately captures the radial and outward pressure depletion in the synthetic gas reservoir.

Figure 5.4 compares the actual reservoir pressure in each grid cell to the values calculated by the NIROM. Each column shows the reservoir pressure in all grid cells at a given report time, normalised to the average pressure in the reservoir, with the leftmost column showing the initial normalised pressure distribution and the columns to the right showing the pressure distribution as time increases. The pressure in all the reservoir grid cells are arranged by natural ordering. The error in these results are negligible and they were obtained using 1 POD basis function, as suggested from using the energy criterion of 99.9% in 3.5. The results presented here were obtained with the linear radial basis function, but similar results were obtained using the other 6 basis functions.

For brevity in the subsequent results, the average of the pressure distribution (and the gas saturation distribution if there is an aquifer) is reported for each reporting time.

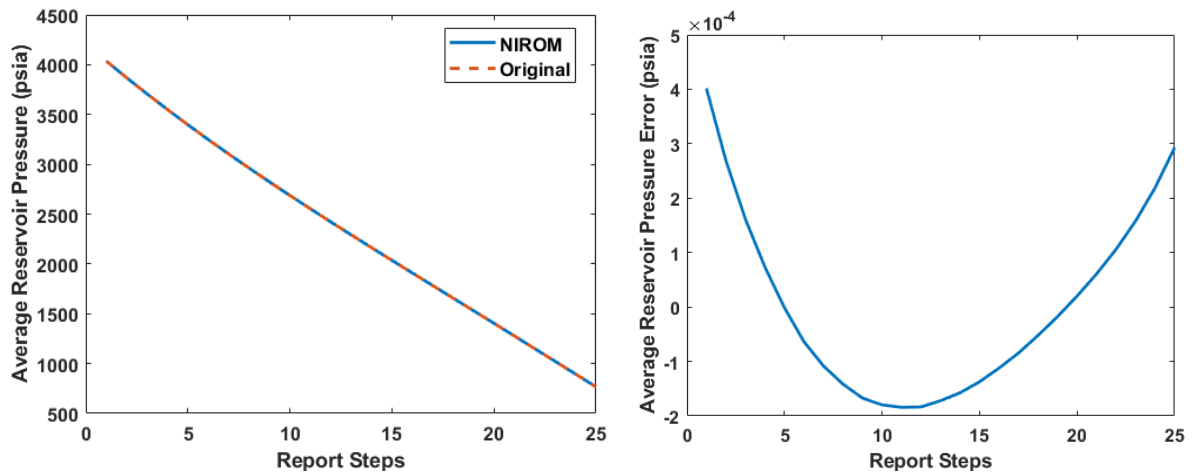


Figure 5.3: The average of the reservoir pressure distribution estimated by the NIROM at each report time in the synthetic gas reservoir at constant production rate control. The NIROM gives a satisfactory match to ECLIPSE with negligible errors. The same result is obtained with all 7 RBFs using 1 POD basis function (as suggested by the energy criterion).

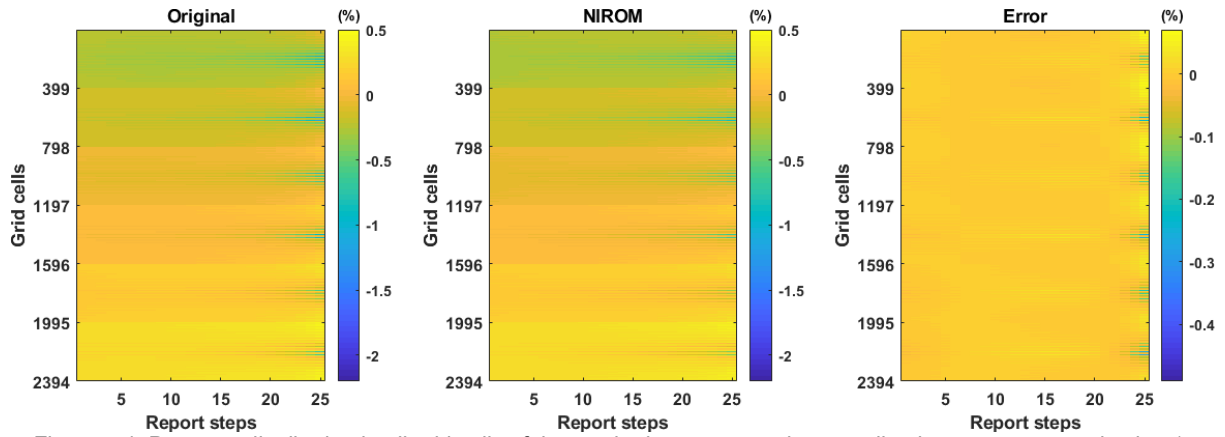


Figure 5.4: Pressure distribution in all grid cells of the synthetic gas reservoir, normalised to zero mean and using 1 POD basis function (as suggested by the energy criterion). The normalised pressure in all 2394 grid cells of the reservoir are plotted as vectors for each of the 25 report steps. The vectors are obtained by a natural ordering of the pressure in all the active grid cells, such that every 399 cells represent a single layer of the reservoir and the well is located in the middle of every 399 cells. Similar results were obtained from all 7 RBFs, showing that the reservoir pressure decreases during constant production rate control. The most errors were experienced in the reservoir layers hosting the production well, as seen by the darker streaks in the middle of each layer.

Figure 5.3 shows the average pressure distribution at each report time of the simulation. NIROM captures the pressure depletion in the reservoir over time as was originally predicted by ECLIPSE.

The ability of the NIROM to estimate the production behaviour of this synthetic reservoir is further illustrated in Figure 5.5 and Figure 5.6. NIROM estimated BHP and cumulative production agree well with the original values obtained from ECLIPSE simulation. Using a minimum energy criterion of 99.9% (3.5) suggested only 1 POD basis function was required for this estimation. However, Figure 5.6 shows that better results, with negligible errors, can be obtained from using more POD basis functions. The solutions shown in this case are similar for all 7 RBF interpolation functions explored.

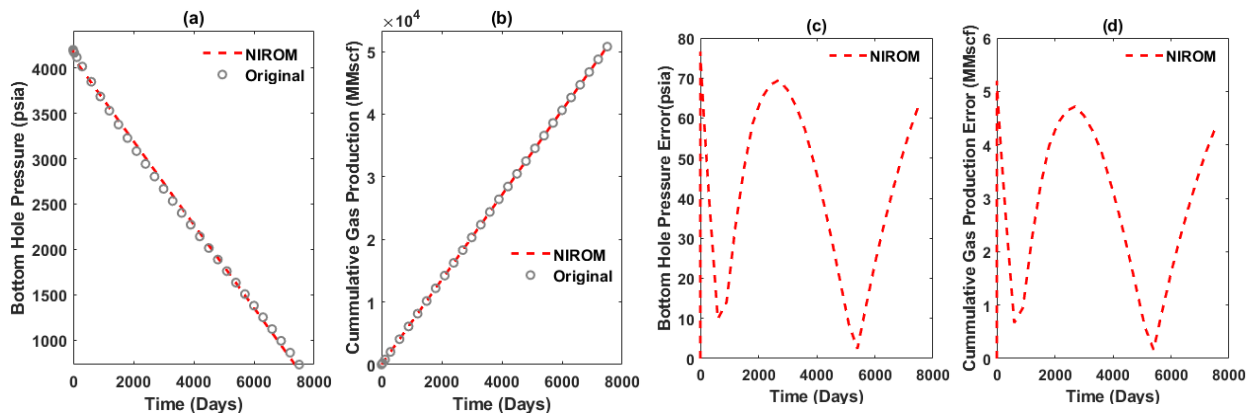


Figure 5.5: Bottom hole pressure and cumulative gas production profiles estimated by the POD-RBF NIROM in comparison to ECLIPSE using 1 POD basis function, as suggested by the 99.9% energy criterion (a & b). Errors in the estimation of the c) well bottom hole pressure were less than 80psi, while errors in the d) cumulative gas production were less than 6MMscf (0.01%) for all 7 RBFs.

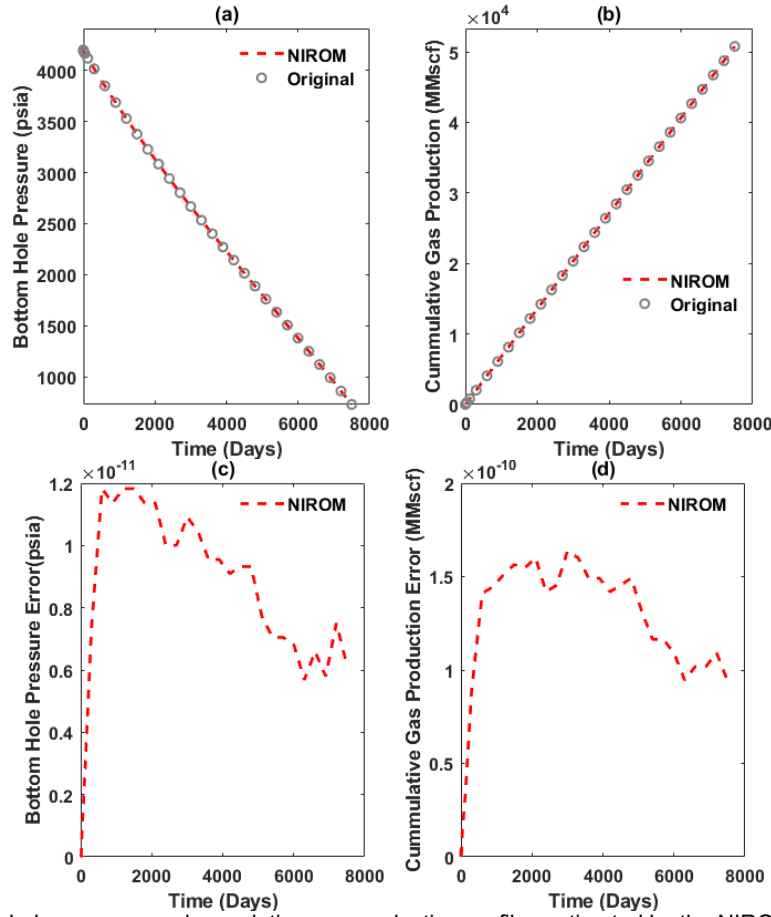


Figure 5.6: Bottom hole pressure and cumulative gas production profiles estimated by the NIROM and predicted by ECLIPSE using 2 POD basis functions (a & b). Adding an extra basis function to the number of basis functions suggested by using the energy criterion gave negligible errors in c & d, for both the decline in bottom hole pressure and the increase in the cumulative gas production. Similar results were obtained with all 7 RBFs.

5.1.1.2 BHP CONTROL

In this case, the average reservoir pressure fell to the BHP of 1000psi after 5 report steps (1500 days). Gas production was practically negligible afterwards and the well shut-in at 3300days. Nevertheless, NIROM RBFs 1 and 5 gave accurate estimates of the reservoir pressure distribution (errors less than $4e-4$), while RBFs 2, 3, 4, 7 and 6 gave maximum absolute errors of 7, 13, 5, 4 and $8E17$ psi respectively (Figure 5.7). This test case was used to evaluate the capability of the NIROM to deal with systems in which there was initially a sudden change in behaviour followed by small changes. Previous analysis by Kostorz et al., (2019)^[37] indicated that the NIROM would not provide good estimates in such cases. For this case, it was observed that since each grid cell of the reservoir has pressure values, even though they were at the constant pressure limit in the neighbourhood of 1000psi, NIROM was able to estimate the grid cell pressure distributions. However, for the production data which are dependent on gas production (for instance gas and water production rates, cumulative productions and well pressures), NIROM could not reproduce any meaningful results when there was no production data available for it to work with (after 3300 days). Hence the snapshot matrix for training the NIROM, as well as NIROM

simulations were truncated for this case to contain only the first 11 report times (3300 days) before the well was shut-in.

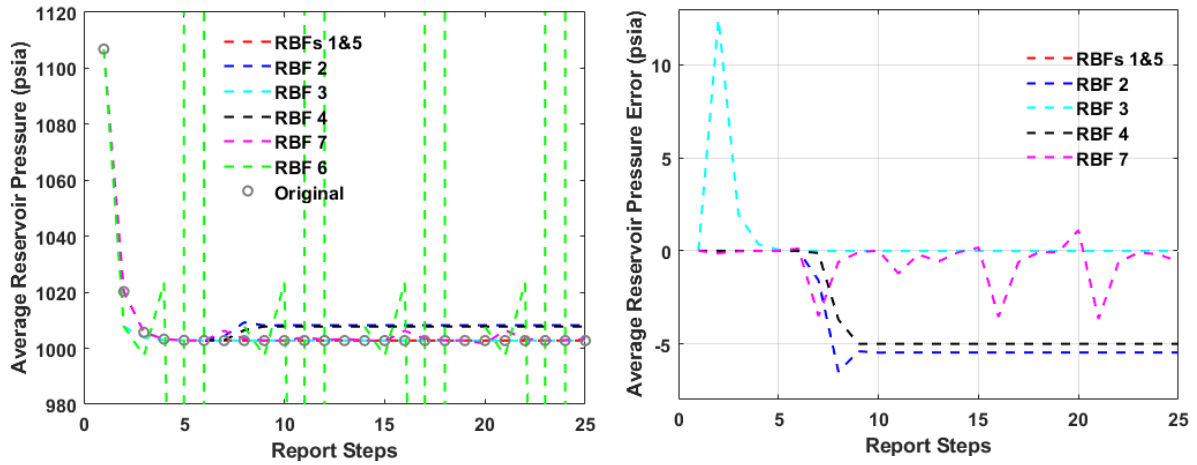


Figure 5.7: The average of the synthetic gas reservoir pressure distribution estimated by the NIROM at each report time when the well was produced by BHP control. These results were obtained using 1 POD basis function (as predicted by the energy criterion). All RBFs (except RBF 6) gave consistent results with ECLIPSE, RBFs 1 and 5 gave the most accurate results.

After truncating the training runs matrix to 3300 days, RBFs 1, 3, 5, 6 and 7 were able to reproduce the sharp changes in production data that occurred over the first 11 report steps using 1 basis function (as suggested by the 99.9% energy criterion). RBFs 2 and 4 gave good estimates up to the 8th report time and then started over predicting (Figure 5.8). Increasing the number of POD basis functions by one gave slightly better NIROM estimates for RBFs 2 and 4 but then RBF 6 developed worse results (Figure 5.9). Hence increasing the number of POD basis functions used in NIROM predictions can be useful for obtaining better results as well as exposing which RBFs are unstable/ inconsistent in their interpolation behaviour.

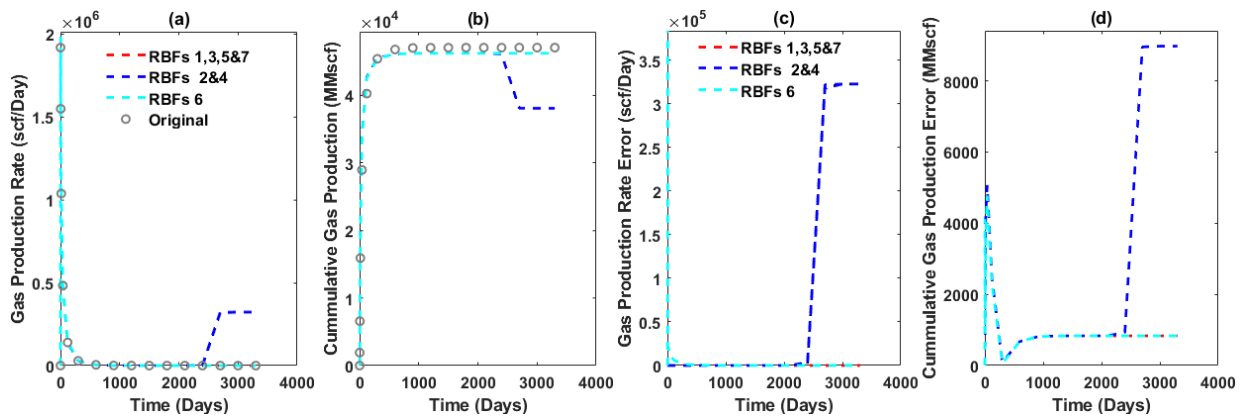


Figure 5.8: Reservoir production data estimated using 1 POD basis function (as suggested by the energy criterion). The NIROM results from RBFs 1, 3, 5, 6 and 7 are consistent with those of ECLIPSE for the production duration of 3300 days (11 report steps) at constant production BHP limit while RBFs 2 and 4 gave satisfactory results for only the first 2400 days.

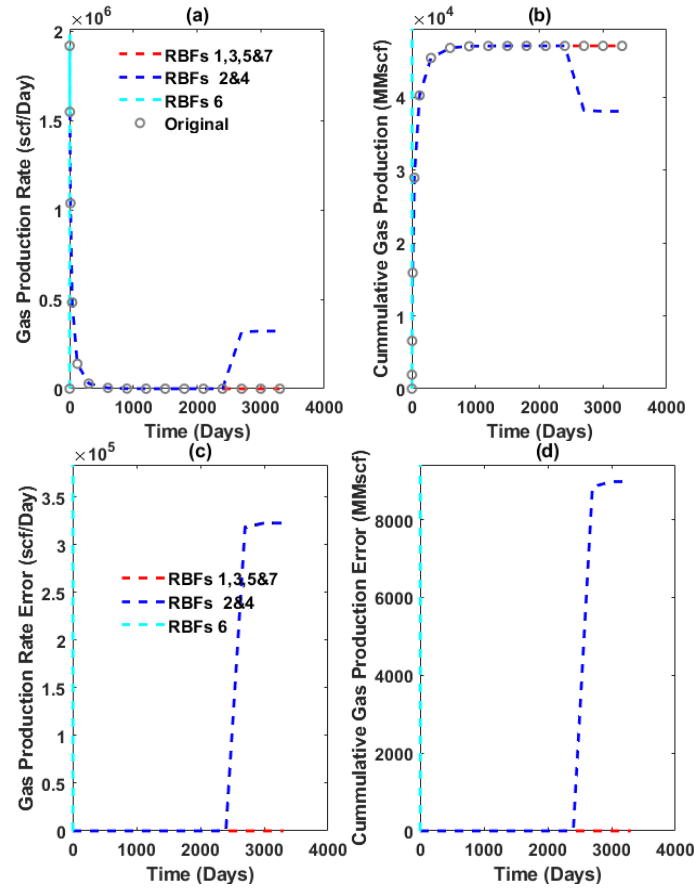


Figure 5.9: Reservoir production data estimated using 2 POD basis functions. By adding an extra basis function, NIROM results are slightly improved for all RBFs except RBF 6 which explodes, with overestimation. RBFs 2 and 4 still give satisfactory results up to 2400 days after which they over-predict the gas production rate and under predict the cumulative gas production.

5.1.2 AQUIFER MODELS

Here, the NIROM is investigated for its ability to capture the water saturation distribution, pressure distribution and the additional production responses that are obtained when there is aquifer influx. For these aquifer cases, the following aquifer production variables are estimated with NIROM: the aquifer influx rate (stb/day), cumulative aquifer influx (stb) and aquifer pressure (psi). For the production well, the water production rate (stb/day), cumulative water production (stb) and water-gas ratio (Mscf/stb) are investigated. This makes a total of 8 production variables that are being estimated with NIROM, i.e. $nx = 8$ for each production profile snapshot.

5.1.2.1 NUMERICAL AQUIFER MODEL

Introducing numerical aquifer support into the synthetic reservoir's base case with BHP limit controlled production allowed the producer to remain open for the 20 years lifecycle that was simulated, with water breaking through in the well. Figure 5.10 and Figure 5.11 show the average of the pressure distribution and the average of the saturation distribution at each report time. The

NIROM correctly estimated the pressure and saturation distribution in the reservoir's grid cells for all the RBFs using 1 and 3 POD basis functions respectively, as suggested by the energy criterion of 99.9%. Figure 5.12 shows the gas saturation distribution in the first layer of the reservoir at 6300 days while Figure 5.13 shows the gas saturation distribution in the entire reservoir with each column representing the gas distribution at each report step. Recall that in these BHP limit controlled production examples of the synthetic gas reservoir model, most of the gas in the reservoir is produced in the first 1500 days of production. Hence by 6300 days (shown in Figure 5.13), the gas saturation in the last 5 layers of the reservoir is negligible.

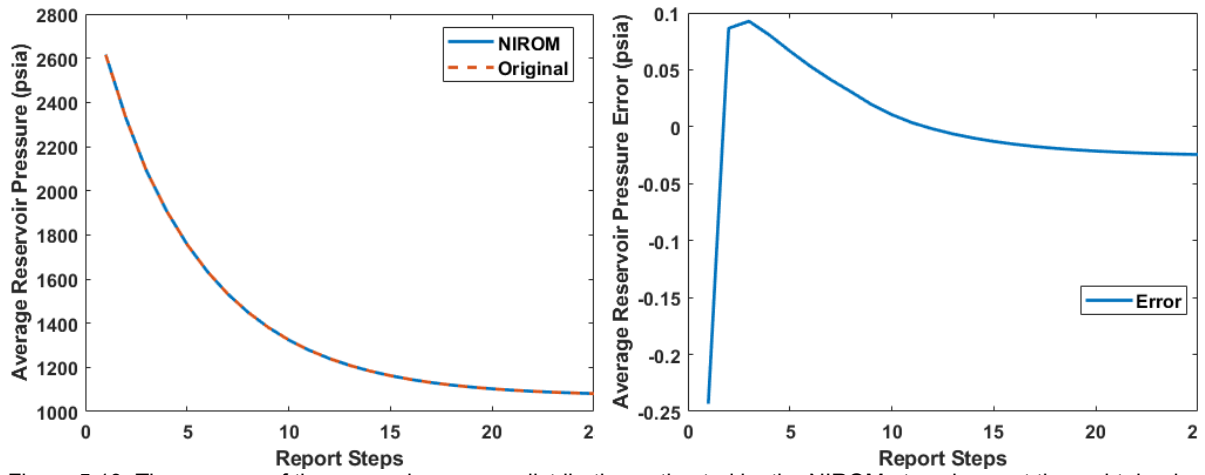


Figure 5.10: The average of the reservoir pressure distribution estimated by the NIROM at each report time, obtained using 1 POD basis function (as estimated by the energy criterion) in a BHP controlled gas well, supported by a numerical aquifer. Aquifer support maintained the average reservoir pressure above the 1000psi BHP limit, resulting in continued gas production for 7500 days. All RBFs gave the same estimates of pressure distribution.

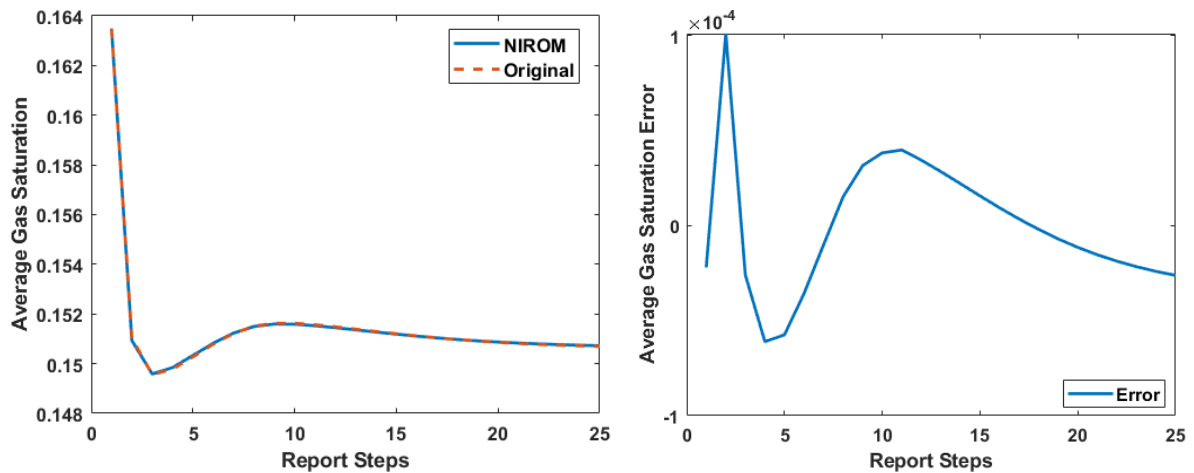


Figure 5.11: The variation over time of the average gas saturation distribution in the synthetic gas reservoir, predicted by NIROM using 3 POD basis functions (the predicted energy criterion). The reservoir is produced at a BHP limit of 1000psi with numerical aquifer support. All 7 RBFs give satisfactory match of the gas saturation decline.

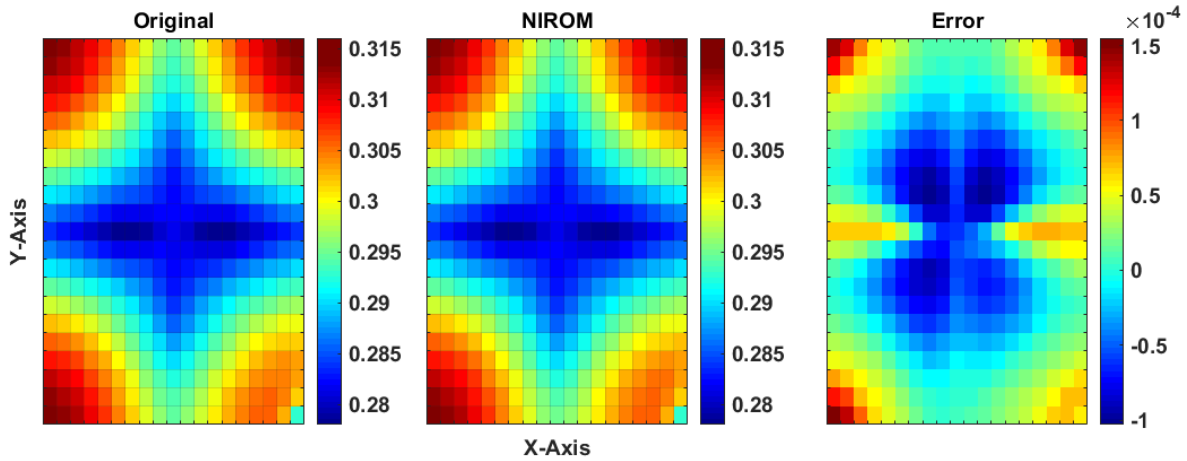


Figure 5.12: Gas Saturation distribution in the first layer of the synthetic gas reservoir at $t = 6300$ days. NIROM adequately captured the pattern of gas depletion in the synthetic gas reservoir with a numerical aquifer.

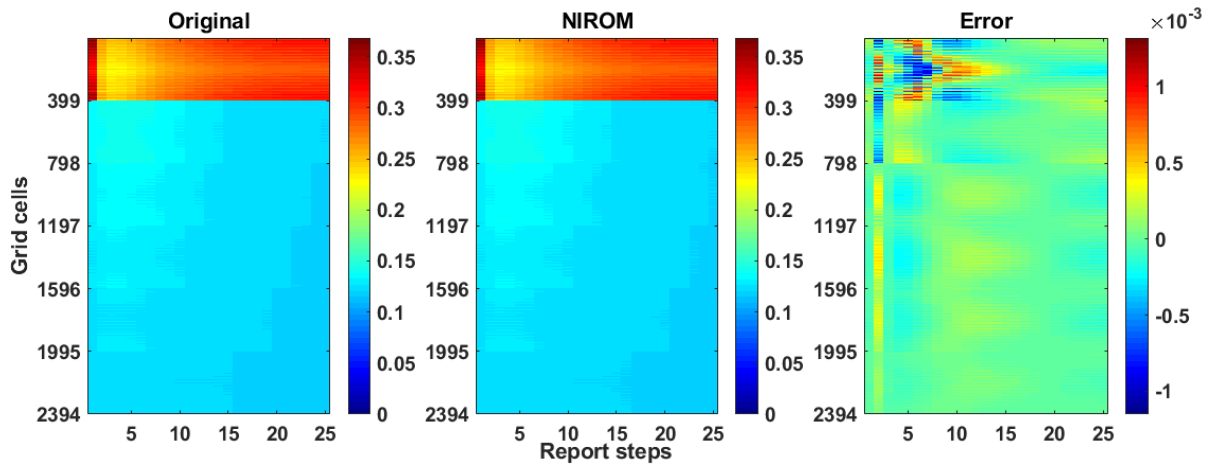


Figure 5.13: Gas saturation distribution in all grid cells of the synthetic gas reservoir with a numerical aquifer. The error in NIROM estimates were less than 1% of the original solutions from ECLIPSE. Although most of the reservoir was already depleted by 6300 days, NIROM adequately captured the saturation distribution across the 6 layers of the synthetic gas reservoir.

For the production data, 3 POD basis functions (the suggested energy criterion) were needed to obtain good NIROM estimations (Figure 5.14), the error profiles are in Figure 5.15. Four (4) POD basis functions can be used to obtain more accurate results as reflected in the error plot of Figure 5.16. And to satisfactorily estimate the water-gas ratio for this numerical aquifer case, 7 POD basis functions were required (Figure 5.17). NIROM results from all RBFs (except RBF 6) gave similar results. RBF 6 overestimated the production data and has not been included in the error plots of Figure 5.15, Figure 5.16, and Figure 5.17b.

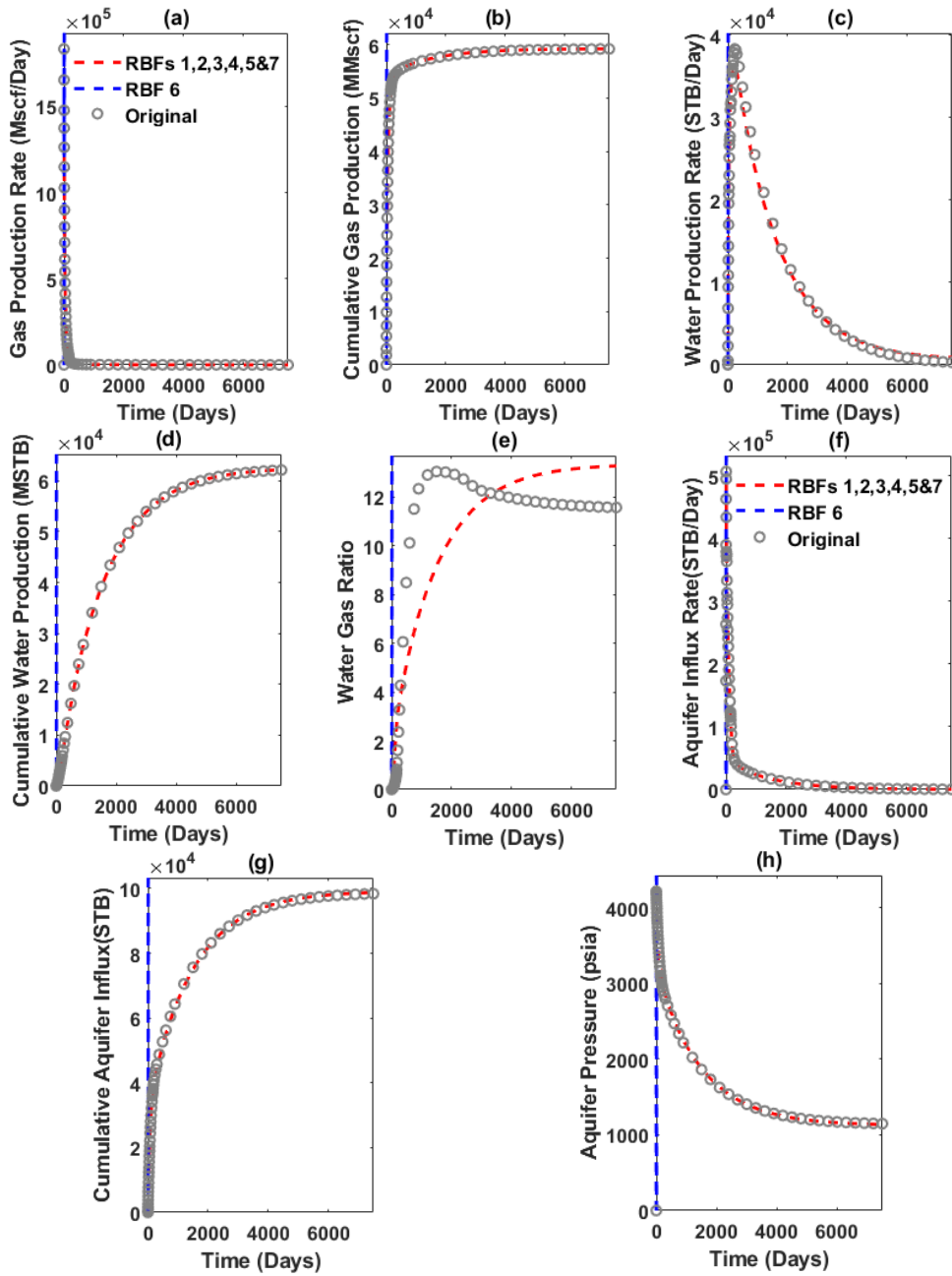


Figure 5.14: Production profiles using 3 POD basis functions (the energy criterion) for the synthetic gas reservoir supported by a numerical aquifer. All RBFs except RBF 6 give satisfactory NIROM estimates of (a) gas production rate (b) cumulative gas production (c) water production rate (d) cumulative water production (f) aquifer influx rate (g) cumulative aquifer influx and (h) aquifer pressure. None of the RBFs gave an accurate estimate of (e) the water gas ratio in the reservoir.

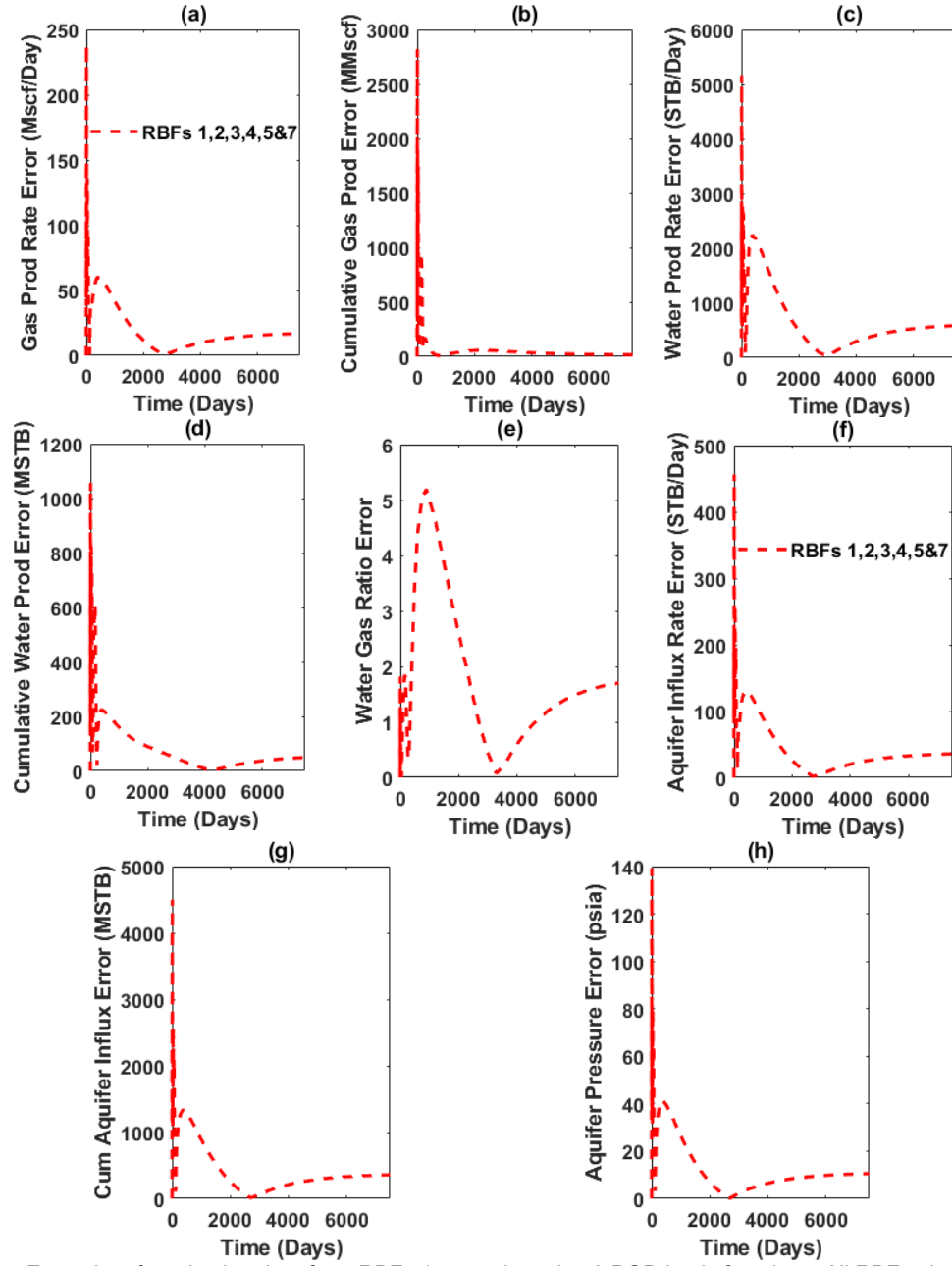


Figure 5.15: Error plot of production data from RBFs 1 - 5 and 7 using 3 POD basis functions. All RBFs give errors that are at least an order of magnitude lower than the actual data for all production profiles, except the water-gas ratio with almost 50% error.

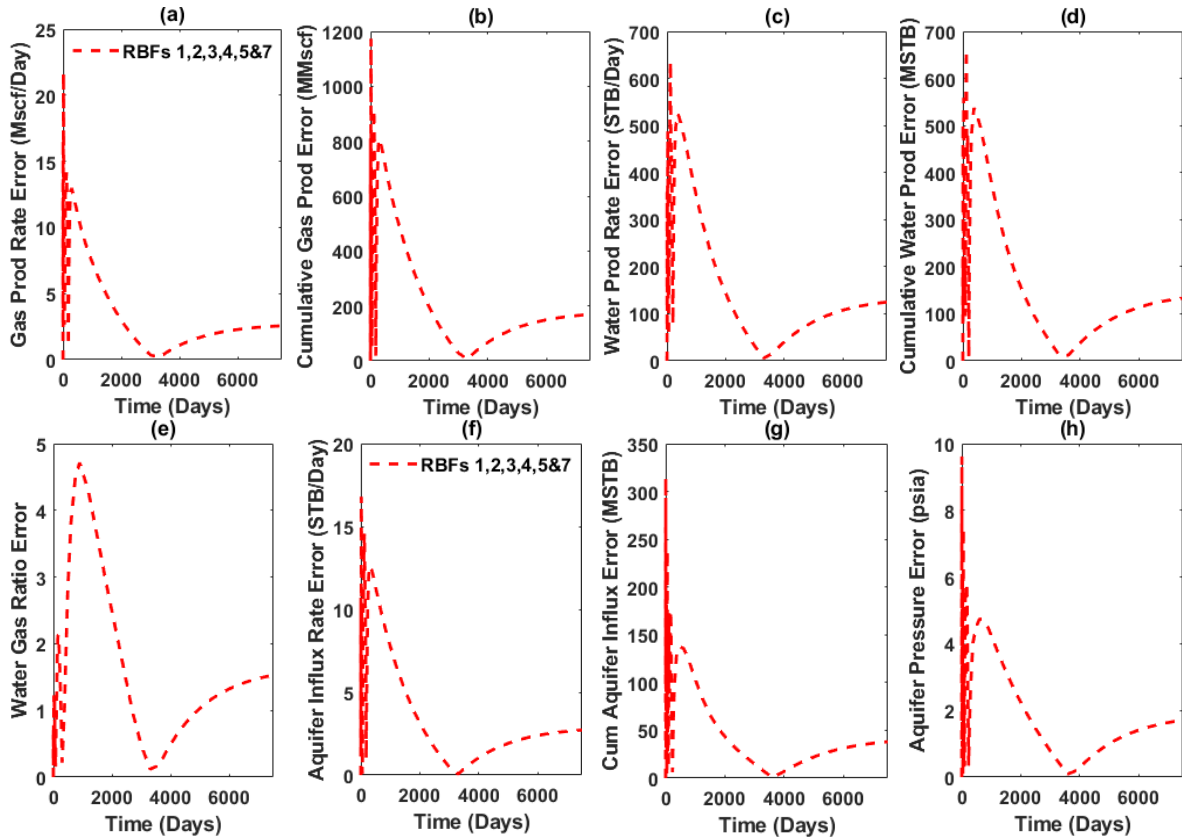


Figure 5.16: Error plot of production data from all RBFs using 4 POD basis functions (energy criterion = 3 POD basis functions). RBFs 1 - 5 and 7 give lower error values by adding an extra basis function to the NIROM simulation. Estimation of the water-gas ratio still shows the most significant error.

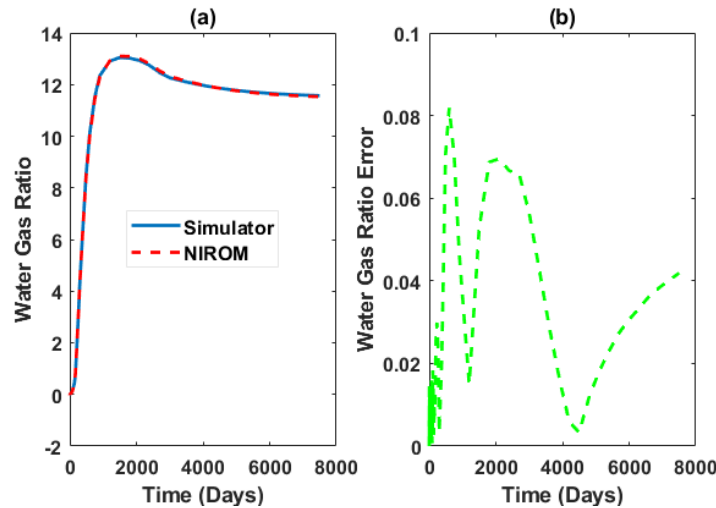


Figure 5.17: NIROM's satisfactory estimation of water gas ratio using 7 POD basis functions (energy criterion = 3 POD basis functions). RBFs 1 - 5 and 7 give the same result.

5.1.2.2 ANALYTICAL AQUIFER MODELS

The Carter Tracy aquifer model provided significant pressure support to the reservoir to maintain production at pressures above 3600psi throughout the production life cycle. All NIROM RBF models satisfactorily estimated the average pressure decline in the reservoir with similar results (Figure 5.18). As for the gas saturation distribution in the reservoir, only RBFs 1, 3 and 6 consistently gave meaningful

and satisfactory estimates (Figure 5.20). For pressure and saturation distribution in the reservoir, 2 POD basis functions gave good results (Figure 5.18 to Figure 5.20). Figure 5.19 shows that NIROM satisfactorily estimated the pressure distribution in all the grid cells of the reservoir.

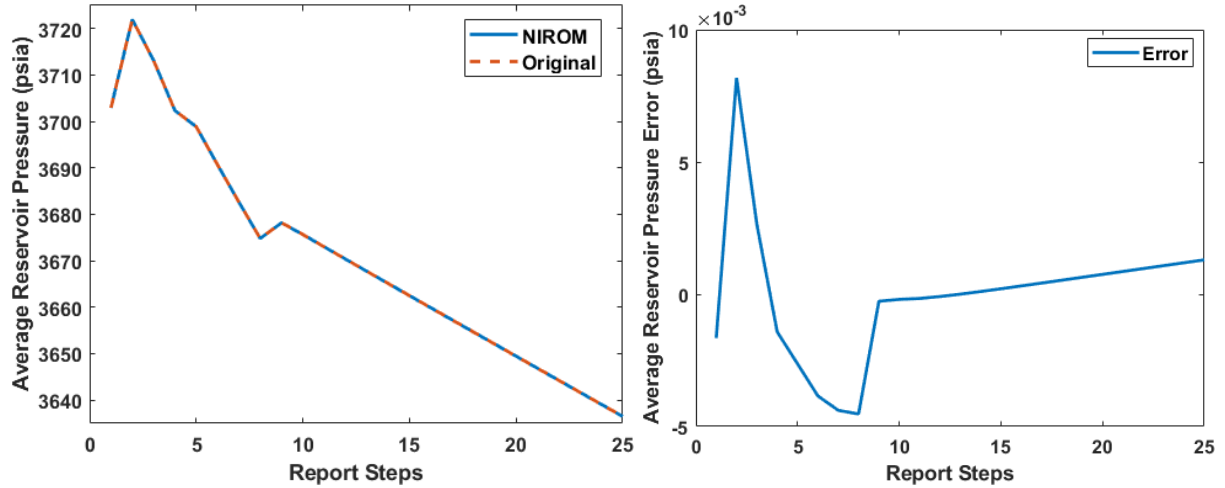


Figure 5.18: Average of the reservoir pressure distribution at each report step of the synthetic gas reservoir with a Carter Tracy aquifer for all RBFs when using 2 POD basis functions (the energy criterion). This analytical aquifer supports reservoir pressure with less than 100 psi pressure depletion throughout production.

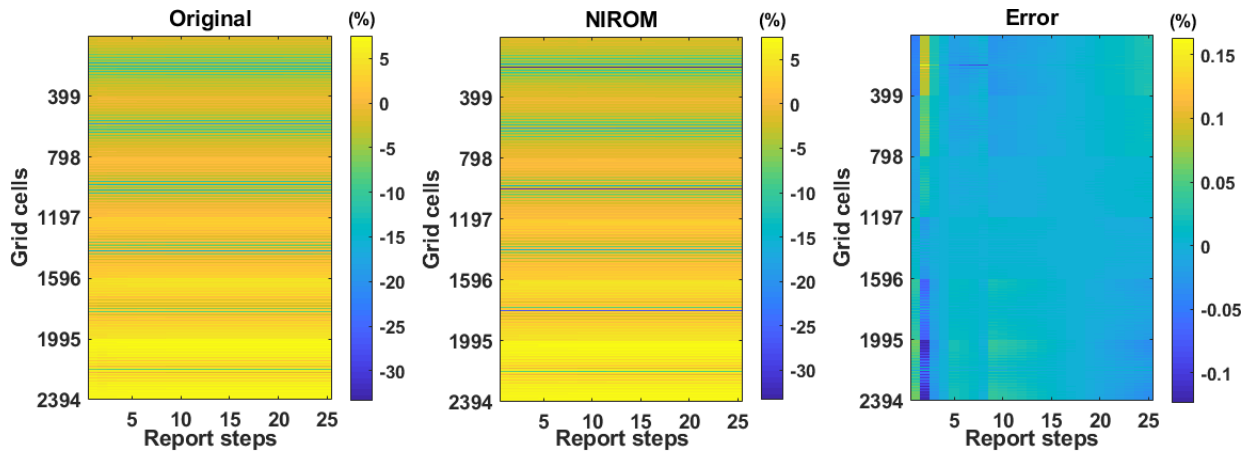


Figure 5.19: Pressure distribution in all grid cells of the synthetic gas reservoir with a Carter Tracy aquifer. Reservoir pressure values were normalised to zero mean and the NIROM was implemented with 2 POD basis functions (as suggested by the energy criterion). The most visible errors in the NIROM predictions were experienced in the reservoir layers hosting the production well, as seen in the darker streak lines in the middle of each layer.

For the production data, Figure 5.21 shows only the results for NIROM's estimate of the parameters that needed the most POD basis functions. NIROM required 6 POD basis functions to give accurate estimates of the cumulative gas production, water production rate and water-gas ratio (mostly water-related data). With 3 POD basis functions (99.9% energy criterion), all the other production data were estimated with satisfactory results. Only RBF 6 gave unsatisfactory results and this is not included in Figure 5.21.

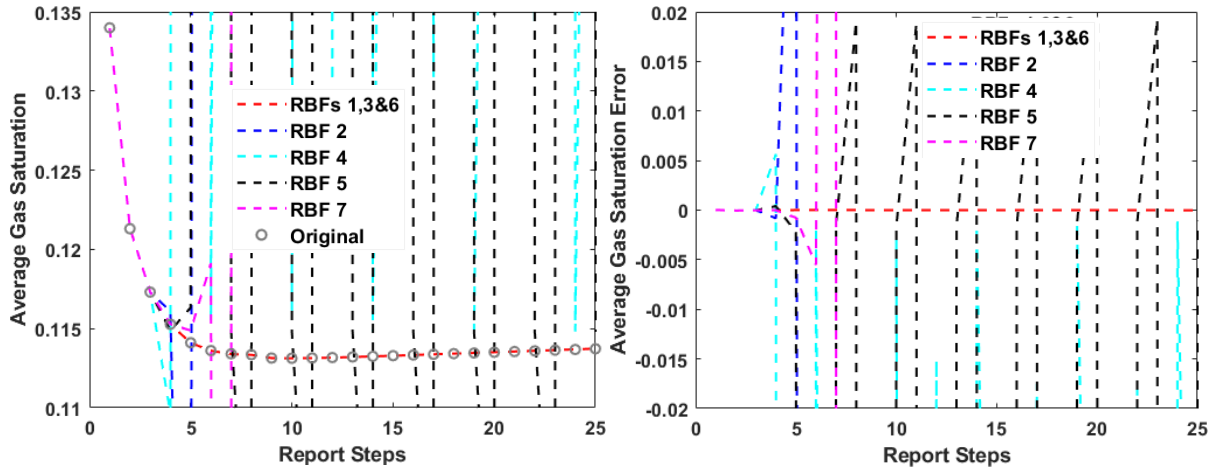


Figure 5.20: Average of the gas saturation distribution at each report time, obtained from the NIROM using 2 POD basis functions for the synthetic reservoir with a Carter Tracy aquifer. Only RBFs 1, 3 and 6 gave reasonable estimates of the gas depletion in the reservoir.

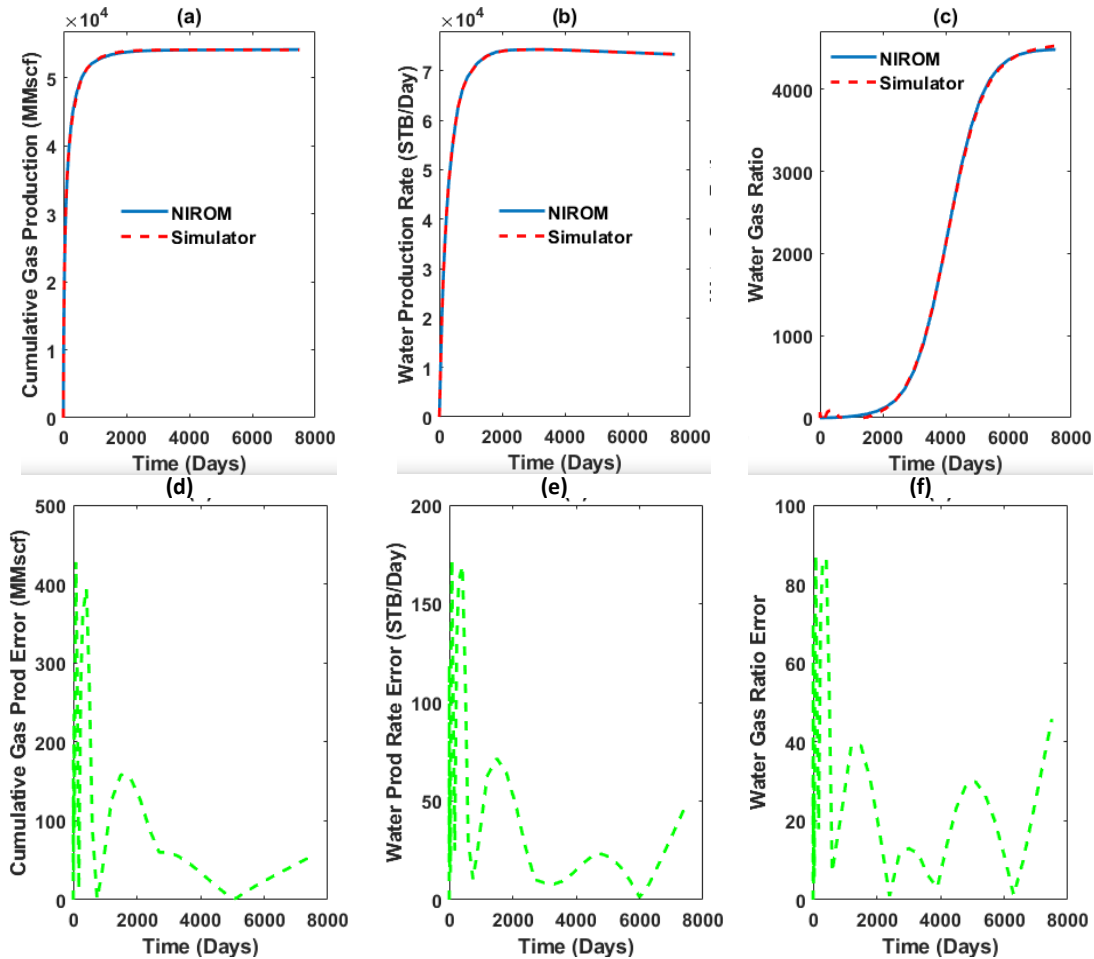


Figure 5.21: Estimated reservoir production profiles using (a & b) 4 POD basis functions and (c) 5 POD basis functions (the energy criterion suggested 3 POD basis functions would be sufficient), for all 7 RBFs. These production data required the most number of POD basis functions to obtain reasonable NIROM estimates. Two of these three production data are water-related.

In the Fetkovich aquifer, pressure decline is similar to that of the numerical aquifer, the reservoir pressure is maintained at just above the well bottom hole pressure limit and production is sustained throughout the 7500days. All RBFs gave good estimates of the average reservoir pressure and the individual grid cell pressure in the reservoir with negligible error (Figure 5.22 and Figure 5.23) while RBFs 1 – 6 gave the best results for the saturation distribution (Figure 5.24). 1 and 3 POD basis functions were used to obtain the results in Figure 5.22 and Figure 5.24 respectively. Similar to the numerical aquifer case, estimating the water gas ratio required the most number of basis functions, 7 basis functions were required instead of the 3 suggested from the energy criterion. For the production data estimation, only the results for NIROM's estimation of water gas ratio with 7 POD basis functions is shown in Figure 5.25. This result is the same for RBFs 1 -5 and 7 while RBF 6 gave a different result to the other RBFs.

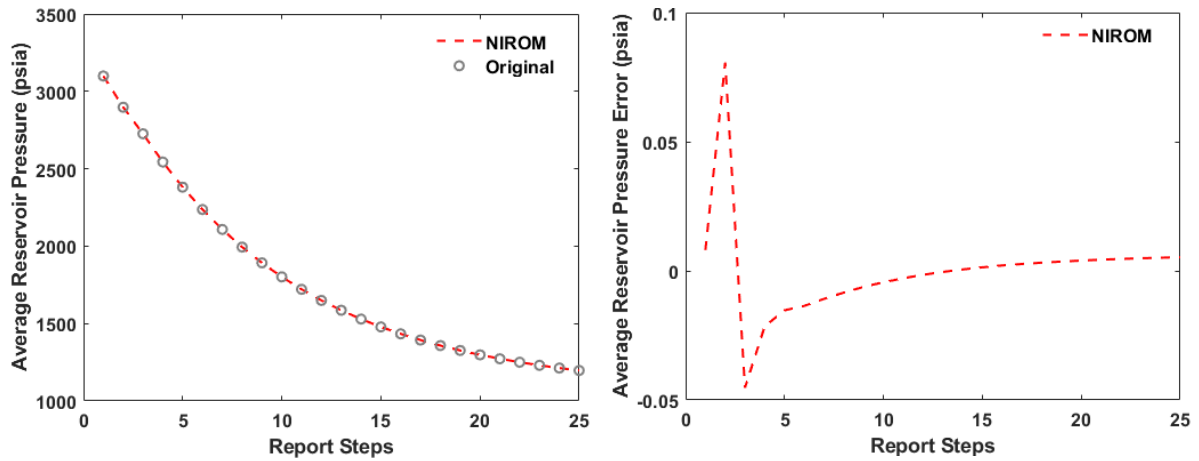


Figure 5.22: Average reservoir pressure distribution in the synthetic gas reservoir with a Fetkovich aquifer using 1 POD basis function (99.9% energy criterion). All 7 RBFs gave similar satisfactory estimates with negligible error.

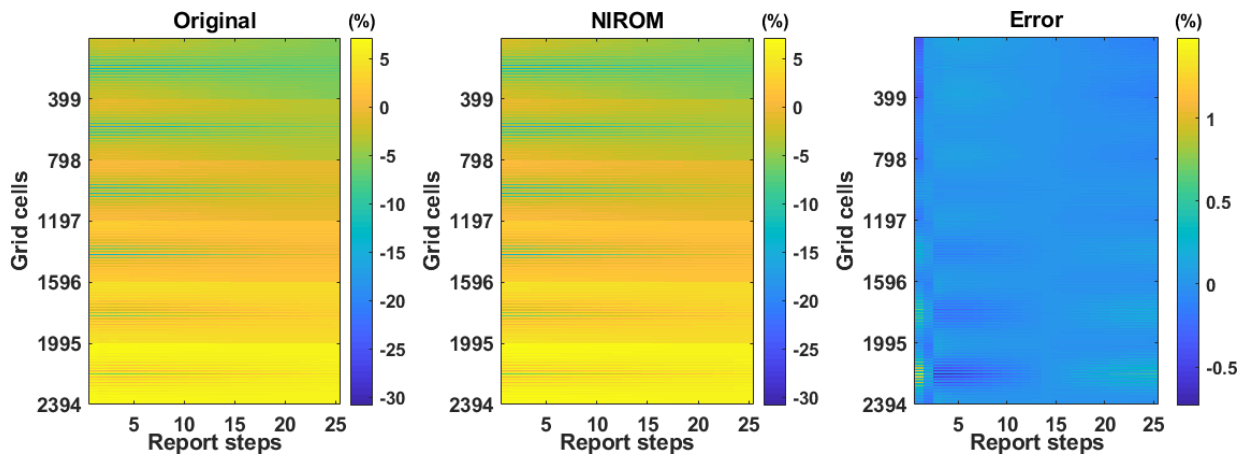


Figure 5.23: Pressure distribution in all grid cells of the synthetic gas reservoir with a Fetkovich aquifer. Reservoir pressure values were normalised to zero mean and the NIROM was implemented with 1 POD basis function (as suggested by the energy criterion).

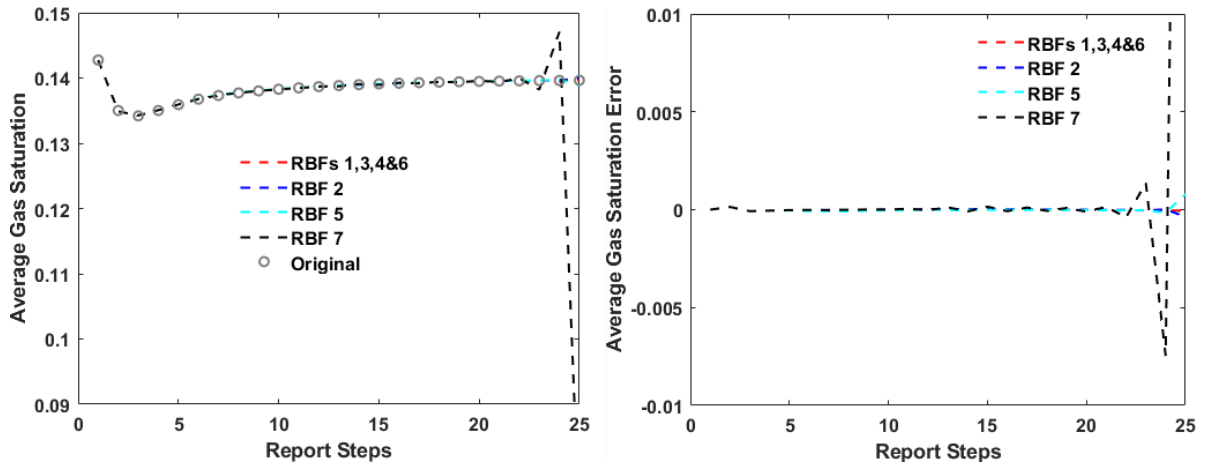


Figure 5.24: Average gas saturation distribution in the synthetic gas reservoir with a Fetkovich aquifer using 3 POD basis functions (99.9% energy criterion). RBFs 1 - 6 give better results than RBF 7 especially after 22 report steps, where RBF 7 gives haphazard results.

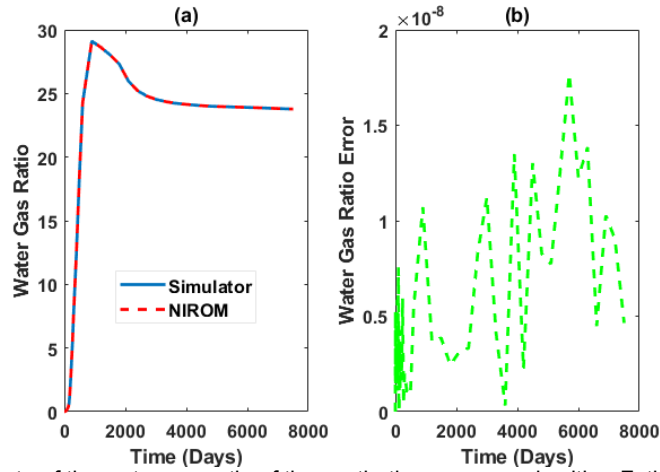


Figure 5.25: NIROM estimate of the water gas ratio of the synthetic gas reservoir with a Fetkovich aquifer using 7 POD basis functions (energy criterion suggested that 3 POD basis functions were required).

5.1.3 GRID REFINEMENT

In these cases, the synthetic gas reservoir is produced by gas rate control for 25 report steps in 300 days intervals. The results shown below are from the linear RBF but similar results were obtained from the other 6 RBFs from section 3.4.

5.1.3.1 LOCAL GRID REFINEMENT (LGR)

Here, local grid refinement has been introduced into the grid cells where the well is located, ($10 \times 11 \times 1:6$). These grid cells were refined to have a local dimension of $5 \times 5 \times 12$. Figure 5.26a shows the pressure distribution in the reservoir's global grid cells, while the pressure distribution in the local grids is shown in Figure 5.26b. In Figure 5.26c as in Figure 5.26a & b, NIROM adequately mimics ECLIPSE in estimating the pressure depletion in both the global and local grid cells in the first layer of the synthetic gas reservoir. Similar results were obtained from all seven RBFs that were considered.

Figure 5.27 shows the pressure distribution in the global and local grid cells of the reservoir in all 6 layers.

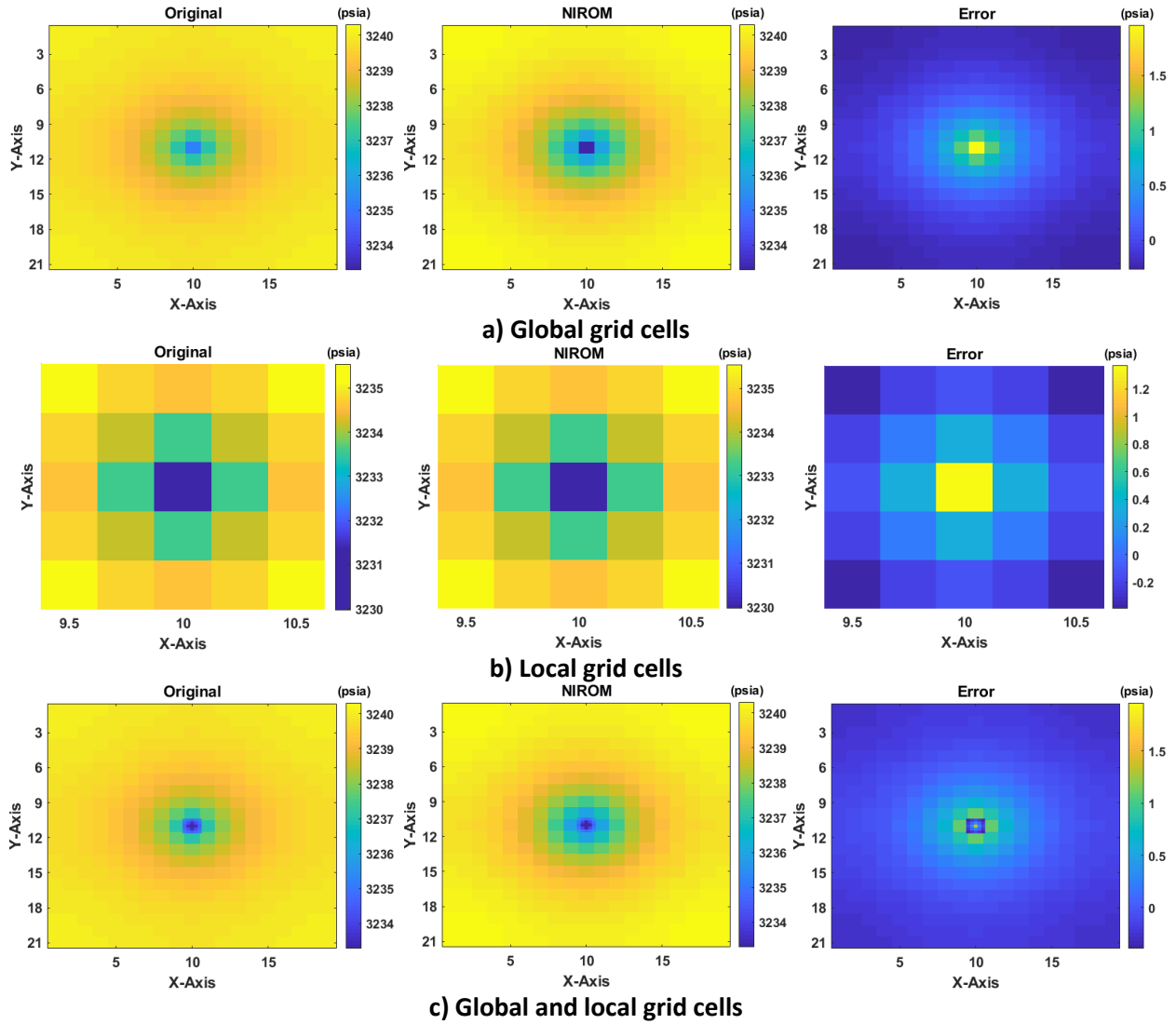


Figure 5.26: Pressure distribution in the first layer of the synthetic gas reservoir at 1800 days, using 1 POD basis function. Pressure distribution in a) the global grid cells only, b) the local grid cells only and c) the global as well as local grid cells. The local grid cells emphasise the pressure distribution in the well's location.

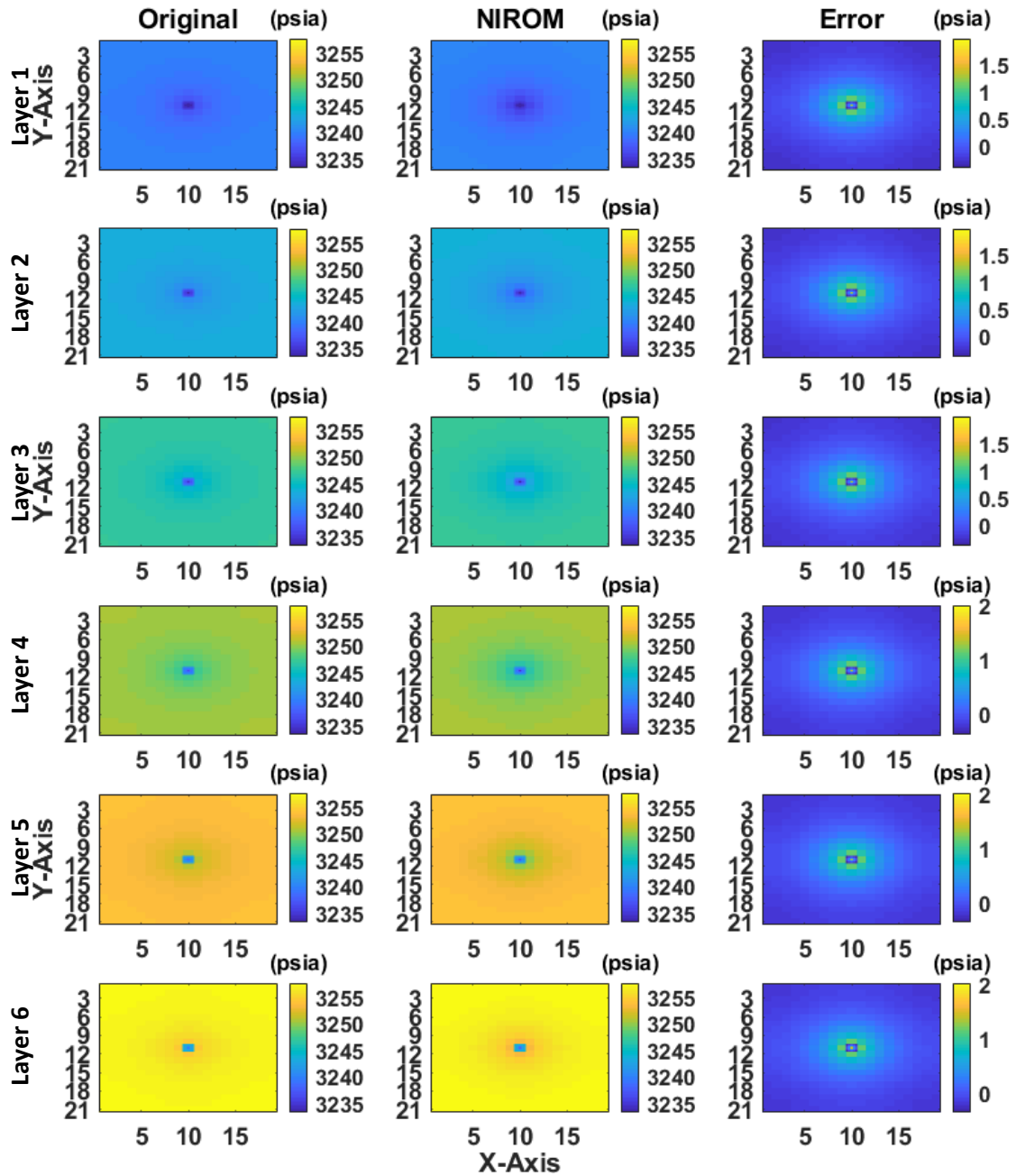


Figure 5.27: Pressure distribution in the global and local grid cells of the synthetic gas reservoir at 1800 days, using 1 POD basis function and across all six layers. NIROM satisfactorily mimics ECLIPSE's solutions for local grid refinement.

5.1.3.2 TARTAN GRID

Figure 5.28 shows that NIROM also does well in estimating the synthetic reservoir's pressure distribution in tartan grid cases. In the six layers of the reservoir, NIROM satisfactorily estimates the reservoir's pressure with errors in the vicinity of zero psi. NIROM also satisfactorily captures the pattern of pressure depletion from the well location outwards towards the boundaries of the homogeneous reservoir.

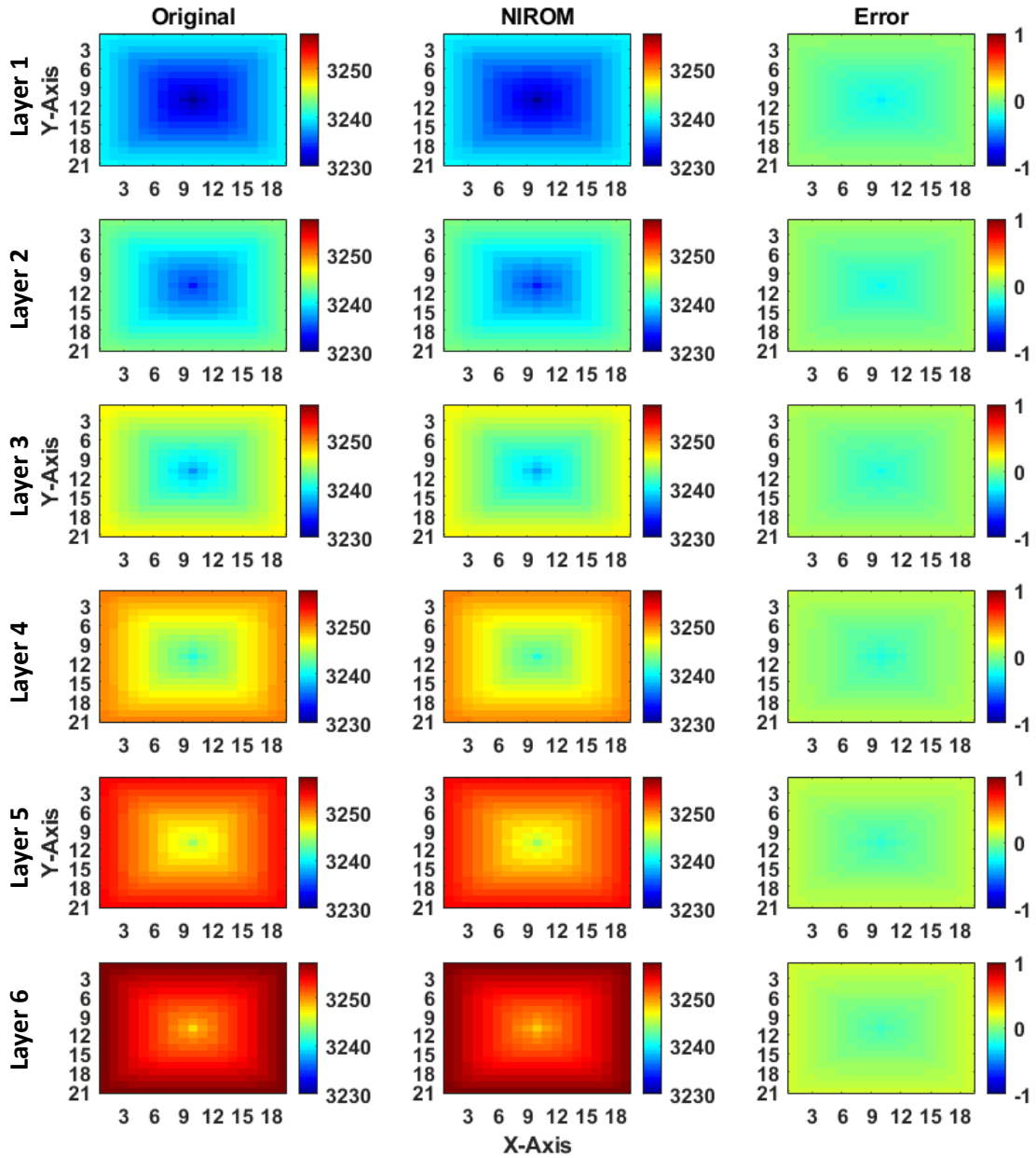


Figure 5.28: Pressure distribution across the six layers in the synthetic gas reservoir with the tartan grid at 1800 days, using 1 POD basis function. NIROM satisfactorily mimics ECLIPSE's solutions from the tartan grid.

5.2 MODIFIED NORNE FIELD

Here, the capability of the NIROM has been demonstrated through the comparison of its simulation with that of ECLIPSE for the modified Norne field model ^[129]. This realistic field case comprises 44927 active grid cells, so the NIROM pressure and saturation predictions in this example are for the 44927 active grid for 7500 days production lifecycle. The reservoir has 3 producers, all produced by BHP limit control, with aquifer pressure support. Figure 5.29 shows the average of the ECLIPSE simulated and NIROM estimated pressure distributions at different times, together with the errors between the NIROM estimates and ECLIPSE results. Figure 5.30 shows the average of the gas saturation distribution estimated by the NIROM and ECLIPSE, together with a graph of the NIROM's errors. In both Figure

5.29 and Figure 5.30, RBFs 3 and 6 do not give satisfactory results when compared to ECLIPSE solutions. Also, 6 and 7 POD basis functions were required to obtain the results in Figure 5.29 and Figure 5.30. Figure 5.31 and Figure 5.32 are the pressure and saturation distributions in all the grids of the heterogeneous gas field throughout its entire production lifecycle (using the linear RBF, RBF 1).

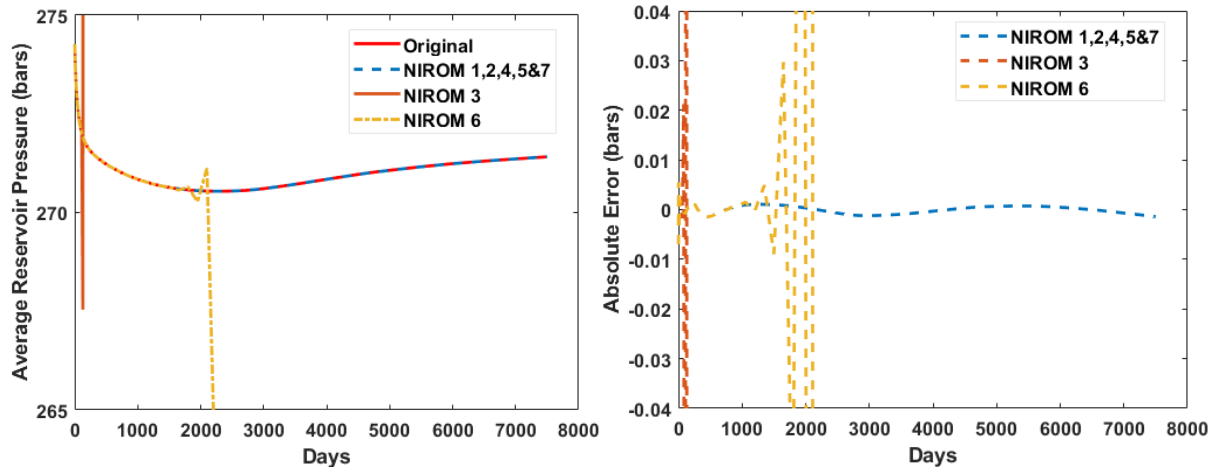


Figure 5.29: The average of the estimated pressure distribution at different report times in the Norne gas field as predicted by the POD-RBF NIROM and ECLIPSE simulation. The NIROM used 6 POD basis functions (as indicated by the energy criterion). RBFs 1, 2, 4, 5 and 7 mimic ECLIPSE with negligible error.

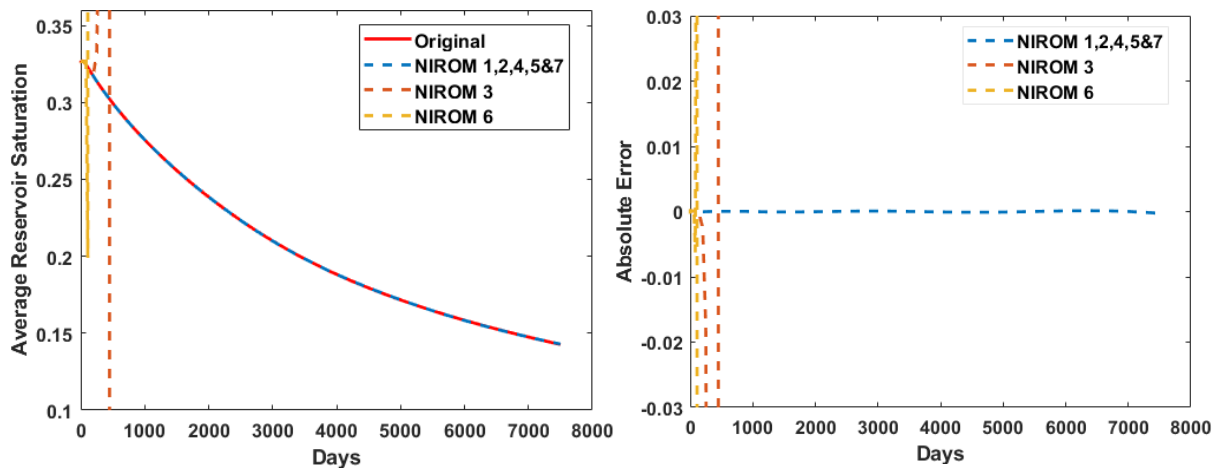


Figure 5.30: The average of the gas saturation distribution at different report times obtained by NIROM using 7 POD basis functions (as suggested by the energy criterion). Compared with ECLIPSE's results, the NIROM satisfactorily matches the gas saturation depletion pattern in the heterogeneous modified Norne field for RBFs 1, 2, 4, 5 and 7. RBFs 3 and 6 deviated from ECLIPSE's results after less than 1000 days.

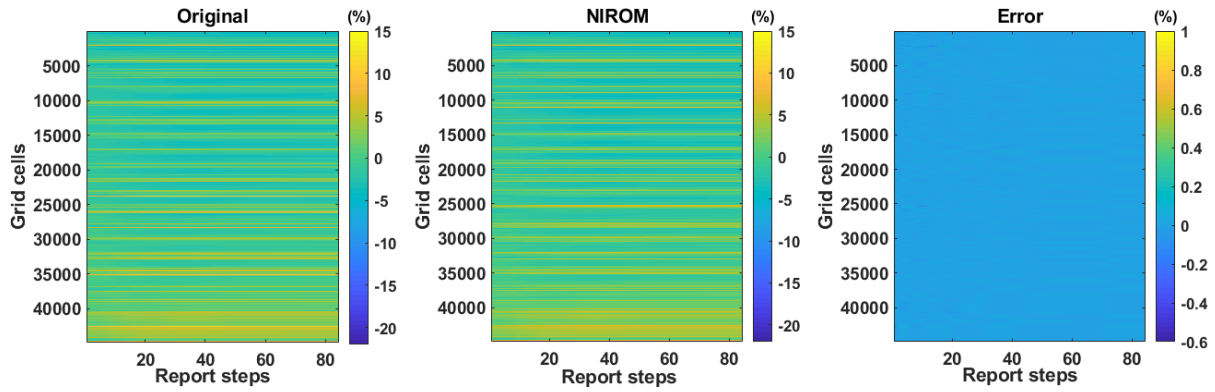


Figure 5.31: Pressure distribution in all grid cells of the modified Norne gas field with pressure values normalised to zero mean. These results were obtained using 6 POD basis functions (the energy criterion) and the linear RBF (RBF 1). NIROM estimations have negligible error.

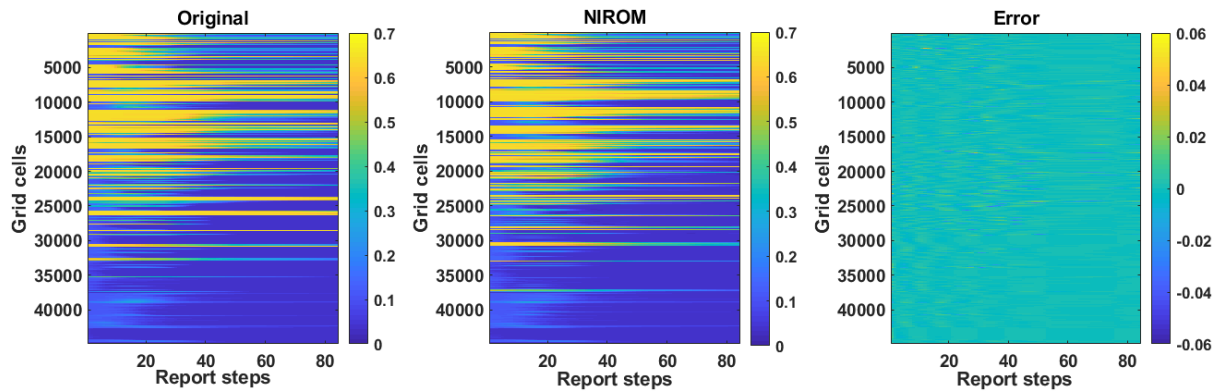


Figure 5.32: Gas saturation distribution in all grid cells of the modified Norne gas field using 7 POD basis functions (the energy criterion) and the linear RBF (RBF 1). The error in NIROM estimates was mostly negligible.

Figure 5.33 to Figure 5.37 show the various production data estimated by the NIROM, with RBFs 1, 3, 5, and 7, together with the ECLIPSE simulation results as well as the associated error plots. The NIROM needs the most POD basis functions to estimate the production water-gas ratio (14 out of a possible 17), however, the other production data (e.g. the total gas production, cumulative field production rate and well production rates from wells 1 and 2) could be estimated to within 1.85e-5%, 1.06%, 7.87% and 0.34% absolute errors respectively using only 4 POD basis functions (Figure 5.35 and Figure 5.36). Adding one more POD basis function to this number decreases these error values to less than $1.3 \times 10^{-3}\%$ (Figure 5.37). RBFs 1, 3, 5, and 7 gave similar results while RBFs 2, 4 and 6 over or underestimate the results and have not been included here.

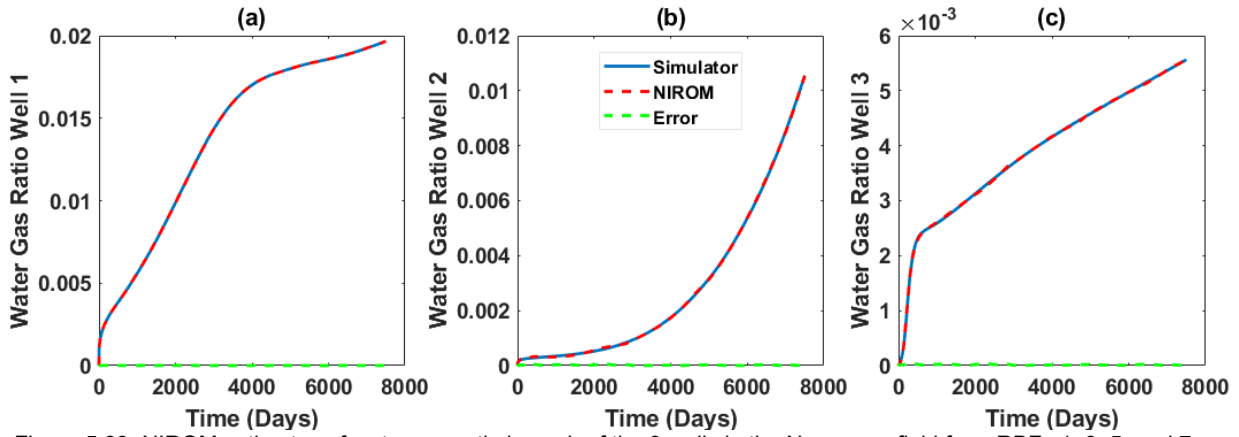


Figure 5.33: NIROM estimates of water gas ratio in each of the 3 wells in the Norne gas field from RBFs 1, 3, 5, and 7. These were obtained using 14 POD basis functions. These water-related variables required the most number of POD basis functions for satisfactory NIROM results.

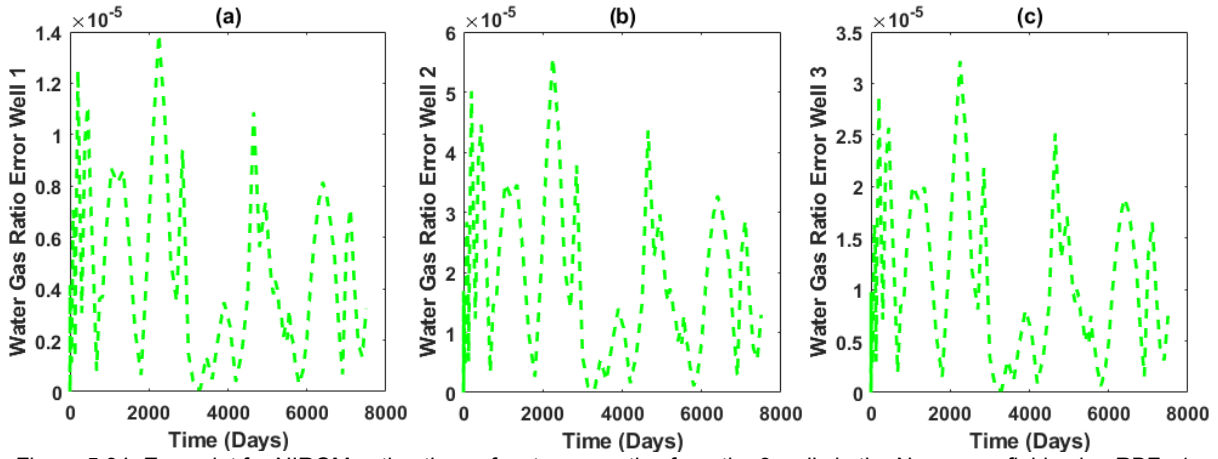


Figure 5.34: Error plot for NIROM estimations of water gas ratios from the 3 wells in the Norne gas field using RBFs 1, 3, 5, and 7. Fourteen (14) POD basis functions were used to obtain these results. Errors are negligible when using a large number of basis functions.

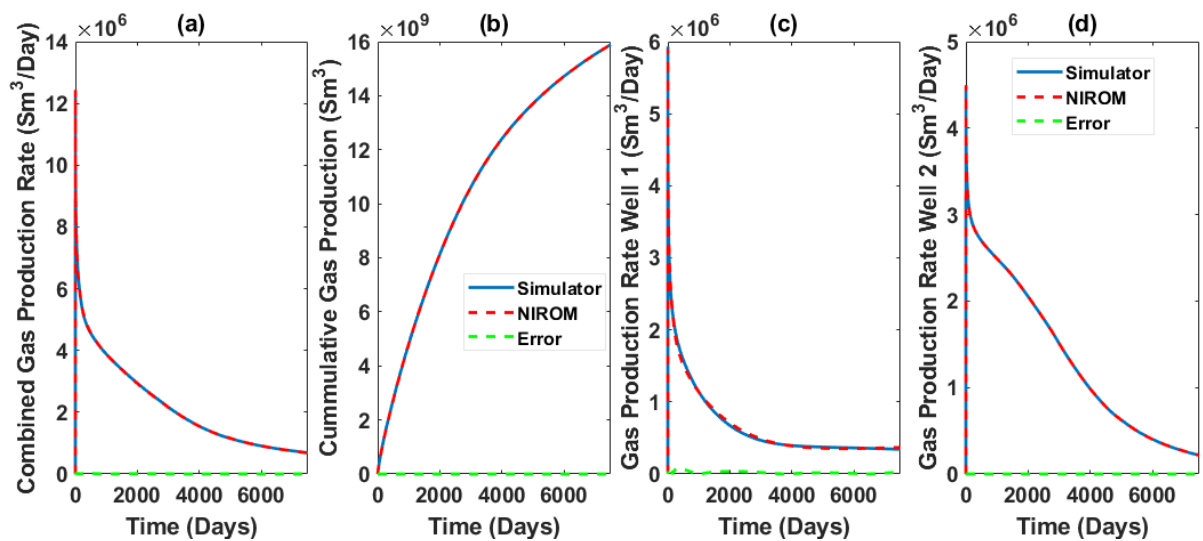


Figure 5.35: Total field gas production rate (far left), cumulative field gas production and gas production rates from wells 1 and 2 in the Norne gas field model as predicted by ECLIPSE and the NIROM. The NIROM used 4 POD basis functions (99.9% energy criterion). RBFs 1, 3, 5, & 7 gave identical satisfactory results.

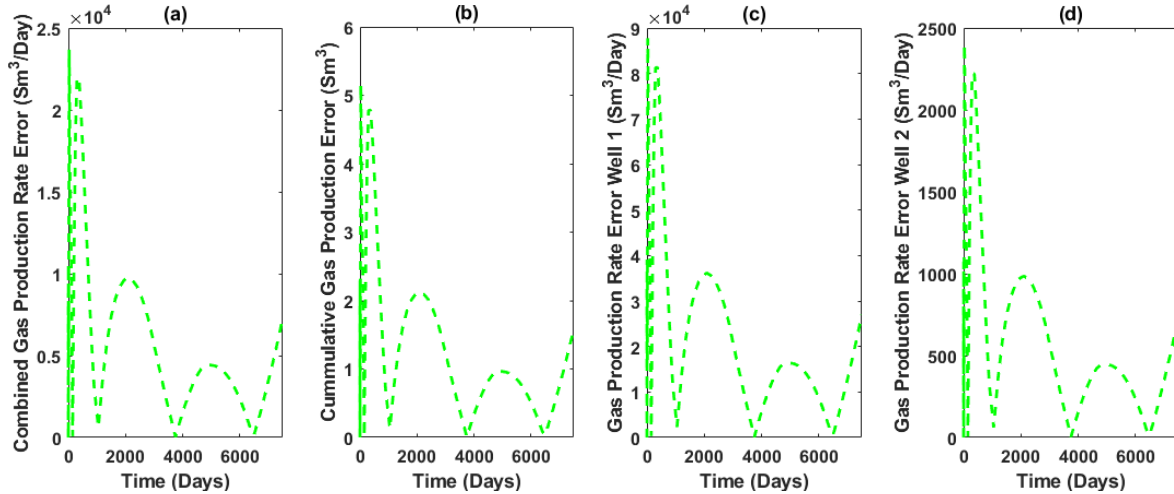


Figure 5.36: Error plots of NIROM's estimation of the production data of the Norne gas field in Figure 5.35, predicted with 4 POD basis functions (99.9% energy criterion). Errors are mostly less than 1% for RBFs 1, 3, 5, & 7.

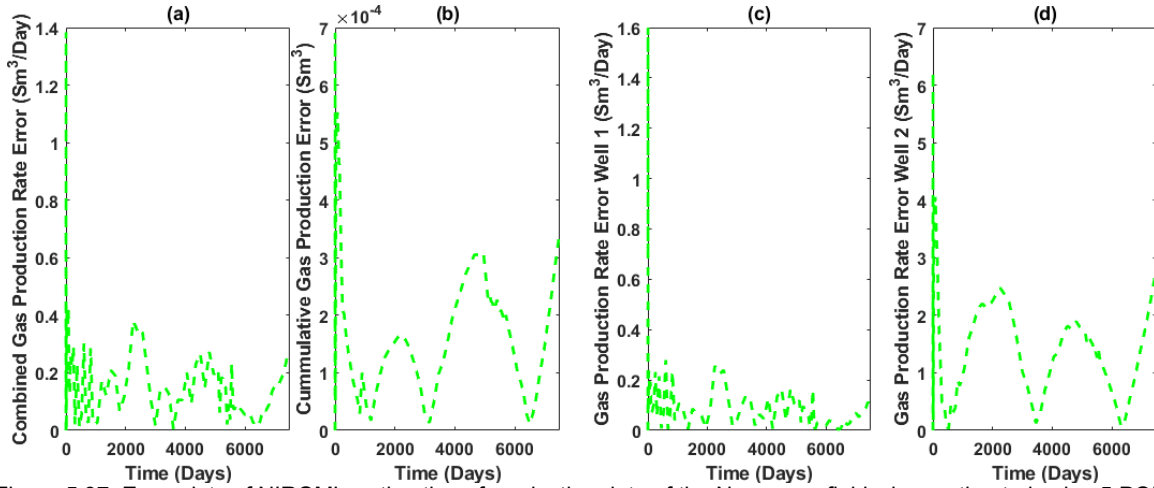


Figure 5.37: Error plots of NIROM's estimation of production data of the Norne gas field when estimated using 5 POD basis functions rather than 4 (99.9% energy criterion). Errors fall to less than $1.41 \times 10^{-3}\%$ when an additional POD basis function is included in the NIROM.

5.3 INTERPOLATION FUNCTION ANALYSIS AND ERROR MEASURES

In this chapter, seven (7) interpolation functions have been used for the different training model scenarios considered. Table 5.1 shows an analysis of the 7 RBFs, which was generated by assigning weights to the RBFs based on their efficiency in each case of estimating the gas reservoir simulations from ECLIPSE. The numbers assigned to RBFs are in the same order as they were listed in chapter 3 (and at the start of this chapter). From Table 5.1, the preferred RBFs have been recommended based on their suitability for each simulation case considered. RBFs 1, 2, 4, 5 and 7 work best for pressure distribution predictions, RBFs 1, 2, 4 and 5 give better estimates for saturation distribution predictions, and for reservoir production variables all the RBFs equally give satisfactory results except for RBFs 2, 4 and 6 which have episodes of instability. Therefore, different RBFs are best suited for different aspects of gas reservoir simulations with NIROM.

On a general note, the recommended RBFs are RBF 1 (linear), RBF 3 (modified thin-plate spline), RBF 5 (Gaussian) and 7 (multi-quadratic), with the linear function being the most consistent and stable RBF for gas reservoir simulation. These recommendations are further confirmed in Table 5.2 and Figure 5.38 where the RMSEs from NIROM estimation of the different production scenarios are shown. From Table 5.2, cases with outlier estimations (overestimation or when NIROM was unable to provide reasonable estimations) are written in red colour. Only RBFs 1 and 2 did not have any red values in Table 5.2, but for Norne field production data estimation, RBF 2 was unstable, so this further confirms that RBF 1 (linear function) is the most preferred RBF for gas reservoir modelling with NIROM.

Table 5.1: Summary of result on best-fit interpolation function(s) for various simulation cases

Case	Interpolation Functions (numbers are consistent with chapter 3)		
	Synthetic model	Norne Model	Recommended RBF
Pressure	1, 2, 3, 4, 5, 6, 7	1, 2, 4, 5, 7	1, 2, 4, 5, 7
Gas Saturation	1, 2, 3, 4, 5, 6	1, 2, 4, 5, 7	1, 2, 4, 5
Production Data	1, 2, 3, 4, 5, 7	1, 3, 5, 7	1, 3, 5, 7
Recommended RBF	1, 2, 3, 4, 5	1, 5, 7	1, 5, 7

Table 5.2: Average RMSE for pressure and saturation distribution estimations with NIROM for all 7 RBFs considered in this work. RBF 1 is the most consistently stable RBF with the least RMSE. The values in red colour are the outliers where NIROM overestimated its simulation or it could not give reasonable estimations.

RBFs		1	2	3	4	5	6	7
Synthetic Gas Reservoir								
Gas Rate Control	Pressure (psi)	0.17	0.17	0.17	0.17	0.17	0.17	0.17
	Production Data	8.7×10^{-11}	3.2×10^{-12}	9.8×10^{-9}	2.6×10^{-12}	2.1×10^{-12}	6.5×10^{-8}	4.9×10^{-10}
BHP Control	Pressure	0.005	0.005	0.005	0.005	0.005	0.005	0.006
Numerical Aquifer	Pressure	2.14	2.14	2.14	2.14	2.14	2.14	2.14
	Gas Saturation	0.0002	0.06	1.25×10^7	0.02	1.37	0.0002	7.40×10^5
Carter Tracy Aquifer	Pressure	0.66	0.66	0.66	0.66	0.66	0.66	0.66
	Gas Saturation	0.0001	0.36	0.0001	6.41	3.37	0.0001	1.81×10^{44}
Fetkovich Aquifer	Pressure	2.16	2.16	2.16	2.16	2.16	2.16	2.16
	Gas Saturation	0.0001	0.0001	0.0001	0.0001	0.0003	0.0001	0.0218
Grid Refinement	Local grid (pressure)	0.84	0.84	0.84	0.84	0.84	0.84	0.84
	Global grid (pressure)	0.58	0.58	0.58	0.58	0.58	0.58	0.58
	Tartan grid (pressure)	0.57	0.57	0.57	0.57	0.57	0.57	0.57
Modified Norne Reservoir with Numerical Aquifer								
Pressure (bars)		0.03	0.03	-	0.03	0.03	-	0.03
Gas Saturation		0.003	0.003	-	0.003	0.003	-	0.003

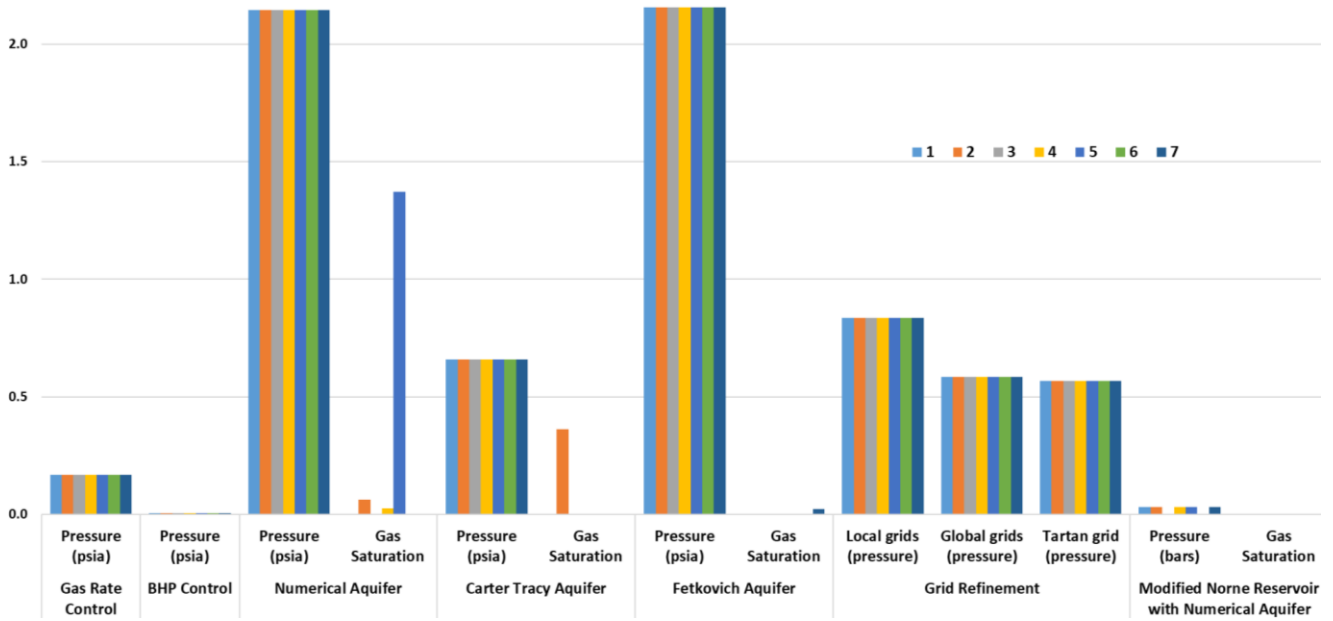


Figure 5.38: Average RMSE for pressure and saturation distribution estimations with NIROM for all 7 RBFs considered in this chapter. NIROM simulations with outliers have not been added to this figure.

5.4 TIME ANALYSIS

Here, the CPU time taken to create and then run the NIROM simulations is compared with the time taken to run the same simulations in ECLIPSE. In this section, the unit of time is seconds as summarised in Table 5.3. This work was executed on a workstation with 40 logical processors, 2.4 GHz base speed and 128GB RAM. The production duration was 7500 days for all the cases except the synthetic gas reservoir case that was produced at constant BHP and no aquifer (it produced for 3300 days). The NIROM gave speedups of between 12 and 60 times ECLIPSE runtime (including the time taken for pre-processing the training data and POD reduction). The actual NIROM simulation times (discounting pre-processing and POD reduction) were between an order of magnitude (for the Norne field case) and 3 orders of magnitude times faster than ECLIPSE. These timings for building and running the NIROM were obtained using an interpreted language (MATLAB) whereas ECLIPSE has been written in a compiled language and carefully optimised over many years of development. So the NIROM should be at least two orders of magnitude faster than ECLIPSE if it was also implemented in a compiled language (as previously found by Kostorz et al. (2019) ^[37]).

Table 5.3: Simulation runtime (in seconds) and average speed ups from NIROM versus ECLIPSE for different gas production cases. Comparison of the cumulative computational time shows up to 60 times simulation speed up with NIROM.

Model	Process	Simulation Cases							
		Synthetic Gas Reservoir							Modified Norne Gas Field
		Gas Rate Control	BHP Control	Numerical Aquifer	Carter Tracy Aquifer	Fetkovich Aquifer	LGR	Tartan grid	
NIROM Runtime	Pre-Processing Data	0.12	0.06	0.14	0.15	0.19	0.30	0.17	10.4
	POD Reduction	0.01	0.005	0.01	0.01	0.01	0.23	0.01	0.19
	Simulation	0.02	0.02	0.01	0.02	0.02	0.002	0.02	0.33
	Total	0.15	0.08	0.16	0.18	0.22	0.54	0.19	10.9
ECLIPSE Runtime		5.34	5.06	8.04	6.98	7.51	18.4	6.28	134.5
Average speed up		36.4	60.1	51.4	39.7	33.7	34.3	33.1	12.3

SUMMARY

In this chapter, the POD-RBF NIROM has been used to provide rapid estimations of both pressure and saturation distributions in 2 different gas reservoir models undergoing depletion, together with their associated production data. The homogeneous synthetic gas reservoir and the heterogeneous modified Norne gas reservoir models have been examined with and without aquifer influx, for different production controls and grid refinement techniques. The accuracy of the NIROM has been compared using different RBFs to estimate the reservoir dynamics in hyperspace, while also assessing the usefulness of the energy criterion advocated by Xiao et al. (2016) ^[38] for estimating the number of POD basis functions required to describe that hyperspace. The NIROM estimates in this chapter were for the same ECLIPSE simulation runs that were used to train the NIROM. However the goal of this work is to test NIROM's performance for new (prediction) cases. Applications of NIROM to gas field modelling of prediction cases are considered in the next three chapters.

From the results that have been discussed in this chapter, it is shown that POD-RBF NIROM can provide good estimates of gas reservoir behaviour over time, both in terms of production data and in terms of estimating pressure and saturation distributions in reservoir models. NIROM's satisfactory performance in modelling the different production scenarios that were considered in this chapter (such as cases with and without aquifer pressure support as well as with different production controls and grid refinements) reflect its usefulness to engineers in practical gas field development planning. For subsequent POD-RBF implementations, the linear radial basis function is the recommended RBF, since it produced reasonable estimates of reservoir behaviour in all the cases that were tested. All the other RBFs resulted in poorer estimates in one or more cases. The 99.9% energy criterion provides a reasonable estimate of the number of POD basis functions required to properly describe the solution of gas reservoir dynamics in hyperspace in many, but not all cases. More POD basis functions are likely

to be required when estimating production data, especially when there is water influx into the reservoir resulting in water being produced.

Furthermore, the NIROM was implemented using an interpreted language (MATLAB), but it is between 12 and 60 times faster than ECLIPSE. The NIROM is likely to provide even greater speed-ups if implemented in a compiled language and especially when the POD reduction is discounted, as would be the case when performing predictions of reservoir behaviour after training the NIROM.

The cases considered in this chapter are summarised in Table 5.4. The linear RBF is used for the rest of the analysis in this work, for predicting new unseen cases that the NIROM is not trained on.

Table 5.4: A summary of the cases that were considered in this chapter and the key findings.

Cases Considered	Summary of Findings
- Replicating specific examples - Evaluated 2 gas reservoirs, including a real field: - homogeneous synthetic gas reservoir - modified heterogeneous Norne gas field - Parameters evaluated: - Production rate - BHP limit - Aquifer models - One Numerical - Two Analytical - Grid refinement - Local grid refinement - Tartan grid	Key Findings
	- Sensitivity analysis on the most suitable radial basis function(s) for NIROM's interpolation – linear RBF was most consistently satisfactory - Sensitivity analysis on the number of POD basis functions that are required to implement NIROM – 99.9% energy criterion may suffice for grid distributed property estimations but not for production variables' estimation
	General Findings
	- Developed NIROM for gas field modelling with ECLIPSE simulation data - Showed that the NIROM is capable of modelling different production dynamics such as sudden changes in pressure, presence of aquifer and different grid refinements - NIROM can satisfactorily reproduce ECLIPSE simulation results with speed-up of at least an order of magnitude.

CHAPTER 6: CONSTANT PARAMETER PREDICTIONS - RESULTS AND DISCUSSION

In this chapter, the POD-RBF NIROM is used to predict the production data and 3D distribution of reservoir pressure and saturations for unseen inputs, using 4 different gas reservoir models. These reservoir models are a homogeneous synthetic single well model, a realistic heterogeneous gas field with 3 wells, a water coning gas reservoir with one producer and a real gas field, Field X. The properties of these reservoirs as well as their simulation parameters were given in Chapter 4. The application of POD-RBF NIROM in this chapter covers the forecasting of the reservoir simulation results for the above listed gas reservoirs, with and without aquifers or shut-in wells, for constant parameter scenarios. The linear RBF was used to learn the reservoirs' dynamics and performance as well as interpolate the reservoirs' performance in prediction cases that the NIROM was not trained on. NIROM prediction cases are within and outside the range of training parameter values and training duration. This chapter also examines the accuracy and speed of the NIROM predictions by comparing them with the results obtained from Schlumberger's ECLIPSE reservoir simulator.

Except for the synthetic gas model, all the other models have more than 10000 grid cells. So to better illustrate the simulation results for pressure and saturation distribution across the grid cells, the average of the reservoir properties at each report step have been plotted for both the NIROM and ECLIPSE. Error plots show the difference between NIROM predictions and ECLIPSE simulations. 'Original' simulation results from ECLIPSE are plotted as small circles while NIROM predictions ('Prediction Q', where Q is the index for that prediction's parameter from the corresponding Tables in Chapter 4) are in dashed lines with the same colour as their corresponding 'Original' result. The range of training runs for each prediction case has been highlighted within greyed backgrounds. For production data, blue solid lines represent results from ECLIPSE (Original), red dashed lines are NIROM prediction results and green dashed lines are plots of the absolute error between ECLIPSE's results and NIROM results.

Furthermore, all simulation cases have been investigated by training NIROM with 5, 10, 15 and/ or 20 training runs between the ranges of training parameters. All the results shown in the following sections are based on 99.9% energy criterion, unless mentioned otherwise.

6.1 SYNTHETIC GAS MODEL

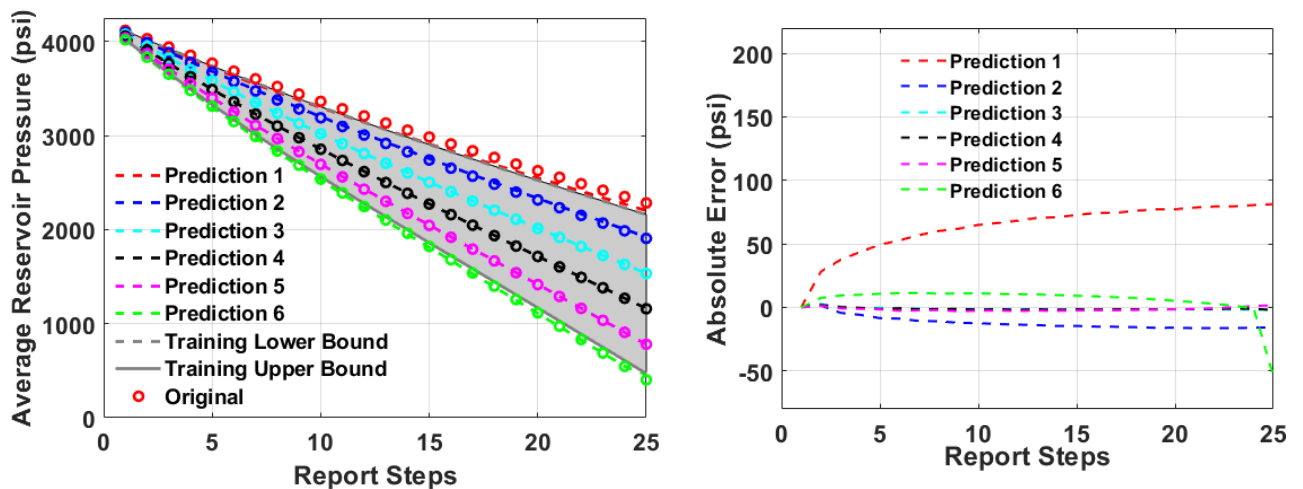
In this section, NIROM has been implemented and tested for four production scenarios (production rate, BHP, porosity and enumeration). Since the synthetic gas reservoir is homogeneous and simple, it presented a suitable example for quickly testing and validating NIROM for the four production

scenarios. Also, the synthetic reservoir's pressure and saturation behaviours are similar, so only the results for pressure distribution are shown here for brevity.

6.1.1 RATE CONTROLLED PRODUCTION

The reservoir pressure experiences a steady decline during constant rate production and the higher the production rates, the steeper the pressure depletion, as expected. The results in Figure 6.1 show NIROM's prediction of the synthetic reservoir's average pressure distribution when the NIROM was trained with 5, 10, 15 and 20 training runs. NIROM predictions of the average reservoir pressure satisfactorily match those of ECLIPSE for predictions within the range of training parameters (predictions 2 to 5), while predictions outside the range of training parameters (predictions 1 and 6) are less accurate. NIROM prediction errors are less than 0.5% within the range of training parameters (predictions 2 to 5). Figure 6.1 also shows that for predictions within the range of training parameters, NIROM's prediction errors are lower as the number of training runs increases, while for predictions 1 and 6, which are outside the range of training runs, NIROM prediction errors increase as the number of training runs increase.

Comparing both predictions outside the range of training runs, NIROM's estimation for prediction 6 are generally better than those for prediction 1, NIROM consistently under-predicted prediction 1. This prediction behaviour for parameters outside the range of training parameters can be attributed to the closeness of a particular prediction case to the range of training parameters. For instance, by visual inspection of Figure 6.1, prediction 6 is closer to the range of training parameters (greyed region) than prediction 1.



a) 5 training runs

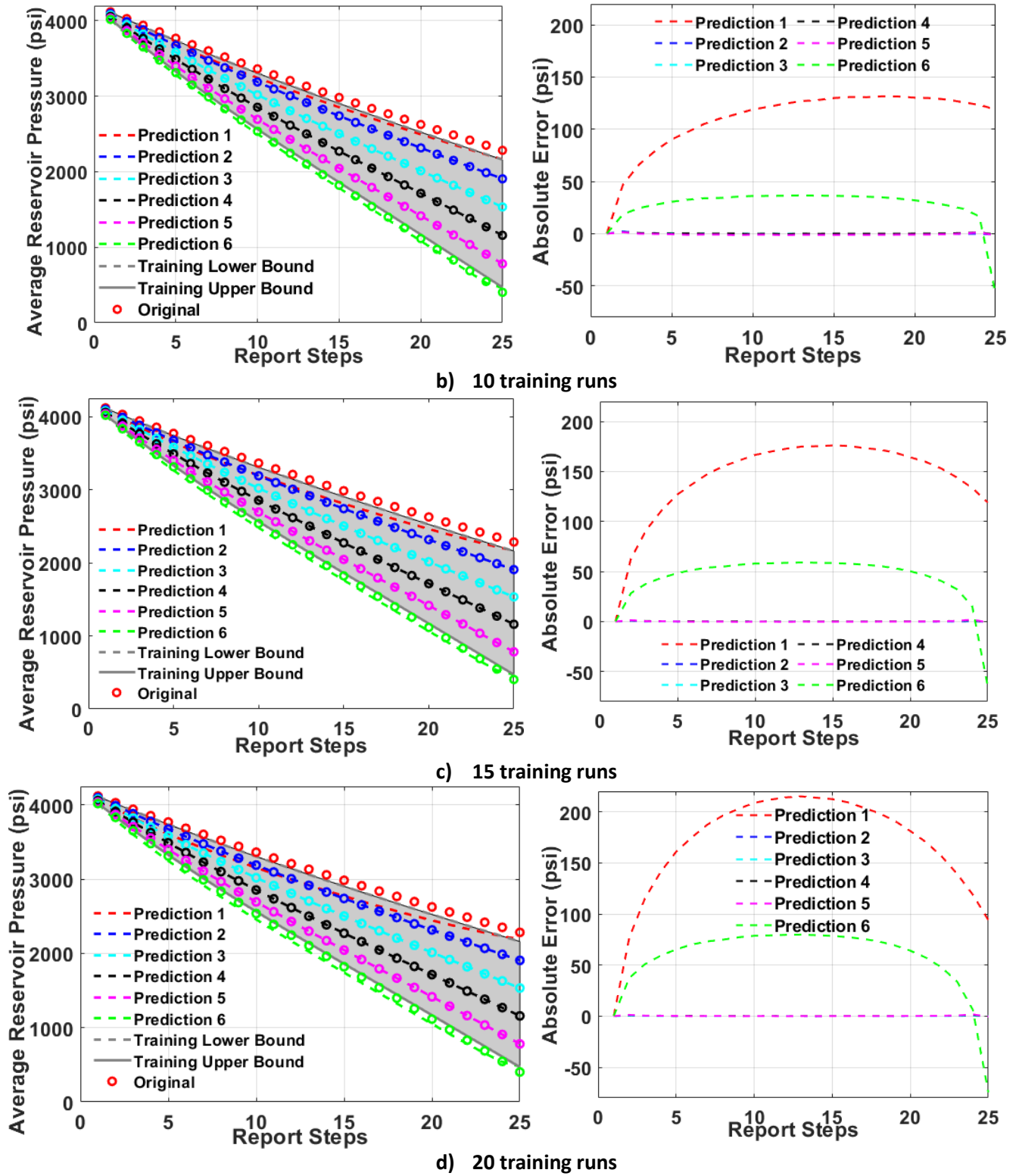


Figure 6.1: NIROM predictions of the average pressure in the synthetic gas reservoir during constant rate production. NIROM predictions 2 to 5 give satisfactory results, while predictions 1 and 6 are worse. Predictions 1 and 6 are outside the range of parameters used in training the NIROM, although NIROM's prediction 6 has less error than prediction 1. These results were obtained using 5, 10, 15 and 20 equally spaced training runs and in most cases, NIROM errors are less than 10psi.

Figure 6.2 shows NIROM predictions of the cumulative gas production and the field pressure from production data. In all the six predictions, the field pressure was more accurately forecasted than the cumulative gas production, but NIROM adequately mimics the patterns of both production variables. NIROM predictions 2 to 5 (within the range of training runs) were better than those outside the range of training runs (predictions 1 and 6). Figure 6.2 shows the results of some of the NIROM predictions when NIROM was trained with 5, 10, 15, and 20 training runs.

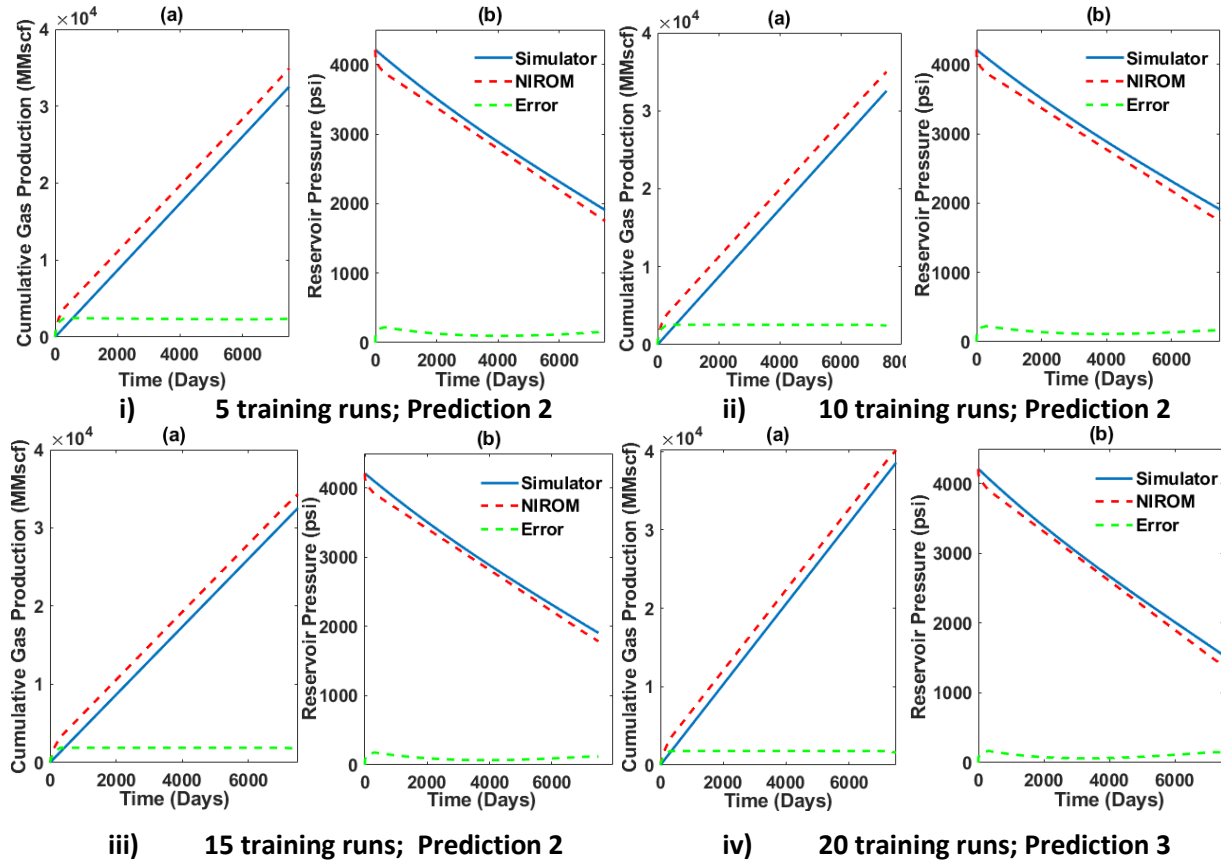


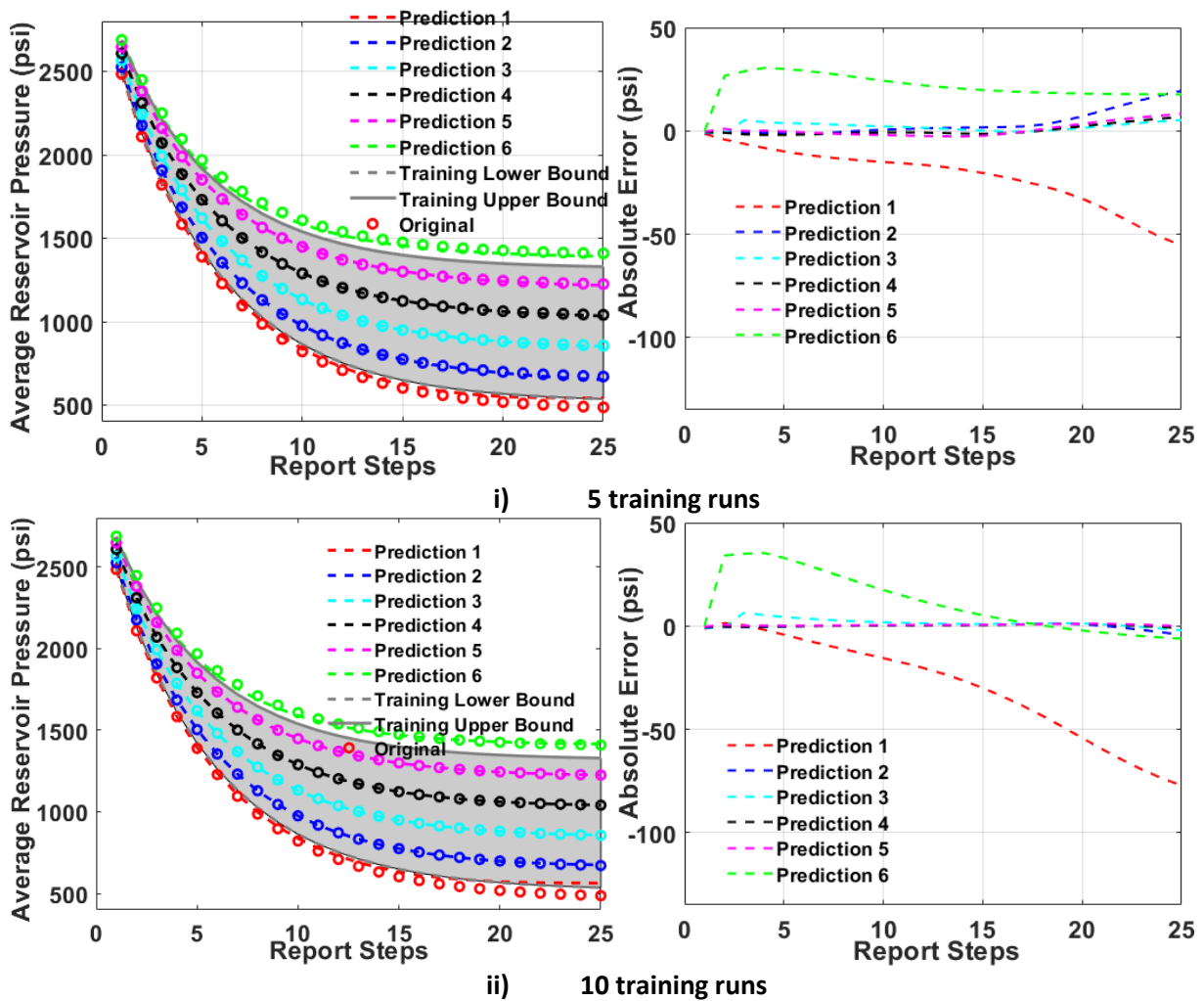
Figure 6.2: NIROM predictions of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir using 1 POD basis function (energy criterion). In these results, the best case was chosen in as an example and NIROM adequately predicts the pattern of pressure depletion and gas production from the reservoir. The NIROM can give better results with an additional POD basis function. Similar results were obtained for predictions within the range of training runs, with less accurate results from predictions 1 and 6. NIROM predictions of the reservoir pressure are more accurate than those of the cumulative gas production.

In the rest of section 6.1, it is shown that NIROM has similar behaviours of 1) predictions within the range of training runs were better than those outside the range of training runs and 2) increasing the number of training runs gave better results for predictions within the range of training runs but worse results for predictions outside the range of training runs. The first point above is in agreement with the work of Kostorz et al (2020)^[99]. In section 6.6 it would be shown that implementing NIROM with more training runs takes longer computational time. This suggests that the choice of the number of

training runs for NIROM implementation involves striking a balance between accuracy and computational speed.

6.1.2 BHP CONTROLLED PRODUCTION

In this section, the NIROM has been trained with BHP values between 450 and 1250psi, with a weak numerical aquifer for pressure support. Here also the best NIROM predictions are for predictions within the range of training parameters, and prediction 6 is better than prediction 1 (Figure 6.3). Nevertheless, there is less than 6% average absolute error from NIROM predictions of the reservoir's pressure distribution, both within and outside the range of training runs. Within the range of training runs, NIROM predictions were estimated with less than 0.4% (Figure 6.3) error. Figure 6.4 shows that NIROM performance was similar for predictions of the production data (cumulative gas production and field pressure). The results in Figure 6.3 and Figure 6.4 were obtained from training NIROM with 5, 10, 15 and 20 training runs.



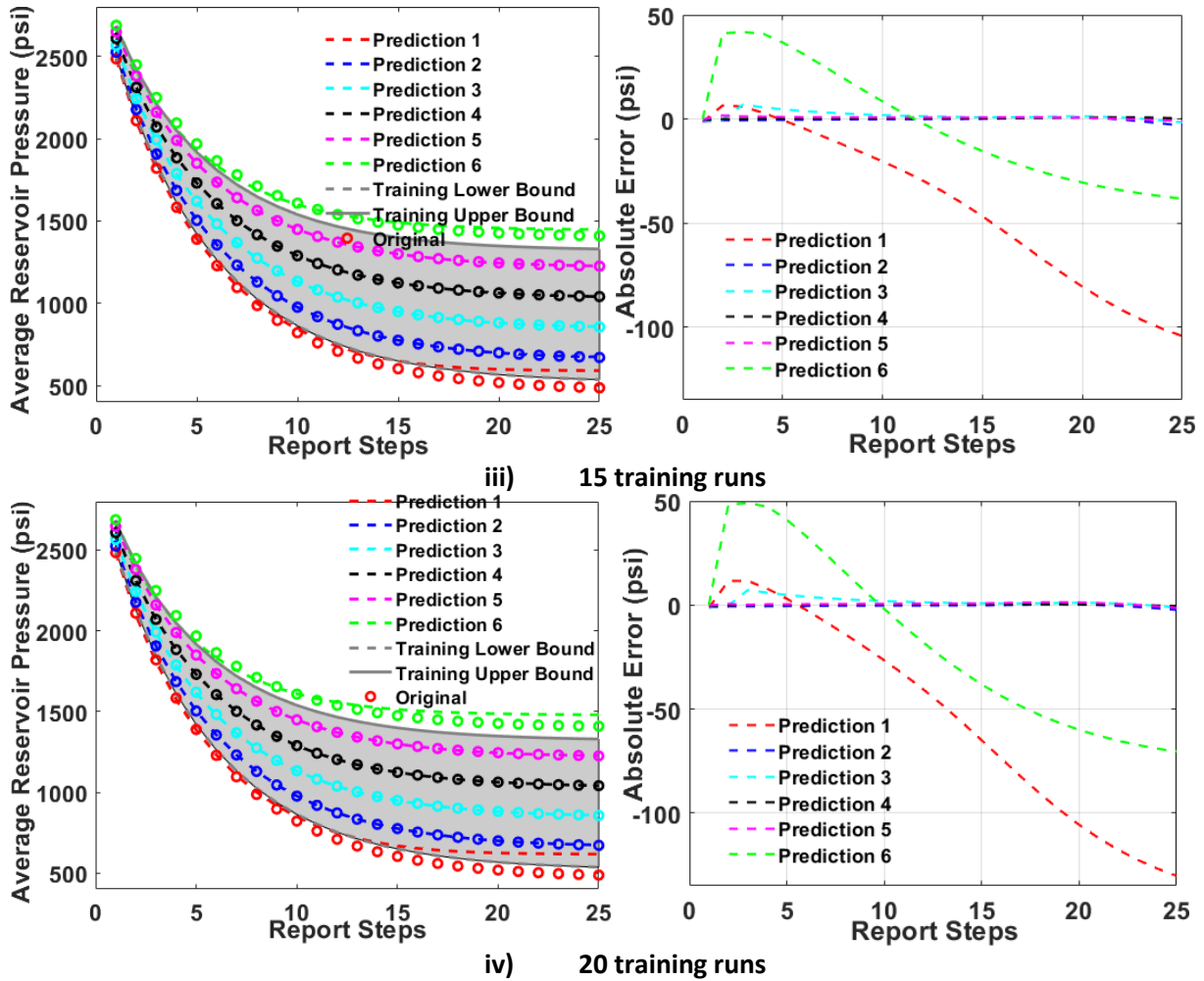


Figure 6.3: NIROM predictions of the average reservoir pressure in the synthetic gas reservoir for BHP controlled production. All 6 predictions were reproduced by the NIROM with absolute errors of less than 6%, even for prediction cases outside the range of training parameters.

In chapter 5 it was shown that producing the synthetic gas reservoir at 1000psi BHP limit meant that most of its GIIP was produced in the first 5 report steps (1500days). It was also shown that with aquifer pressure support, gas production could continue until the end of 20 years but at less than 10% of the starting production rate. This behaviour is also seen in Figure 6.4b where the initial production rate of the synthetic gas reservoir was more than 1000MMscf but the rate reduced to less than 10MMscf before 1000 days, with a gradual decline to less than 100MMscf at the end of 20 years. As was seen in chapter 5, Figure 6.4 also shows that NIROM trained with the optimal number of training runs (99.9% energy criterion) could not accurately predict the reservoir's production data beyond the first few report steps (less than 500days). The change in both variables (cumulative production and production rate), i.e. their "simulator" values, were negligible after the initial 500 days, and the POD-RBF NIROM was unable to accurately predict the low/ "negligible" values. This was resolved in chapter 5 by adding extra POD basis functions to the NIROM.

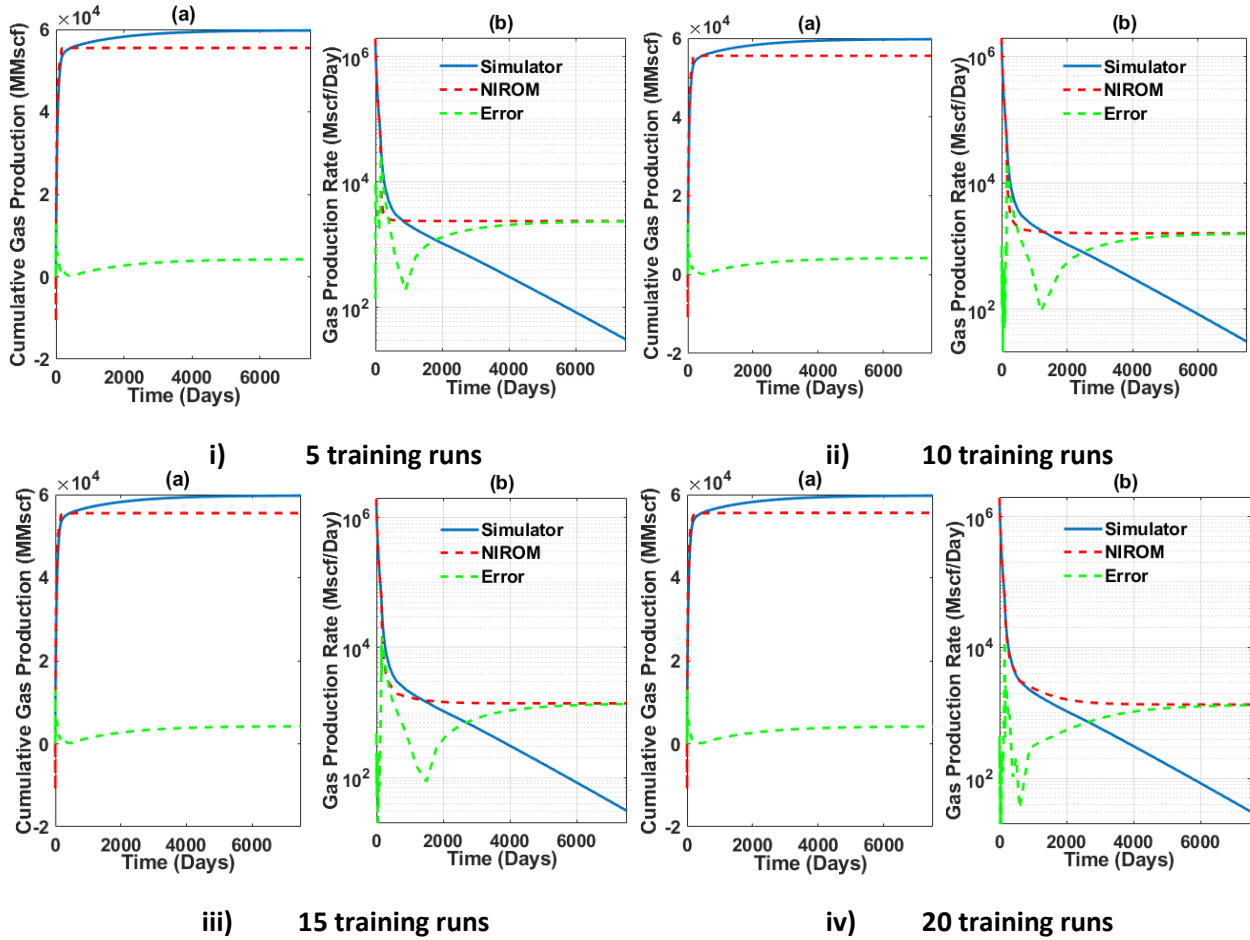


Figure 6.4: NIROM's prediction 3 of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir produced by BHP limit control. NIROM adequately predicts the pattern of pressure depletion and gas production from the reservoir and can give better results with an additional POD basis function. Similar results were obtained for predictions 1, 2, 4 to 6, although predictions 1 and 6 were the least accurate.

6.1.3 UNCERTAIN CONSTANT POROSITY CASE

In this section, the POD-RBF NIROM was evaluated for predicting simulation solutions for the synthetic gas reservoir with different constant porosity values from Table 4.2. Overall, NIROM prediction of cases within the range of training runs were better than those outside the range of training runs (Figure 6.5 and Figure 6.6). From Figure 6.5, prediction 6 gave worse results than prediction 1, even though they were both outside the range of training runs. NIROM's prediction 1 satisfactorily mimicked the pattern of ECLIPSE simulation for the low porosity case (0.2 porosity), since the lower boundary of the training runs had a similar pressure decline behaviour in its later time. The errors for NIROM's predictions 2 to 5 were mostly less than 0.2% (Figure 6.5). As for the earlier synthetic gas reservoir scenarios, increasing the number of training runs gave better results for predictions 2 to 5 but worse results for predictions 1 and 6 (Figure 6.5). Figure 6.6 shows the best NIROM predictions of cumulative gas production and the reservoir pressure using 99.9% energy criterion. NIROM adequately mimics the patterns of ECLIPSE simulation results.

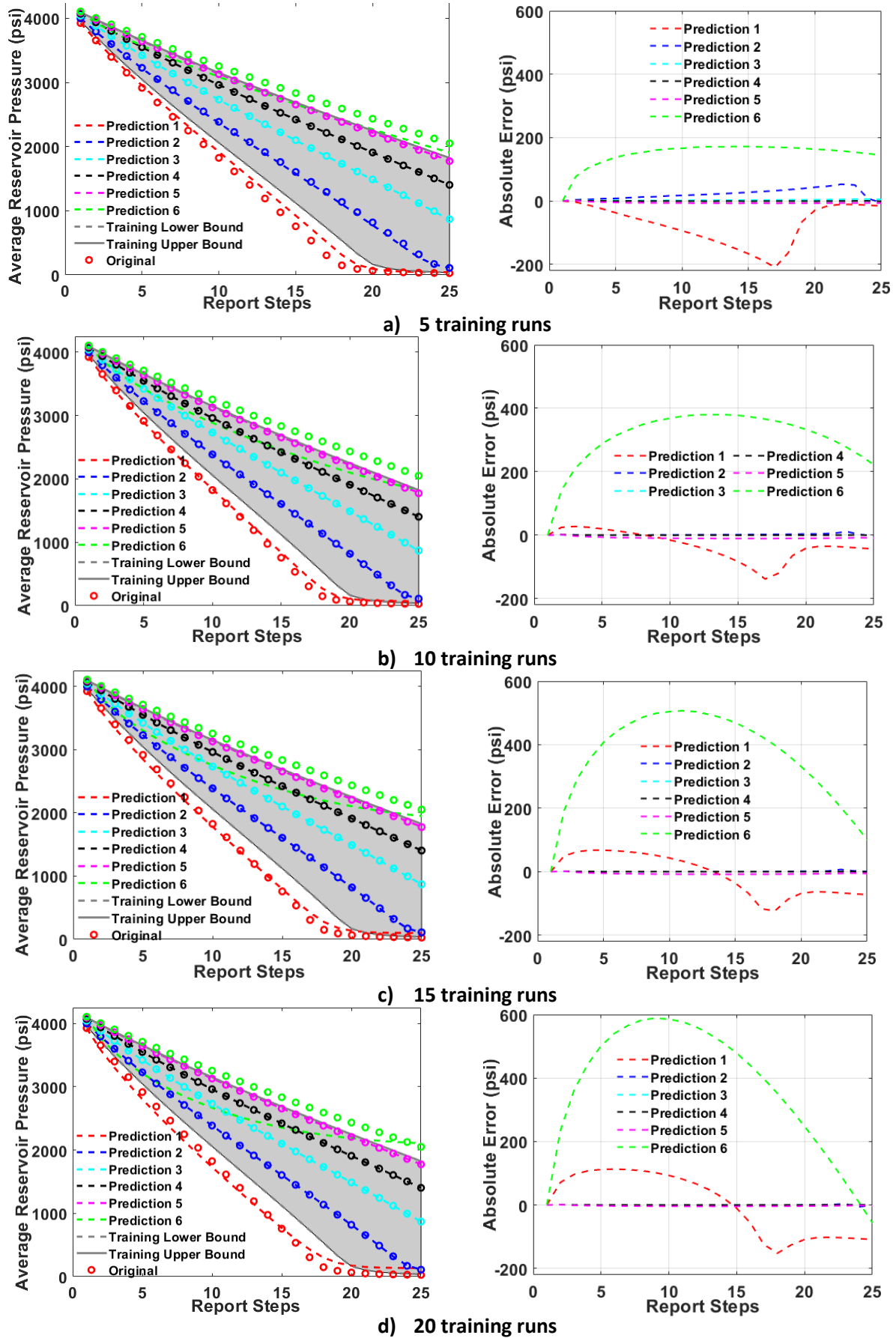


Figure 6.5: NIROM predictions of the average reservoir pressure in the synthetic gas reservoir with different porosity values. Predictions 2 to 5 were satisfactorily reproduced by the NIROM while the predictions outside the range of

training parameters (predictions 1 and 6) were less satisfactory. Despite the low porosity value in prediction 1 NIROM adequately predicted the pattern of pressure depletion in the reservoir.

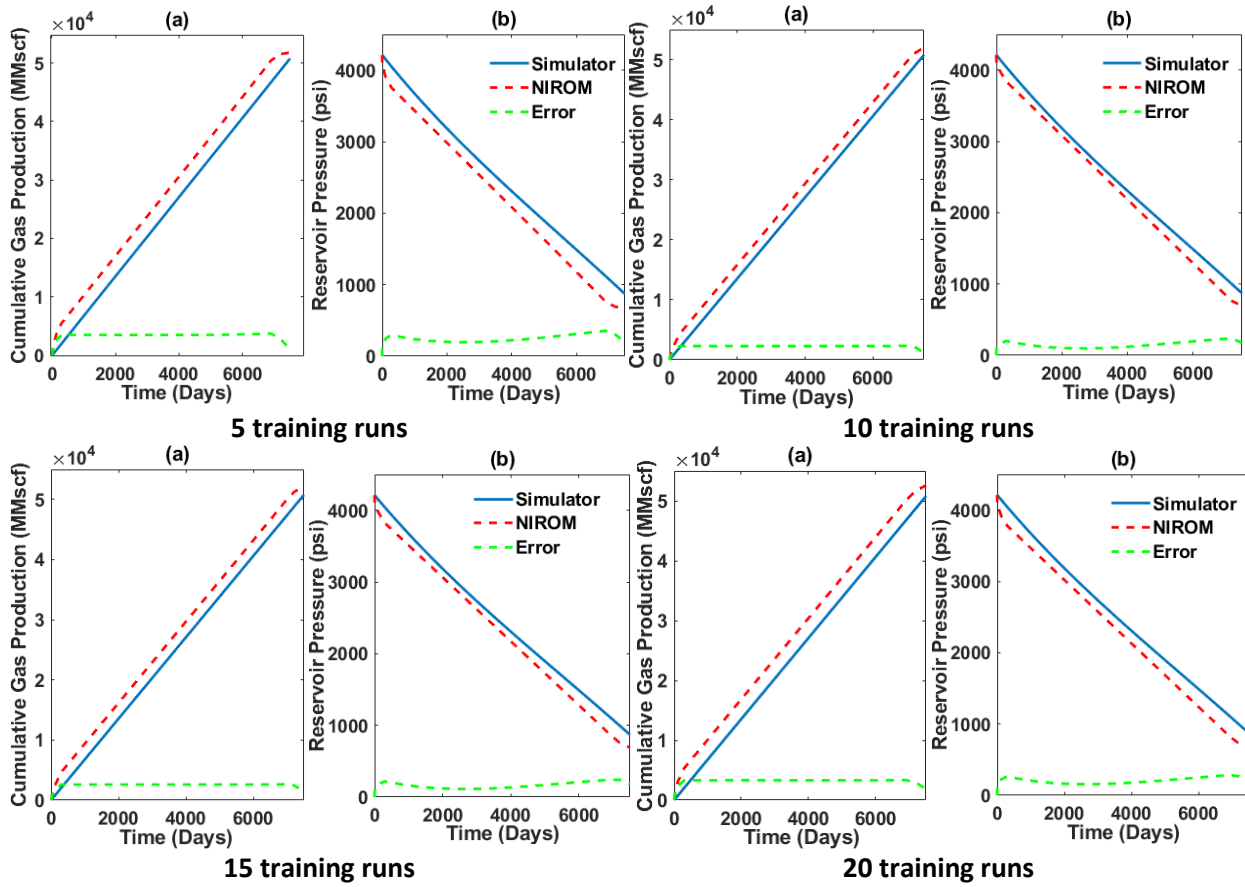


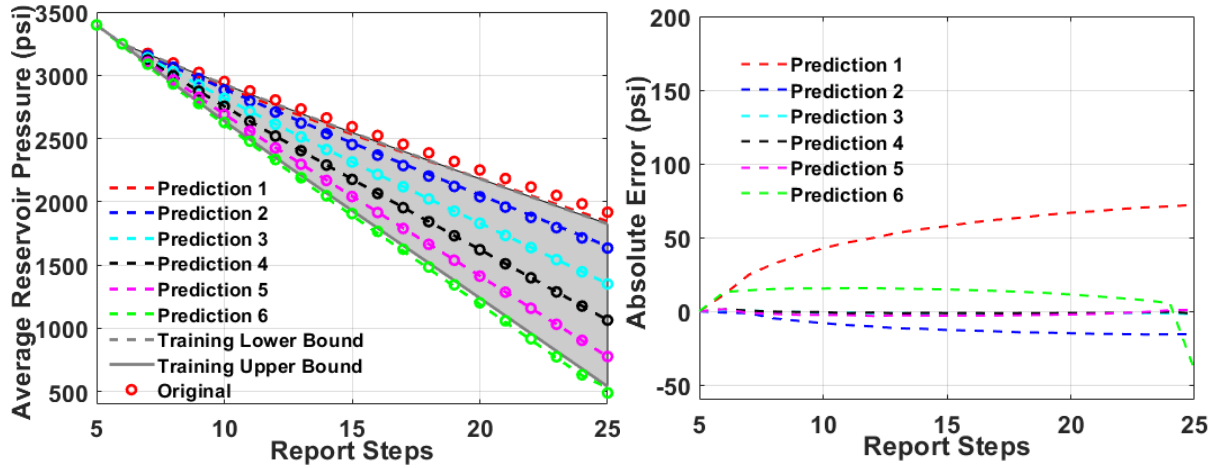
Figure 6.6: NIROM's prediction 3 of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir with different constant porosity values. NIROM adequately predicts the pattern of pressure depletion and cumulative production from the reservoir.

6.1.4 INITIALIZATION BY ENUMERATION

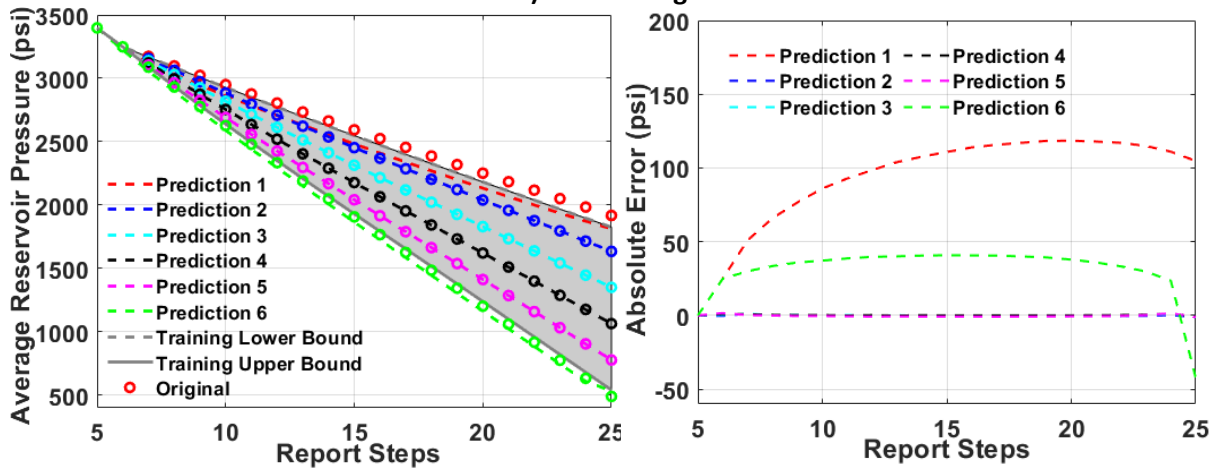
This case is similar to the constant production rate case, except that the synthetic gas reservoir has been initialised from a restart file (enumeration) rather than the typical initialisation by equilibration. After a reservoir's dynamic model has been built and history matched, subsequent predictions of the reservoir's performance are carried out based on the history matched dynamic model. But instead of running the full history of the reservoir's performance before commencing each prediction run, engineers save some time by initialising the prediction simulations from a restart file of the reservoir's production history, a process called enumeration. It was important to test NIROM on these cases to determine if NIROM can tolerate predictions that depend on information from restart files.

In this section, the same prediction parameters as in the rate controlled production (section 6.1.1) were used for NIROM predictions. The base case reservoir was run to obtain restart files for 6 report steps so that the predictions by enumeration commenced from the end of the 5th report step to the 25th report step (1501 days to 7500 days). The results in Figure 6.7 and Figure 6.8 are similar to those

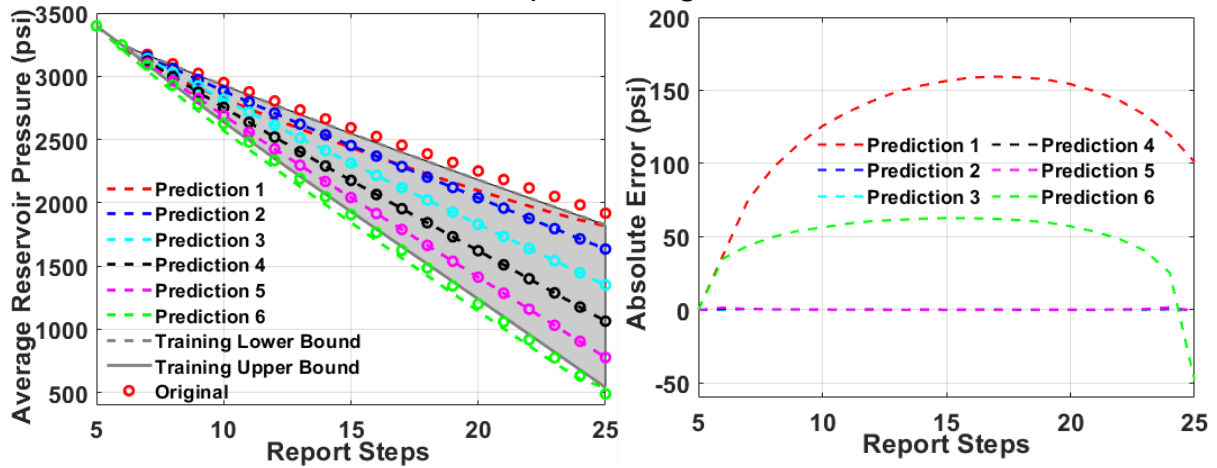
established from the other results of the synthetic gas reservoir, i.e. better results within the range of training parameters that get better with more training runs and worse results outside the range of training runs that get worse with more training runs. Here also NIROM satisfactory predicts the distribution of pressure as well as the production data. Prediction 6 which is closer to the range of training runs has less error than prediction 1 (Figure 6.7).



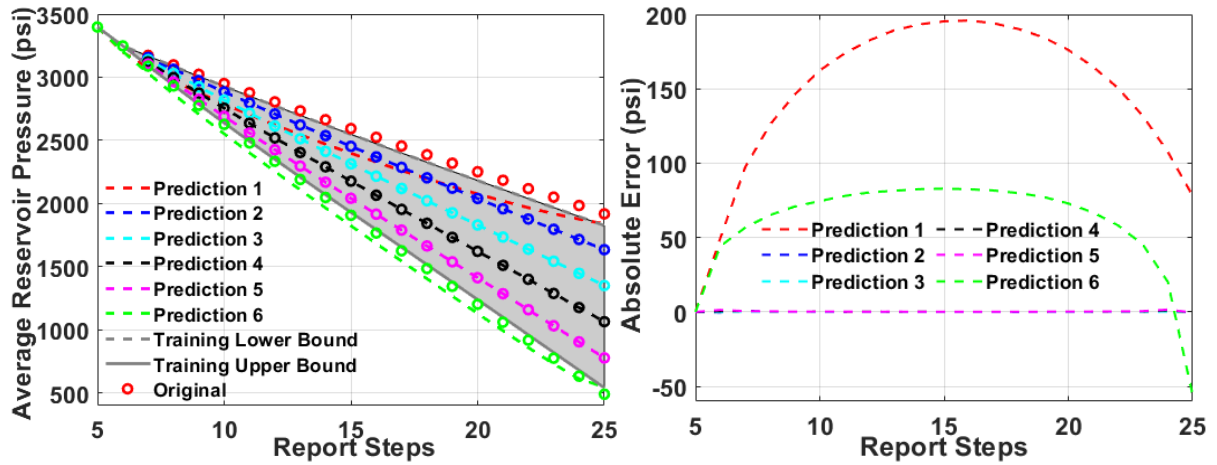
i) 5 training runs



ii) 10 training runs



iii) 15 training runs



iv) 20 training runs

Figure 6.7: NIROM predictions of the average pressure in the synthetic gas reservoir during constant rate production that was initialised by enumeration. NIROM gives satisfactory predictions of cases within the range of training parameters using 5, 10, 15 and 20 equally spaced training runs. Predictions outside the range of training parameters are less accurate.

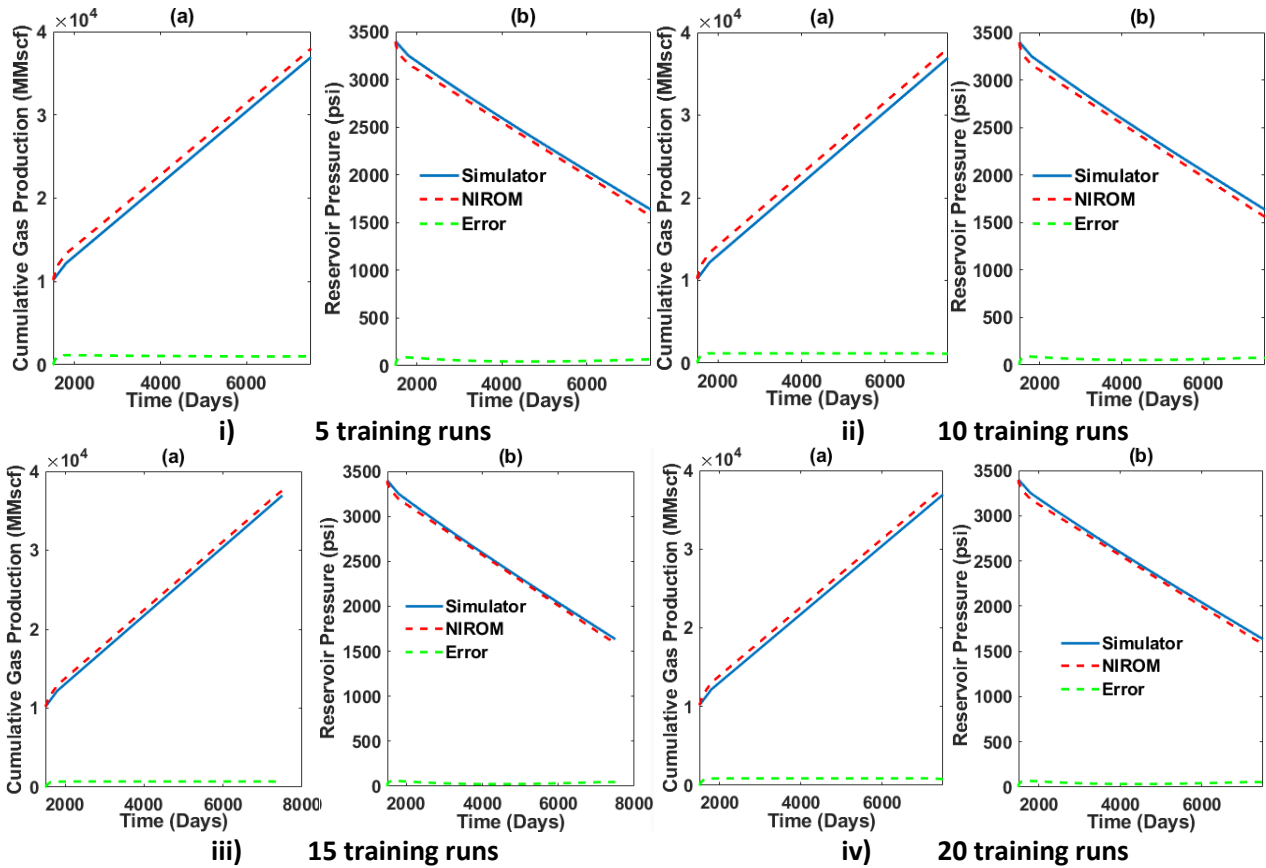


Figure 6.8: NIROM predictions of the cumulative gas production and the reservoir pressure for the synthetic gas reservoir when initialised by enumeration. NIROM adequately predicts the reservoir pressure and cumulative gas production from the reservoir. The best NIROM prediction (prediction 2) for 5, 10, 15, and 20 training run cases are shown in (i) to (iv) above. Similar results were obtained for other predictions within the range of training runs and less accurate results from predictions 1 and 6.

The general deduction from all the cases that were considered in this section (6.1) is that the NIROM satisfactorily predicts gas reservoir performance of unseen constant parameters scenarios, and it performs better during interpolation of unseen cases than during extrapolation.

6.2 HETEROGENEOUS GAS RESERVOIR (MODIFIED NORNE FIELD)

In this modified Norne gas field case, NIROM predictions have been investigated outside of the range of training parameters and training duration. Given that the reservoir is heterogeneous, NIROM predictions of both pressure and saturation distribution in the reservoir are shown below, as averages at every report step (annually).

6.2.1 RATE CONTROLLED PRODUCTION

For the prediction of the average reservoir pressure and saturation distribution in Figure 6.9 and Figure 6.10 respectively, predictions 1 and 6 (outside of the range of training runs) have values that are quite close to the boundaries of the training parameters, hence the minimal error in their predictions (less than 0.2bar). However in the 4 extra report steps (1500 days) that NIROM was not trained for, there is an increase in prediction errors, especially for prediction 6. This demonstrates that the most satisfactory application of NIROM for predicting reservoir performance is not only restricted to being within the range of training parameters but similar restrictions apply to the prediction durations. NIROM works best for predictions of the same or shorter durations as the training runs. This experience is similar to the pressure and saturation distribution results of the Norne field with a constant rate or BHP production control (Figure 6.9 to Figure 6.12).

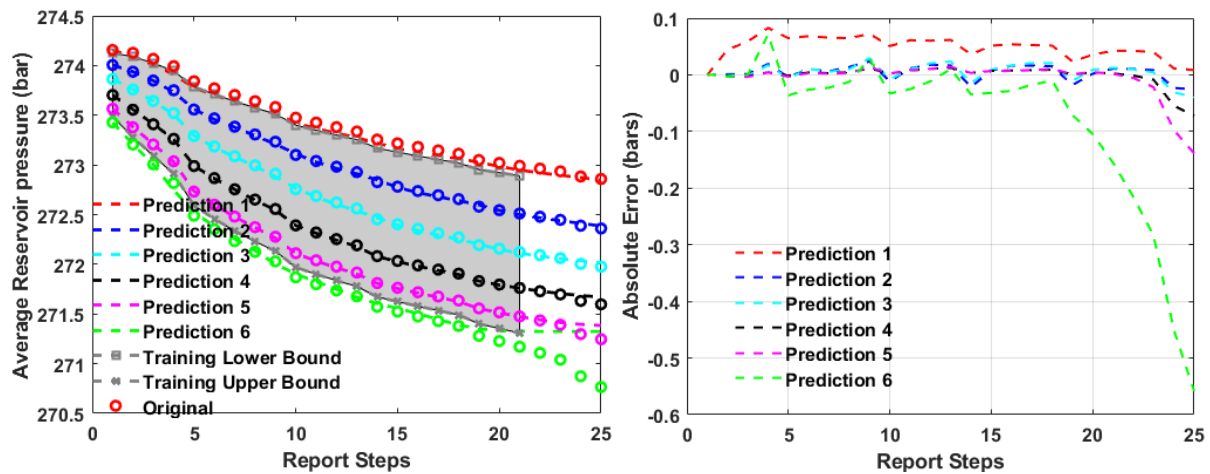


Figure 6.9: NIROM predictions of the average reservoir pressure across all the grid cells of the Norne gas reservoir. NIROM reproduces the average reservoir pressure in the reservoir with errors of less than 0.5bars, although predictions are less accurate in the last 4 report steps that NIROM was not trained on. Prediction 6 gives the highest error of all 6 prediction cases, as it is the farthest from the range of training runs.

For the average pressure distribution results in Figure 6.9, 5 POD basis functions (energy criterion) were needed to construct the NIROM and in Figure 6.10, the average gas saturation plot across the reservoir, 13 POD basis functions (energy criterion) were needed to construct the NIROM. Since more POD basis functions were required to capture the gas saturation distribution behaviour, this suggests

that capturing the changes in the gas saturation distribution in the reservoir is more complex than capturing the pressure distribution changes. The NIROM was trained with 10 training runs.

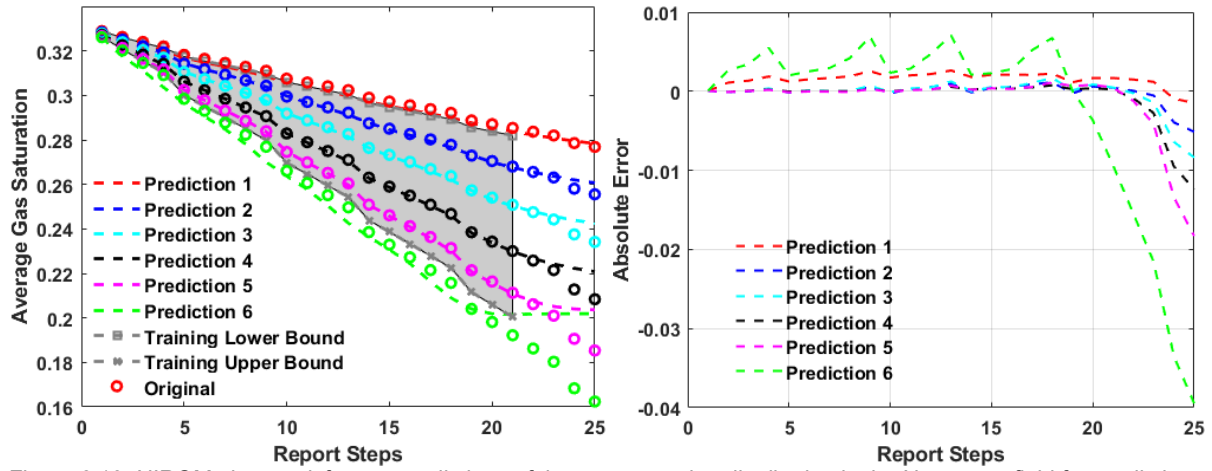


Figure 6.10: NIROM gives satisfactory predictions of the gas saturation distribution in the Norne gas field for predictions of similar production duration and range of training parameters. Outside the training runs' duration and range of parameters, NIROM predictions are less accurate. Prediction 6 has the most error in prediction and especially after the initial 21 report steps, because it is the most extreme prediction case.

6.2.2 BHP CONTROLLED PRODUCTION

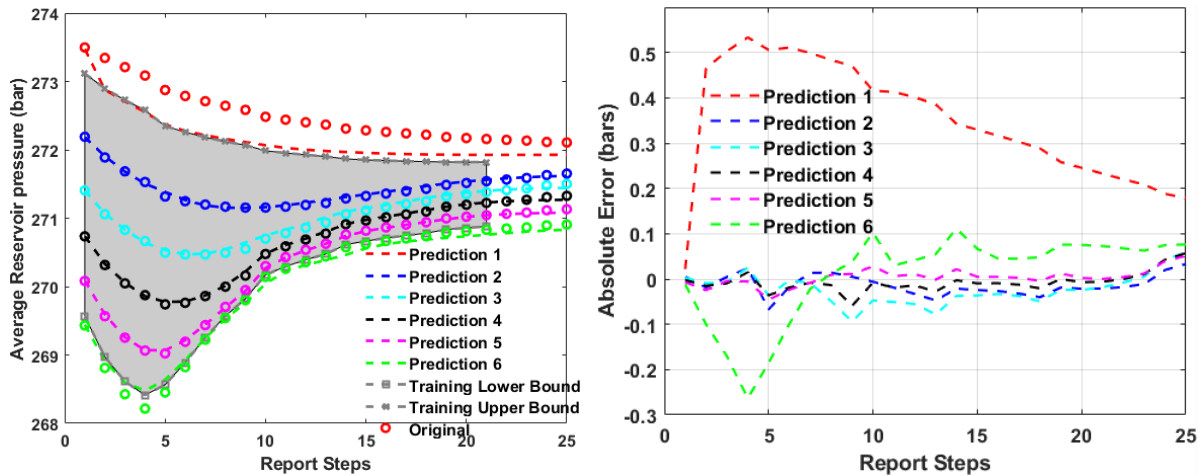


Figure 6.11: Average reservoir pressure distribution at different report steps of the NIROM trained on 10 equally spaced training runs. Prediction 1, the farthest from the range of training runs, has the highest prediction error during the time interval for which the model was trained, and the least error deviation outside of the training duration of the NIROM. All the other errors are negligible, mostly less than 0.5%.

Figure 6.11 and Figure 6.12 are plots of the average pressure and saturation distributions in the modified Norne field produced at constant BHP limits, respectively. In both figures, prediction 6 is close to the lower boundary of the training parameters while prediction 1 is not close to the upper boundary of the training data. It was noticed that although both NIROM predictions are satisfactory (less than 0.6bar absolute error), yet prediction 1 which is farther from the range of training parameters, exhibits more error than prediction 6. This reinforces what was indicated in prior results

that NIROM gives better prediction results for parameters outside but within close proximity to the range of training runs, than in predictions for parameters father away from the range of training runs. Seven (7) POD basis functions (99.9% energy criterion) were required to obtain the NIROM results in Figure 6.11 (for pressure distribution) while 16 POD basis functions (99.9% energy criterion) were required for Figure 6.12 (for gas saturation distribution). In Figure 6.12 it was observed that NIROM predictions were least accurate during the 4 extra report steps that NIROM was required to predict. Otherwise, NIROM predictions were mostly less than 0.1bar for pressure distribution and less than 0.005 for saturation distribution.

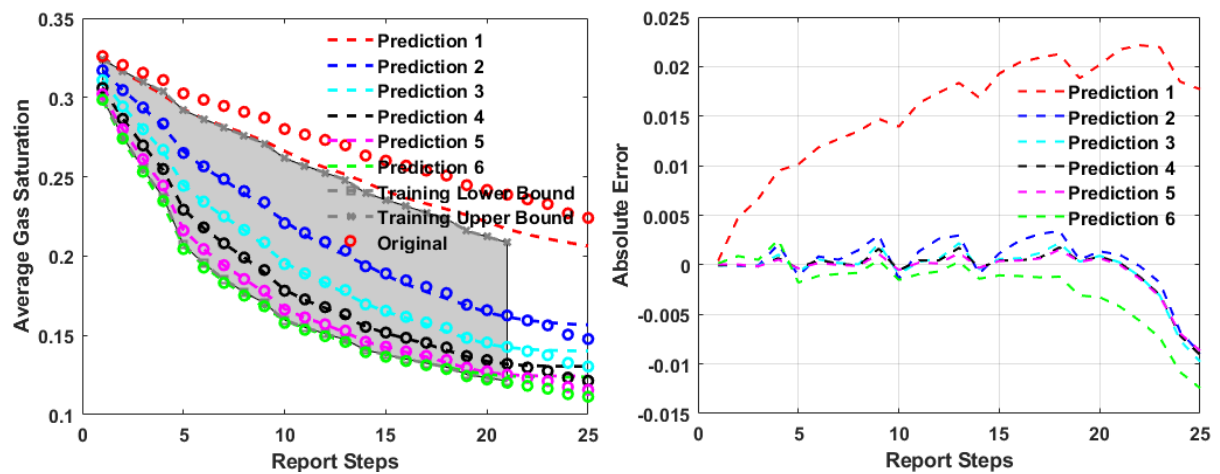


Figure 6.12: Average gas saturation in the Norne field model produced by constant BHP control with 10 training runs. Prediction errors are negligible, especially in the duration for which the NIROM has been trained.

Eighteen (18) other production performance data were predicted with NIROM. These production data include individual well and cumulative reservoir BHPs as well as gas and water production rates and ratios (Figure 6.13). These data are not reported for specific grid cells but are a continuous array of values across the reservoir or well(s). 9 POD basis functions were required to attain satisfactory predictions of these data, even though the energy criteria discussed in section 3.3 anticipated that NIROM should require 3 POD basis functions. The most difficult production variables to predict with NIROM were the water-related variables especially the water-gas ratios, Figure 6.13 (even though their values are small in comparison to other production data). For those production variables, NIROM predictions got better as more POD basis functions were used to train the NIROM (Figure 6.14). This is in agreement with Xiao (2016) ^[38], who proposed that better NIROM results can be obtained by increasing its number of POD basis functions.

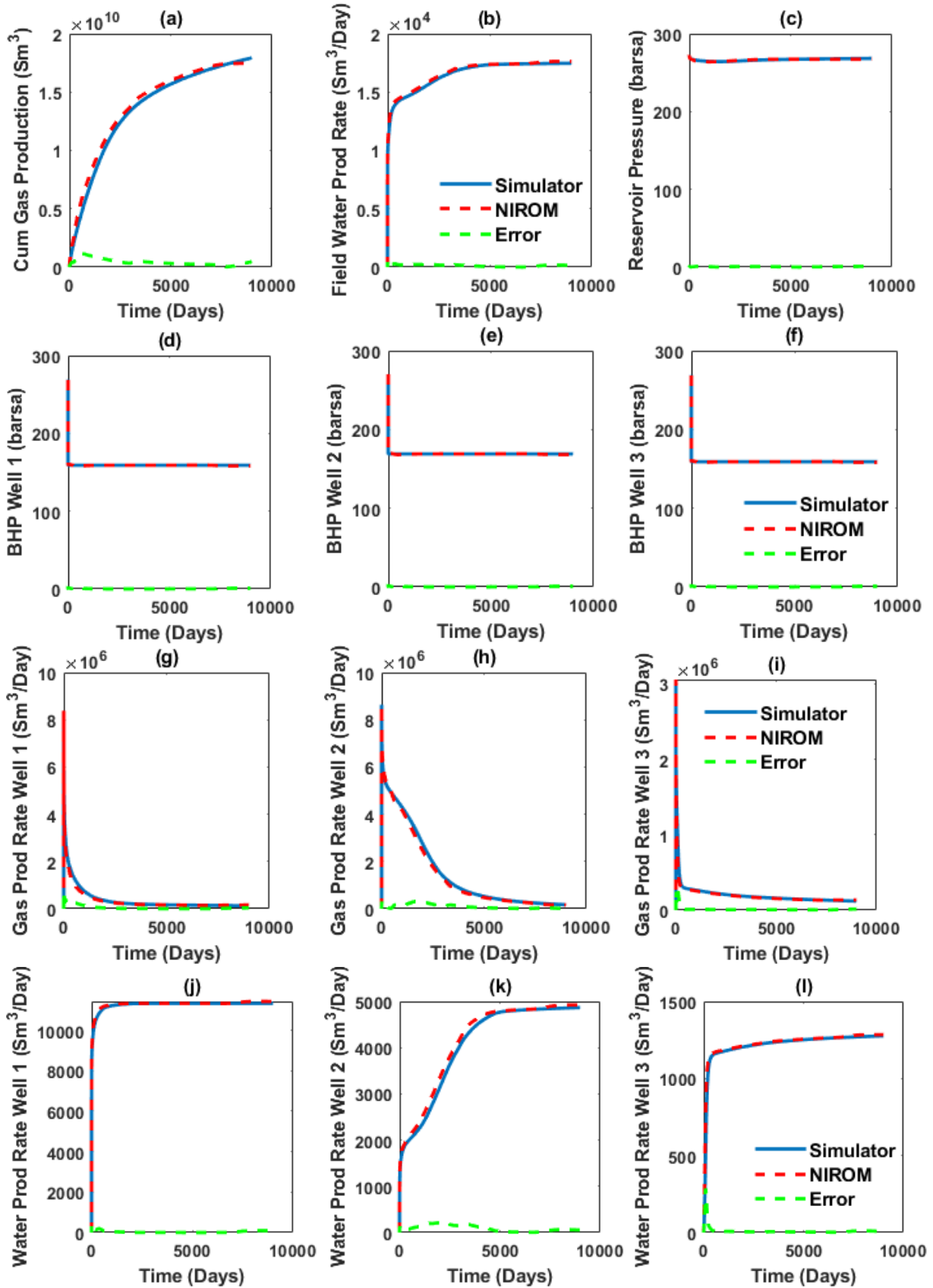


Figure 6.13: NIROM predictions of well and field gas and water production rates as well as BHP for prediction 5 of the modified Norne Field. Satisfactory results were obtained with NIROM trained on 10 training runs and 9 POD basis functions (energy criterion = 3 POD basis function). Similar results were obtained for predictions 2, 3 and 4, which are within the range of training runs and relatively less accurate results from predictions 1 and 6.

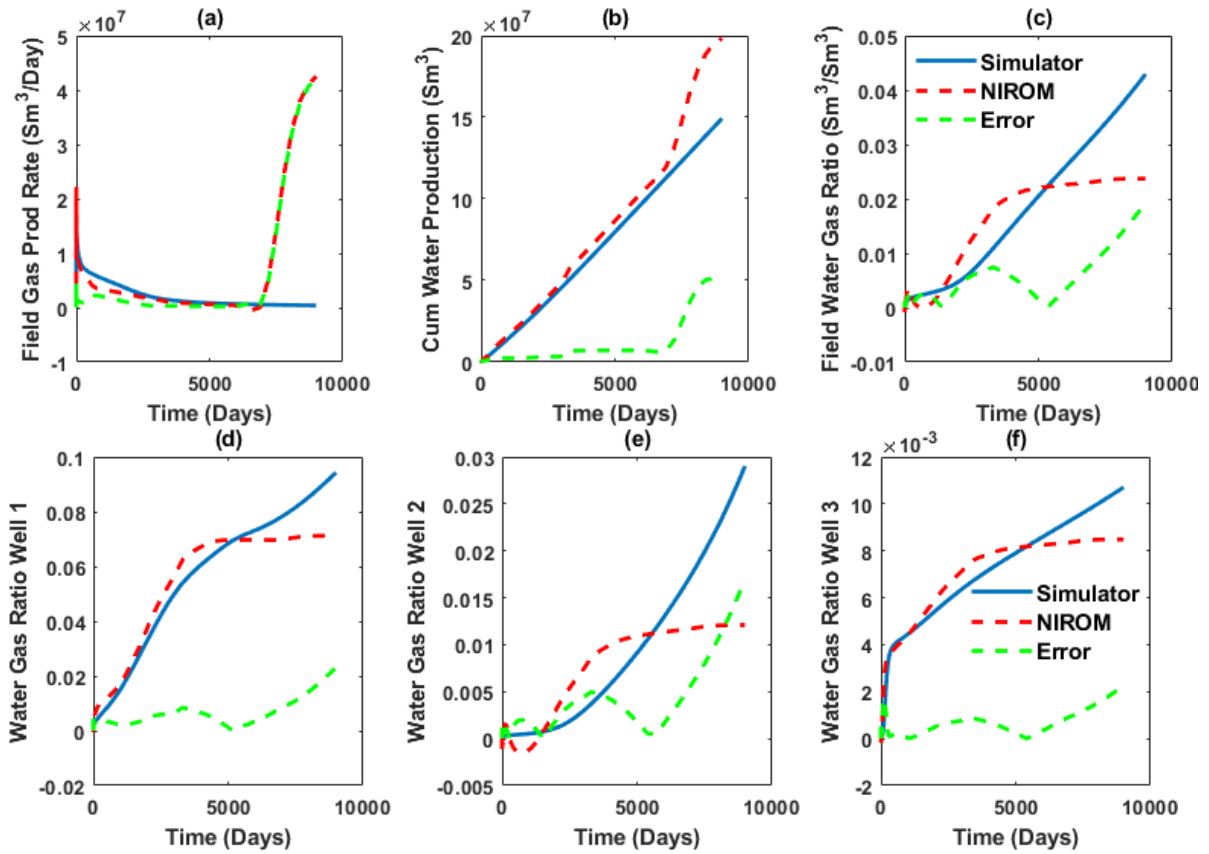


Figure 6.14: NIROM predictions of water gas ratios in all 3 wells and field gas production for prediction 5. These predictions were mostly for water-related production variables and do not adequately match those of ECLIPSE. This result is based on 9 POD basis functions, (energy criterion = 3 POD basis function).

From the water-related production data results in Figure 6.14c - f in addition to the production data results seen earlier in Figure 6.4b, it was observed that NIROM's inability to satisfactorily predict these production data could be linked to the magnitude of these production data, which are negligibly small compared with the other production data. This suggests that the NIROM is unable to cope with prediction cases where there is very small change in data over time.

6.3 WATER CONING GAS RESERVOIR

6.3.1 RATE CONTROLLED PRODUCTION

From Figure 6.15 and Figure 6.16, it is shown that NIROM's predictions within the range of training parameters are satisfactory, with errors of less than 3psia for pressure and 0.003 for saturation. Outside the range of training parameters, predictions are not satisfactory. NIROM was trained with 10 training runs for these predictions. For the pressure distribution predictions (Figure 6.15), 2 POD basis functions (99.9% energy criterion) were required and for gas saturation distribution predictions, 88 POD basis functions (99.9% energy criterion) were required. The high number of POD basis functions required to predict the reservoir's gas saturation distribution is linked to the water influx and water coning activities in the vicinity of the well. Commercial simulators typically require more

computational time to simulate the intricate dynamics of water-related scenarios (such as water flooding or water coning into the production well) hence similar challenges are attributed to NIROM's need for more basis functions in predicting the gas saturation across the reservoir.

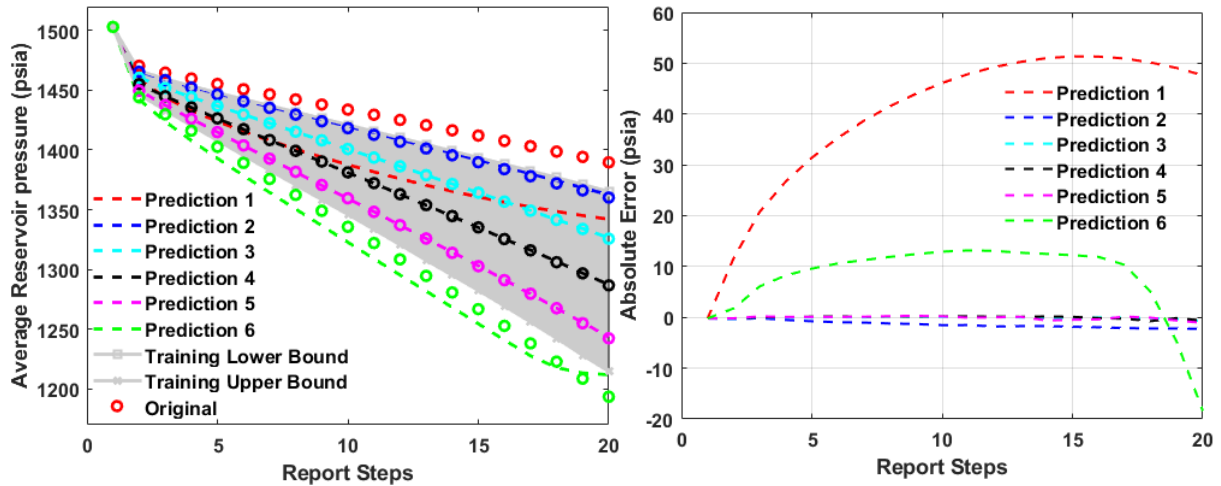


Figure 6.15: Average reservoir pressure predictions from NIROM match those from ECLIPSE, with negligible error for predictions 2 to 5 that are within the range of training parameters, while predictions outside the range of training parameters have larger errors. NIROM was trained with 10 training runs and 2 POD basis functions.

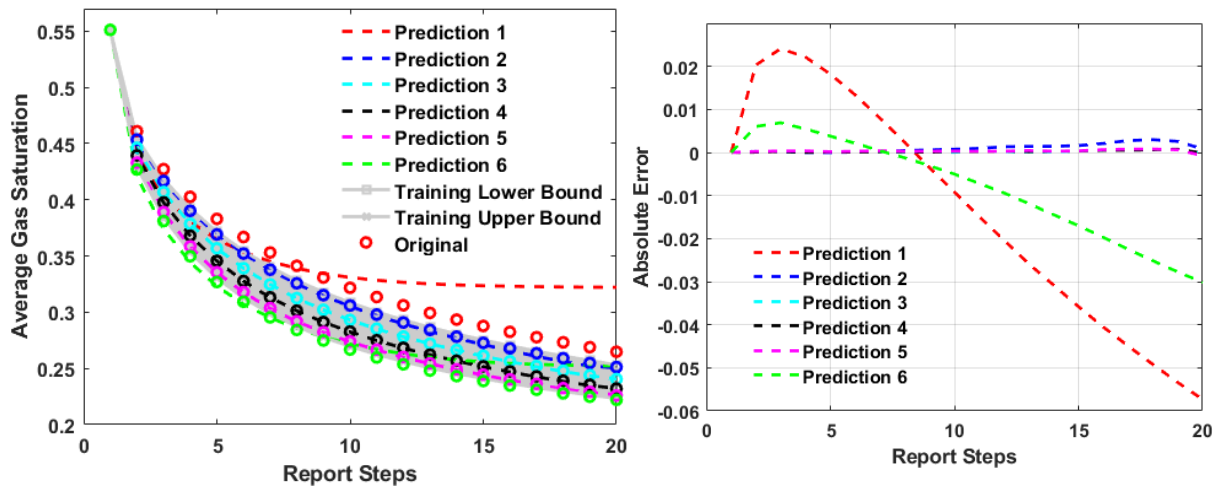


Figure 6.16: Average gas saturation distribution in the water coning model shows negligible error for predictions within the range of training parameters (predictions 2 to 5), and more errors outside the range of training parameters (predictions 1 and 6). NIROM errors are mostly less than 10% in comparison to ECLIPSE. NIROM adequately captured the non-linear decrease of gas saturation distribution in the reservoir with 88 POD basis functions.

Figure 6.17 and Figure 6.18 are plots of the gas saturation and pressure distribution in the water coning gas reservoir respectively, for prediction 3. Each column of both plots shows the distribution of the reservoir property across over 12000 grid cells of the reservoir at the beginning of every year. Figure 6.17 shows that the error in NIROM's prediction of the gas saturation distribution is mostly negligible while Figure 6.18 shows that the error in NIROM's prediction of the pressure distribution in

the reservoir was less than 1%. In both cases, it can be seen that NIROM satisfactory mimics ECLIPSE's simulation.

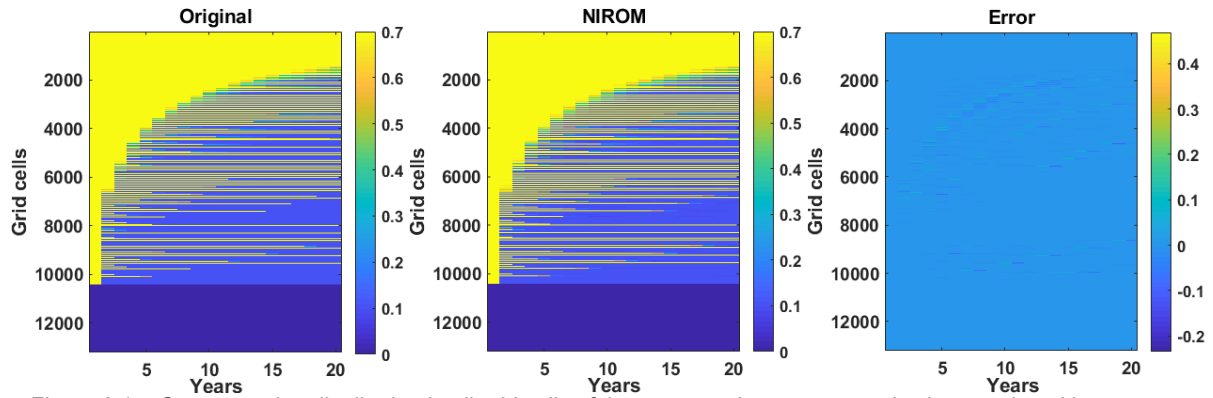


Figure 6.17: Gas saturation distribution in all grid cells of the water coning gas reservoir when produced by constant rate control. The error in this NIROM prediction (prediction 3) was generally negligible.

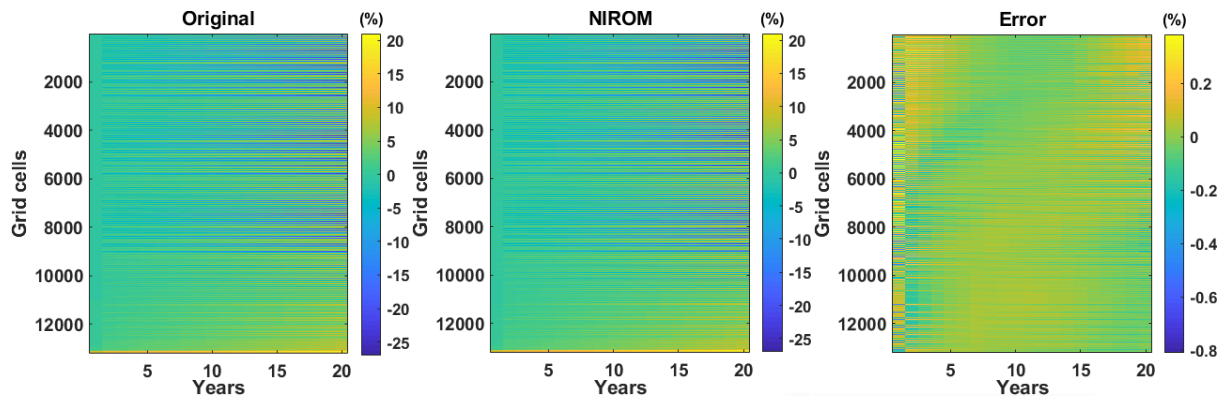


Figure 6.18: Pressure distribution in all grid cells of the water coning gas reservoir, normalised to zero mean and using 2 POD basis functions (as suggested by the energy criterion). The error in this NIROM prediction 3 was less than 1% of the original solutions from ECLIPSE.

For the other production performance data that NIROM was used to predict, 4 POD basis functions were required to obtain satisfactory NIROM predictions (Figure 6.20). This is in contrast to the number of basis functions computed by the 99.9% energy criterion (1 POD basis function). Nevertheless, both the satisfactory results (Figure 6.20) and those from using the energy criterion (Figure 6.19) are included below. A comparison of the improvements from Figure 6.19 to Figure 6.20 further demonstrates the possible benefit of increasing the number of basis functions used in NIROM simulations.

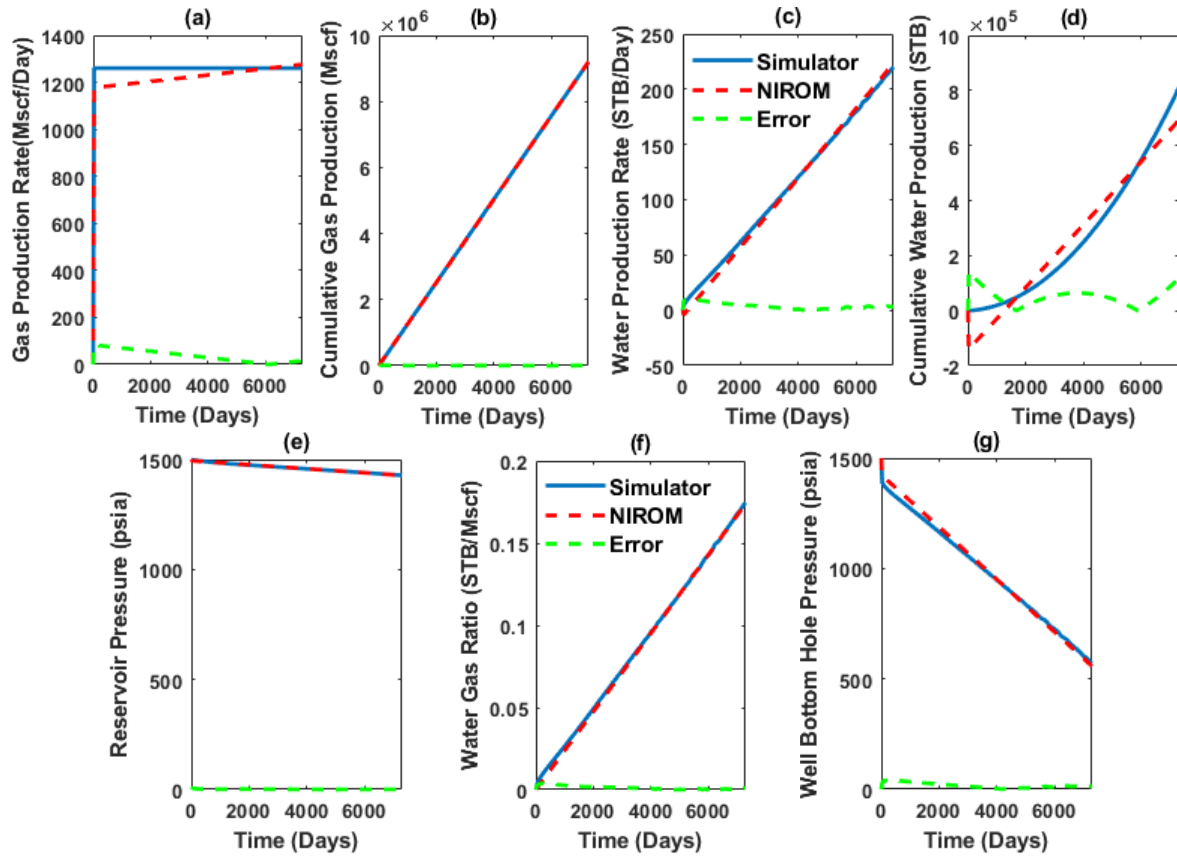
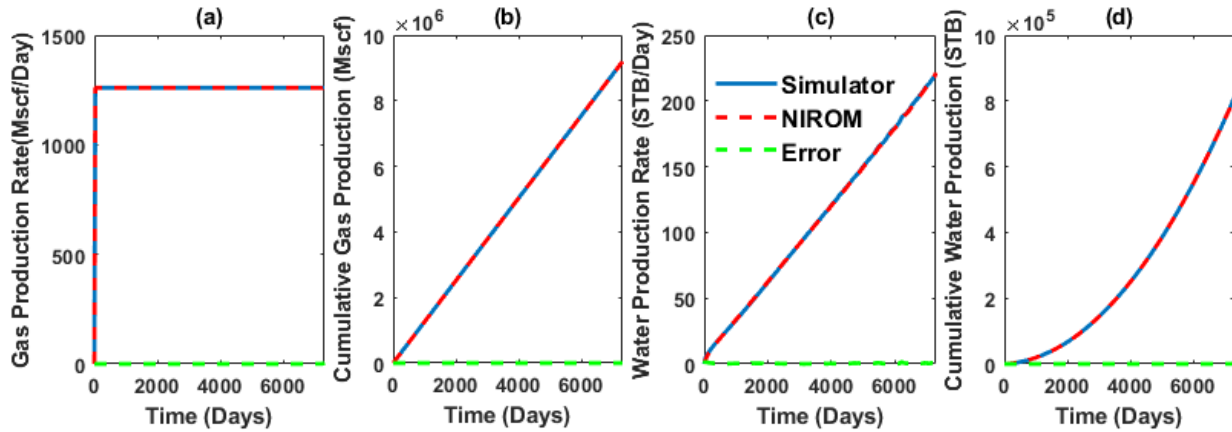


Figure 6.19: NIROM's prediction 4 of different reservoir production data using 1 POD basis function (99.9% energy criterion) and 10 training runs. NIROM predictions of cumulative gas production and average reservoir pressure are the most accurate followed by those of gas production rate, water gas ratio, well BHP and water production rate, which were less satisfactory. Prediction of the cumulative water production was the least accurate. Adding 2 more POD functions gave a satisfactory prediction of all the production variables (Figure 6.20).



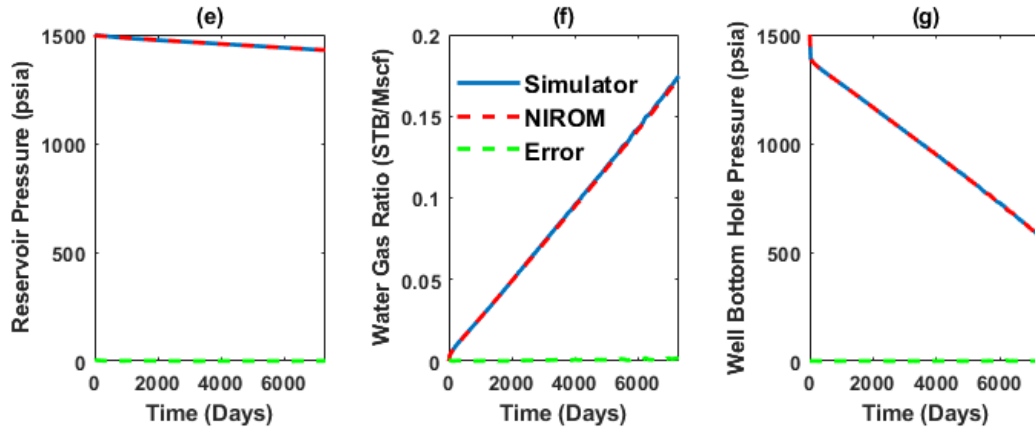


Figure 6.20: NIROM's prediction 4 of different reservoir production data with a satisfactory match. These results were obtained using 4 POD basis functions (99.9% energy criterion = 1 POD basis function) and 10 training runs. Adding 2 more POD functions gave accurate estimates of all the water and gas production variables.

6.3.2 BHP CONTROLLED PRODUCTION

Here, as in other cases, predictions within the range of training parameters have the least errors while those outside the range of training parameters have the highest error values (Figure 6.21 and Figure 6.22). From the average gas saturation predictions of Figure 6.22, it is again seen that prediction 6 which is much further out of the range of training parameters exhibits more prediction errors in comparison to prediction 1. Furthermore, it was observed that 2 POD basis functions (the energy criterion) were required to obtain satisfactory predictions of the average pressure distribution in the reservoir while 133 POD basis functions (the energy criterion) were required to satisfactorily predict the gas saturation distribution in the reservoir. The NIROM results in Figure 6.21 and Figure 6.22 were trained with 10 runs.

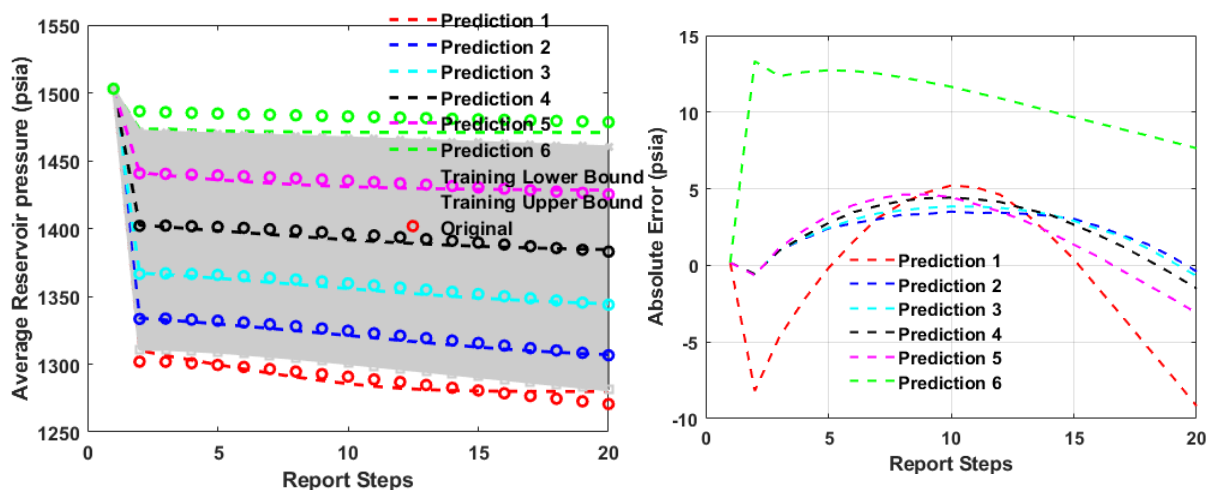


Figure 6.21: NIROM predictions of average pressure distribution in a water coning model produced at constant BHP, using 2 POD basis functions. Predictions within the range of training parameters have lower error margins than those outside the range of training parameters. However, errors are mostly less than 1% from ECLIPSE's solution.

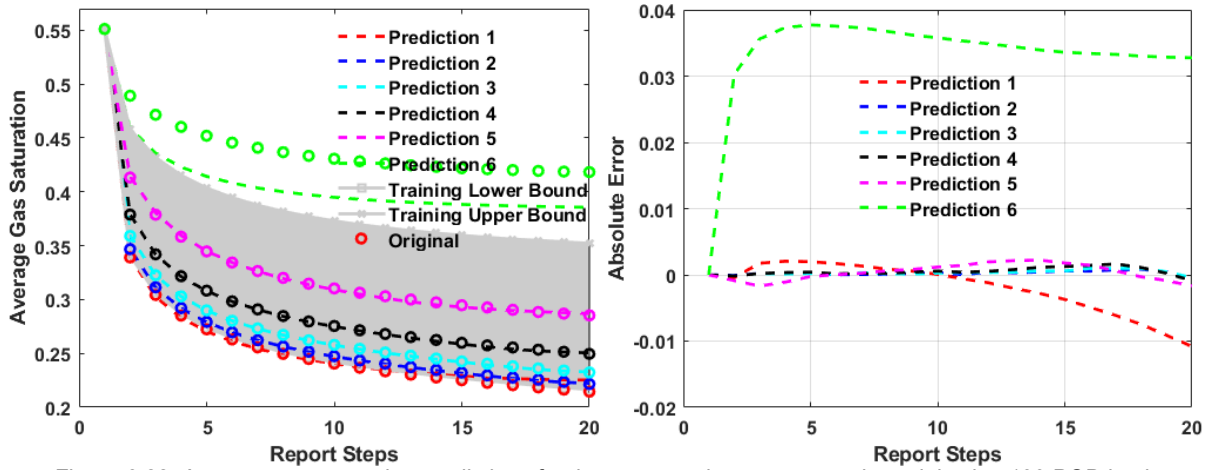


Figure 6.22: Average gas saturation predictions for the water coning gas reservoir model using 133 POD basis functions. Here the reservoir was produced by BHP control, and unlike for the rate controlled production example where both extrapolation predictions (1 and 6) were least accurate, prediction 6 was least accurate in this example. Prediction 6 had a much larger error margin than the other predictions cases that had less than 10% prediction error and were comparatively more accurate.

6.4 FIELD X

In these examples, the NIROM was trained with the constant production values/ parameters in Table 4.10 (page 92). Figure 6.23 shows that NIROM predictions within the range of training runs are better than those outside the range of training runs. The predictions within the range of training runs also improved as the number of training runs increased from 5 to 20 while the predictions outside the range of training parameters got worse as the number of training runs increased. This is in agreement with all the constant parameter examples in previous sections of this chapter. This behaviour is the same for the average of the saturation distribution in Figure 6.24.

NIROM satisfactorily predicted most of Field X's production data with the optimal number of POD basis functions i.e. energy criterion = 3 (Figure 6.25). Four of the five least accurately predicted production variables are water-related (water production and aquifer influx rates and their totals). A similar behaviour was observed for the water coning and Norne gas field, suggesting that the NIROM finds it more challenging to predict water-related production variables. This behaviour could be because of the relatively lower rate of change of water production compared to the change in other variables over time, or perhaps this behaviour is in agreement with the work of Kostorz et al. (2020)^[99] where the NIROM could not accurately capture the onset of water influx/ the shock front in their water flooding scenario. The least accurate NIROM predictions can however be improved by adding more POD basis functions to their implementation. Figure 6.25 shows NIROM predictions of prediction 2 trained with 5 runs. Training NIROM with 5 runs gave the best result for the production variables, even though this was not the case for the pressure and gas saturation distribution predictions shown in Figure 6.23 and Figure 6.24.

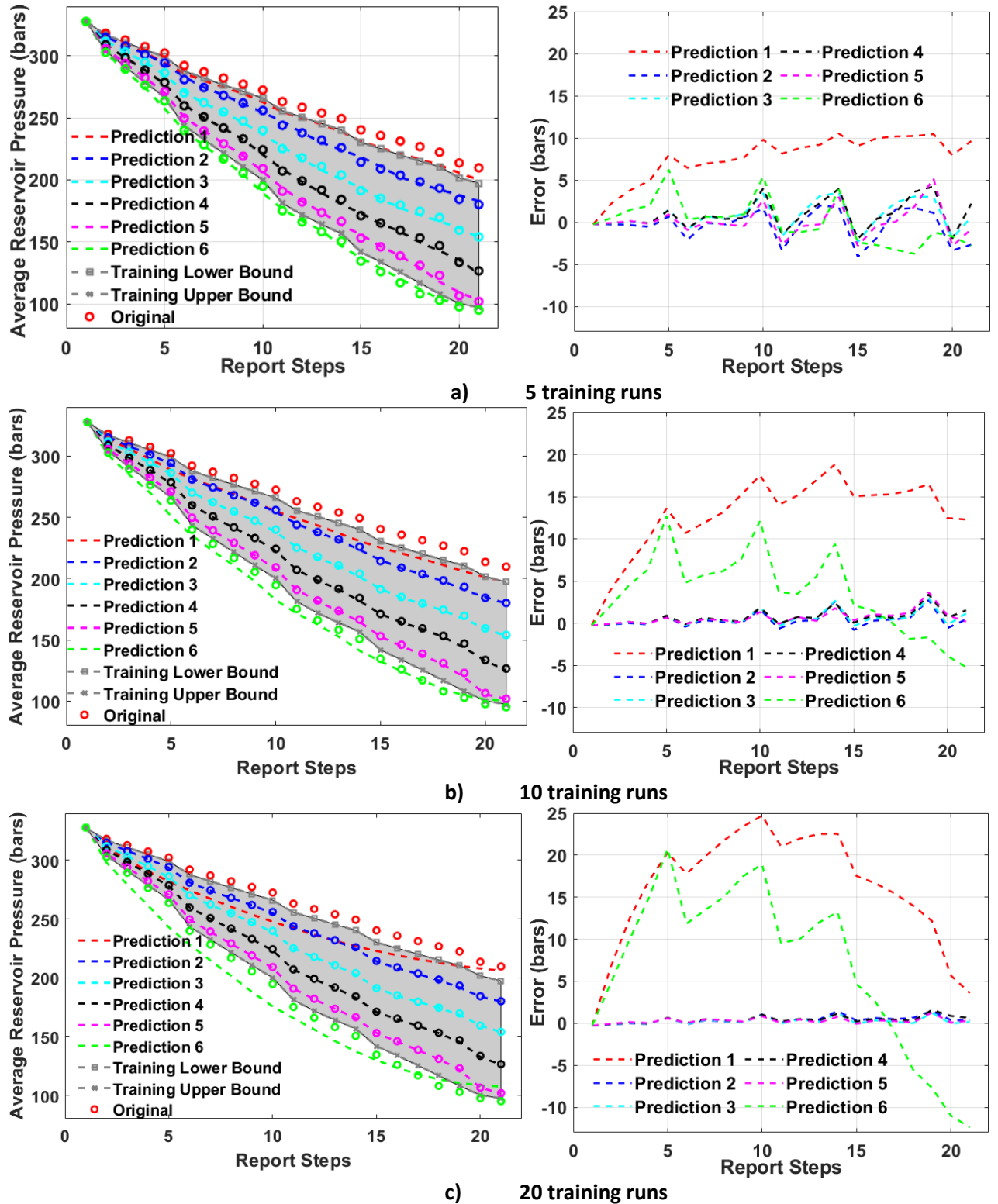


Figure 6.23: NIROM's prediction of the average pressure distribution in Field X when it was produced at different constant production rate values. Predictions 2 to 5 were always more accurate than predictions 1 and 6 (which were outside the range of training parameters). NIROM's predictions of cases 2 to 5 which are within the range of training parameters improved as the number of training runs increased, while NIROM's prediction of cases 1 and 6 got worse as the number of training runs increased from 5 to 20.

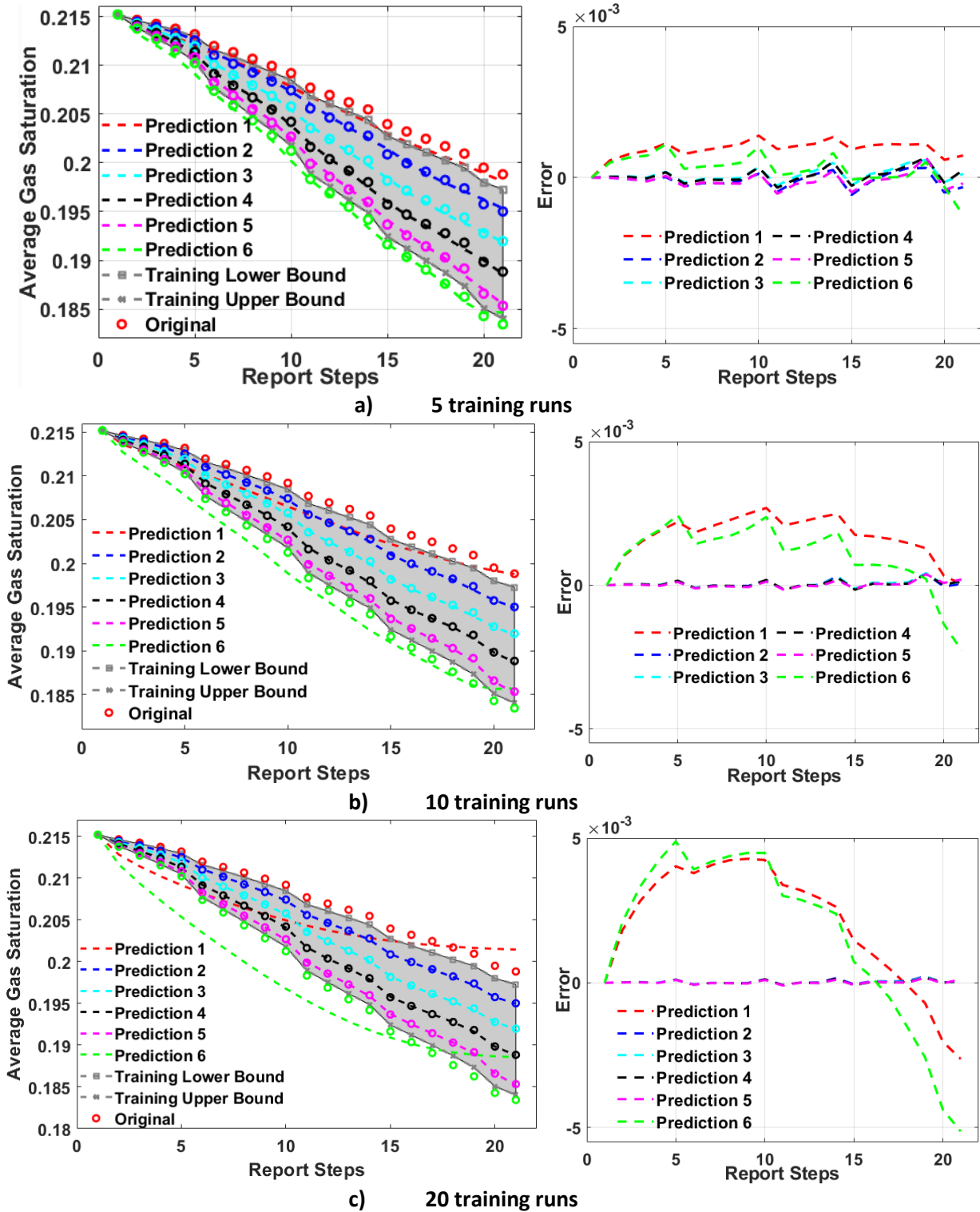
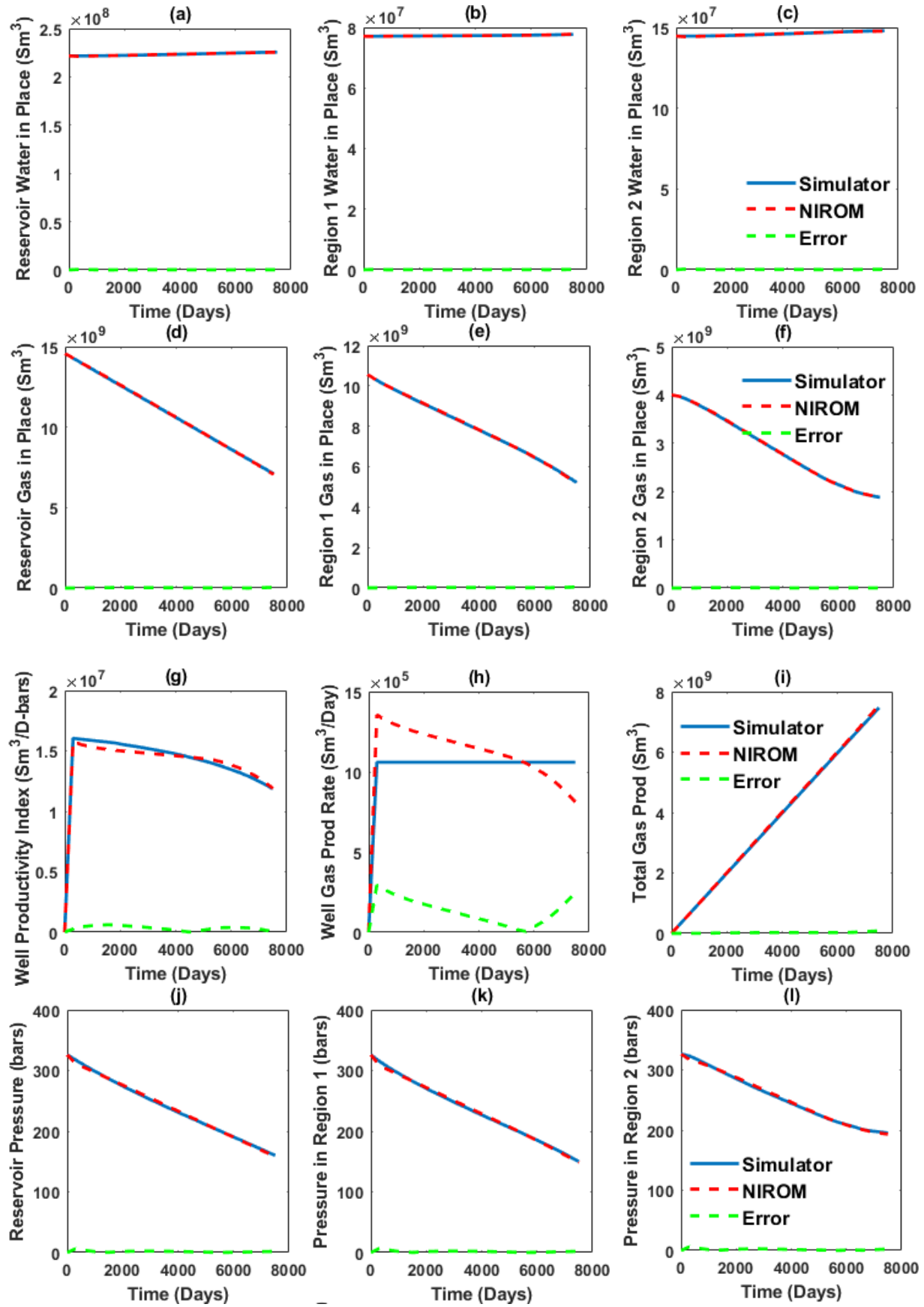


Figure 6.24: NIROM's predictions of the average gas saturation distribution in Field X at every report step. In these cases, Field X was produced at different constant production rate values. Here also, predictions 2 to 5 (the interpolation cases) are more accurate than predictions 1 and 6 (the extrapolation cases). However, the prediction cases that were within the range of training parameters get better as the number of training runs increased while the extrapolation cases got worse as the number of training runs increased.



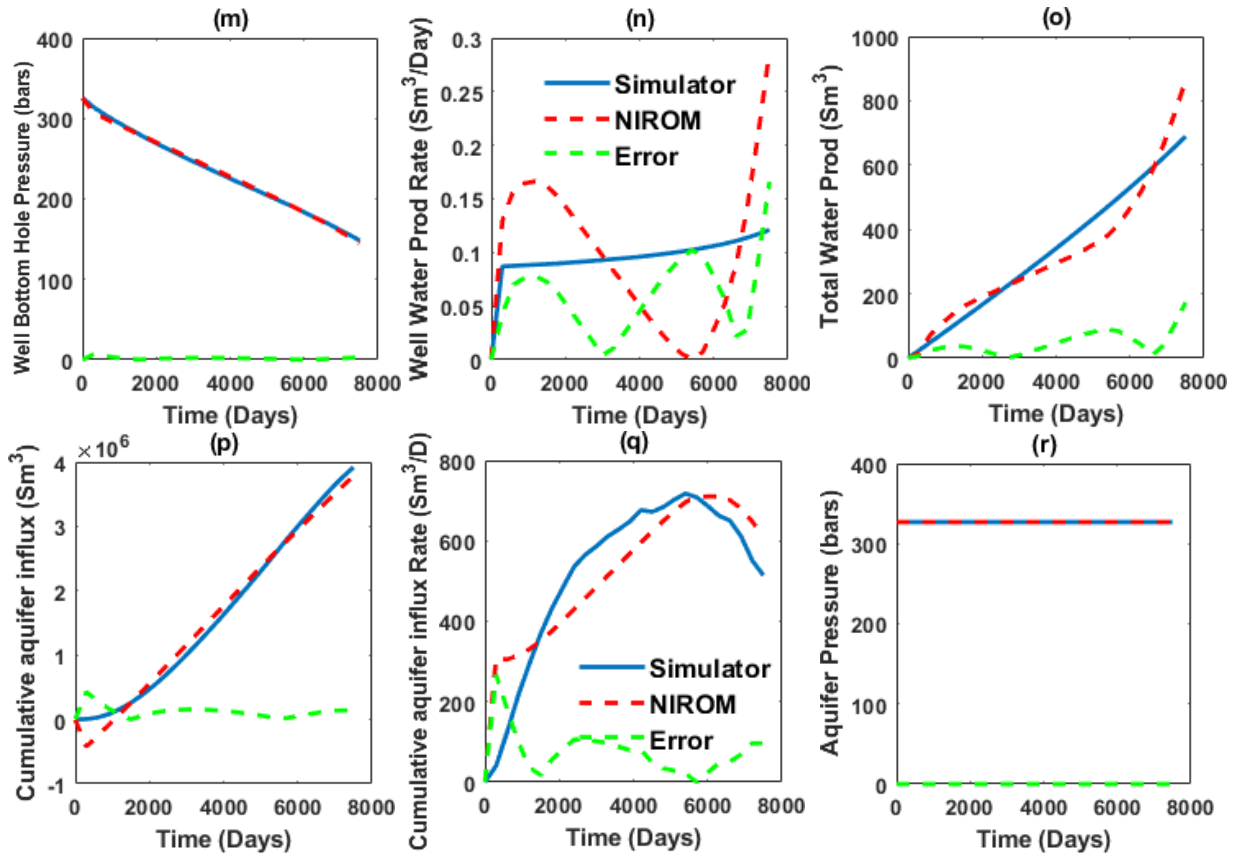


Figure 6.25: NIROM's prediction 2 of Field X's production variables using 5 training runs. The water production and aquifer influx variables were the least accurately predicted variables. Otherwise, most of the other NIROM predictions were satisfactory.

From the NIROM prediction results that have been shown in this chapter, it can be seen that implementing the NIROM based on the 99.9% energy criterion of 3.5 (optimum number of basis functions) gives satisfactory predictions for the grid distributed properties. However for the production variables, more basis functions were required to obtain satisfactory results than were suggested by the energy criterion, especially when the progressive change in a production variable has small values with increasing production time. Also recall from chapter 3 that the training data for the time series production variables are typically less than those of the grid distributed reservoir properties. For instance for the modified Norne gas field that was discussed in section 6.2, each report step had over 44000 data points (grid cells) for the grid distributed properties and it had 18 data points for the 18 production variables that were modelled. However these fewer data points come from different variables with different units and range of values. The data is normalised to unit values, nonetheless, the wide range of variation in the production dataset may still be contributing to why the NIROM requires more POD basis functions to accurately predict some of the production variables. Hence it is proposed that the NIROM be implemented with the full POD basis functions for the production data. This would have minimal impact on the NIROM's total runtime, particularly during POD projection and back-projection, since they have fewer data points and the NIROM would give

more accurate results. Besides, the NIROM implementation framework always computes the full and optimal number of POD basis functions that are required for predicting each property that is modelled with the NIROM (grid property or production data), whether or not they are eventually used for implementing the NIROM. Thus it is proposed that the POD-RBF NIROM can be suitably implemented with the optimal number of POD basis functions for predictions of the grid distributed properties and the full rank POD basis functions can be used for predicting the production variables.

6.5 ERROR ANALYSIS

In this section, the average percentage RMSEs (based on Equation 3.20) in all 6 predictions of each simulation case that has been considered (both within and outside the range of training runs) are provided for comparative analysis. This analysis highlights NIROM's suitability for different production scenarios as well as the most suitable number of training runs for each case. Figure 6.26 shows the average RMSEs in the 4 constant parameter prediction cases of section 6.1 and Figure 6.27 shows the average percentage RMSEs for the constant parameter predictions in sections 6.2 to 6.4 (modified Norne field, water coning model and Field X). The RMSEs shown in this section are also based on implementing NIROM predictions with the optimal number of POD basis functions. Although the gas saturation distribution in the synthetic gas reservoir was not shown in section 6.1, their average percentage RMSEs are shown below.

From Figure 6.26 and Figure 6.27, it can be seen that 5 training runs can suffice for satisfactory NIROM predictions. However a few exceptions, especially the production data predictions, show that training NIROM with few training runs does not always lead to the best overall predictions. Some of these exceptions are in the prediction of the production data in the constant production rate case and gas saturation distribution in the constant BHP limit case. A possible reason for these exceptions could be the extent of over- or under-predictions that occur in predictions outside the range of training parameters (predictions 1 and 6). It was earlier highlighted that with an increasing number of training runs, predictions within the range of training parameters get better, while predictions outside the range of training runs get worse. This means that for any given number of training runs, if the errors in NIROM's predictions 1 and 6 are significantly more than the improvement in predictions 2 to 5 (within the range of training runs), then the average RMSE for that case will be worse. For instance in Figure 6.1, whereas the maximum error from the 5 training runs case was less than 100psi, the maximum error from the 20 training runs case was more than 200psi. The exaggerated errors in predictions 1 and 6 of the 20 training runs case therefore led to a higher average RMSE for that case. Nevertheless, these results still show that NIROM can give satisfactory predictions of the grid distributed reservoir properties with a minimal number of training runs.

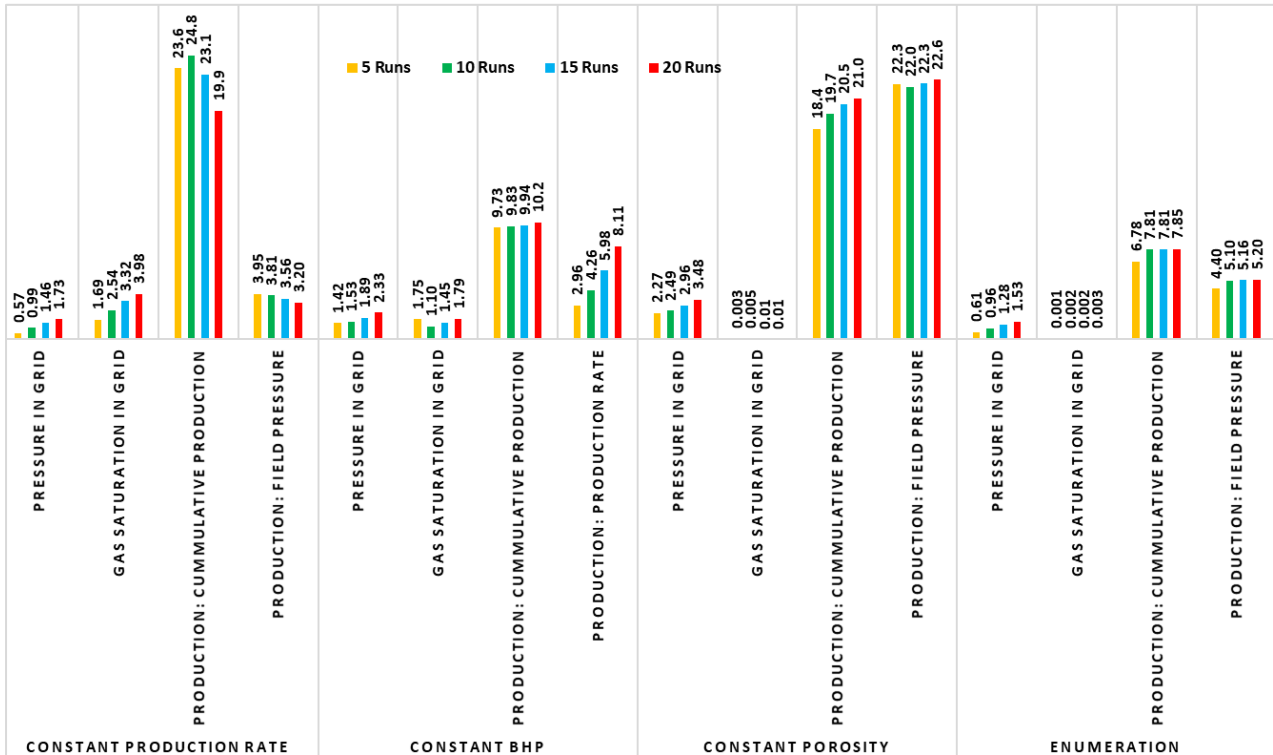


Figure 6.26: Average percentage RMSEs of all predictions in each constant parameter case of the synthetic gas reservoir model. NIROM predictions of grid distributed pressure and gas saturation give smaller errors than those of the production data that require more than the optimum number of POD basis functions to obtain satisfactory results.

The average percentage RMSEs for the prediction of pressure distributions in all 4 reservoir models are less than 4% (Figure 6.26 and Figure 6.27), reflecting that NIROM works well for forecasting the grid distribution of pressure in different types of gas reservoirs produced with constant parameters. For the production data, there are higher NIROM errors since a single average value is computed across several production variables (including water-related variables that have larger errors). This is in addition to the compounded errors for predictions outside the range of training runs and considering that the RMSEs in Figure 6.26 and Figure 6.27 were estimated based on the optimum number of POD basis functions required to implement NIROM. Recall that in previous sections (6.1.2 and 6.1.3) it was shown that accurate predictions of production data required more POD basis functions than were computed from the energy criterion. With 10 training runs, the average percentage RMSE in the prediction of the gas saturation can be kept below 10%.

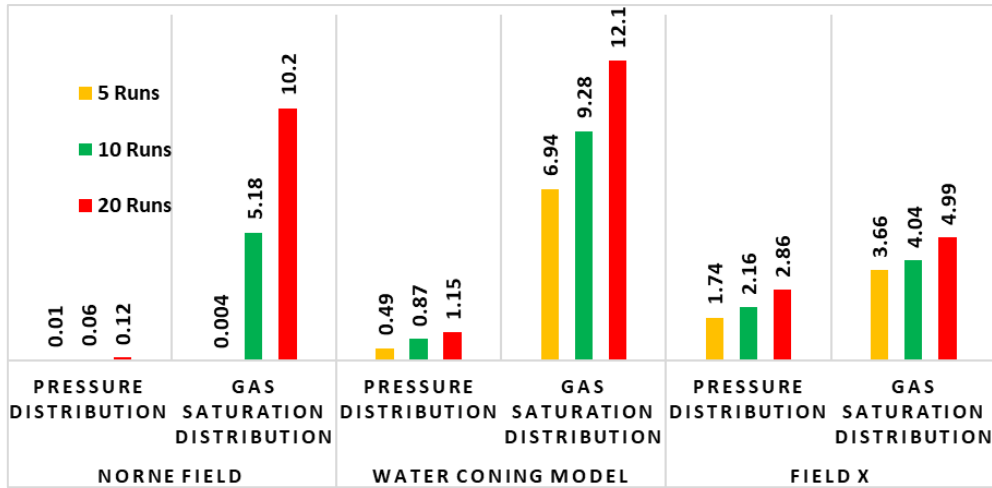


Figure 6.27: Average of the percentage RMSE for NIROM predictions for the modified Norne field and water coning reservoir models. NIROM predictions of the pressure distribution in the reservoirs had lower errors than those for the gas saturation distribution. NIROM implemented with 5 training runs consistently gave the least RMSEs.

Table 6.1: Number of POD basis functions required for NIROM predictions. The listed values represent the energy criterion except where it is mentioned otherwise.

Reservoir Model	Cases	Number of POD basis functions required versus energy criterion	
		Production Rate	BHP Limit
Modified Norne Field	Production Control		
	Pressure in grid	5	7
	Gas Saturation in grid	13	16
	Production data	9	(energy criterion = 3)
Water Coning Reservoir	Production Control		
	Pressure in grid	2	2
	Gas Saturation in grid	88	133
	Production data	4	(energy criterion = 1)
Field X	Production Control		
	Pressure in grid	2	
	Gas Saturation in grid	85	
	Production data	3	(energy criterion = 1)

For the synthetic gas reservoir, one POD basis function was sufficient for making all the NIROM predictions, i.e. pressure and saturation distribution in all grid cells in addition to production variables. Table 6.1 shows the number of POD basis functions that were required to make NIROM predictions for the other 3 reservoirs that were considered in this work. In the cases shown in Table 6.1, NIROM required the highest number of POD basis functions for implementing predictions of the gas saturation distribution in the respective reservoirs and fewer POD basis functions to learn the underlying pressure depletion mechanism. This could be likened to the work of Kostorz et al. (2019) ^[37] where it was shown that increasing the number of POD basis functions used for implementing the NIROM may improve the accuracy of predicting complex production scenarios like the water saturation distribution of a propagating shock front.

Nonetheless, the average number of training runs used in this work are fewer than those that were used in Kostorz et al. (2019) ^[37]. Kostorz et al. (2019) ^[37] used between 200 and 600 training runs for their test case which was a 2D oil reservoir undergoing gas flooding. The different reservoir dynamics that are considered in their work and this work (pressure depletion gas reservoirs), in addition to the different scales of magnitude of the test cases in both works is a possible reason for such a difference in the number of training runs required for the NIROM to learn the systems' dynamics and make satisfactory predictions. Further investigation is required to establish this.

6.6 TIME ANALYSIS

Having seen that NIROM gives satisfactory predictions of reservoir performance in different reservoir models, analysing the speed at which NIROM produces these predictions is the other crucial aspect of this work. It is particularly important to evaluate NIROM's speedup for practical simulation cases with a large number of simulation grid cells and/ or for reservoirs with complex dynamics such as the heterogeneous reservoirs and the water coning model. In this section, the average time spent in running simulations with NIROM and ECLIPSE is compared in tables (Table 6.2 to Table 6.4) and with stacked column charts (Figure 6.28 to Figure 6.34). ECLIPSE gives a single value for its runtime, but this section provides the breakdown of the time that was spent in executing the different aspects of NIROM's simulation (Figure 6.28 to Figure 6.34). CPU time has a unit of seconds in this work, which was executed on a workstation with 40 logical processors, 2.4 GHz base speed and 128GB RAM.

6.6.1 SYNTHETIC GAS RESERVOIR

From Table 6.2, NIROM simulations are between 8 and 17 times faster than ECLIPSE simulations of the homogeneous synthetic gas reservoir, depending on the number of training runs used to train the NIROM. Table 6.2 along with Figure 6.28 and Figure 6.29 show that most of the time spent in running the NIROM was in the pre-processing steps, when extracting training or prediction data from ECLIPSE output files into the NIROM's programming interface. The next step that accounts for the most NIROM simulation time was during POD reduction. POD involves matrix inversion, which is computationally intensive for large data. So, the more data is used to train the NIROM (in terms of model size/ number of grid cells and/ or the number of training runs and report steps) the more time is spent in the POD reduction step. Figure 6.28 and Figure 6.29 show a progressive increment in POD reduction time as the number of training runs increased. The computation of interpolation weights is also dependant on the amount of data available for training the NIROM (number of the training runs). Hence the progressive increase in time taken to compute the interpolation weights as the number of training

runs increased. The main time spent for NIROM interpolation of new cases is significantly smaller than the total NIROM runtime, an average of 0.01 second.

Table 6.2: Summary of time taken to implement the constant production controlled (rate and BHP limit) cases of the synthetic gas reservoir. These values correspond to the plots in Figure 6.28 and Figure 6.29.

Production Control	Rate				BHP limit			
Number of Training Runs	5	10	15	20	5	10	15	20
Extracting Training Data	0.15	0.14	0.15	0.12	0.16	0.13	0.16	0.15
Mean Extraction	0.03	0.04	0.06	0.07	0.02	0.03	0.06	0.07
POD Reduction	0.05	0.10	0.17	0.25	0.05	0.10	0.17	0.22
Parameterisation & Normalization	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Computing Interpolation weights	0.03	0.03	0.04	0.07	0.02	0.05	0.09	0.13
Total NIROM Training Time	0.27	0.32	0.43	0.51	0.26	0.32	0.49	0.57
Extracting Prediction Data	0.15	0.15	0.14	0.15	0.15	0.15	0.16	0.15
Normalising Initial Condition Data	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
NIROM Interpolation	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Total NIROM Time	0.42	0.47	0.58	0.67	0.43	0.49	0.66	0.74
Total ECLIPSE Time	5.51	5.51	5.51	5.51	7.41	7.41	7.41	7.41
Speed-up (Ratio)	13.1	11.7	9.53	8.19	17.4	15.2	11.2	10.0

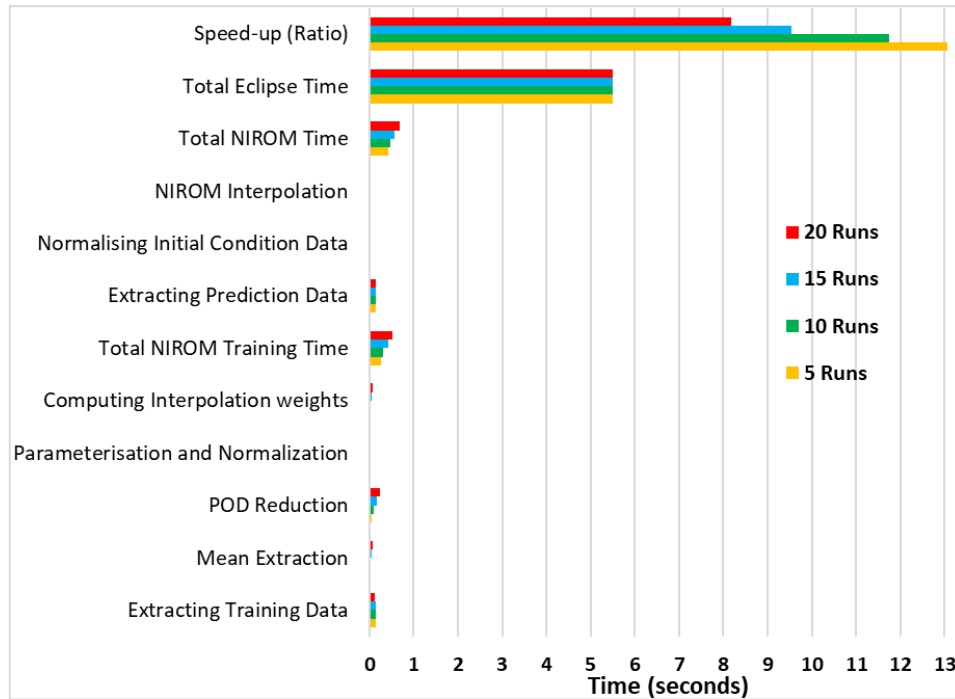


Figure 6.28: Summary of the time taken to implement the constant production rate cases of the synthetic gas reservoir. NIROM speedups of between 8 and 13 times the speed of ECLIPSE are achievable.

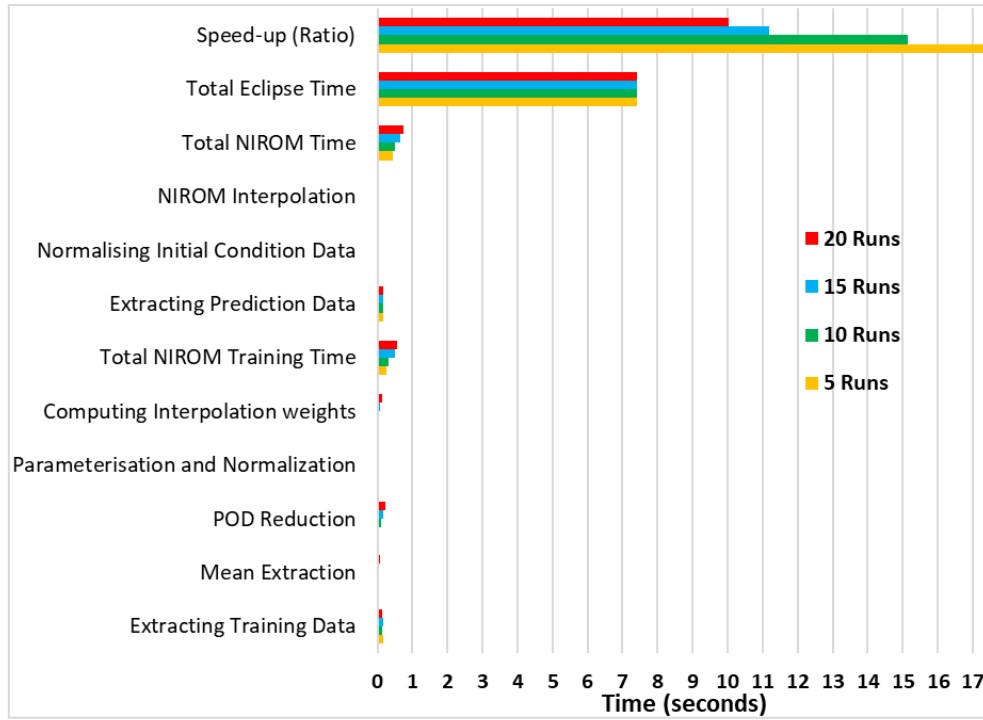


Figure 6.29: Summary of the time taken to implement the constant BHP cases of the synthetic gas reservoir. NIROM speedups of between 10 and 17 times the speed of ECLIPSE are achievable.

Table 6.3: Summary of the time taken to implement the constant porosity and initialisation by enumeration for constant production rate cases of the synthetic gas reservoir. These values correspond to Figure 6.30 and Figure 6.31.

Production Control	Porosity				Enumeration (Rate)			
Number of Training Runs	5	10	15	20	5	10	15	20
Extracting Training Data	0.14	0.13	0.13	0.13	0.1	0.1	0.12	0.12
Mean Extraction	0.02	0.03	0.05	0.07	0.02	0.03	0.04	0.07
POD Reduction	0.04	0.09	0.17	0.24	0.03	0.07	0.12	0.19
Parameterisation and Normalization	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Computing Interpolation Weights	0.02	0.03	0.04	0.06	0.02	0.02	0.03	0.05
Total NIROM Training Time	0.24	0.29	0.4	0.51	0.17	0.23	0.33	0.44
Extracting Prediction Data	0.14	0.14	0.13	0.14	0.1	0.11	0.11	0.11
Normalising Initial Condition Data	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001
NIROM Interpolation	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Total NIROM Time	0.39	0.44	0.54	0.67	0.28	0.35	0.44	0.55
Total ECLIPSE Time	5.76	5.76	5.76	5.76	5.18	5.18	5.18	5.18
Speed-up (Ratio)	14.7	13.1	10.6	8.65	18.4	14.9	11.8	9.42

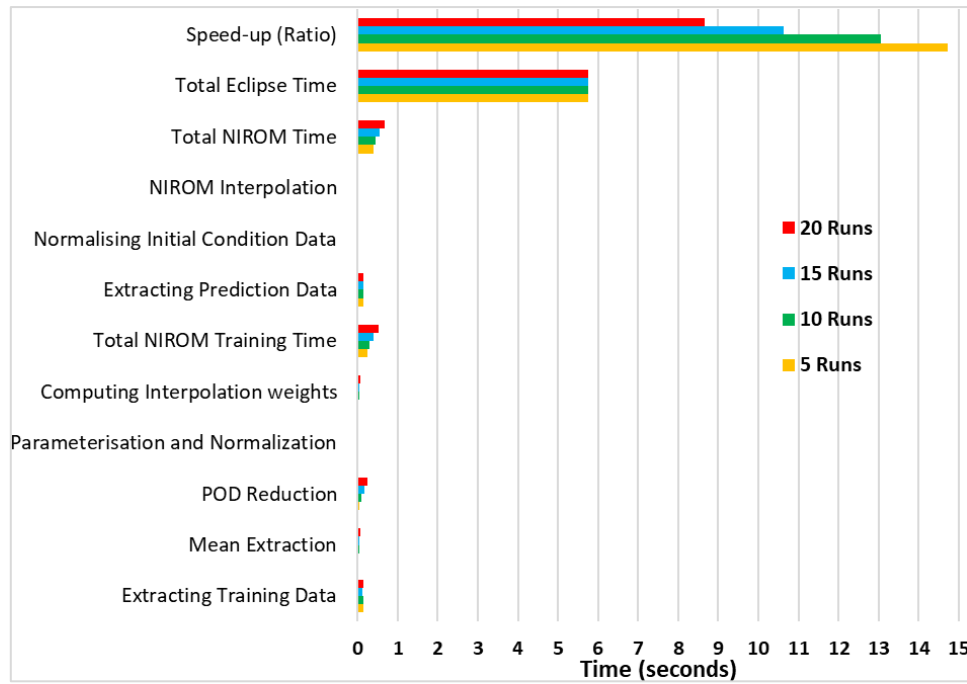


Figure 6.30: Summary of the time taken to implement the constant porosity cases of the synthetic gas reservoir. NIROM speedup of between 8 and 14 times ECLIPSE are achievable.

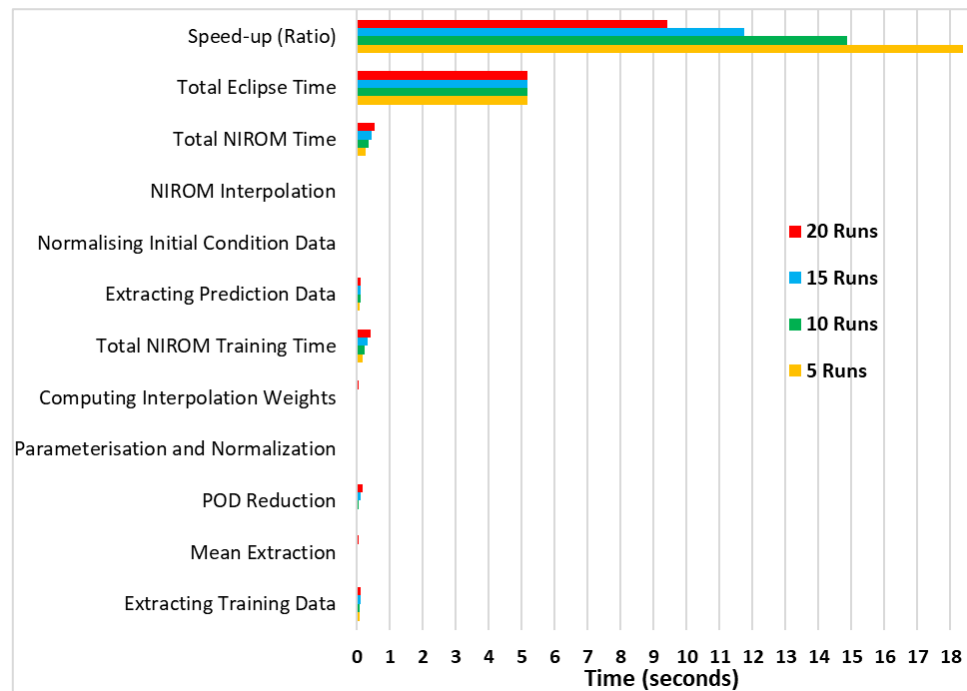


Figure 6.31: Summary of the time taken to implement initialisation by enumeration for constant production rate cases of the synthetic gas reservoir. NIROM speedups of between 9 and 18 times ECLIPSE are achievable.

6.6.2 MODIFIED NORNE GAS FIELD AND WATER CONING GAS RESERVOIR

From Table 6.4, NIROM is 16 to 30 times faster than ECLIPSE for the heterogeneous Norne field model and up to 2 orders of magnitude faster for the water coning model. The breakdown of the time that NIROM spent in executing simulations for these models also confirms that the time required for POD reduction increases as the number of training runs increases. The time spent in POD reduction

represented the most significant individual factor influencing the speedup that is obtainable with the NIROM. Besides the time spent for POD reduction, most of the time used in NIROM simulations were spent in extracting training data from ECLIPSE into the NIROM framework and also in extracting prediction data into the NIROM (Figure 6.32 and Figure 6.33). For the modified Norne field the actual time spent on NIROM's interpolation is 0.03 second, while NIROM interpolates the water coning model simulations in 0.05 to 0.07 second. This means that apart from the time spent on training the NIROM, up to 4 orders of magnitude speedup is possible with NIROM.

Table 6.4: Breakdown of the average time taken to implement constant parameter simulations of the modified Norne field and the water coning reservoir model. POD reduction time increases as the number of training runs increase. NIROM achieves at least an order of magnitude speedup for both models.

Reservoir Model	Modified Norne Field			Water Coning Reservoir		
Number of Training Runs	5	10	20	5	10	20
Extracting Training Data	0.89	0.97	0.94	0.18	0.18	0.19
Mean Extraction	0.35	0.66	1.21	0.11	0.21	0.35
POD Reduction	1.12	1.79	3.52	0.21	0.45	0.70
Parameterisation and Normalization	0.02	0.03	0.06	0.01	0.02	0.03
Computing Interpolation Weights	0.02	0.04	0.08	0.12	0.47	1.09
Total NIROM Training Time	2.40	3.48	5.81	0.63	1.33	2.36
Extracting Prediction Data	1.18	1.24	1.16	0.22	0.18	0.18
Normalising Initial Condition Data	0.003	0.003	0.003	0.003	0.004	0.003
NIROM Interpolation	0.03	0.03	0.03	0.05	0.07	0.07
Total NIROM Time	3.62	4.76	7.0	0.90	1.58	2.61
Total ECLIPSE Time	111.4	111.4	111.4	254	254	254
Speed-up (Ratio)	30.8	23.4	15.9	281.4	160.4	97.2

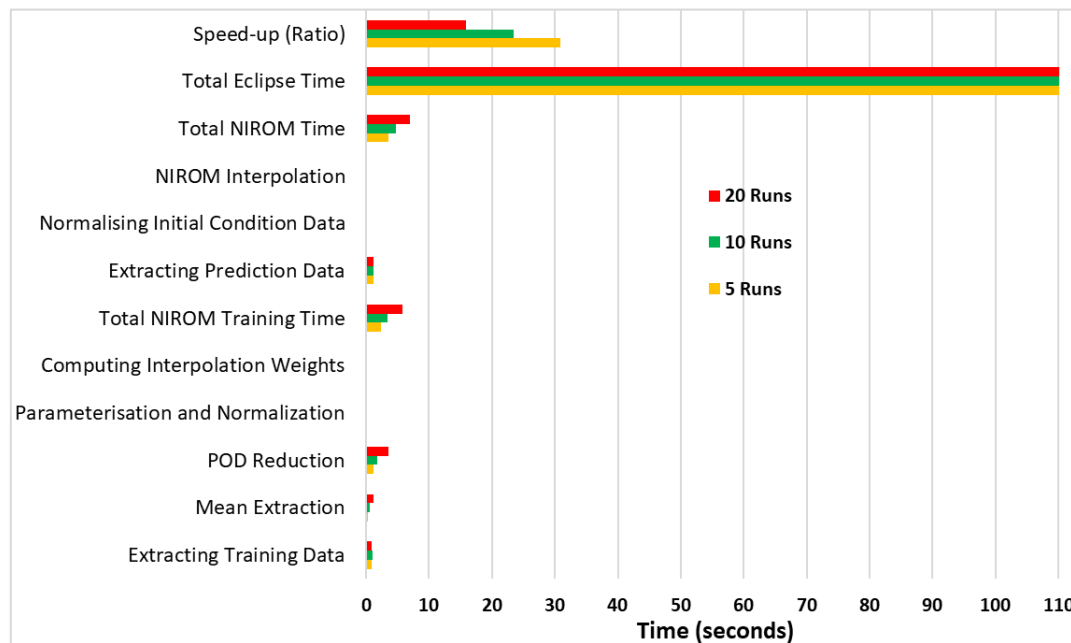


Figure 6.32: A breakdown of the average time taken to implement NIROM simulations of the modified Norne field. Time spent on POD reduction is proportional to the number of training runs. NIROM is 15 to 30 times faster than ECLIPSE.

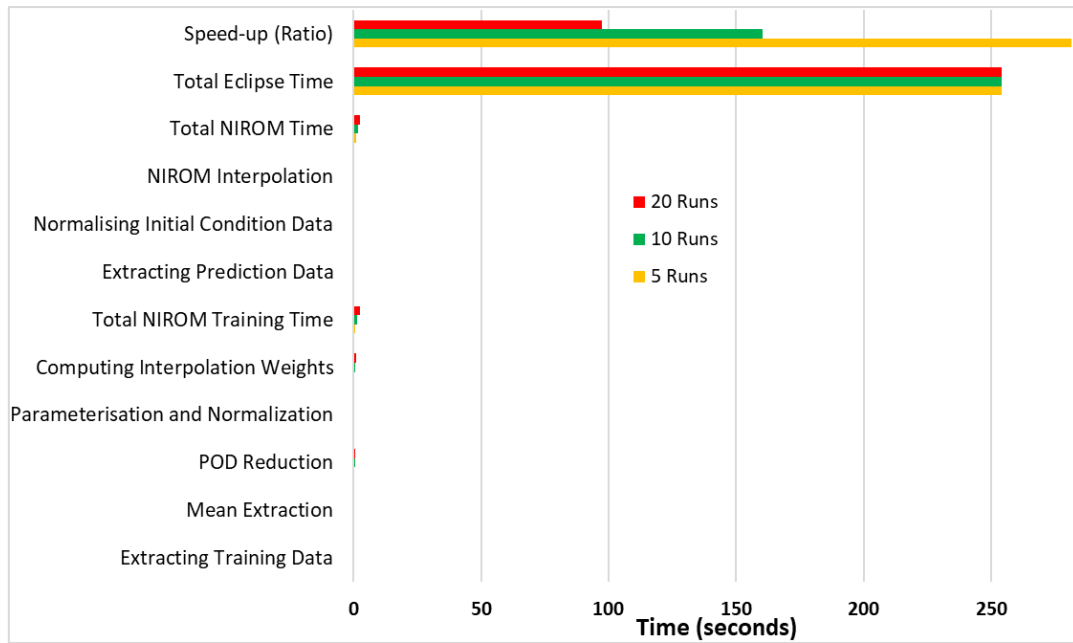


Figure 6.33: Breakdown of the average time taken to implement NIROM for the water coning reservoir. Time spent in actual NIROM prediction/ interpolation is negligible, leading to significant speedups of between 97 and 281 times ECLIPSE's run time.

6.6.3 FIELD X

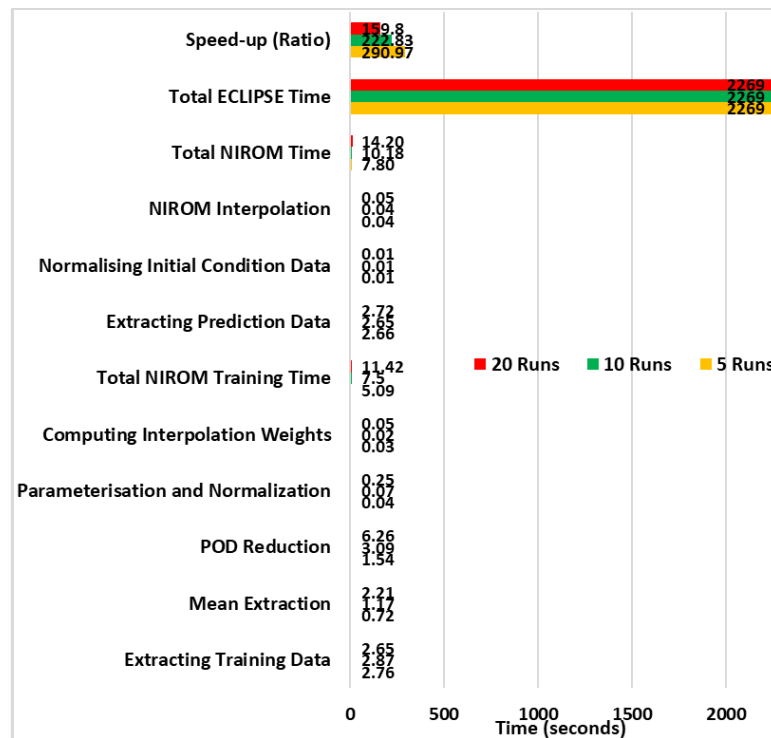


Figure 6.34: The breakdown of the average time spent during NIROM predictions of constant rate cases of Field X. NIROM is between 160 and 290 times faster than ECLIPSE in simulating this real gas reservoir case.

Figure 6.34 shows the breakdown of the time taken to run NIROM for the constant production rate cases of Field X. From Figure 6.34, it is seen that whereas it takes ECLIPSE over 30 minutes to complete a full simulation run of Field X, NIROM predictions can be implemented in under 15 seconds for the

constant rate predictions. Consistent with the findings for the other reservoirs that were evaluated, most of the time spent in implementing NIROM is spent on POD decomposition, followed by data extraction from ECLIPSE output files.

6.6.4. ANALYSIS OF SIMULATION SPEEDUP

Table 6.5 shows a summary of NIROM's speed-up in all the reservoir cases that were considered in this chapter. From Table 6.5 it can be seen that NIROM had higher speed-ups for the predictions in all the other gas reservoirs than the synthetic gas reservoir. These other gas reservoirs have more realistic and complex production scenarios such as being heterogeneous or experiencing water coning, scenarios that are often time-consuming to model in ECLIPSE. Therefore with an average NIROM runtime for these more complex reservoir models, the overall NIROM simulation speedup will be higher because the corresponding ECLIPSE runtime is higher than average.

Table 6.5: NIROM's simulation speed-up for the constant parameter test cases considered in this chapter.

Reservoir Model	Synthetic Gas Reservoir								
Production Control	Rate Controlled Production					BHP Limit Controlled Production			
Number of Training Runs	5	10	15	20		5	10	15	20
Speed-up (Ratio)	13.1	11.7	9.53	8.19		17.4	15.2	11.2	10
Production Control	Porosity					Enumeration (Rate)			
Number of Training Runs	5	10	15	20		5	10	15	20
Speed-up (Ratio)	14.7	13.1	10.6	8.65		18.4	14.9	11.8	9.42
Reservoir Model	Modified Norne Field		Water Coning Reservoir			Field X			
Number of Training Runs	5	10	20	5	10	20	5	10	20
Speed-up (Ratio)	30.8	23.4	15.9	281.4	160.4	97.2	291.0	222.8	159.8

Another factor that affects NIROM's speed-up is the amount of data that was available for training/building the NIROM. For the three gas reservoirs, besides the synthetic reservoir, (Table 6.4 and Figure 6.34), as well as in other cases that would be seen in chapters 7 and 8, it is observed that NIROM's total runtime increased as the amount of simulation data in the reservoir model increased. The amount of data that is available for training the NIROM is in turn dependent on the number of grid cells in the reservoir model as well as the frequency at which the reservoir's simulation solution is outputted from the conventional simulator (report steps). In chapter 3, it was shown that the total amount of data available for training the NIROM is a product of the number of grid cells, number of report steps and number of training runs ($nx \times nt \times ns$), Equation 3.8. Hence NIROM's total training time for Field X with 125000 grid cells was 5 seconds while its total training time for the Norne field and the water coning field (with 44000 and 12000 grid cells respectively) were 2.4 and 0.63 seconds for the cases with 5 training runs respectively.

NIROM is written in an interpreted language whereas ECLIPSE is in a compiled language and has been improved through several iterations since its inception. Hence NIROM has room for more computational speed up if it is implemented in a compiled language and when its training time is discounted from its total runtime. Other specialised techniques for speeding up the NIROM could include, implementing NIROM with hardware (or cloud resources) that offer greater computational capacity such as GPUs in addition to investigating other methods for fast data reduction or matrix inversion. Some less specialised approaches to speeding up NIROM is to build it with the least number of report steps that are sufficient for executing a particular study/ project and with only the active grid cells (data points) of a reservoir. If possible, NIROM can be implemented for smaller portions of interest in a gas reservoir model (sector models), this would be faster to implement than when NIROM is implemented for the full reservoir model.

The optimum NIROM implementation would require that the engineer strikes a balance between the amount of data that goes into the NIROM (this affects simulation runtime) and the NIROM's accuracy, since this depends on the number of training runs for constant parameters cases.

SUMMARY

In this chapter, the POD-RBF NIROM approach has been investigated for the simulation of grid distributed reservoir properties (pressure and saturation) and other well or field specific production data, for constant parameter prediction cases. The accuracy of NIROM predictions was compared to similar simulation runs from ECLIPSE for 4 gas reservoir models: a homogeneous synthetic gas reservoir, a heterogeneous gas reservoir that is similar to the Norne oil field in geology and structure, a water coning gas reservoir and a real gas Field X from Woodside Petroleum. NIROM was evaluated for predicting the performance of these reservoirs for various scenarios of different production rates, BHP limits, porosity values, production by enumeration and equilibration. This chapter also provided insight into the extent of computational speedup that is achievable with NIROM for predictions of the mentioned gas reservoirs' performances.

Through the different examples that have been evaluated in this chapter, it has been shown that NIROM is at least an order of magnitude faster than ECLIPSE for constant parameter predictions. It was also shown that NIROM gives reasonable accuracy for gas reservoir predictions of unseen or new simulation cases, provided that NIROM has been trained within the range of parameters and duration that bounds the prediction runs. Simulation accuracy of more than 99% is obtainable with a computational speedup of up to 2 orders of magnitude than ECLIPSE. It was further shown that NIROM can be constructed with satisfactory accuracy with as few as 5 training runs. Increasing the number of

training runs could improve NIROM's predictions, but this comes at the expense of the simulation speedup. For predicting the production variables of the gas reservoirs that were considered, it was seen that the number of POD basis functions used to construct the NIROM is crucial to obtaining accurate predictions. Predicting individual well or field specific production data such as production rates and cumulative productions, especially for production data that experience negligible change with time, required the most number of POD basis functions for satisfactory results. Nonetheless, the full rank number of POD basis functions can be used NIROM's simulation of the production data since they are fewer than the grid distributed data and would have less impact on NIROM's overall computational time, especially for the realistic reservoirs.

Table 6.6 is a summary of the cases that have been evaluated in this chapter, as well as the relevant findings.

Table 6.6: A summary of the prediction cases and key findings in this chapter.

Cases Considered	Findings
• Constant Parameter Prediction Examples • 4 gas reservoirs were evaluated, including 2 real fields ○ homogeneous synthetic gas reservoir ○ modified heterogeneous Norne gas field ○ homogeneous water coning gas reservoir ○ Real gas field X from Woodside • Parameters evaluated ○ production rate ○ BHP limit ○ Porosity ○ Enumeration	Key Findings
	• NIROM performs best when the training runs are within a similar range of parameters and production duration as the prediction runs • With increasing number of training runs, NIROM gives better predictions within the range of training runs and worse predictions outside the range of training runs • 10 training runs is a good balance between accuracy and speed of NIROM performance
	General Findings
	• Demonstrated the NIROM's predictive capacity for gas field modelling, based on different constant production constraints • 99.9% energy criterion is sufficient for determining the number of POD basis functions that are needed for NIROM predictions of spatially distributed reservoir properties • NIROM with the full rank POD basis functions should be used for predicting time series production data for the best results • NIROM gives up to 2 orders of magnitude speed-up for gas performance modelling • NIROM does not cope well with simulation cases that experience little to no property changes over time

CHAPTER 7: VARYING PARAMETER PREDICTIONS - RESULTS AND DISCUSSION

In this chapter, the POD-RBF NIROM has been applied to modelling the performance of more realistic time-varying production scenarios. All production wells were initially set to rate controlled production with a bottom hole pressure (BHP) or tubing head pressure (THP) limit. In other words, rate controlled production switches to BHP or THP control when ECLIPSE estimates that a well is unable to meet the intended production rate.

However, both the well production rates and well pressure limit are assigned as the parameters for training and prediction runs for the entire production lifecycle. Two types of varying production profiles are considered for the 20 years production lifecycle of the test gas fields (synthetic gas reservoir, modified Norne gas field and Field X). One production profile has seasonally varying rates, these are divided into at least 240 report steps (30 days intervals). The other varying production profile has less variability and is designed to have 75 report steps (100 day interval), with the production rate changing after at least 1000 days, approximately triennially. The production profiles for the prediction and training runs of all test examples are provided accordingly.

POD-RBF NIROM was used to model the pressure and gas saturation distribution in all the active grid cells of the three gas reservoirs, up to 125,000 grid cells for Field X, as well as the corresponding production variables, i.e. gas and water production rates, cumulative production and ratios. For ease and brevity of showing the results of NIROM predictions of the grid distributed reservoir properties (pressure and gas saturation), the average of the property value across all grid cells of the reservoir have been plotted at each report step. The optimal number of POD basis functions (the equivalent of 99.9% energy criterion, 3.5) have been used to generate all the examples shown in this chapter and the linear RBF was used to learn and interpolate the NIROM's predictions. NIROM's performance for predicting new unseen cases was then compared against similar simulation solutions from ECLIPSE, for both accuracy (errors, absolute errors or % errors and RMSEs) and speed.

7.1 SYNTHETIC GAS RESERVOIR MODEL

7.1.1 SEASONALLY VARYING PRODUCTION PROFILE

This example consists of mixed production control cases. For instance in Figure 7.1, two of the six predictions (predictions 1 and 2) have experienced a switch in their production control before the end of the simulation lifecycle. The other 4 prediction cases are produced by varying only their production rates throughout the simulation lifecycle. This case has 252 report steps at intervals of 30 days and NIROM simulations have been implemented with 5, 10, and 20 training runs (Figure 7.4).

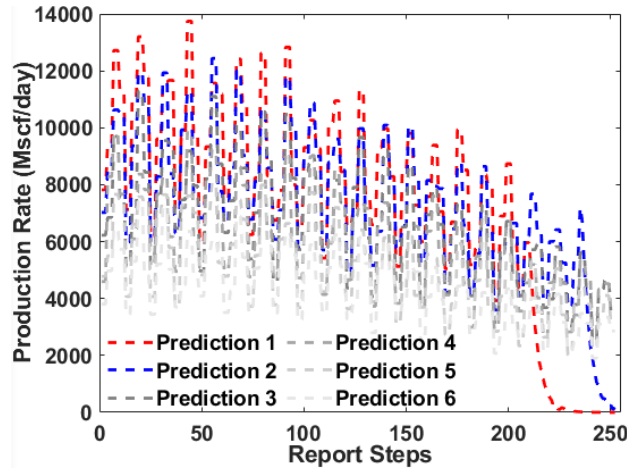


Figure 7.1: A seasonally varying production profile with mixed production control. In this example, predictions 1 and 2 switched to BHP limit control before the end of the production lifecycle. The other predictions are on varying rate control throughout the production duration. NIROM simulations of this case are shown in Figure 7.2 and Figure 7.7, below. The corresponding training runs for this prediction example are in Figure 7.4.

The results in Figure 7.2 show that NIROM can satisfactorily predict the pattern of pressure distribution in the synthetic gas reservoir with a varying production profile and mixed production controls (Figure 7.1). The NIROM trained with 5, 10 and 20 training runs (Figure 7.4) gives varying degree of accuracy of predictions. Cases with mixed production controls generally experienced the most prediction errors, especially at the time interval where the production control was switched (i.e. predictions 1 and 2 in Figure 7.2). Nevertheless, NIROM gives satisfactory predictions with most errors lying within the vicinity of $\pm 10\%$. NIROM predictions with 10 training runs gave more accurate predictions than when NIROM was trained with 5 or 20 training runs. It is suspected that the accuracy of the NIROM's performance in these cases is linked to how much of the varying prediction profiles were captured within the range of training runs' profiles. This was checked by generating another set of training runs with non-mixed production controls (Figure 7.5), i.e. varying production rates only, and using them to predict the cases in Figure 7.1 (with mixed controls). In this case, the resulting predictions are shown in Figure 7.6. The NIROM predictions in Figure 7.6 exhibit significantly larger errors than those in Figure 7.2, since they were trained with non-mixed controls and did not learn the behaviour of the mixed control systems of the prediction production profiles. This emphasises the need for training NIROM with the range of parameters and similar scenarios that would be considered during its prediction runs.

Creating the right range of training parameters or scenarios can be tricky to achieve in the varying parameter cases since their production profiles are generated based on different multipliers with a random combination of rates and durations. Nevertheless, care should be taken to ensure that NIROM is trained on the most suitable and closest range of training runs, for satisfactory predictions.

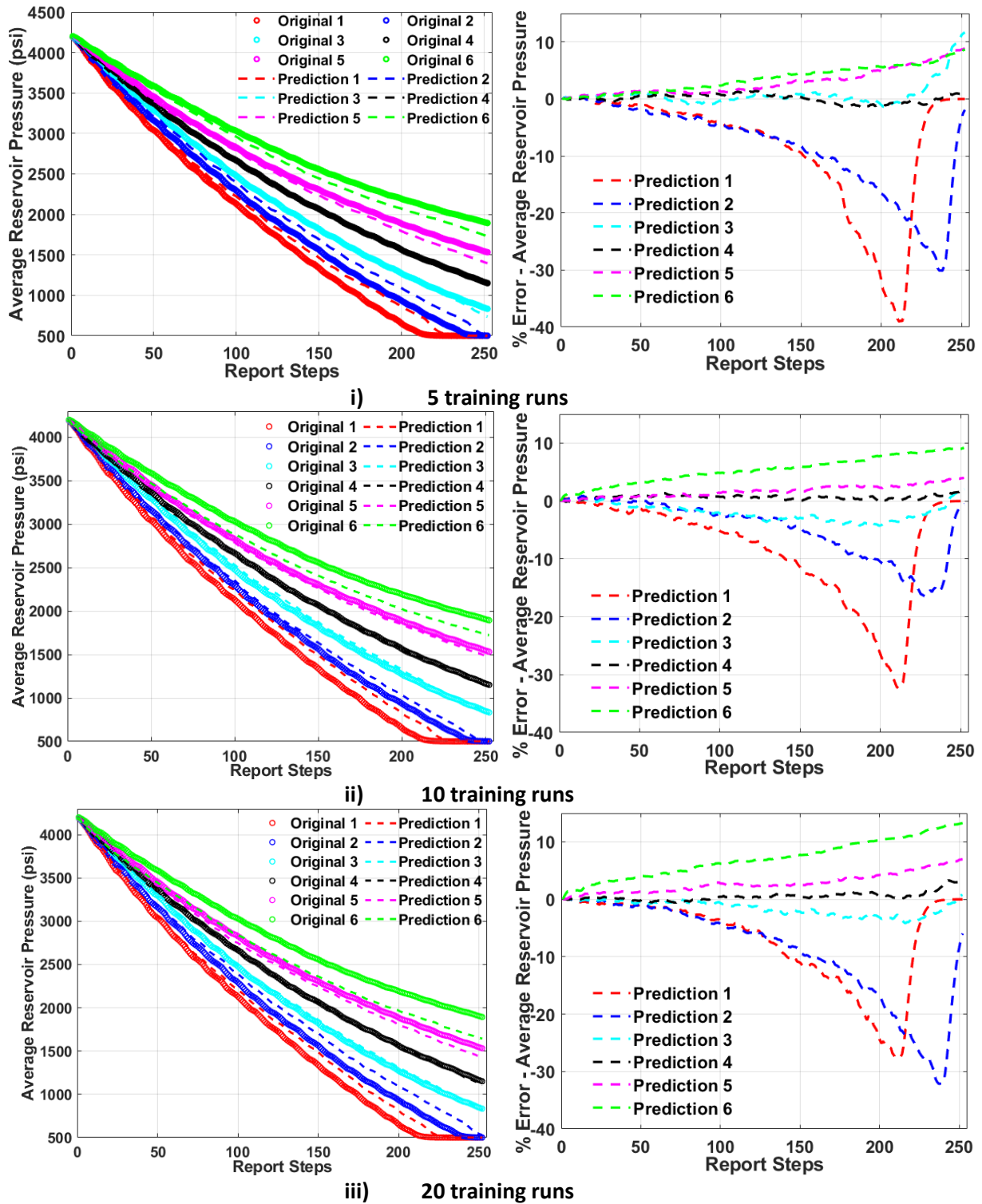


Figure 7.2: Average pressure distribution in the synthetic gas reservoir when produced with the varying seasonally production profile in Figure 7.1. Prediction with 10 training runs gives the least prediction errors overall. Nevertheless, NIROM prediction errors are mostly less than 10%. The corresponding training runs for these predictions are in Figure 7.4. NIROM's predictions 1 and 2 are the least accurate because they have mixed production controls. The error in these two cases were especially worse during the interval when their production control changed from gas rate to BHP limit.

Figure 7.3 shows NIROM's prediction of the normalised pressure distribution in all the grid cells of the synthetic gas reservoir for predictions 2 and 4 and confirms that NIROM's predictions are satisfactory. NIROM's prediction error is less than 2% even for prediction 2 that had mixed production controls.

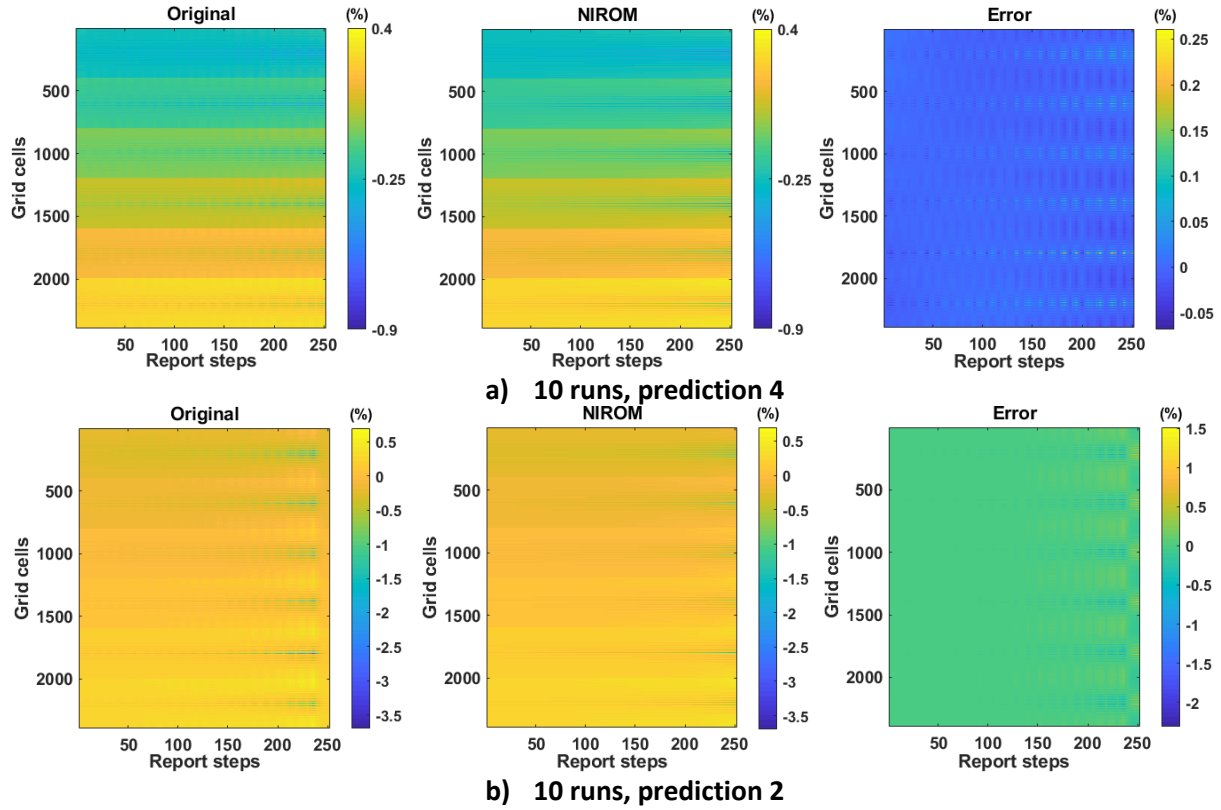
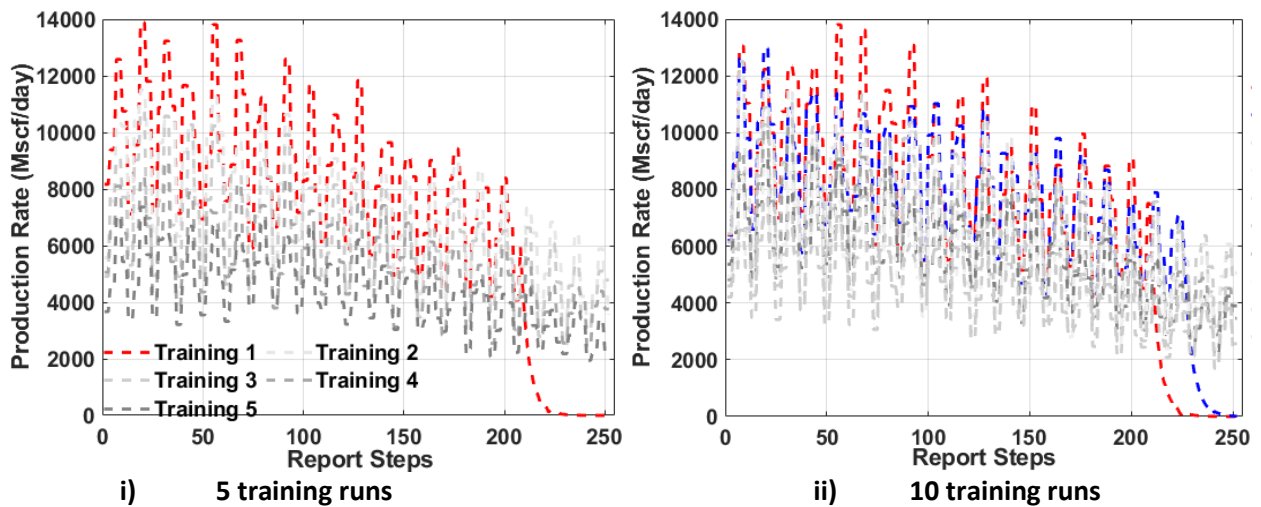


Figure 7.3: NIROM's predictions 2 and 4 of the pressure distribution in all the grid cells of the synthetic gas reservoir when produced by the seasonally varying production profile and mixed production controls of Figure 7.1. These pressure values have been normalised to zero mean. The error in these NIROM predictions is mostly negligible even for prediction 2 after its production control switched to BHP limit control.



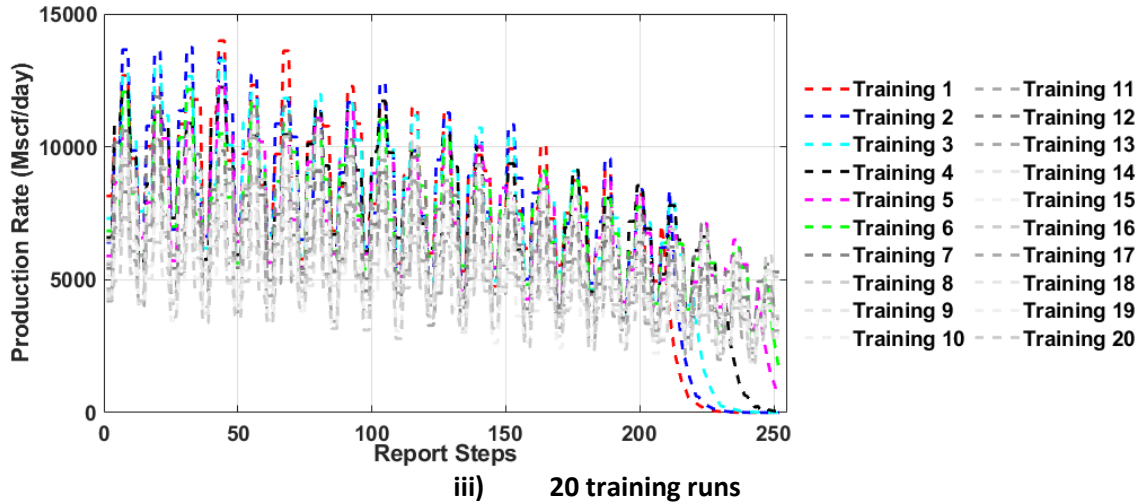


Figure 7.4: Production profiles for 5, 10 and 20 training runs corresponding to the results in Figure 7.2. Each of these training runs has at least one case of mixed production control, hence their suitability for the prediction cases in Figure 7.1, as seen in the results in Figure 7.2 and Figure 7.3.

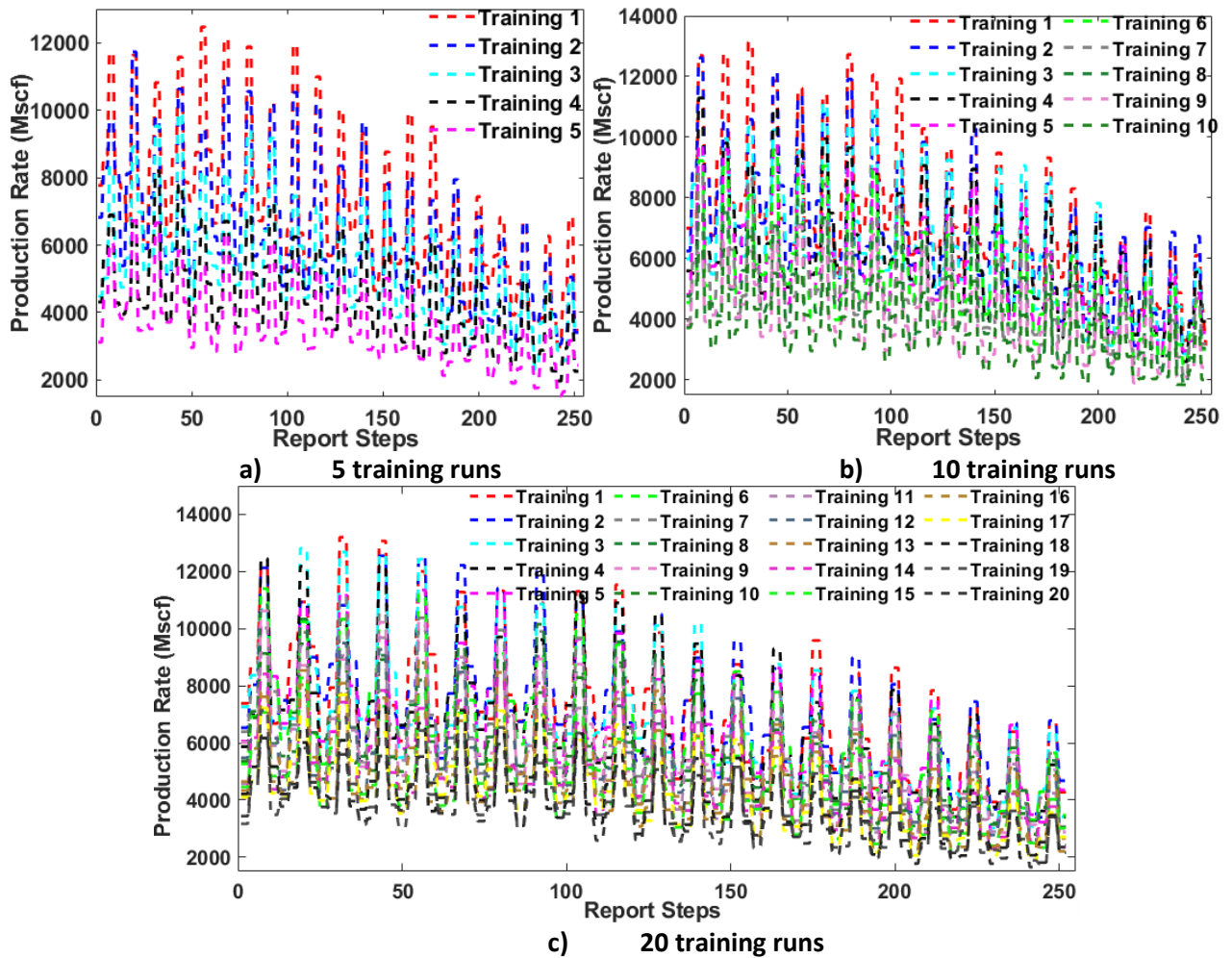


Figure 7.5: Production profiles of 5, 10 and 20 training runs corresponding to the results in Figure 7.6. These training runs are all varying production rates with no mixed production control, hence their erroneous predictions (Figure 7.6) of predictions 1 and 2 of Figure 7.1.

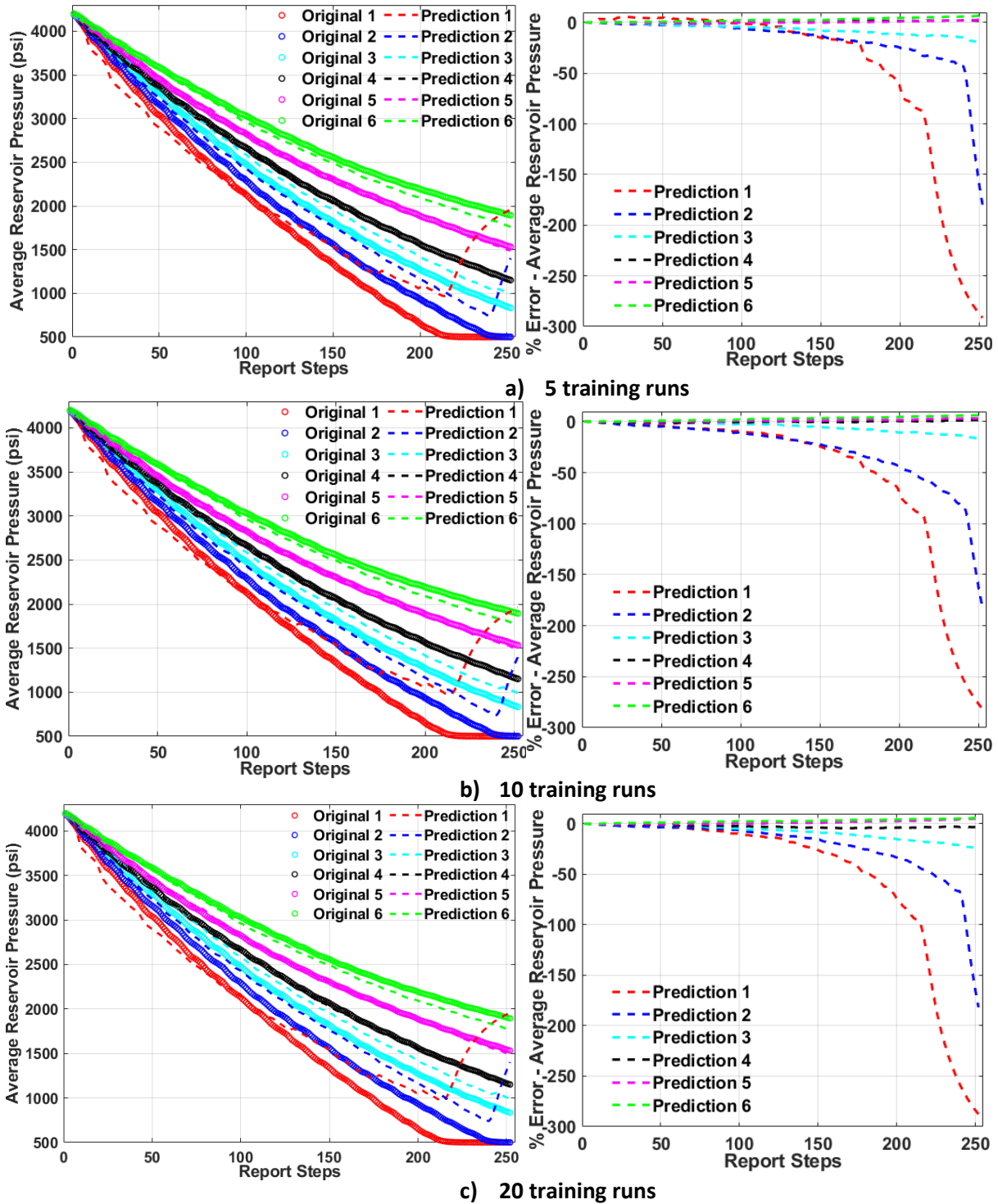


Figure 7.6: NIROM prediction of the average pressure distribution in the synthetic gas reservoir when NIROM was trained with non-mixed production controls (Figure 7.5), whereas the prediction cases have mixed production controls (Figure 7.1). Hence the NIROM has significantly over-predicted the average reservoir pressure during the BHP limit controlled durations of predictions 1 and 2. This emphasises the need for training the NIROM within similar conditions to the cases it is required to make predictions for.

With the minimum number of POD basis function (1 POD), NIROM gives satisfactory predictions of the production variables that are investigated for the synthetic gas reservoir (Figure 7.7). Figure 7.7 shows the results of NIROM's prediction 1 trained with 5, 10 and 20 training runs. In Figure 7.7, there is almost no difference in NIROM predictions with 5, 10 or 20 training runs.

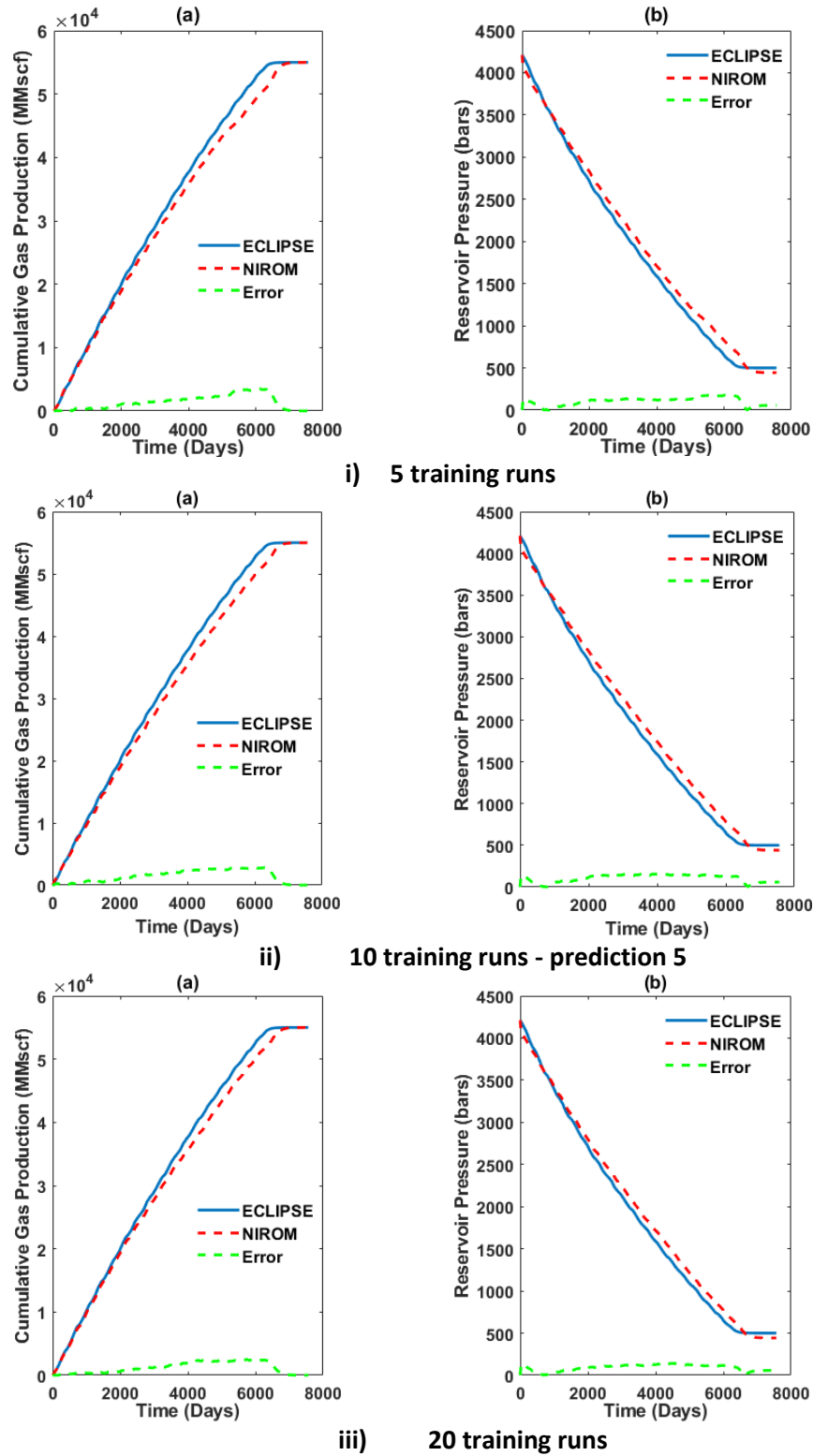


Figure 7.7: NIROM prediction of the production data (cumulative production and reservoir pressure) for prediction 1. Using the optimum number of POD basis functions (1 POD basis function), NIROM satisfactorily identifies the intervals when the production control switched to the BHP limit. Increasing the number of training runs did not have a significant effect on the accuracy of NIROM's prediction.

7.1.2 TRIENNIALLY VARYING PRODUCTION PROFILE

Here the production profile is also time-varying like in the seasonally varying production profile, however, there are fewer variations in the production parameters within the 20 years production lifecycle. This example demonstrates NIROM's capability for modelling a different variant of the varying parameter production scenario. Figure 7.8 shows the production profiles for the NIROM predictions in Figure 7.9. The 5, 10 and 20 training runs used to train NIROM for the predictions in Figure 7.9 are shown in Figure 7.10. NIROM predictions in Figure 7.9 have errors of less than 10% for all predictions except prediction 1. From Figure 7.8, it can be seen that the production profile for prediction 1 has production rates that are above the range of the production rates of the other predictions. For instance for the 10th and 25th report steps and after the 40th report step, prediction 1 was produced at or above 7500Mscf/day whereas the other prediction cases were below 7000Mscf/day throughout. In comparison to the production profiles for the training runs (Figure 7.10), it can also be seen that not many of the training runs had their production rates above 7500Mscf/day. Since the range of training production profiles was mostly below 7500Mscf/day, the high errors in prediction 1 can be linked to its production profile being, for the most part, outside the range of training parameter values. This reiterates the necessity of training the NIROM within the range of the prediction runs.

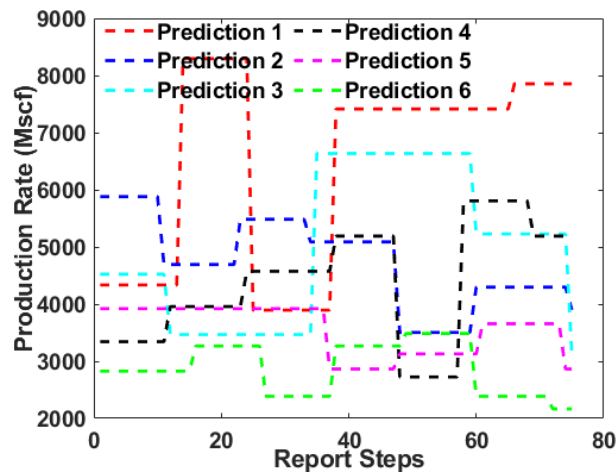


Figure 7.8: Production profile for the triennially varying parameter case of the synthetic gas reservoir. The frequency at which the production rates vary is after 1000 – 1500 days (10 - 15 report steps) interval.

Furthermore, in this triennially varying parameter example, training NIROM with 10 training runs gave the best results for predictions 2 to 6 (Figure 7.9ii). This is probably more associated with the even spread of the training data of the 10 training runs case, than with the number of training runs. Figure 7.9 also shows that NIROM recognized the intervals when there were changes in production rate, NIROM reflected these changes by satisfactorily mimicking the pattern of the solutions from ECLIPSE.

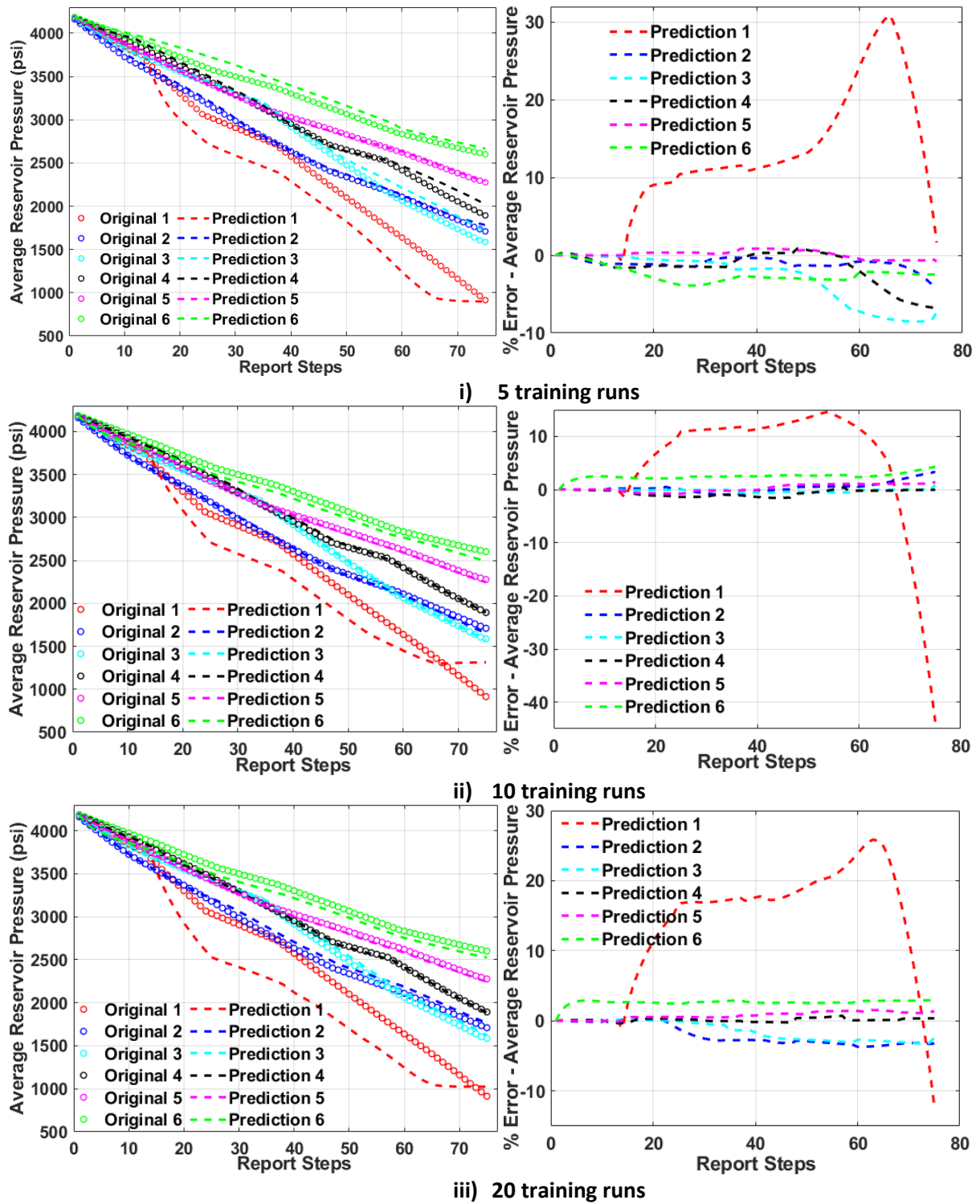


Figure 7.9: NIROM predictions of the average pressure distribution in the synthetic gas reservoir produced with a triennially varying production profile. NIROM predictions have less than $\pm 10\%$ error (mostly less than $\pm 5\%$) for all predictions except prediction 1.

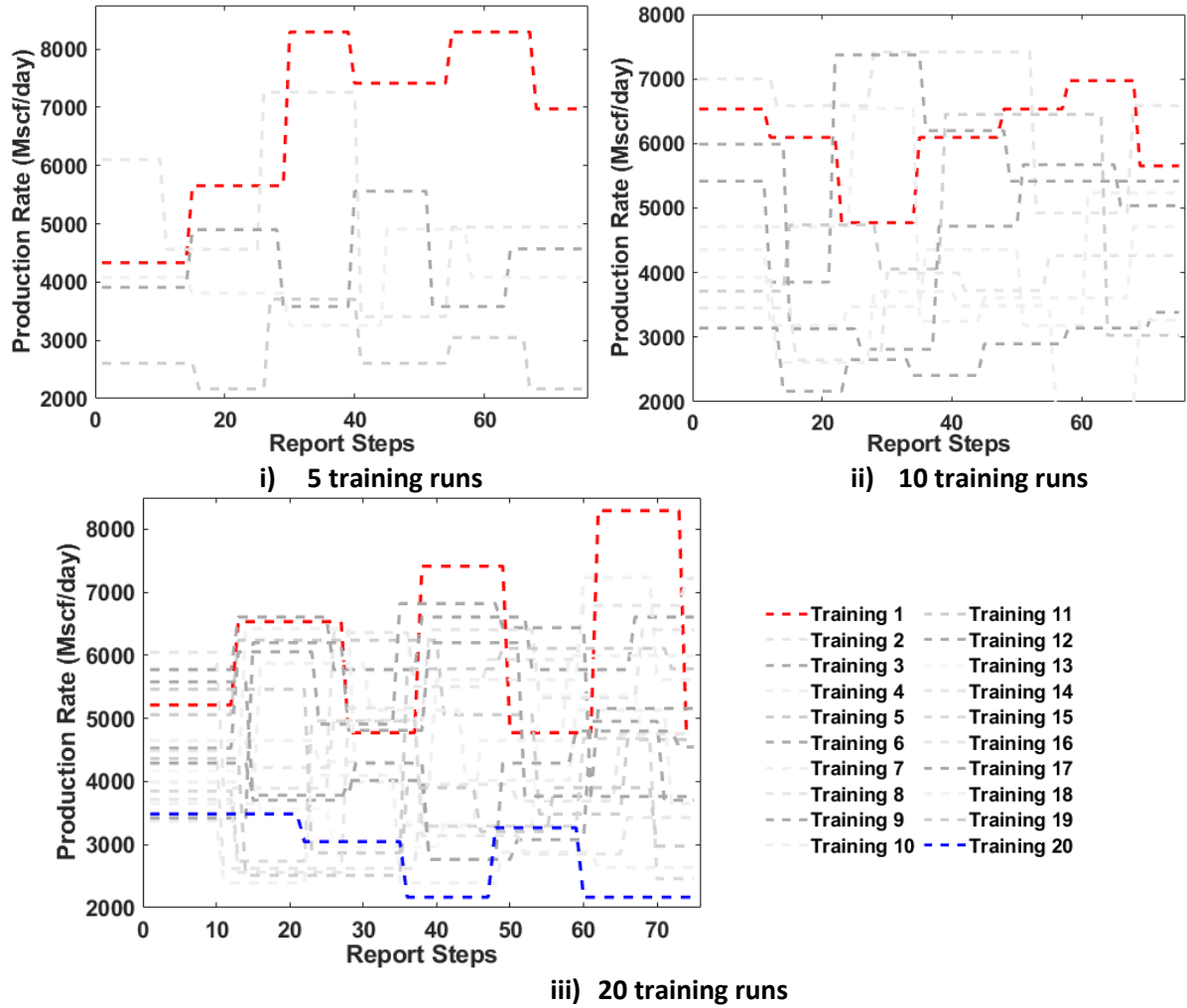


Figure 7.10: Production profiles of the simulation runs used to train NIROM for the triennially varying production rate cases. The training runs are spread between 2000 and 8000Mscf/day.

Having noted that NIROM gave better prediction results for predictions 2 to 6, Figure 7.11 evaluates NIROM prediction 1's performance for the production variables that were modelled for the synthetic gas reservoir cases, the cumulative gas production and reservoir pressure are presented. In Figure 7.11, NIROM starts off giving good predictions of the reservoir's performance before the first change in production rate (around 1 to 1300 days), but afterwards NIROM predictions are not accurate. Although the NIROM tends to identify the intervals when there was a change in production rate and tries to mimic the pattern of the reservoir's performance at the intervals when the rates varied. NIROM's predictions did not improve as the number of training runs increased.

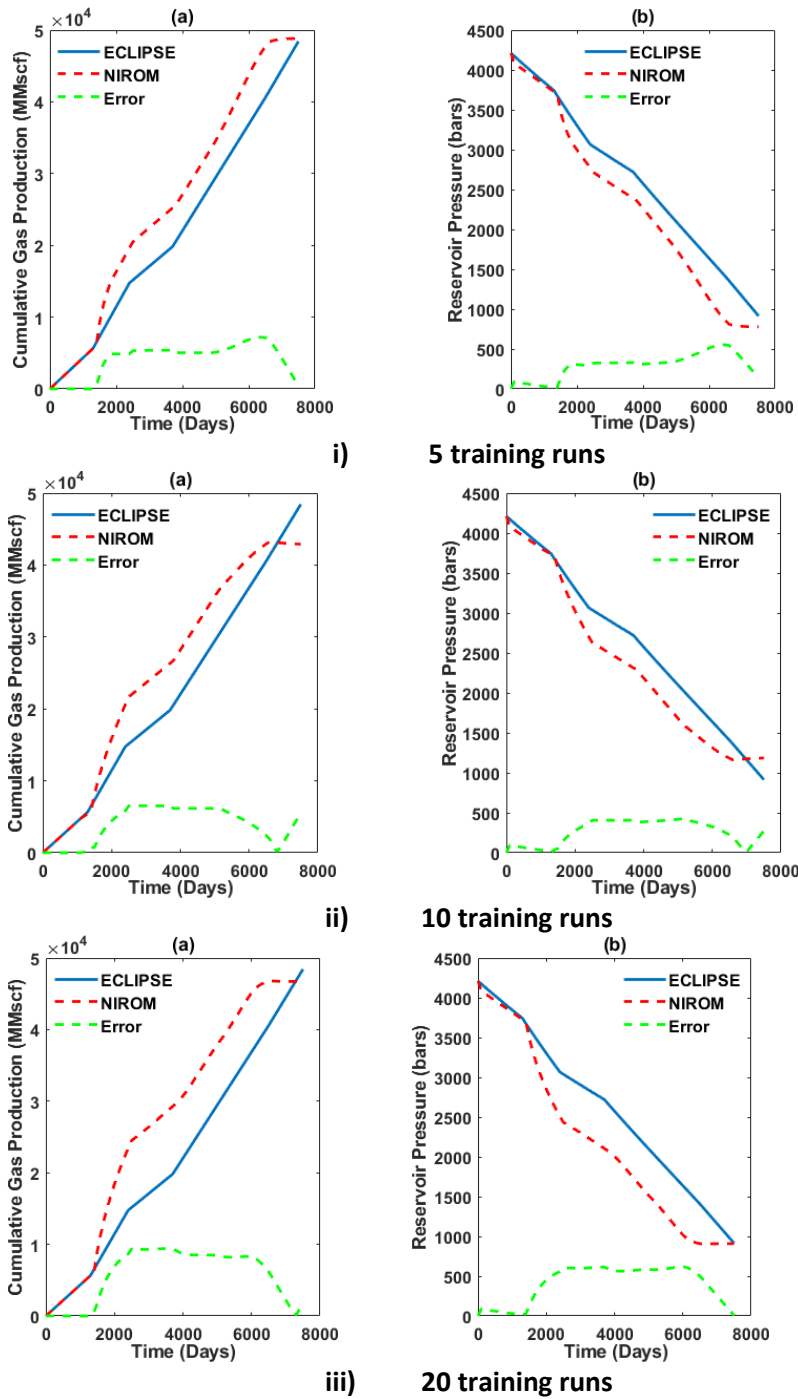


Figure 7.11: NIROM prediction of the cumulative gas production and reservoir pressure for the triennially varying prediction 1. NIROM was able to mimic the pattern of the reservoir's performance and had the most errors when it was trained with 5 training runs.

7.2 MODIFIED NORNE GAS FIELD

Given that the seasonally varying production profile is similar and more representative to the production profiles that are seen in the petroleum industry, the Norne field is evaluated for only the seasonally varying mixed production control scenarios. Figure 7.12 shows the seasonally varying production profiles of the prediction cases that were considered for evaluating the NIROM for the heterogeneous Norne field.

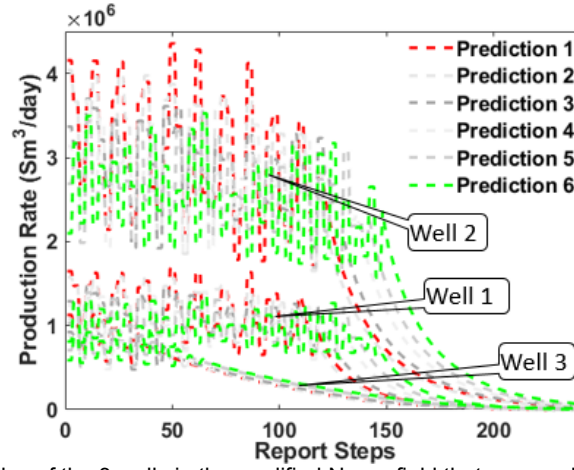


Figure 7.12: Production profiles of the 3 wells in the modified Norne field that are used to test NIROM's prediction of varying production profiles with mixed boundary conditions. In all 3 wells, production control (boundary condition) was switched to BHP limit control after some period of being produced at varying production rates.

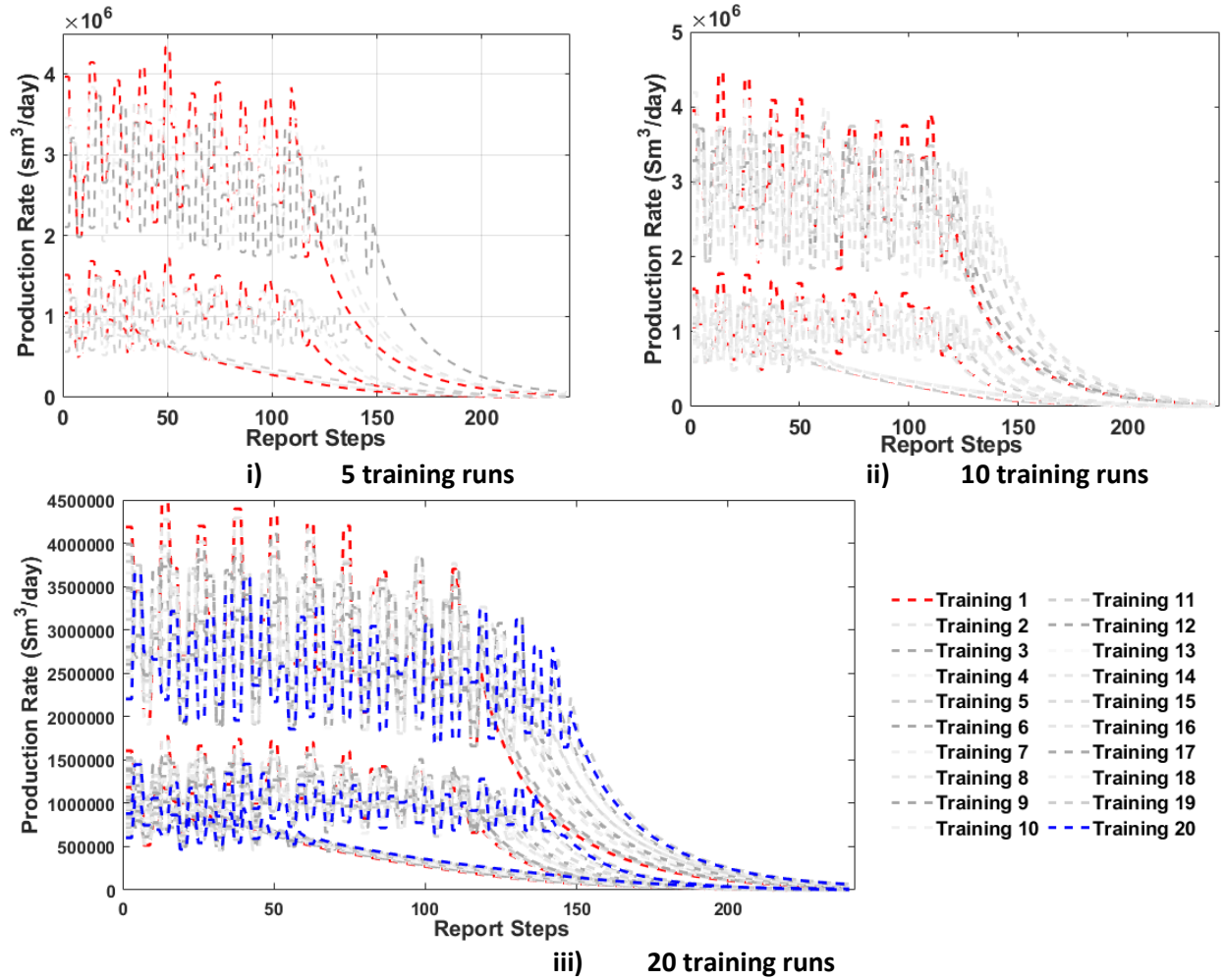


Figure 7.13: Production profiles for the training runs used to train NIROM for the results in Figure 7.14 to Figure 7.16. These training runs have mixed production profiles like the prediction runs in Figure 7.12.

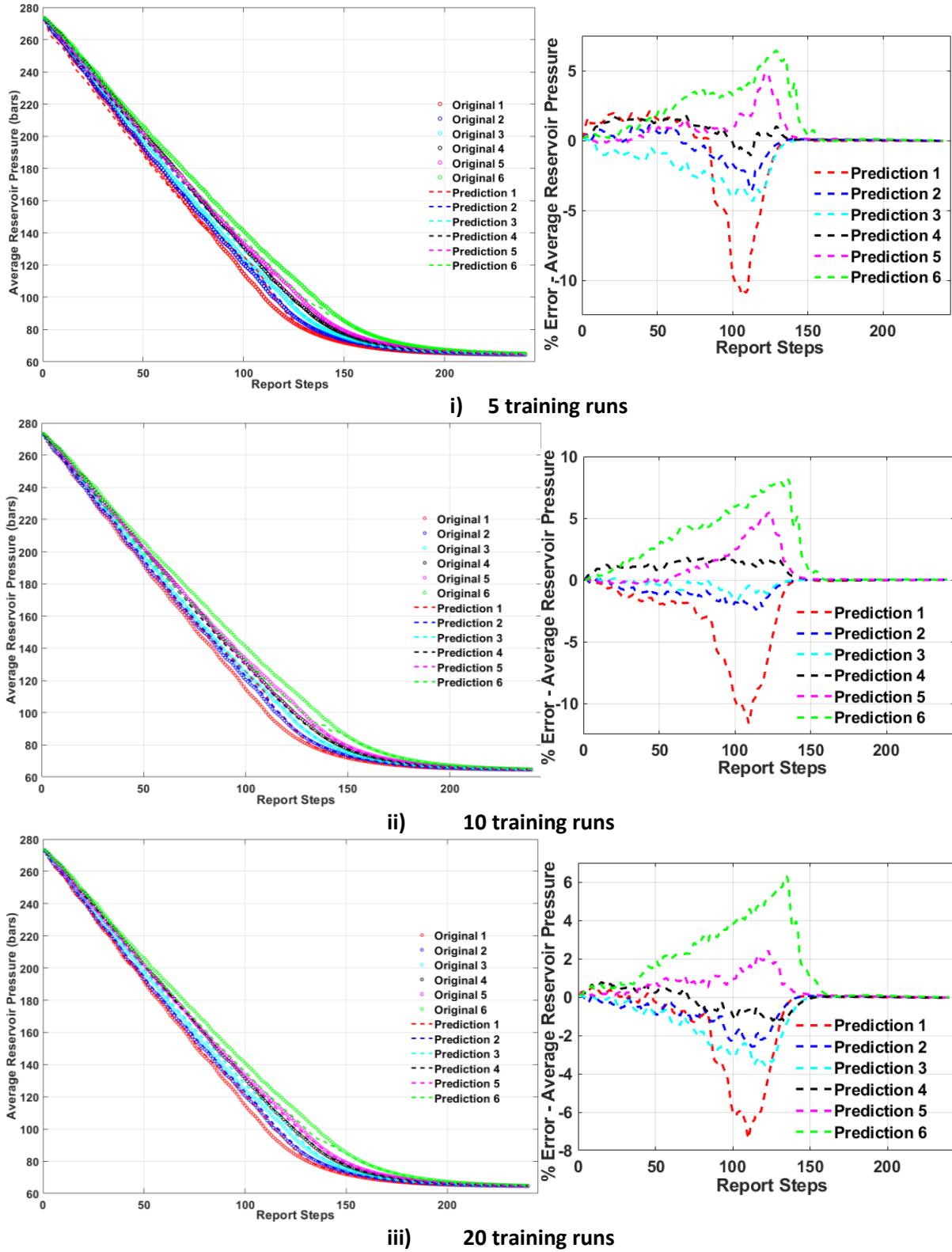


Figure 7.14: NIROM predictions of the average pressure in the modified Norne gas field for the seasonally varying production rate and mixed production control scenarios of Figure 7.12.

NIROM predictions of the average pressure distribution in the reservoir are satisfactory (Figure 7.14), as in all the other cases that have been seen in this chapter. Figure 7.15 and Figure 7.16 show NIROM's satisfactory prediction of the pressure and gas saturation distribution in all the grid cells of the gas

field. Figure 7.14 and Figure 7.15a, both show that NIROM's prediction 1 was least accurate between the 100th and 130th report steps. Generally, Figure 7.14 shows that NIROM predictions of predictions 2 to 5 are the most accurate with errors within the range of $\pm 5\%$, while those of predictions 1 and 6 are as high as $\pm 10\%$. For this example, the POD-RBF NIROM was implemented with the training runs in Figure 7.13. Figure 7.16 shows that NIROM's prediction of the gas saturation distribution in all the grid cells of the gas field has errors that are mostly in the vicinity of zero.

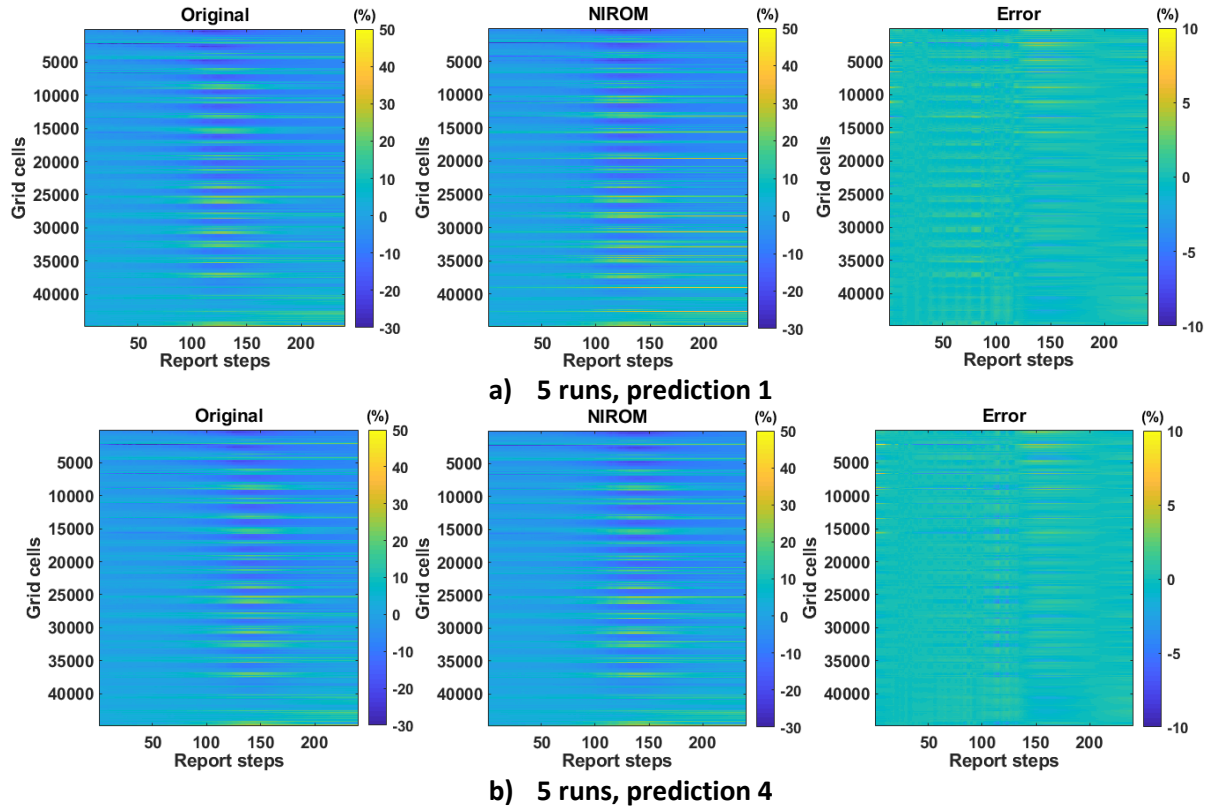


Figure 7.15: Pressure distribution in all grid cells of the modified Norne gas field, normalised to zero mean (predictions 1 and 4). In both cases, the NIROM's predictions show the most errors between the 100th and 130th report step corresponding to the region with the highest error in Figure 7.14, i.e. the time interval when the production control changed. Nevertheless, NIROM's prediction of the pressure across the reservoir is satisfactory.

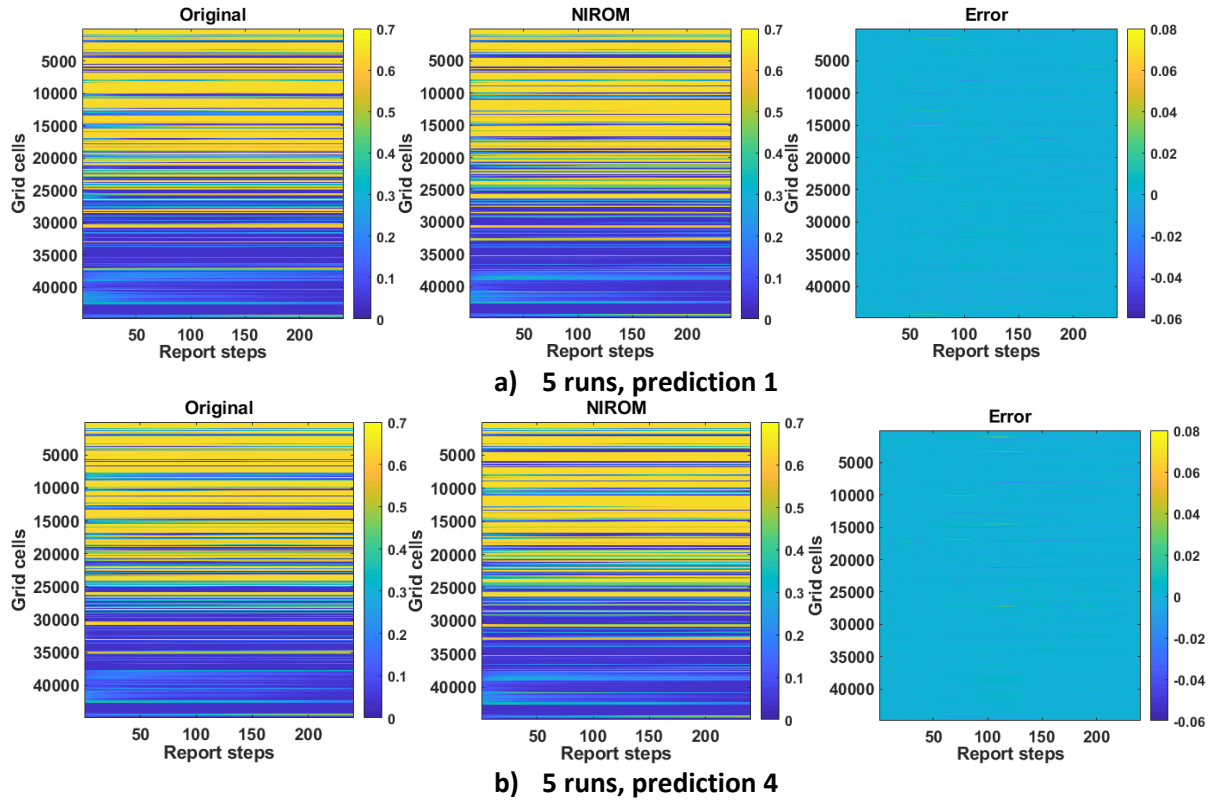
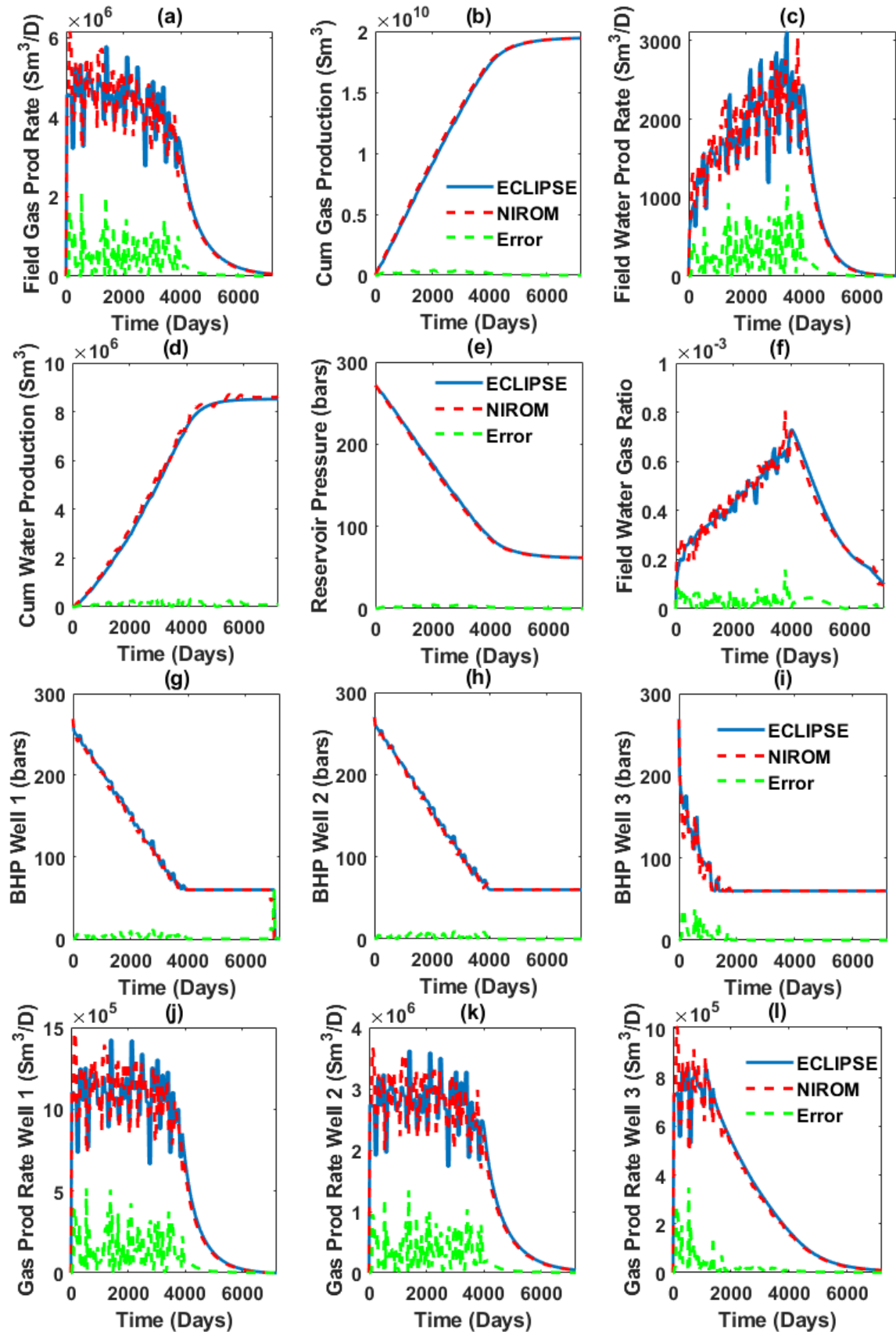


Figure 7.16: NIROM's prediction 1 and 4 of the gas saturation distribution in all the grid cells of the modified Norne gas field. NIROM satisfactorily captured the pattern of gas saturation distribution in the gas field, with negligible error. The interval when production control changed to BHP limit is not noticeable in these images.

For brevity, Figure 7.17 only contains results of NIROM predictions of the production variables of the Norne field when NIROM was trained with 20 training runs. Similar results were obtained from using 5 or 10 training runs, which were all obtained by using 9 POD basis functions (energy criteria = 1 POD basis function). Figure 7.17 shows that NIROM gives satisfactorily predictions of most of the production variables and adequately captures the pattern of varying production rates (gas and water) in the 3 wells and the entire field (Figure 7.17 – a, c, j-o).

Like in the constant parameter cases, it is also seen here that NIROM predictions of the production variables required more POD basis functions than were suggested by the 99.9% energy criterion. A quick solution for determining the number of POD basis functions that would suit this particular gas field when making predictions of production variables would be to train the NIROM with the full POD basis functions. This was mentioned in chapter 6, where it was discussed that the total amount of data that is used for predictions of production variables is significantly less than the amount of data that is used for predictions of the spatial distribution of reservoir properties, hence making production data prediction with the full basis functions will have a negligible effect on NIROM's total runtime.



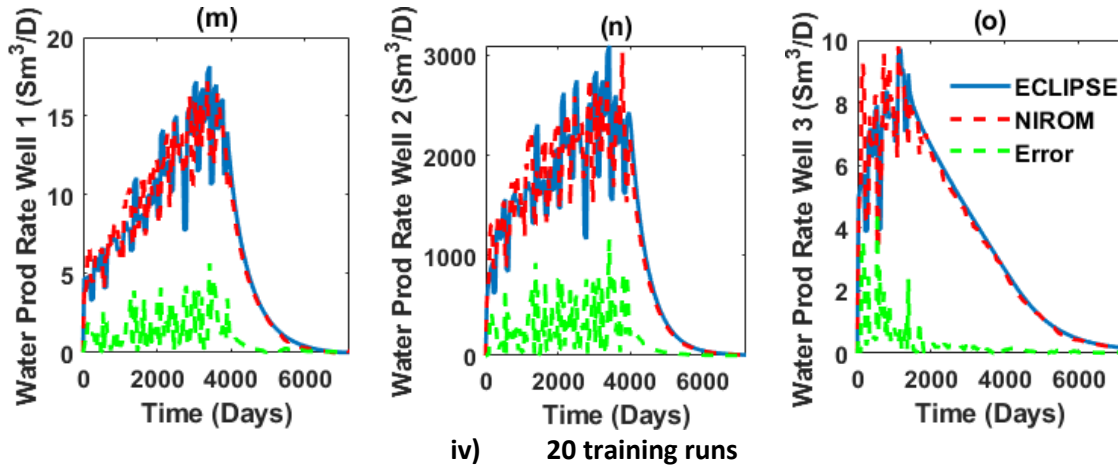


Figure 7.17: NIROM prediction 4 of production variables of the modified Norne gas field trained with 20 runs and 9 POD basis functions (energy criterion = 1 POD basis function). NIROM adequately predicts most of the production variables and satisfactorily captured their performance pattern.

7.3 FIELD X

7.3.1 BHP LIMIT

Figure 7.18 shows that the prediction runs that were evaluated in this case were evenly spread, with the maximum rate being around $2.25 \text{ sm}^3/\text{day}$ for predictions 2 to 6. Prediction 1 had higher production rates than the other 5 predictions in most instances, but Figure 7.19 and Figure 7.20 show that NIROM gave good predictions of this case, as well as the others. This was expected since the production profiles of the training runs in Figure 7.22 shows that the training runs had similar patterns and covered a similar range of values. For the average pressure distribution results in Figure 7.19 and the average gas saturation distribution in Figure 7.20, the NIROM with 10 training runs gave the most accurate predictions. In Figure 7.19, the maximum error in NIROM predictions was just over 6%, 5% and 4% for the cases with 5, 20 and 10 training runs, respectively.

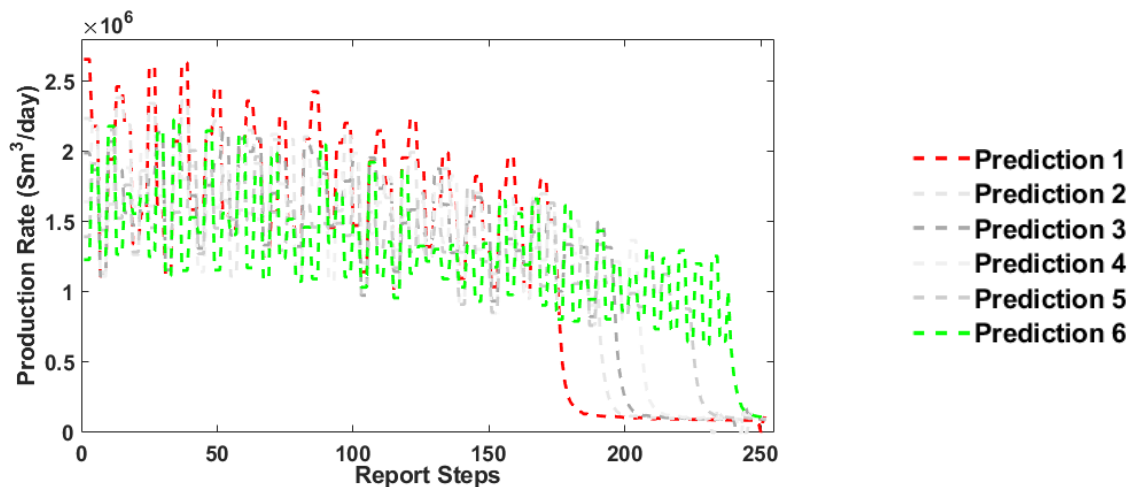


Figure 7.18: Production profiles for the prediction runs of the seasonally varying parameter example of Field X, based on a BHP limit of 100bars. All 6 production profiles show that the field switched to BHP control during its production lifecycle, with instances of the producer shutting in after 230 report steps.

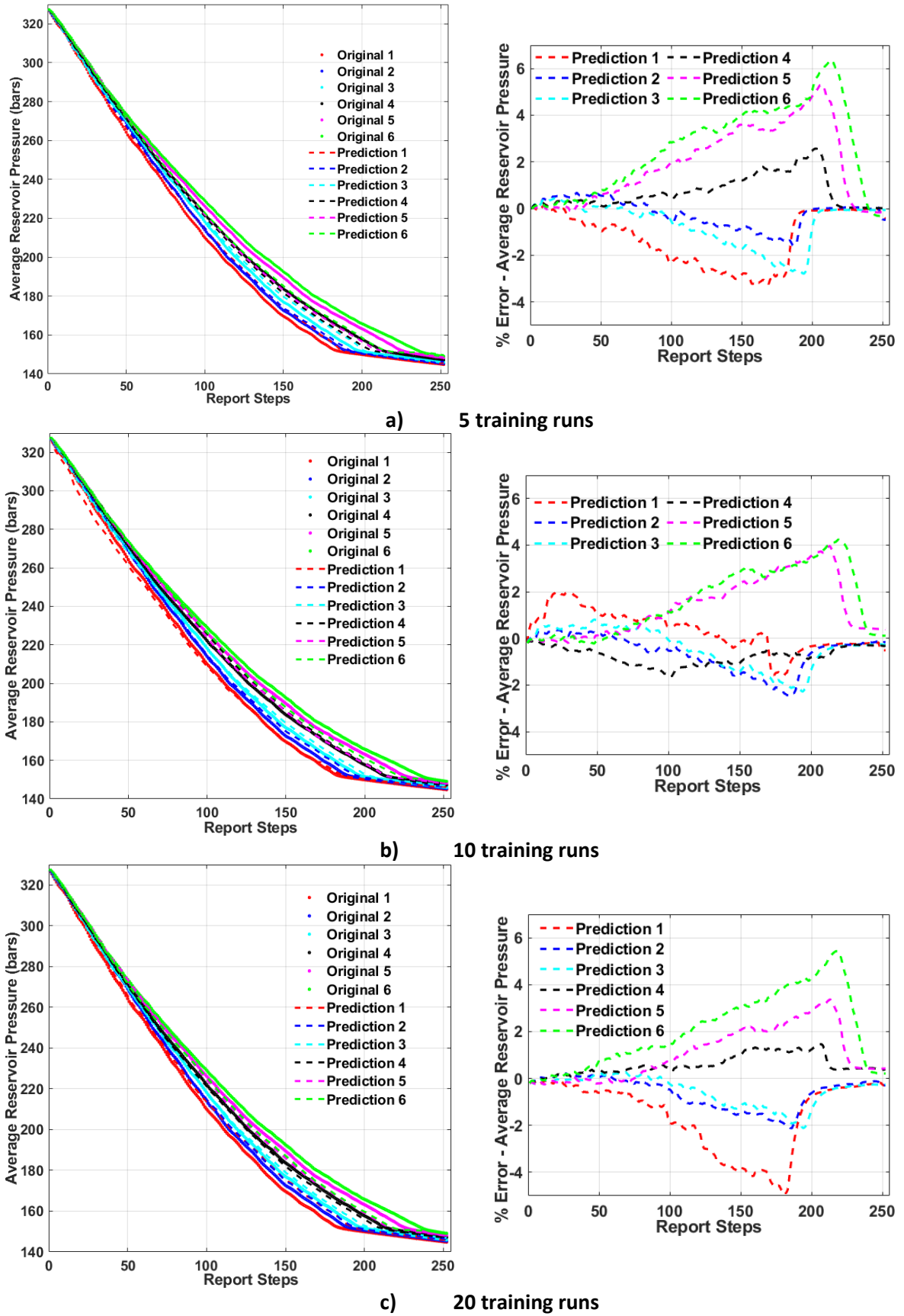


Figure 7.19: NIROM's prediction of the average pressure distribution in Field X produced with a seasonally varying production profile and BHP limit. Predictions with 10 training runs were the most accurate with a maximum error of just over 4%. Predictions with 5 and 20 training runs have a maximum error of just over 6%.

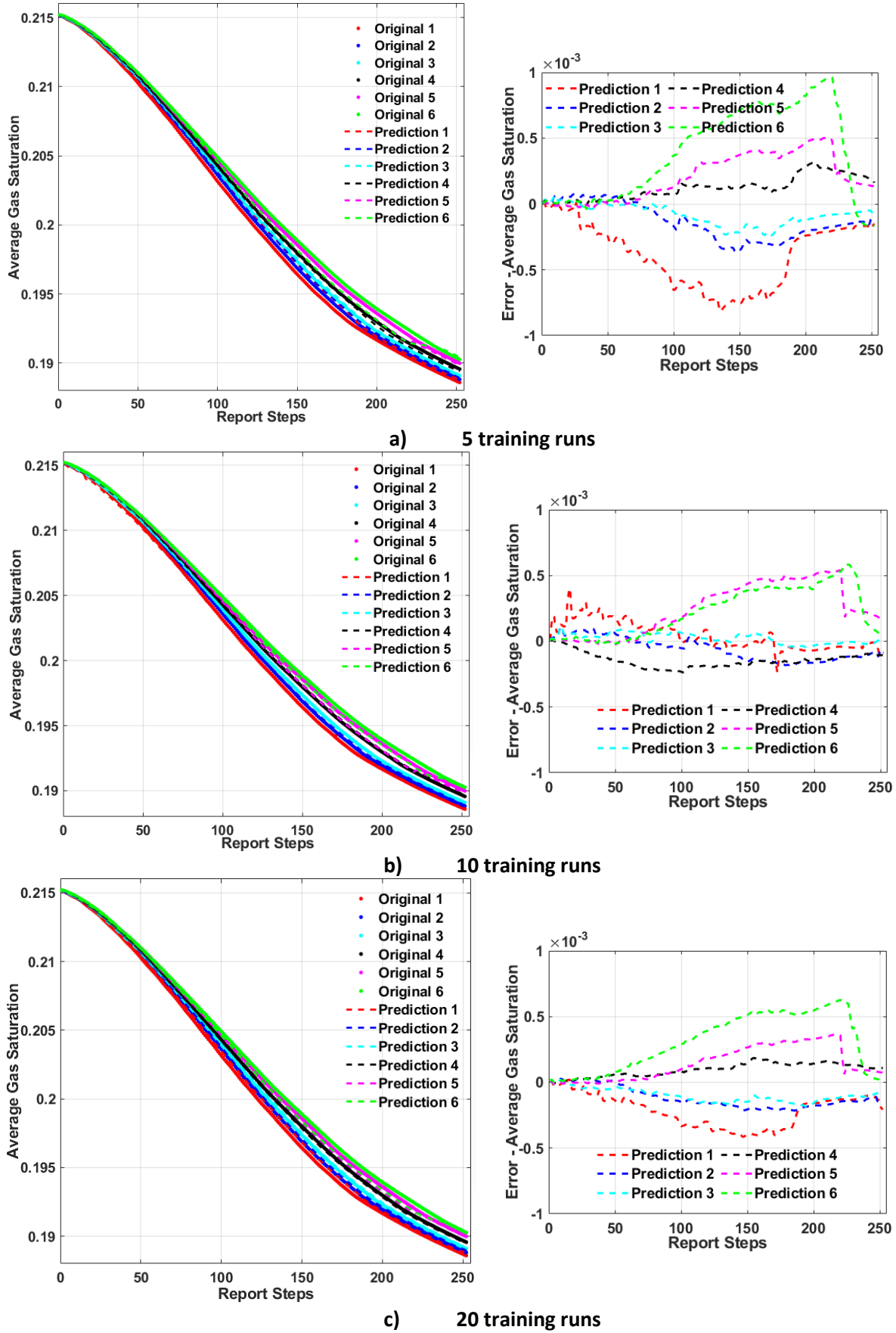


Figure 7.20: NIROM's prediction of the average saturation distribution in Field X produced with a seasonally varying production profile and BHP limit of 100bars. Predictions with 10 training runs were the most accurate, but no case had errors greater than 0.001.

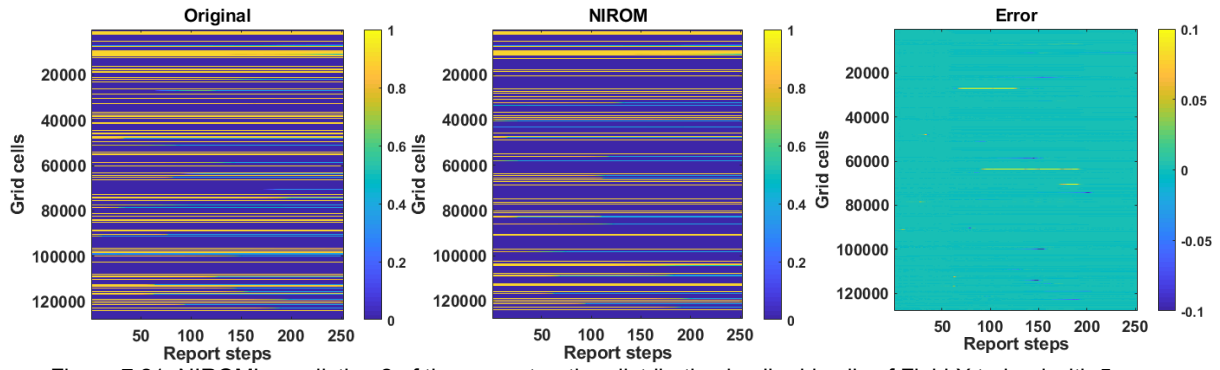


Figure 7.21: NIROM's prediction 2 of the gas saturation distribution in all grid cells of Field X trained with 5 runs. NIROM satisfactorily captured the pattern of gas saturation distribution in the gas field, with negligible error.

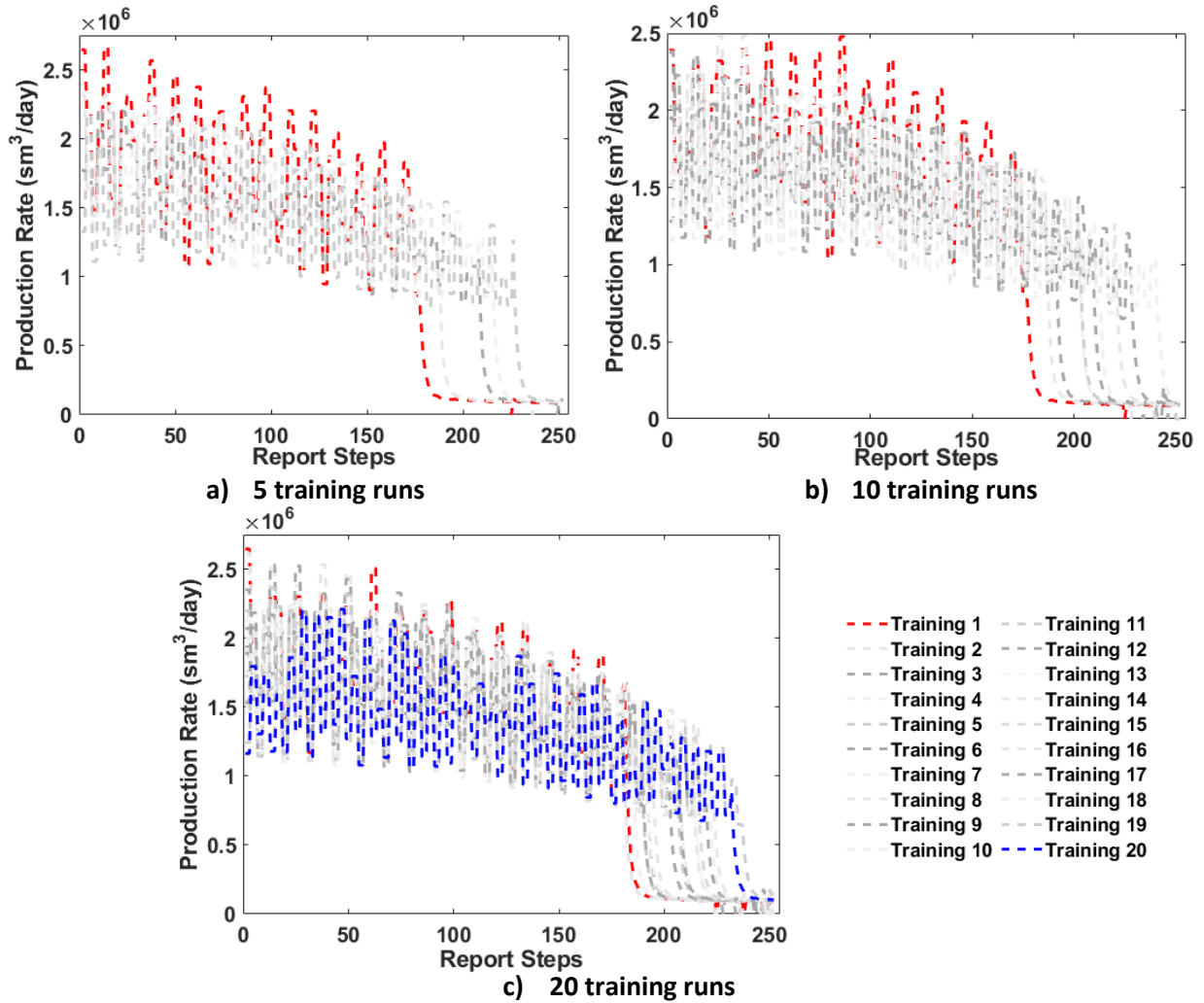
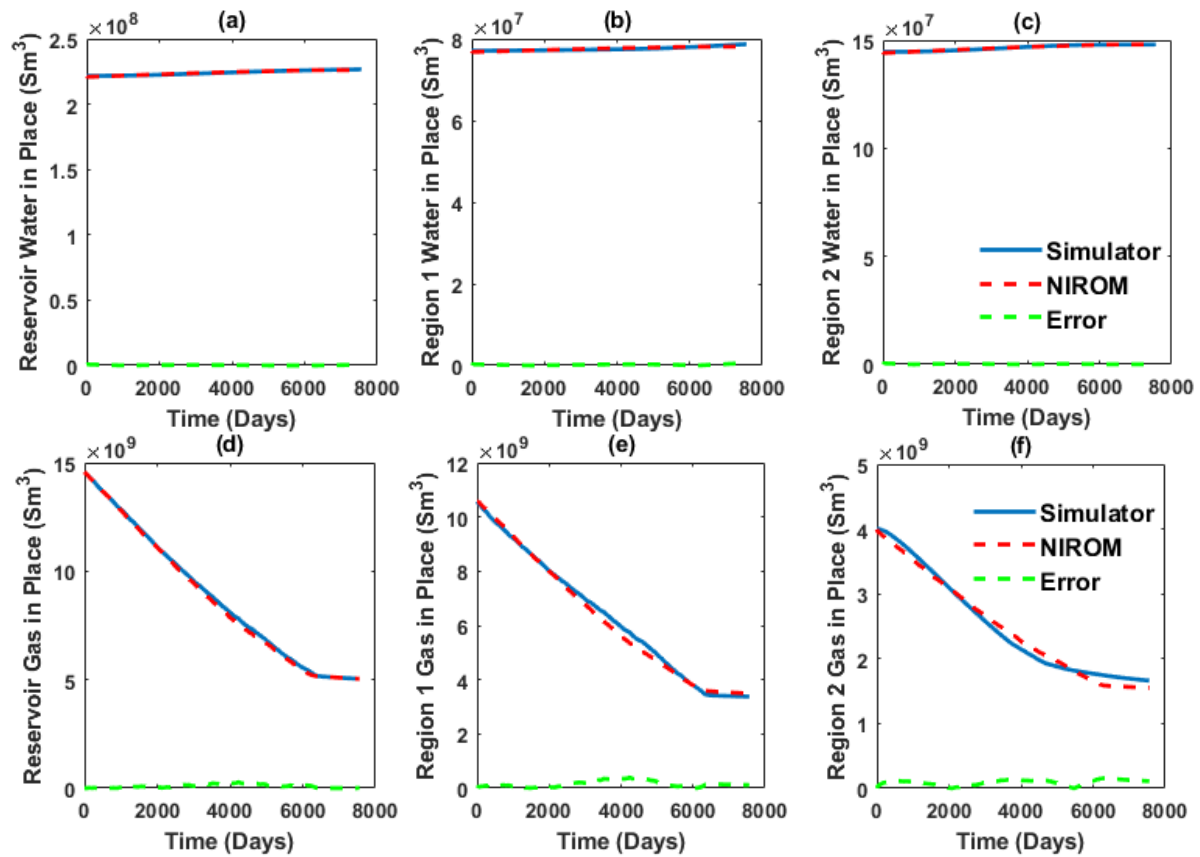


Figure 7.22: Production profiles used to train the NIROM for the prediction runs in Figure 7.18. The training runs span a similar range of parameters as the prediction runs, with some values being as high as prediction 1 of Figure 7.18. Given that these training runs cover the range of prediction runs, NIROM predictions in Figure 7.19 to Figure 7.21 and Figure 7.23 are satisfactory.

In Figure 7.20, the range of values for the average gas saturation distribution in Field X is just slightly greater than 0.215, which can be considered to be a small change for a practical gas reservoir. However, Figure 7.21 shows the gas saturation distribution across all the grid cells of Field X at any given report step, for prediction 2 of this mixed production control example. From Figure 7.21, it can

be seen that the gas saturation distribution in Field X is such that there are extreme gas saturation values of close to 0 and 1. Hence the low range of saturation values in Figure 7.20 is because Figure 7.20 only shows the average gas saturation distribution across the reservoir at every report step. The full gas saturation distribution plot in Figure 7.21 confirms NIROM's satisfactory prediction of the reservoir property in each grid cell of Field X and complements Figure 7.20.

NIROM's prediction of some of the production data for Field X is shown in Figure 7.23. NIROM satisfactorily predicted most of the production variables (pressure in the different regions of Field X and cumulative production), but found it challenging to predict the varying production rates (of gas, water and the aquifer influx). The results in Figure 7.23 were obtained using the optimum number of POD basis functions, but as shown in the Norne field example (Figure 7.17), better results can be obtained by adding more (or the full) POD basis functions to the NIROM's implementation. This is further illustrated in section 7.3.2.



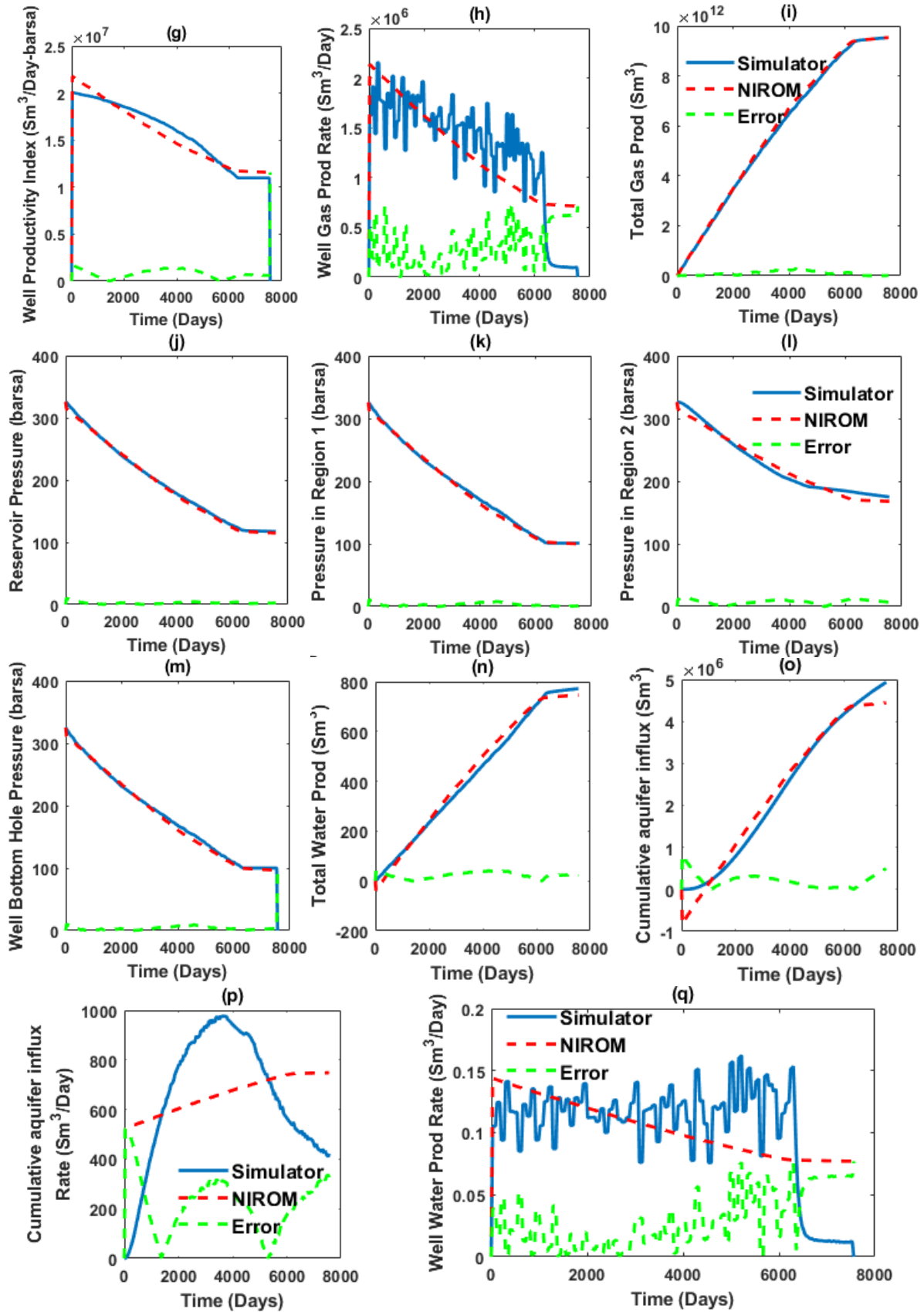


Figure 7.23: NIROM's prediction 4 of Field X's production variables using 5 training runs. With the optimum number of POD basis functions, NIROM did not accurately predict the time-varying production rates of Field X. However, NIROM satisfactorily predicts more than half of the production variables that were considered.

7.3.2 THP LIMIT

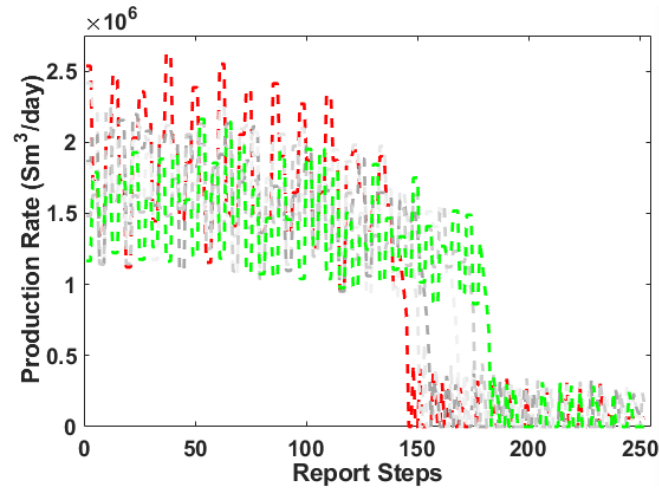


Figure 7.24: Production profile for the prediction runs of the seasonally varying parameter cases of Field X with THP limit. From the 150th report step, the production profiles for Field X were switching between THP limit control and episodes of shut-in due to a water gas ratio limit imposed on the reservoir.

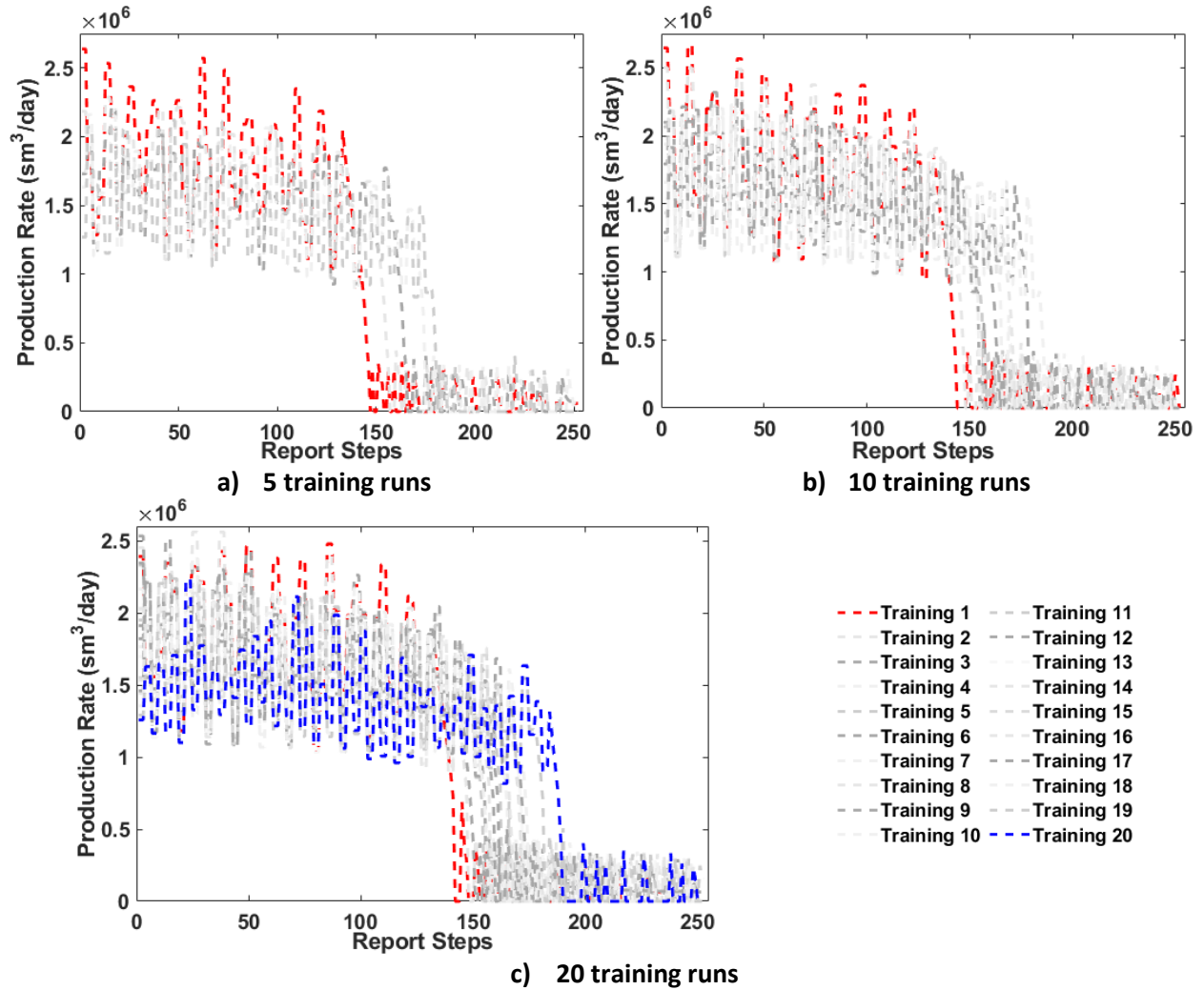


Figure 7.25: Production profiles of the training runs used for the NIROM predictions in Figure 7.24. These training runs have a similar shut-in and re-opening behaviour as the prediction runs in Figure 7.24, hence the satisfactory prediction results for these training cases as shown in Figure 7.26 to Figure 7.28.

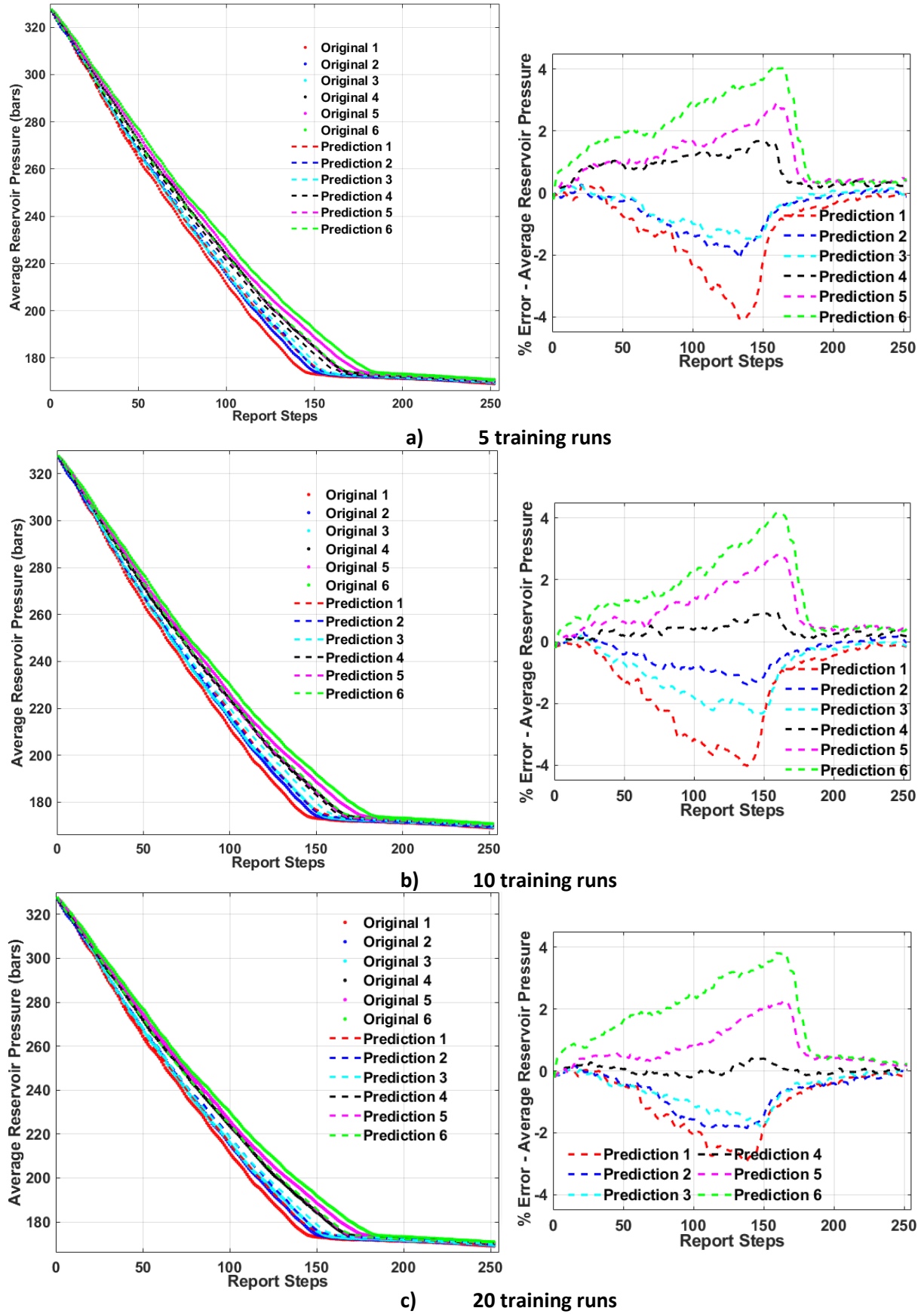


Figure 7.26: NIROM predictions of the average pressure distribution in Field X for the seasonally varying production profiles with THP limit control in Figure 7.24. NIROM predictions with 5, 10 and 20 training runs are satisfactory with at most 4% error.

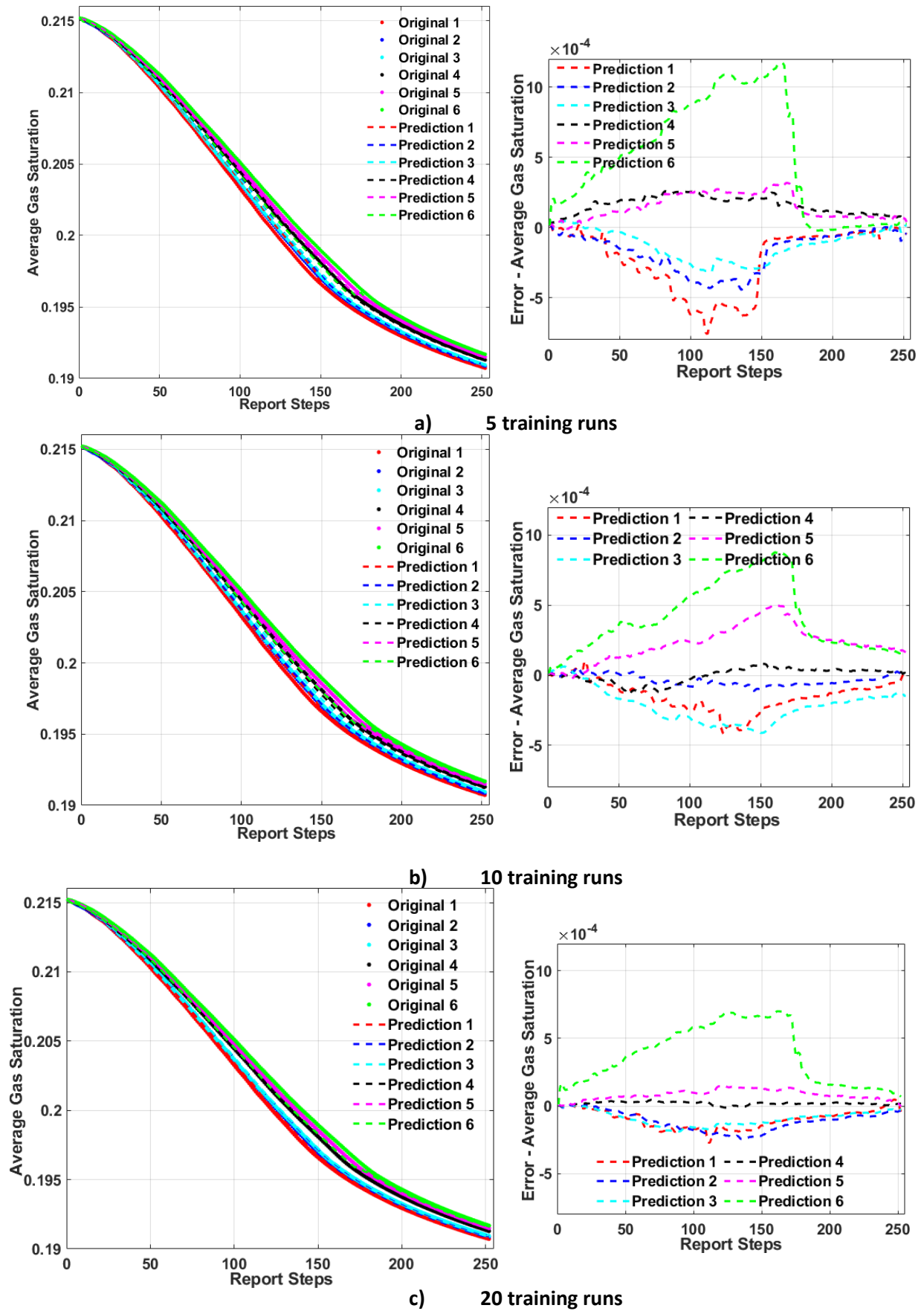
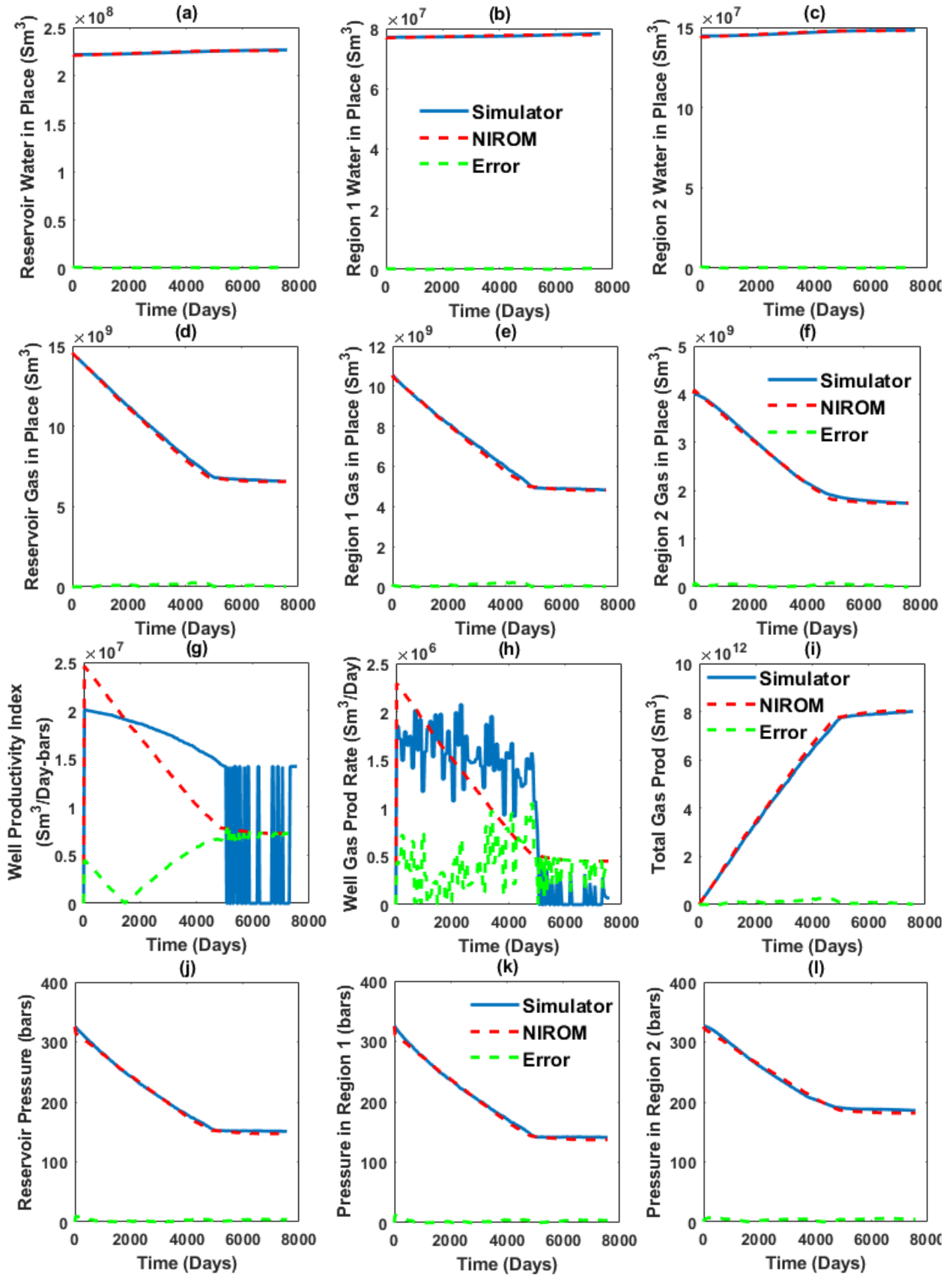


Figure 7.27: NIROM's satisfactory predictions of the average gas saturation distribution in Field X when it was produced by a seasonally varying production profile with THP limit. Training NIROM with 20 runs gave the best results.



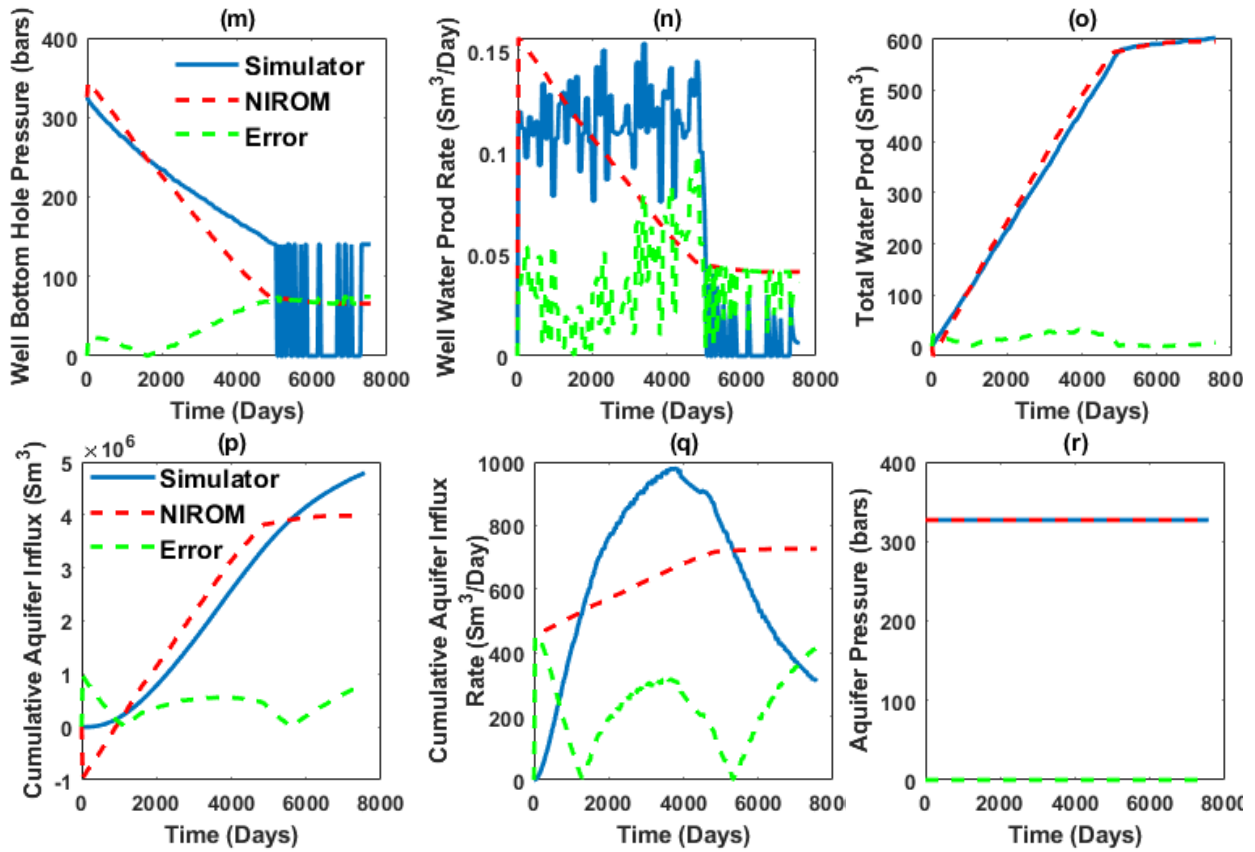


Figure 7.28: NIROM's prediction 4 of the production variables of Field X when trained with 20 runs. NIROM satisfactorily predicted most of the production variables with the optimum number of training runs but found it challenging to capture the pattern of the varying production rates, and especially after the production control began to switch.

The results for this case were based on the production profile of Figure 7.24. From Figure 7.24 it can be seen that once the production control of the field switched to THP limit, the production profile was characterised by episodes of well shut-in and restart until the end of the production lifecycle. The producer was persistently being shut-in because, in addition to all the other properties of the reservoir that were discussed in chapter 4, the reservoir had a water gas ratio limit of $5.61 \times 10^{-6} \text{sm}^3/\text{sm}^3$. The producer was however set to be re-opened if future well tests showed lower water gas ratios in Field X. Perhaps this intermittent production behaviour of the Field X model could have been minimised by using shorter report times or tweaking some ECLIPSE parameters that affect its numerical instability. Nevertheless, NIROM is only as good as the simulation results from which it is trained so if the simulation results have numerical errors in them (as may be the case in this example), then so will the NIROM predictions. Notwithstanding, NIROM was able to satisfactory forecast the field's performance as shown in (Figure 7.26 to Figure 7.28) since it was trained with runs that had similar shut-in and restart production profiles (Figure 7.25).

In Figure 7.26 the error in NIROM predictions of the average reservoir pressure per report step was a maximum of 4% and in Figure 7.27 the error in NIROM predictions of the average gas saturation

distribution was mostly less than 10×10^{-4} . In Figure 7.26 and Figure 7.27, the best NIROM results were obtained with 20 training runs. In Figure 7.26, it was also observed that the NIROM did not always accurately capture the time when the production control switched to THP limit. NIROM's highest prediction errors occurred during the periods when the production control was switched. For example in the 20 training runs cases, the NIROM suggested that the switch in production control occurred 7 days in advance for prediction 6 (which had the highest absolute error), while in prediction 1 (with the next highest error) the NIROM suggested that the switch occurred 5 days after ECLIPSE. This shows that even though the NIROM generally gives satisfactory predictions in the varying production case with THP limit, the NIROM is sensitive to the change in production control and care should be taken to ensure that simulation errors are within acceptable limits.

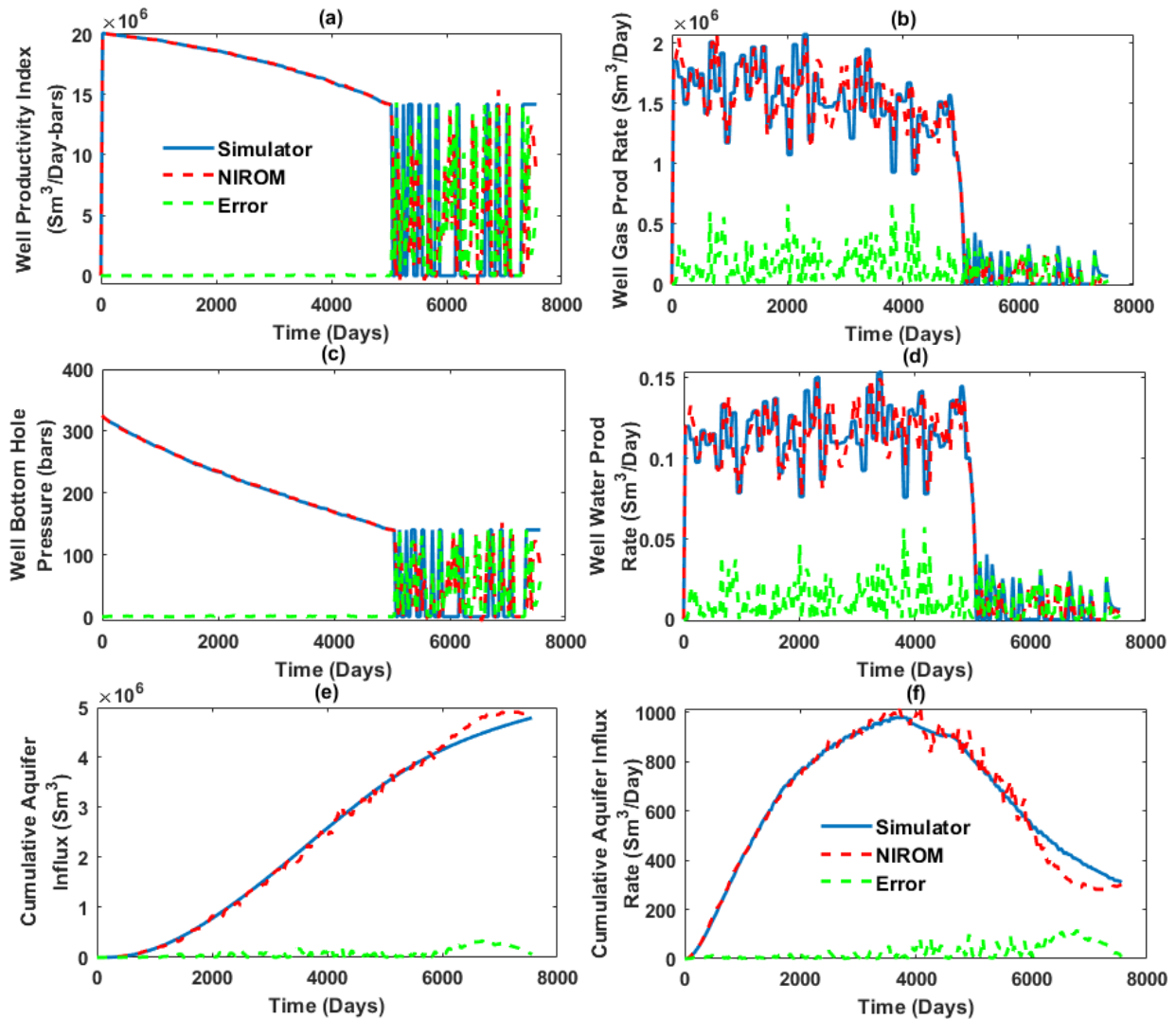


Figure 7.29: NIROM's prediction 4 of some production variables of Field X using the full POD basis functions. NIROM better predicts the varying production rates and BHP when more POD basis functions were added to its implementation.

For the production data predictions in Figure 7.28, NIROM gave satisfactory predictions of most of the production variables except the production rates (gas, water and aquifer) and well BHP. Adding more POD basis functions to the prediction gave better results as shown in Figure 7.29. Figure 7.29 shows that with the full set of POD basis functions, NIROM satisfactorily captures the pattern of the varying production rates and well BHP.

NIROM's prediction of the production data in Figure 7.11, Figure 7.23 and Figure 7.28 were either satisfactory or within the same order of magnitude as the values from ECLIPSE results. These results showed that even when NIROM predictions were not accurate, NIROM was able to mimic the pattern of ECLIPSE's results, although in most cases adding more POD basis functions gave better NIROM predictions. Overall, NIROM's best prediction of production variables is obtained when the NIROM is implemented with the full POD basis functions of the training data, which has a negligible effect on NIROM's overall runtime. And even when NIROM's results are not accurate, the different practical examples that have been considered in this chapter show that they can still give engineers an idea of a reservoir's behaviour and the order of magnitude of values that can be expected from the reservoir's production performance.

7.4 ERROR ANALYSIS

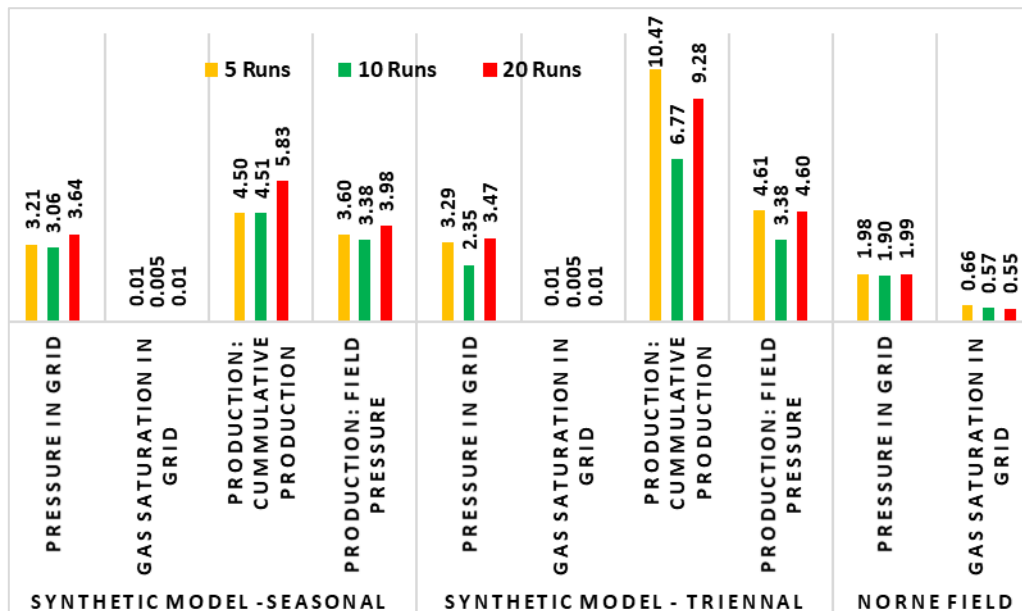


Figure 7.30: Average percentage RMSEs of NIROM when implemented for predicting the performance of the synthetic gas reservoir and the Norne gas field, produced with varying production profiles and mixed boundary conditions. NIROM prediction of grid distributed pressure and saturations have lower RMSEs than its prediction of production variables, based on the optimal number of POD basis functions. There is also no consistent correlation between the number of training runs and the accuracy of NIROM predictions.

Figure 7.30 shows the average percentage RMSEs (Equation 3.20) for all the NIROM predictions of the synthetic gas reservoir and Norne field when produced with time-varying production profiles. These errors were computed based on the optimal number of POD basis functions required to generate NIROM predictions. For the synthetic gas reservoir, the average RMSEs for both the seasonally varying and the triennially varying profiles are shown below. A comparison of all RMSE values in Figure 7.30 shows that NIROM had lower errors when predicting the distribution of pressure or saturation in the reservoirs than for the production variables. However as has been highlighted in previous sections, NIROM predictions of production variables can be improved by increasing the number of POD basis functions applied to those predictions. It is recommended that the full POD basis functions are used to implement the NIROM for predicting production variables.

From Figure 7.30, it can also be seen that NIROM's accuracy does not always improve when the number of training runs is increased. It was earlier shown that having a more evenly spread range of training runs, that cover the range of prediction parameters, has a better impact on the quality/accuracy of NIROM predictions than having more training runs. Overall, NIROM predictions of varying parameter production cases based on the optimal number of POD basis functions has RMSEs that are predominantly less than 5% for the synthetic gas reservoir and less than 2% for the heterogeneous Norne field.

NIROM predictions of the reservoir property distribution in the real gas reservoir (Field X) have less than 3% average RMSEs, despite being produced with varying production profiles and mixed production controls (Figure 7.31). 3% error is within an acceptable error margin in industry applications, therefore NIROM is a suitable tool for gas reservoir performance modelling. The average RMSEs in Figure 7.31 also show no clear relationship between the number of training runs used to train NIROM and NIROM's accuracy. As explained before, the accuracy of NIROM's predictions is more attributed to the extent of spread of the training runs rather than their quantity. If the number of training runs covers a similar interval or range as the prediction runs, then NIROM predictions will be satisfactory with 5 or 20 (fewer or more) training runs.

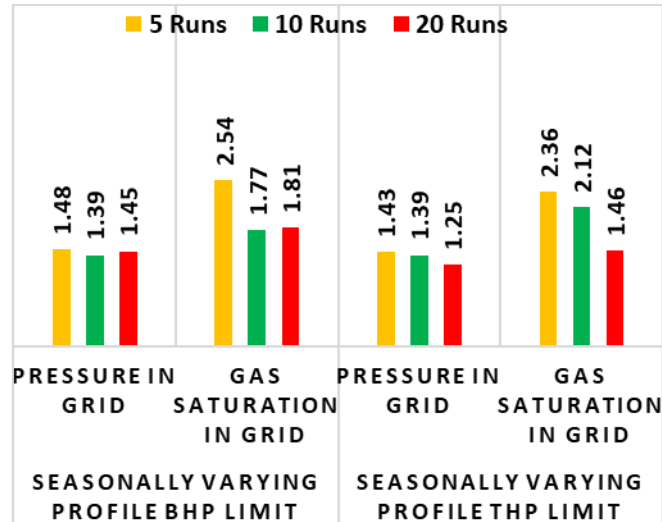


Figure 7.31: Average RMSEs for NIROM predictions of the pressure and saturation distribution in all grid cells of Field X for a) constant parameter cases b) seasonally varying parameter cases with BHP limit and c) seasonally varying production profile with THP limit.

For the 3 gas reservoirs that were considered in this varying parameter chapter, it has been shown that the best NIROM prediction results are obtained when the NIROM is used for predicting production scenarios that are similar to the production scenarios that it was trained with. This behaviour was consistent for the different practical production scenarios that were evaluated in this chapter (even for the real gas field – Field X) and for the ranges of parameters that were evaluated. This shows that NIROM would replicate this same behaviour for other ranges of parameters or production scenarios that an engineer might be interested in considering during field development planning, i.e. as long as the NIROM is trained with similar production scenarios as the prediction cases, then the NIROM would give satisfactory results.

This chapter showed the satisfactory performance of the POD-RBF NIROM for time-varying production scenarios with mixed production controls. Some examples of NIROM's prediction of the full spatial distribution of pressure and saturation in the gas reservoirs were shown in Figure 7.3, Figure 7.15, Figure 7.16, Figure 7.21, which complemented their respective average pressure and saturation distribution results. NIROM's satisfactory prediction in these cases highlight its relevance for fast prediction of spatially distributed reservoir properties in real gas reservoirs and for real production cases that engineers consider during field development planning. These results also highlight NIROM's advantage over the other conventional reservoir modelling approaches such as the material balance equations which are spatially unaware and unable to provide information about the pressure distribution in heterogeneous reservoirs, at different production times.

7.5 TIME ANALYSIS

7.5.1 SYNTHETIC GAS RESERVOIR

Figure 7.32 and Figure 7.33 show the breakdown of the time spent in generating the NIROM results for the seasonally and triennially varying production cases of the synthetic gas reservoir. Table 7.1 shows the values from which Figure 7.32 and Figure 7.33 were generated. Consistent with the constant parameter cases, most of the time take to implement the NIROM is spent on extracting data (training and prediction data) as well as during POD reduction. For the seasonally varying production profile (Figure 7.32), a significant amount of NIROM time was also spent in normalising the data for the case with 20 training runs and in computing the interpolation weights. Besides, when the NIROM was trained with 10 or 20 runs, NIROM was slower than ECLIPSE. This lack of speed up in NIROM predictions trained with 10 or 20 seasonally varying training runs is attributed to the increased number of reports steps in these cases. Whereas NIROM was implemented with a maximum of 25 report steps in the constant parameter cases, 3 to 10 times more report steps are required to create representative time-varying parameter cases (especially in the seasonally varying case). This equates to 3 to 10 times more data and therefore longer time spent extracting the data and implementing POD reduction. However, if the time spent in these supplementary steps are discounted, the actual time spent in NIROM interpolation/ prediction is negligible (less than 0.2 seconds).

Table 7.1: Breakdown of the time take to run the varying parameter cases of the synthetic gas reservoir model, in seconds. These results are plotted in Figure 7.32 and Figure 7.33.

Production Profile	Seasonally Varying			Triennially Varying		
Number of Training Runs	5	10	20	5	10	20
Extracting Training Data	7.27	8.83	9.21	0.95	1.00	0.90
Mean Extraction	0.10	0.20	0.42	0.04	0.07	0.11
POD Reduction	1.25	6.48	7.29	0.49	0.55	1.82
Normalization and Computing Interpolation Weights	0.37	1.33	6.54	0.24	0.18	0.51
Total NIROM Training Time	8.98	16.84	23.45	1.72	1.80	3.34
Extracting Prediction Data	6.67	8.11	8.46	0.92	0.92	0.92
NIROM Interpolation	0.06	0.09	0.15	0.02	0.02	0.02
Total NIROM Time	15.7	25.0	32.1	2.66	2.74	4.29
Total ECLIPSE Time	23.6	23.6	23.6	9.47	9.47	9.47
Speed-up (Ratio)	1.50	0.94	0.74	3.56	3.45	2.21

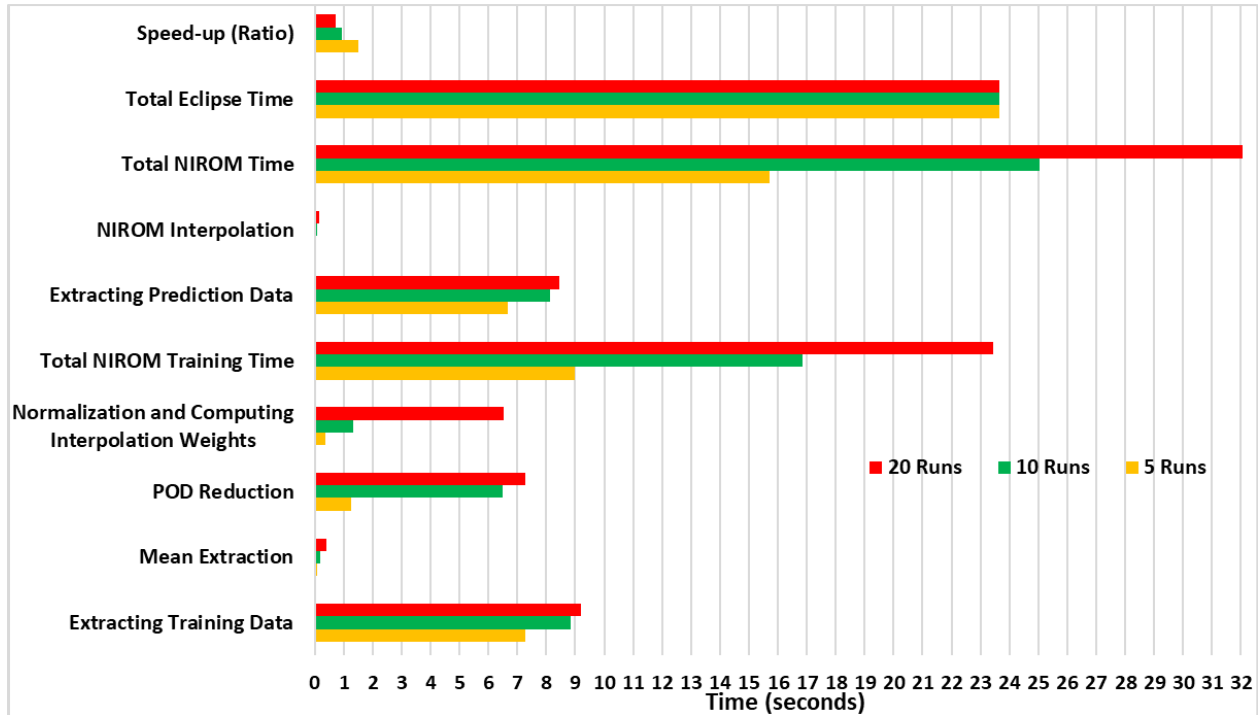


Figure 7.32: The breakdown of time taken to train NIROM for the seasonally varying production profile of the synthetic gas reservoir, as well as the speed up from NIROM predictions. Significantly more time is taken during POD reduction for the cases with 10 and 20 training runs than with 5 training runs. The NIROM built with 10 or 20 training runs were slower than ECLIPSE.

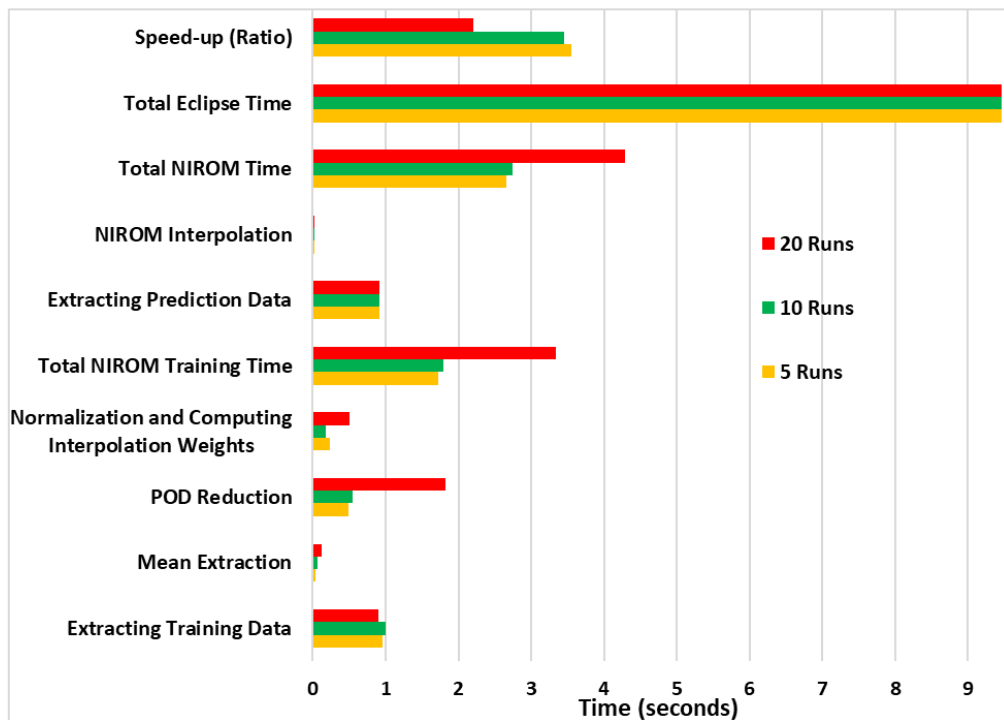


Figure 7.33: Breakdown of the average time taken to implement NIROM for the triennially varying production profile of the synthetic gas reservoir. NIROM is 2 to 3.5 times faster than ECLIPSE for these cases. Training NIROM with 20 runs gives the least speed-up since it involves more data. The time spent in the actual NIROM interpolation/ prediction is negligible compared to the time spent in building the NIROM and if the NIROM is built in a compiled language it would be even faster.

In the seasonally varying case (Figure 7.32), NIROM is a maximum of 1.5 times faster than ECLIPSE when trained with 5 training runs. In the triennially varying case, NIROM has speedups of 2 to 3.5 times the ECLIPSE run time (Figure 7.33). Since the triennially varying cases have fewer report steps and therefore less data than the seasonally varying case, they have higher NIROM speedups as shown in Table 7.1. This shows that similar to constant parameter cases, NIROM's implementation speed is dependent on the amount of data it is trained on, which has also been shown to depend on the number of report steps in the conventional simulator's output data. Hence reiterating the importance of building NIROM with the least amount of report steps (training data) that is sufficient for the particular reservoir simulation study.

7.5.2 MODIFIED NORNE GAS FIELD

Table 7.2 and Figure 7.34 further emphasise the observations in the previous section, that with more simulation data NIROM's speed becomes slower. The Norne field has over 44000 grid cells and therefore requires more time for extracting simulation data from ECLIPSE into NIROM and also for POD reduction. Table 7.2 shows that the time required for POD reduction when the NIROM was trained with 20 runs was more than four times the time required for POD reduction of the case with 10 training runs. Nevertheless, NIROM was found to be slightly faster than ECLIPSE when it was trained with 20 runs (1.2 times faster). This further reiterates the computational time cost of training NIROM with several runs.

Table 7.2: Average time taken to implement NIROM for the seasonally varying production profile of the modified Norne gas field. NIROM is a maximum of 2 times faster than ECLIPSE.

Number of Training Runs	5	10	20
Extracting Training Data	84.8	82.4	83.4
Mean Extraction	2.05	4.28	7.65
POD Reduction	8.49	26.33	114.00
Normalization & Computing Interpolation Weights	0.29	1.27	4.08
Total NIROM Training Time	95.6	114.3	209.1
Extracting Prediction Data	85.14	86.94	88.77
NIROM Interpolation	0.21	0.26	0.34
Total NIROM Time	180.9	201.5	298.2
Total ECLIPSE Time	362.2	362.2	362.2
Speed-up (Ratio)	2.00	1.80	1.21



Figure 7.34: Breakdown of the average time taken to implement NIROM predictions for the seasonally varying production parameter cases of the modified Norne field with mixed production controls. Less than 0.2% of the NIROM implementation time is spent on the actual NIROM predictions/ interpolation.

7.5.3 FIELD X - TIME ANALYSIS

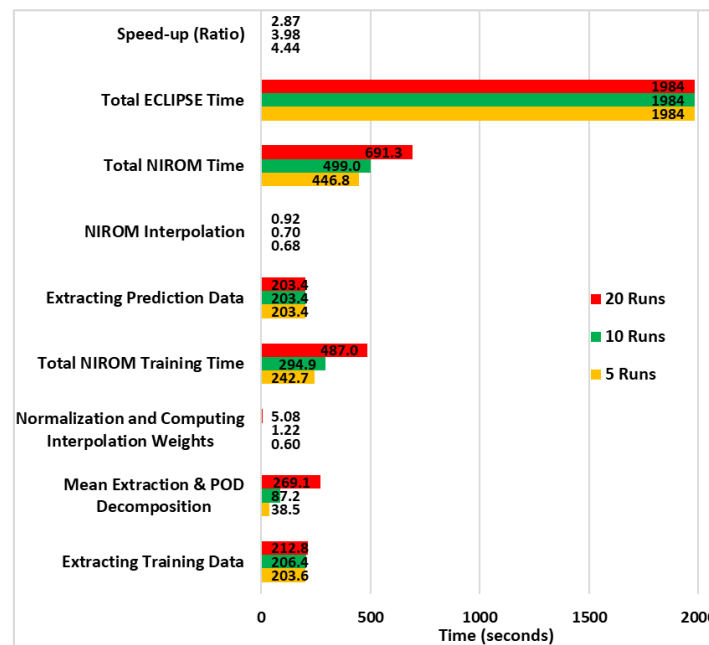


Figure 7.35: The breakdown of the average time taken for NIROM predictions of the seasonally varying cases of Field X. NIROM is between 2 and 4 times faster than ECLIPSE, although less than 1 second is spent on the actual NIROM interpolation step. This shows that NIROM has room for significantly more speedup if it is implemented in a compiled language like ECLIPSE.

Figure 7.35 shows the breakdown of the time taken to implement NIROM for the varying parameter cases of Field X. From Figure 7.35 it is seen that whereas it takes ECLIPSE over 30 minutes to complete a full simulation run of Field X, NIROM predictions can be implemented in about 11.5 minutes for the seasonally varying cases. NIROM implementation takes more time for Field X than for the Norne field, in the seasonally varying cases, since Field X has about 3 times more grid cells/ data than the Norne field. Nevertheless, NIROM is at least 3 times faster than ECLIPSE for this real gas reservoir example. Here also, a significant amount of the time spent in implementing NIROM is spent on POD reduction and when training or prediction data were extracted from ECLIPSE output files into the NIROM framework. Figure 7.35 also shows that the time spent in the POD reduction is more than doubled as the number of training runs increased from 5 to 10 to 20. The amount of time spent in generating interpolation weights also significantly increased as the number of training runs increased.

From Table 7.1, Table 7.2 and Figure 7.35, it can be seen that NIROM's total simulation runtime increased as the reservoir model size increased, i.e. NIROM is slower in the seasonally varying case than in the triennially varying production case of the synthetic gas reservoir, as well as with increasing number of training runs. This again highlights the dependence of NIROM's speed on the size of the dynamic model (number of grid cells, report steps and training runs).

7.5.4. ANALYSIS OF SIMULATION SPEEDUP

NIROM's speed-up for the three gas reservoirs that were considered in this chapter are shown in Table 7.3. Comparing the seasonally varying cases of the 3 reservoirs, it is seen that the more practical (heterogeneous) reservoir models had higher speedups than the homogeneous synthetic reservoir since they would normally be time consuming to simulate in ECLIPSE. A similar performance was observed in chapter 6, where NIROM's speedups for the water coning reservoir were an order of magnitude higher than for the heterogeneous gas reservoir. So NIROM is significantly faster than ECLIPSE for simulating complex/ realistic gas reservoirs and scenarios.

Table 7.3: NIROM's simulation speed-up for the varying parameter test cases considered in this chapter. NIROM's speed-up depends on the amount of data that was used to train the NIROM (model size) as well as the complexity of the reservoir model (homogenous or heterogeneous).

Reservoir Model	Synthetic Gas Reservoir					
Production Profile	Seasonally Varying			Triennially Varying		
Number of Training Runs	5	10	20	10	20	20
Speed-up (Ratio)	1.50	0.94	0.74	3.56	3.45	2.21
Reservoir Model	Modified Norne Field			Field X		
Production Profile	Seasonally Varying					
Number of Training Runs	5	10	20	5	10	20
Speed-up (Ratio)	2.00	1.80	1.21	4.44	3.98	2.87

SUMMARY

In this chapter, NIROM has been evaluated for modelling the performance of 3 gas reservoirs: a homogeneous synthetic gas reservoir, a heterogeneous gas reservoir and Field X which is a real gas field provided by Woodside Petroleum. NIROM was implemented for predicting these reservoirs' performance for time-varying production profiles, each case was evaluated with 5, 10 and 20 training runs. Triennial and seasonally varying production profiles were considered, with BHP or THP production control limit. NIROM was applied to predicting the spatial distribution of pressure and gas saturation in these gas reservoirs, as well as some of their production variables.

NIROM gives better predictions of the pressure distribution in the reservoirs than it does for the gas saturation distribution. For the production variables, NIROM requires more than the optimal number of POD basis functions to give a satisfactory prediction of the varying production rates (for gas, water and the aquifer). NIROM's satisfactory predictions are associated with how much of the range of prediction scenarios are covered by the scenarios of training runs with which the NIROM is implemented, i.e. a NIROM trained without mixed production controls will not satisfactorily predict a reservoir's performance for scenarios with mixed production controls.

The seasonally varying production profile is more representative of industry gas production profiles and its simulations are time-consuming since they involve more report steps/ simulation data. For these cases, NIROM gives satisfactory predictions with less than 3% average RMSE and speed-ups of 1.2 to 4 times ECLIPSE runtime for the 2 heterogeneous reservoir models. The triennially varying production profile cases are faster than the seasonally varying cases (since they have less data) but they are less representative of industry scenarios. However, recall that NIROM is written in an interpreted language whereas ECLIPSE is written in a compiled language and has been improved through several iterations since its inception. Hence NIROM has room for more computational speed up if it is implemented in a compiled language and when its training time is discounted. **Error! Reference source not found.** is a summary of the cases that have been evaluated in this chapter, as well as the key findings.

Table 7.4: A summary of the prediction cases and key findings in this chapter

Cases Considered	Summary and Findings
1) Time-Varying Parameter Examples a. Seasonally varying production profiles b. Triennially varying production profiles 2) Mixed production controls: production rate with BHP or THP limit 3) 3 gas reservoirs were evaluated, including 2 real fields a. homogeneous synthetic gas reservoir b. modified heterogeneous Norne gas field c. Real gas Field X from Woodside	Key Findings
	1) Demonstrated NIROM's predictive capacity for gas field modelling, based on different time-varying production constraints 2) The size and properties of a dynamic model affect the extent of NIROM speed-up that can be obtained for the model, including the number of grid cells and report steps in the model, as well as the complexity of the reservoir model (homogeneous or heterogeneous) and how many training runs are used to implement the NIROM.
	General Findings
	3) NIROM performs best when the training runs are of similar production scenarios as the prediction runs 4) The number of training runs used to train the NIROM is not as important as point 3 above 5) NIROM gives speed-up of at least 1.2 times ECLIPSE's runtime and still has room for speed improvement

CHAPTER 8: POD-AE RESULTS AND DISCUSSION

In this chapter, POD-AE is applied to modelling the grid pressure distribution in the modified Norne gas field. The gas saturation distribution and production data, such as production rates or cumulative production, are not considered since the main goal in this chapter was to analyse the performance of POD-AE for gas field modelling. Also, the more practical varying production profile scenarios are evaluated, for cases with and without shut-in wells and with BHP production control limit. Both the triennially and seasonally varying production profiles are considered. The accuracy of predictions from POD-AE was then compared to POD-RBF, in addition to the breakdown of their runtime and speedup. The time analysis in this chapter only accounts for the time spent in training the models and using them for forecasts (interpolation), i.e. it does not account for the time spent during data extraction from ECLIPSE (training and prediction runs), mean extraction and POD reduction.

In contrast to previous chapters, the examples in this chapter have been trained with 50 training runs (both POD-AE and POD-RBF examples). Trial and error showed that 50 training runs were needed to allow the autoencoder to learn the field's dynamics. Fortunately, the ECLIPSE runtime for the Norne gas field is just over six minutes, making it suitable for quick analysis and validation with 50 training runs. The POD-AE model was validated with an additional new dataset (one of the 6 prediction runs) and the other 5 prediction runs were later forecasted with POD-AE and POD-RBF. For consistency, the same training and prediction data are used for both POD-AE and POD-RBF. Some results from training the Norne field with less than 50 training runs are available in Appendix B.

In this chapter also, the average of the reservoir pressure as computed from the pressure distribution in all active grid cells of the gas field are shown per report step. An example of the full pressure field prediction is also shown to give some perspective of both models' ability to predict grid distributed dynamic data (Figure 8.11). In all the results in this chapter, simulation results from ECLIPSE are in circles, while predictions from POD-AE and POD-RBF models are in solid, dashed, dashed-dot or dotted lines. Lines representing the same prediction cases have similar colours for clarity. The POD-AE was tested for models with 1 to 5 encoder layers and a similar number of decoder layers, but for brevity, only the most satisfactory results are presented in this chapter. For the POD-RBF NIROM, predictions from using the full rank POD basis functions as well as those from the optimal number of POD basis functions are also compared.

8.1 CASES WITH SHUT-IN WELLS

In these cases, one or 2 wells were shut-in when production could not be continued at the set BHP limit of 60bars. POD-AE and POD-RBF predictions of the reservoir's pressure depletion are given in Figure 8.2 and Figure 8.4, while their corresponding production profiles are given in Figure 8.1a & b. The line colours for the 5 predictions in Figure 8.2 and Figure 8.4 correspond to the same predictions in Figure 8.1. Figure 8.1a shows the production profiles for the seasonally varying example and Figure 8.1b is the production profile for the example with triennially varying production profiles.

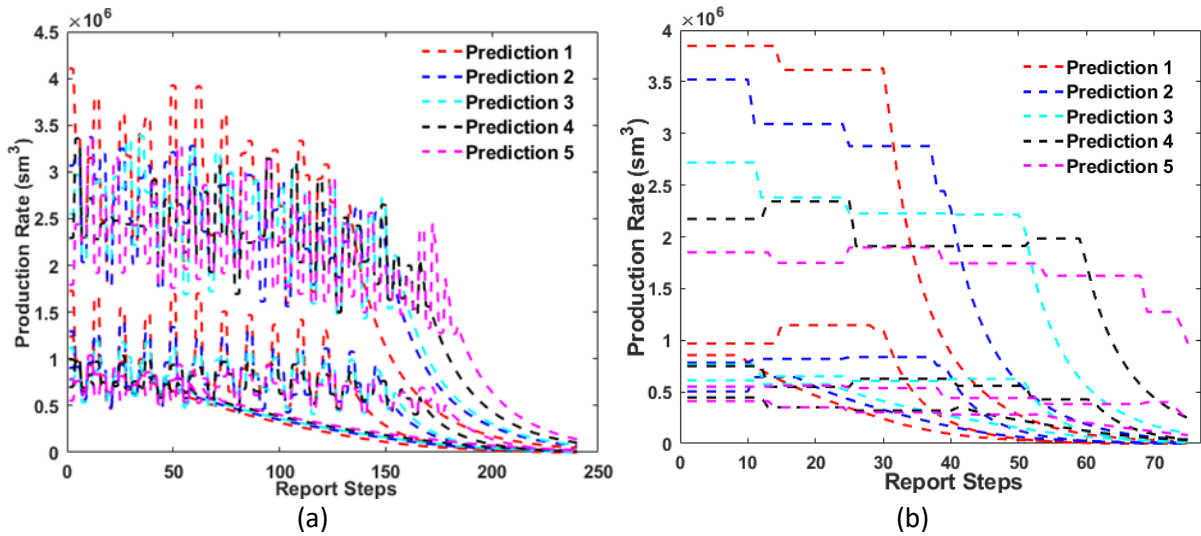
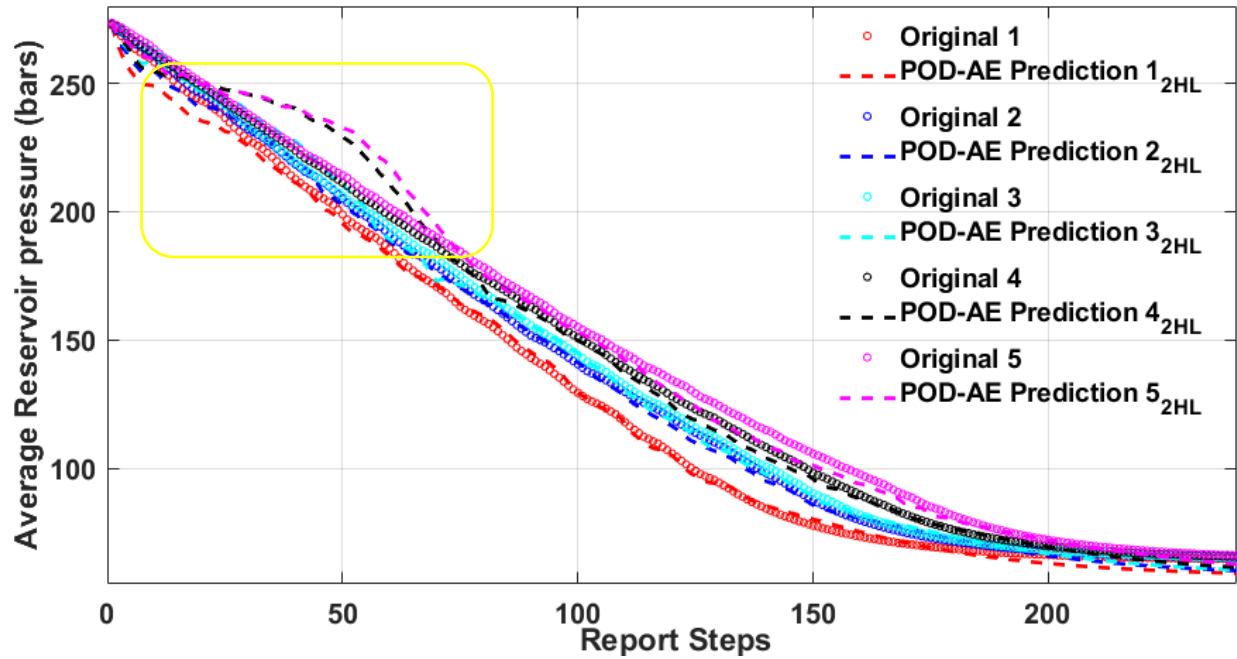


Figure 8.1: Production profiles for the (a) seasonal and (b) triennially varying prediction cases of the modified Norne gas field with shut-in wells. These production profiles are for the prediction results in Figure 8.2 & Figure 8.4 respectively, with matching line colours.

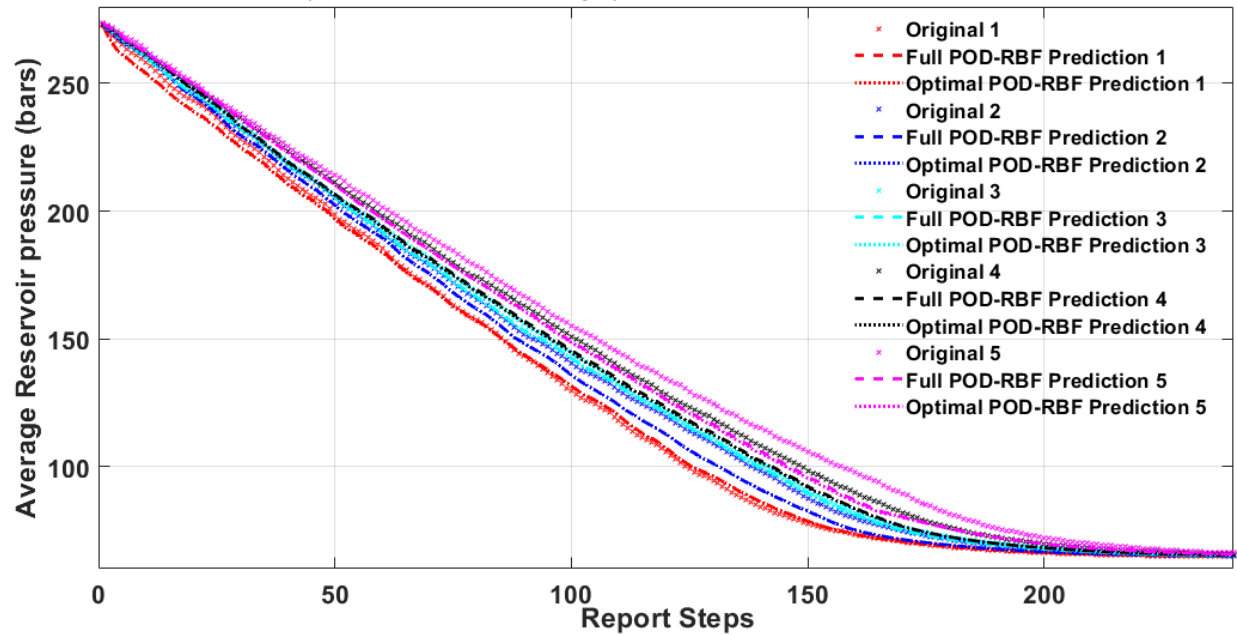
Figure 8.2 and Figure 8.4 show that both POD-AE and POD RBF models give satisfactory predictions and capture the pattern of the reservoir's pressure distribution. This is emphasised in Figure 8.3, where the percentage error in the predictions from both models were in most cases less than 10%. POD-RBF predictions from the full rank and optimal rank POD basis functions (Figure 8.2b) are similar, with overlapping results, and they gave better estimates than the POD-AE results in (Figure 8.2a). POD-RBF results are also consistent in performance throughout the production lifecycle.

In Figure 8.2a, the POD-AE model with 2 hidden layers (HLs) was initially unstable, especially for predictions 4 and 5, but gave better predictions after 75 report steps. This period of unstable prediction is shown in the yellow rectangle in Figure 8.2a. In Figure 8.4a, the POD-AE models with 2 and 5 hidden layers also show some unstable predictions, especially for predictions 4 & 5, however, they give a reasonable estimate of the pattern of pressure depletion in the reservoir. Figure 8.2a and Figure 8.4a show results from the POD-AE model with 2 and 5 hidden layers but similar results were

obtained for models with 3 and 4 hidden layers, as they also satisfactorily captured the pattern of pressure depletion in the reservoir with slightly less accurate estimates.



(a) POD-AE predictions of the average pressure distribution in the Norne field



(b) POD-RBF predictions of the average pressure distribution in the Norne field

Figure 8.2: Predictions of the average pressure distribution in all grid cells of the modified Norne field when trained with seasonally varying production rates and (a) the POD-AE model with 2 HLs or (b) POD-RBF models with the full and optimal number of POD basis functions. Predictions from both POD-AE and POD-RBF are generally satisfactory with minute errors and especially for predictions 1 to 3. In (a), the POD-AE with 2HLs gave unstable predictions for the 4th and 5th predictions in the first 75 report steps encircled with a yellow rectangle. Also, predictions with both POD-RBF, i.e. (b), look identical, overlapping each other in most cases, and they give more consistent and stable predictions throughout the production lifecycle than the POD-AE prediction in (a).

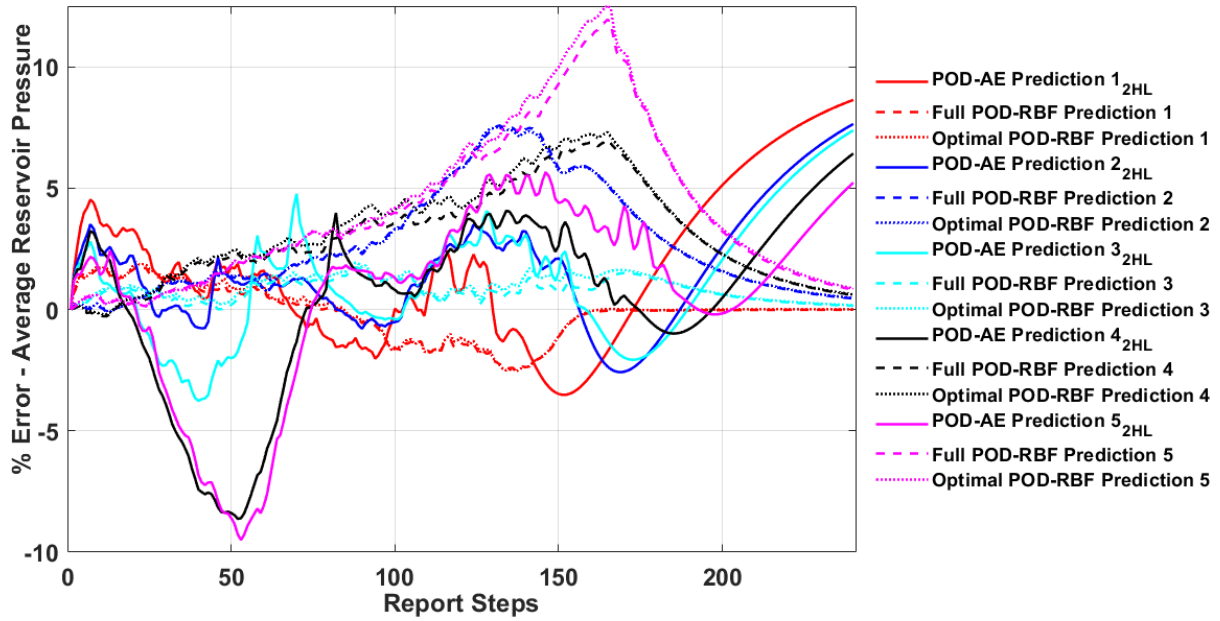
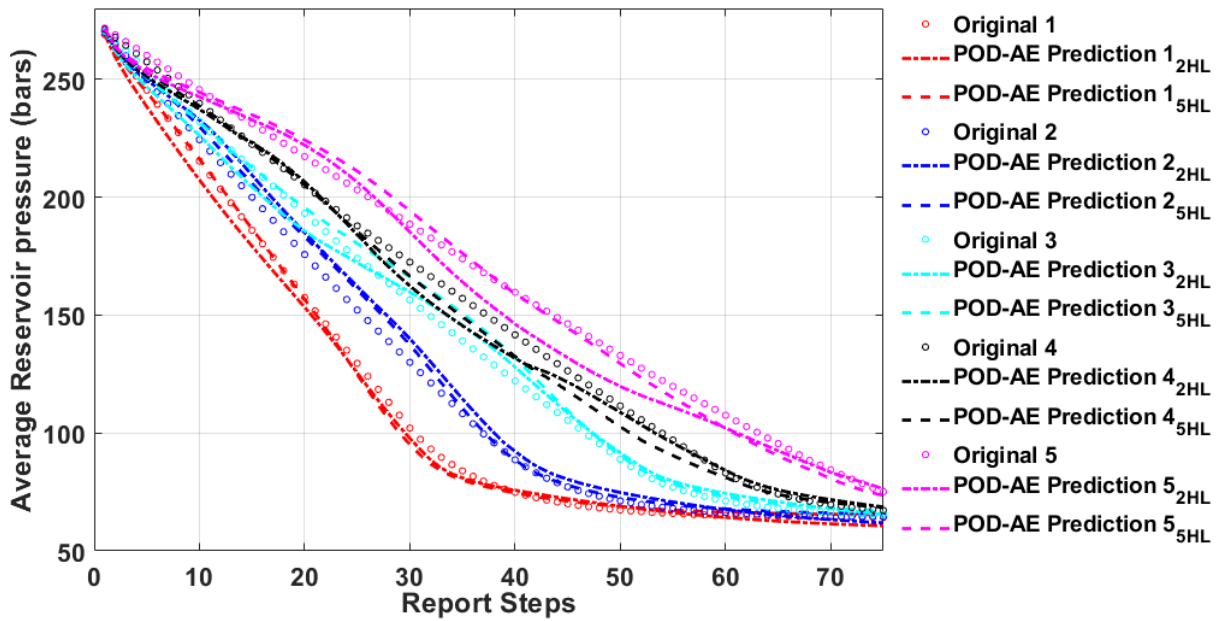
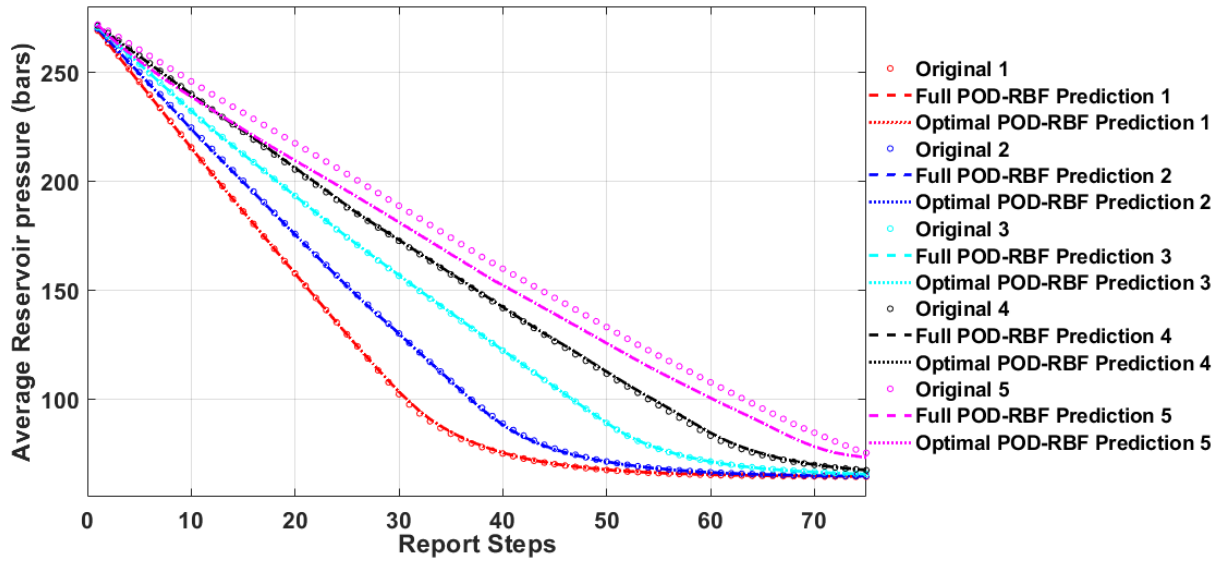


Figure 8.3: Error plot for the predictions of the pressure distribution in the modified Norne gas field using POD-AE model with 2 HLs as well as POD-RBF models with full and optimal (99.9% energy criterion) number of POD basis functions. The POD-AE model has the highest errors, especially for predictions 4 and 5, while the errors from both POD-RBF models have less variance and lower values. The full rank POD-RBF model has similar or lower errors than the optimal POD-RBF model.



(a) POD-AE predictions of the average pressure distribution in the Norne field



(b) POD-RBF predictions of the average pressure distribution in the Norne field

Figure 8.4: Predictions of average pressure distribution in all grid cells of the modified Norne gas field when trained with the triennially varying production profile. Predictions from (a) POD-AE and (b) POD-RBF models satisfactorily match those from ECLIPSE. The POD-RBF solutions, i.e. (b) are more consistent and accurate than the results in Figure 8.4a. Figure 8.4b also shows that the results from both POD-RBF models are identical.

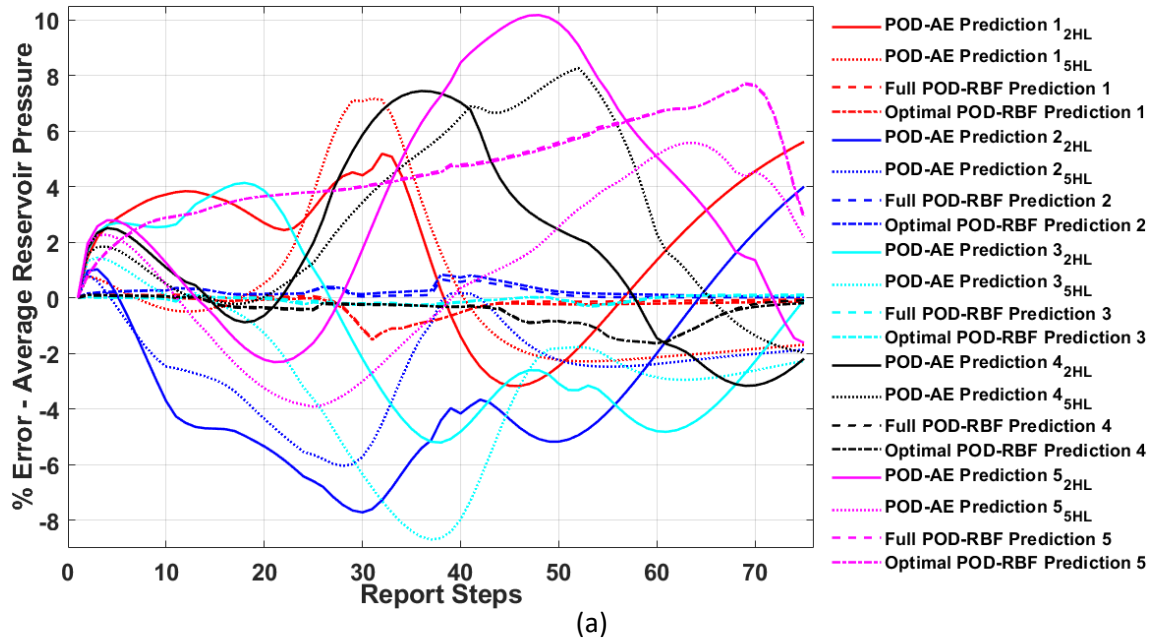


Figure 8.5: Percentage error in the predictions of the average pressure distribution in all grid cells of the Norne gas field when trained with a triennially varying production profile (Figure 8.1b). Both POD-RBF and POD-AE models give satisfactory predictions with less than 10% error. The POD-RBF solutions are more consistent and accurate (mostly less than 2% error). Results from the full and optimal rank POD-RBF models are mostly identical.

As earlier mentioned, most of the POD-AE predictions in Figure 8.2 and Figure 8.4 were unstable at some point during the 20 years production period that was modelled. This can also be seen in the wider range of error values for the POD-AE model in Figure 8.5 (up to 10% error). The reason for these over and under-predictions of the POD-AE models is not clear from the cases that were considered and given this project's timeframe, it is an area for further research. Nonetheless, the POD-AE model

can provide engineers with a reasonable idea of the pattern of the reservoir's pressure performance. However, the error in the POD-RBF models is less than 2% in most cases. These results highlight the superiority of the POD-RBF over the POD-AE for gas field modelling, since the results from the POD-RBF are more consistent and accurate than those of the POD-AE models. The POD-RBF is also less complicated for non-specialist engineers to implement and troubleshoot for sources of error, since it does not have several hyper-parameters or require fine-tuning of the hyper-parameters like most machine learning methods.

8.2 CASES WITHOUT SHUT-IN WELLS

In these examples, the production profiles have been generated such that no well gets shut-in during the production lifecycle. For the case with the triennially varying production profile, this meant that the reservoir was produced from only 2 of the 3 wells (Figure 8.6b) since the location of the third well made it prone to shutting in. Here also, the colour map for the prediction example in Figure 8.6 is the same as those used in Figure 8.8 to Figure 8.11.

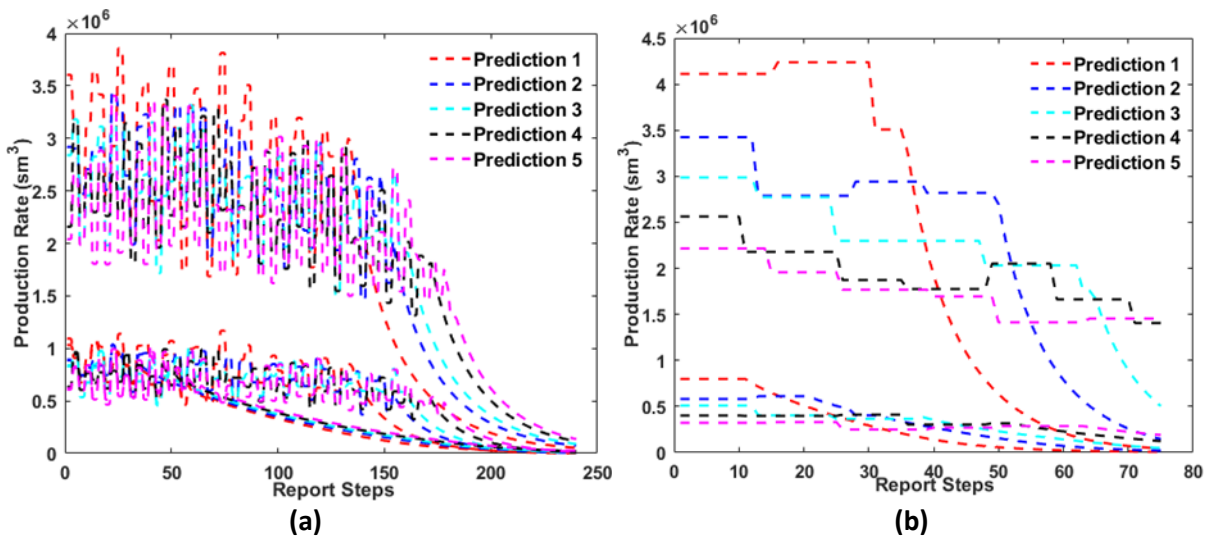


Figure 8.6: Production profiles for prediction runs of (a) seasonal and (b) triennially varying cases with no shut-in wells. These are the production profiles for the results in Figure 8.8 to Figure 8.11, with matching line colours. The seasonal profile is for production from 3 wells while the triennial profile is for production from 2 wells.

Figure 8.7 shows an example of the pressure data for all the active grid cells of the modified Norne gas field. These are the prediction 3 results from the POD-AE with 2 and 5 hidden layers as well as POD-RBF with optimal and full POD basis functions of Figure 8.6a. Figure 8.7 shows satisfactory predictions from all 4 POD-based models at the individual grid cell level. Figure 8.7 also highlights the improvement in the accuracy of the POD-RBF predictions with the full versus optimal rank of the POD basis functions. This difference is not very obvious when the average of these values are shown in Figure 8.9. Comparing the result in Figure 8.7a & b to those in Figure 8.8 also show that the POD-AE

predictions of Figure 8.8 were most accurate after the 150th and 170th report steps for the model with 2 and 5 hidden layers respectively.

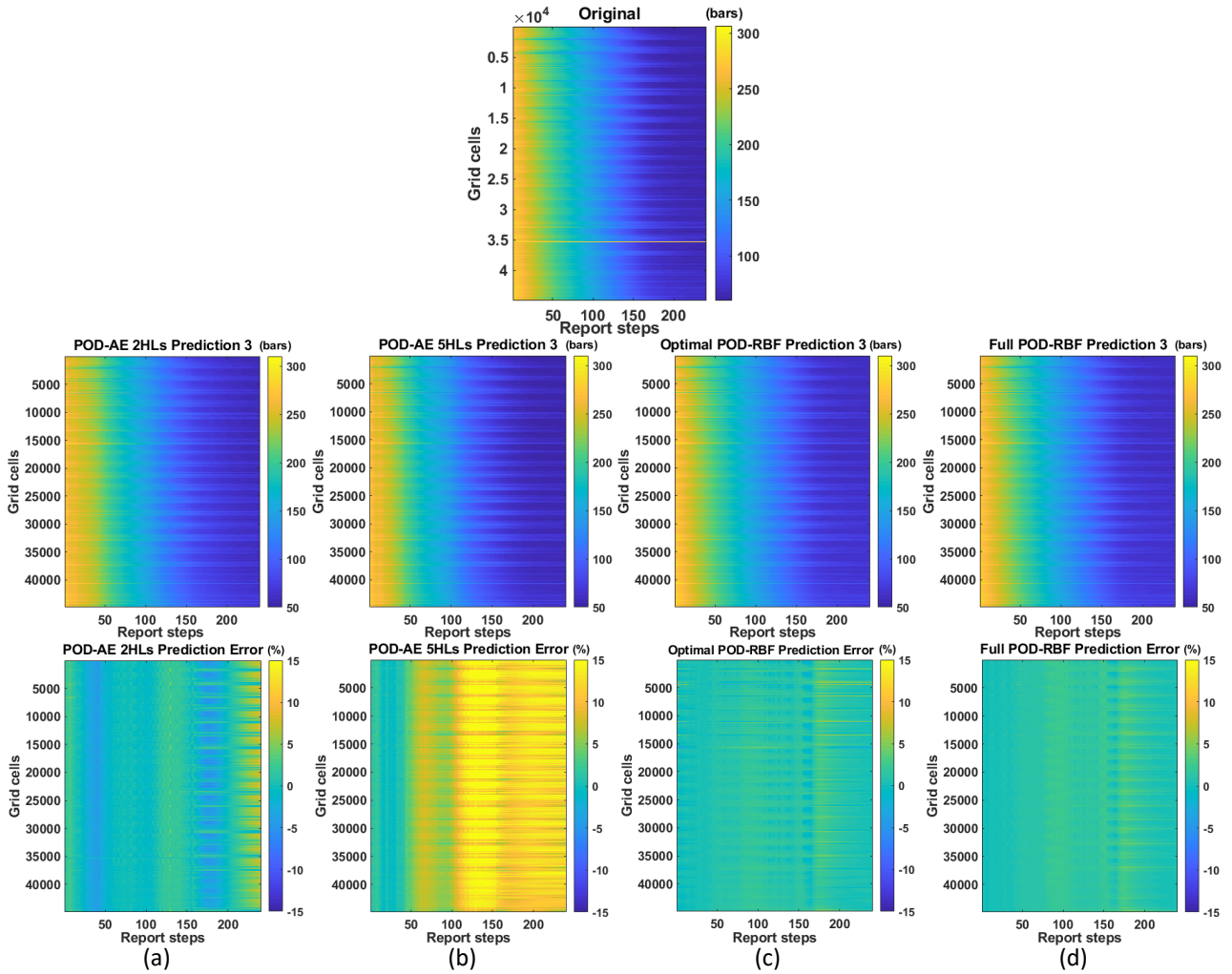


Figure 8.7: Predictions of the pressure distribution in all grid cells of the modified Norne gas field for prediction 3 of the seasonally varying production profile with no shut-in wells. All the models (POD-AE with 2 and 5HLs as well as POD-RBF with optimal and full POD basis functions) gave satisfactory predictions of the reservoir pressure distribution in over 44000 active grid cells. Predictions from the POD-RBF models (c & d) were better than those from the POD-AE models (a & b). The improvement in predictions with the full rank POD-RBF model is also obvious in these results.

Figure 8.8 and Figure 8.9 show results from POD-AE and POD-RBF predictions of the modified Norne field simulation cases with the seasonally varying production profile of Figure 8.6a. They have been plotted separately for clarity. In Figure 8.8, the POD-AE models with 2 and 5 hidden layers gave predictions that were not always accurate but matched or captured the pattern of pressure depletion in the reservoir. Predictions from the POD-AE model with 5 hidden layers mostly under-predicted the pressure distribution in the reservoir but evenly, while predictions from the POD-AE with 2 hidden layers were slightly unstable in the initial 80 report steps, especially for predictions 4 and 5. The POD-RBF predictions in Figure 8.9 show more stable and satisfactory predictions from the full and optimal rank POD-RBF, their results are similar and are only distinguishable from Figure 8.7c & d. Figure 8.10

shows the percentage error in these predictions and confirms that the POD-RBF models with less than 10% average error are better than the POD-AE model with more than 10% errors.

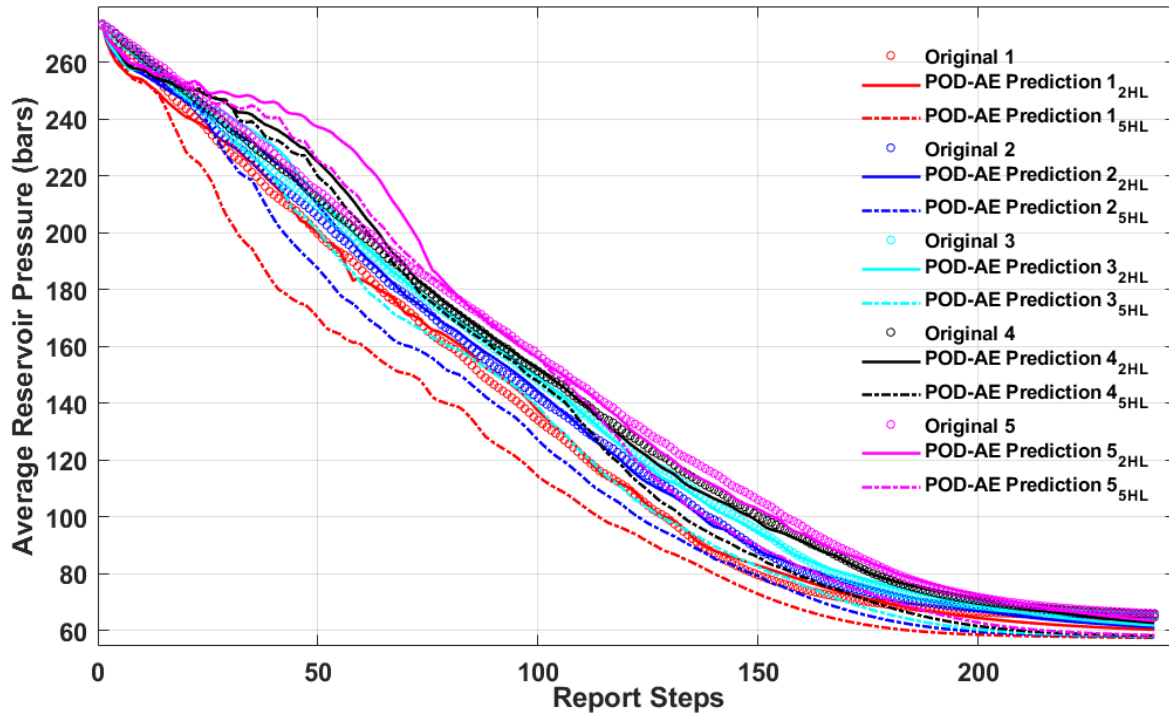


Figure 8.8: Predictions of the average pressure distribution in all grid cells of the modified Norne gas field when trained with seasonally varying production rates and POD-AE models with 2 and 5 HLs. The POD-AE models gave predictions that satisfactory matched the pattern of pressure depletion in the reservoir, with some over- and under-predictions for 2 and 5 HLs respectively.

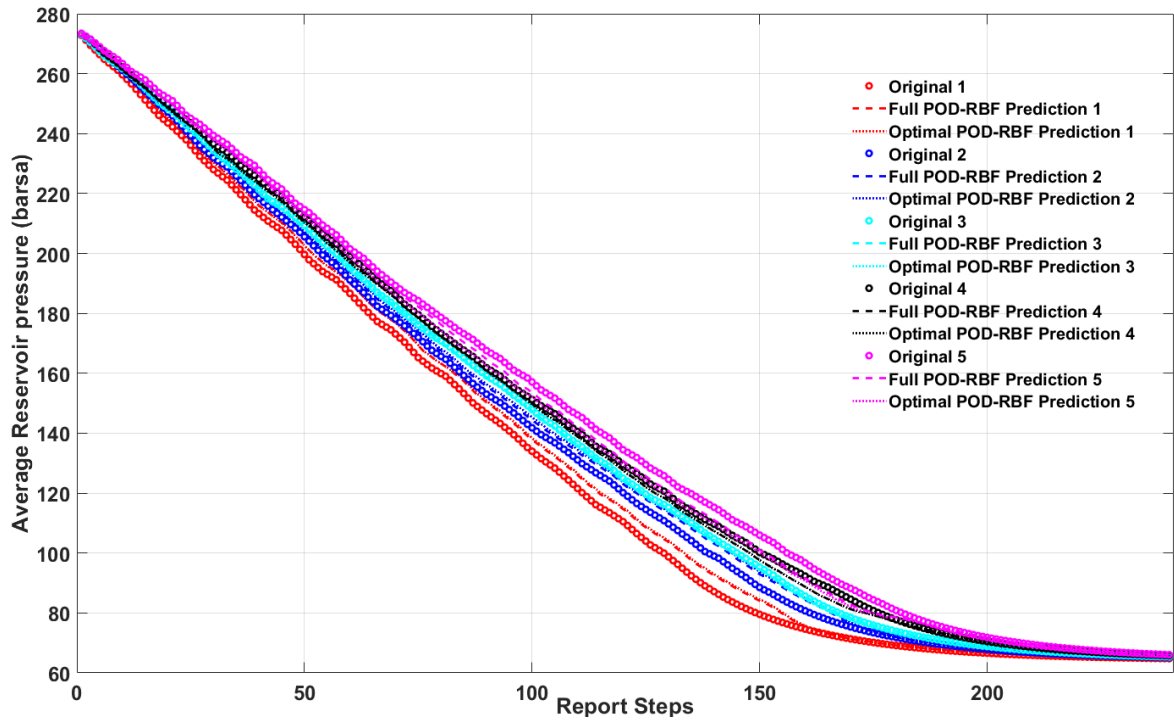


Figure 8.9: POD-RBF predictions of the average pressure distribution in all grid cells of the modified Norne gas field when trained with seasonally varying production rates. The results from the full and optimal rank POD-RBF models are similar and consistent throughout the prediction duration.

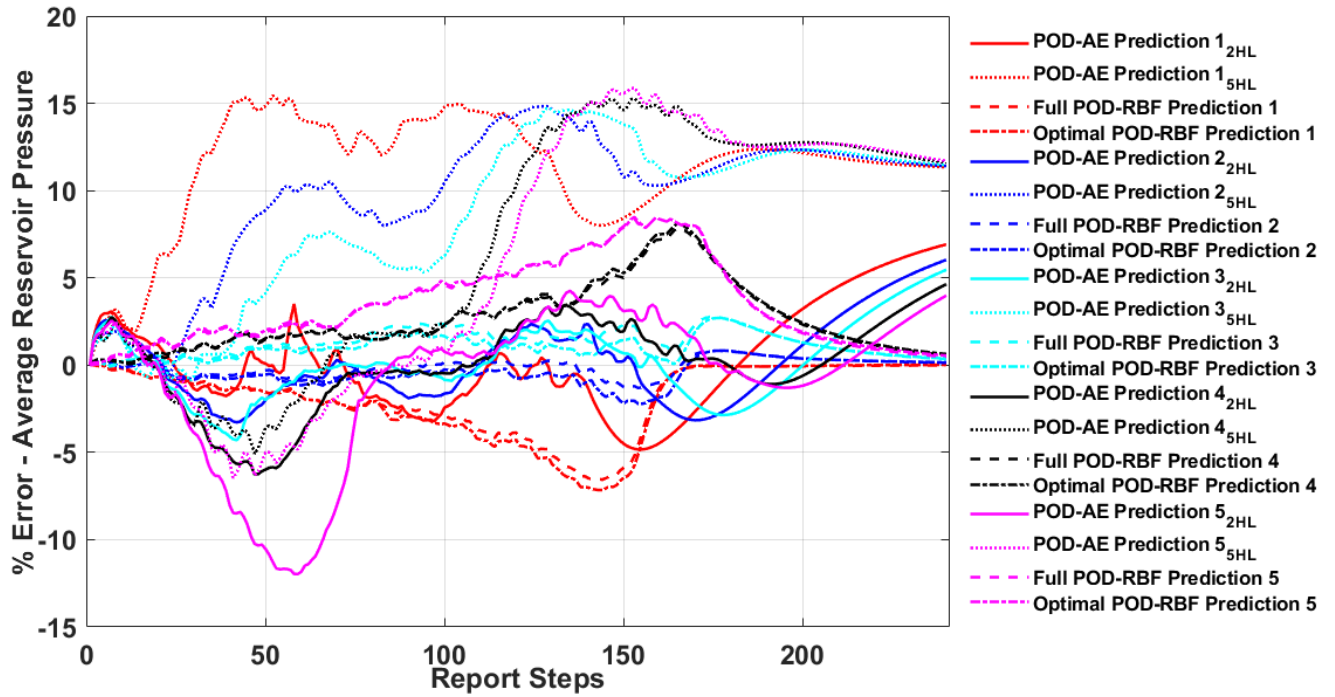


Figure 8.10: Percentage error in the predictions of the average pressure distribution in all grid cells of the Norne gas field when trained with seasonally varying production profiles for POD-AE models with 2 and 5 HLs as well as the full and optimal POD-RBF models. The POD-AE models with 2 and 5 HLs over and under predict the field's pressure distribution, respectively. The full and optimal rank POD models give similar results with lower prediction errors.

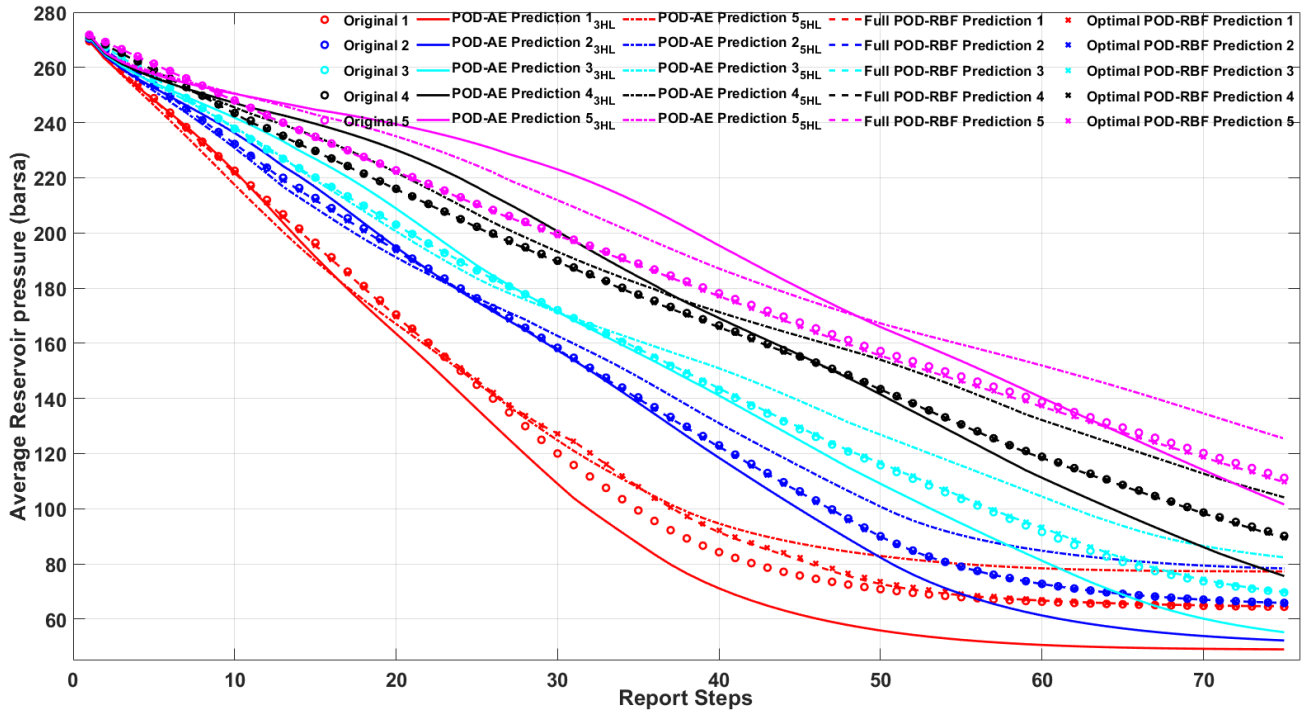


Figure 8.11: Predictions of average pressure distribution in the modified Norne gas field with production rates varying after at least 2.5 years. The POD-AE models are trained with 3 and 5 HLs and satisfactorily capture the pattern of pressure depletion in the gas field while predictions from the POD-RBF models are more accurate, with similar results from the full and optimal rank POD basis function POD-RBF NIROMs.

Figure 8.11 shows results from both POD-AE and POD-RBF models for the triennially varying production profile of Figure 8.6b. The POD-RBF results were more consistent than the results from the

POD-AE models. Similar to the other cases seen so far, both POD-RBF results are also identical and more accurate than the results from the POD-AE models. The percentage errors for the predictions in Figure 8.11 are shown in Figure 8.12 where it can be seen that both POD-RBF models have less than 2% errors for all predictions except prediction 1, while the POD-AE models had higher errors up to |23%|. Figure 8.12 also shows that the POD-RBF models were more stable, with a consistent error band than the POD-AE models which had a more inconsistent spread of errors. The consistency and accuracy of the POD-RBF method make it a better technique for gas field modelling than the POD-AE model.

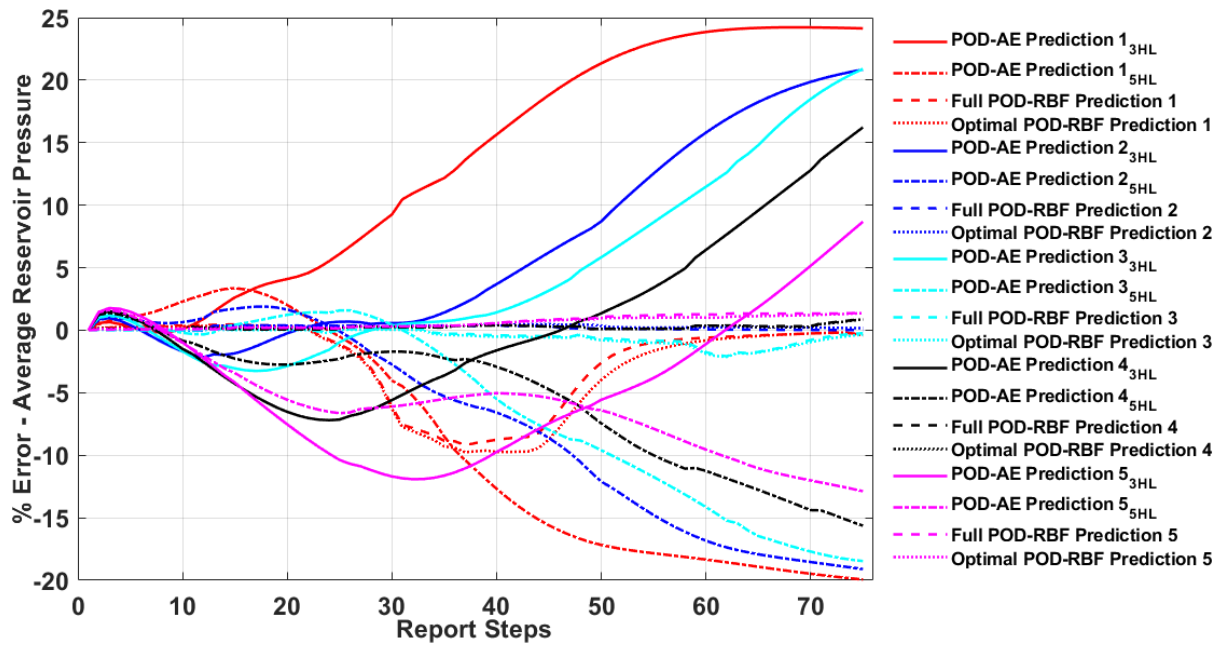


Figure 8.12: Percentage errors in the predictions of the pressure distribution in the modified Norne gas field using POD-AE models with 3 and 5 HLs as well as POD-RBF models with full and optimal (99.9% energy criterion) number of POD basis functions. POD-AE models have the highest errors while the errors from both POD-RBF models are less than 2%, except for prediction 1. The full rank POD-RBF model has almost the same or slightly lower errors than the optimal POD-RBF model.

8.3 ERROR ANALYSIS

In this section, the accuracy of predictions from both POD-AE and POD-RBF NIROMs are compared with the metrics: average error in pressure predictions and average RMSE of predictions (Equation 3.20). NIROM predictions are compared to the results of the same simulation cases in ECLIPSE. The results shown in this section are an average for the 5 prediction examples that were evaluated for each case. Where a model over- or under-predicts different prediction examples of the same case, the average error is not a true reflection of the model's performance since negative and positive values cancel out. The RMSE is a better and more consistent assessment of the errors than the average error. This is shown in Figure 142 and Table 28, where for example the POD-AE 2 hidden layers model that visibly shows the highest errors in Figure 8.2 of the seasonally varying case with shut-in wells shows

the lowest average error value of 0.77bars and the highest RMSE value of 5.03bars. For this reason, the prediction cases in Table 28 are compared based on their RMSE values.

Table 8.1: Average errors and RMSEs for pressure prediction cases (in bars). POD-RBF models show lower RMSEs than POD-AE models. And POD-RBF cases without shut-in wells have lower RMSEs (are more accurate) than cases with shut-in wells.

Model	Cases with shut-in wells							
	Seasonally varied				Non-seasonally varied			
	POD-AE 2HL	Full NIROM	Optimal NIROM		POD-AE 2HL	POD-AE 5HL	Full NIROM	Optimal NIROM
Average Error	0.77	2.52	2.72		1.28	-0.32	1.27	1.27
Average RMSE	5.03	3.77	4.14		5.83	4.66	3.18	3.36
Model	Cases without shut-in wells							
	Seasonally varied				Non-seasonally varied			
	POD-AE 2HL	POD-AE 5HL	Full NIROM	Optimal NIROM	POD-AE 3HL	POD-AE 5HL	Full NIROM	Optimal NIROM
Average Error	-0.74	1.19	-0.07	0.05	1.96	-6.73	-0.17	-0.21
Average RMSE	5.05	11.81	2.25	2.90	10.06	9.25	1.88	2.86

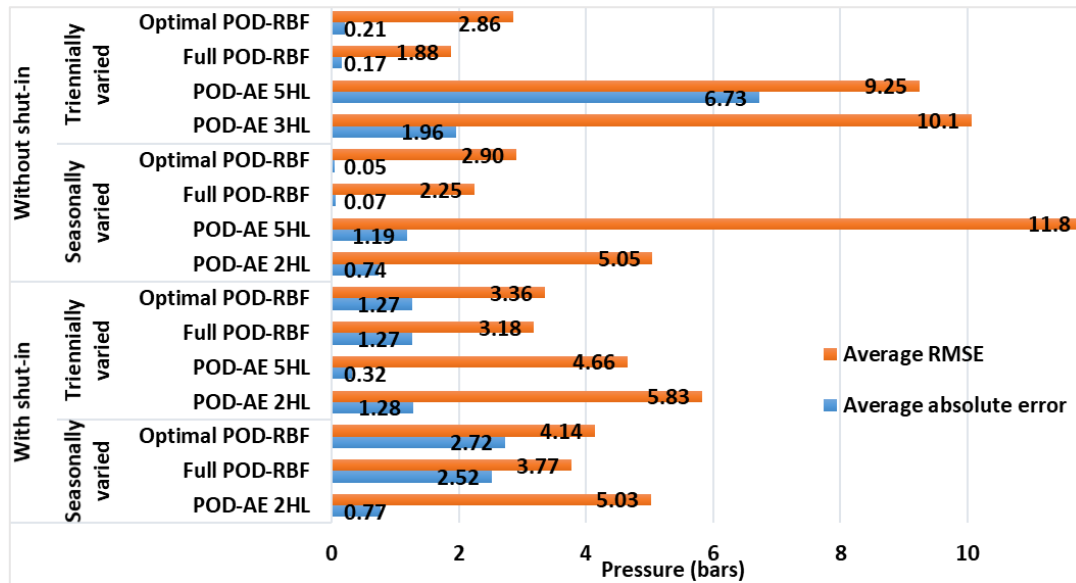


Figure 8.13: Average absolute errors and RMSEs for predictions of the pressure distribution in the modified Norne field using POD-AE and POD-RBF NIROMs. Prediction cases with shut-in wells are less accurate than those without shut-in wells. POD-RBF predictions are more accurate than POD-AE predictions.

The results in Table 8.1 agree with and further highlight the results shown in the earlier sections, i.e. the RMSEs for the POD-AE models are always higher than those for the POD-RBF models. In all the cases that were evaluated, POD-RBF models have less than 5bars average RMSE while the POD-AE models have as high as 11.8bars average RMSE. The RMSE values in Table 8.1 also show the slim margin of error between POD-RBF predictions with the full rank POD data and using the truncated/

optimal POD basis functions. The POD-RBF with full basis functions was always more accurate than the POD-RBF with the optimal number of POD basis functions. The POD-RBF models also gave better predictions in the cases without shut-in wells than in the cases with shut-in wells, as seen in their lower RMSEs. This shows that although the POD-RBF models could approximate the behaviour in the (more complicated) shut-in conditions they did not accurately predict this performance. The chart in Figure 8.13 shows the absolute values of the average errors and average RMSEs of the result in Table 8.1, where it is shown that the POD-RBF models are consistently lower than those of the POD-AE. Figure 8.13 also shows that in the cases with shut-in wells, the errors in POD-AE predictions are almost double or more than the errors from POD-RBF predictions. This further reflects the superior accuracy of the POD-RBF models in comparison to those of the POD-AE models.

8.4 TIME ANALYSIS

From Table 8.2 and Figure 8.14, it is seen that implementing gas field simulations with the POD-RBF models is faster than with POD-AE models. The POD-RBF implementation for the triennially varying production profiles was at least an order of magnitude faster than for the data-intensive seasonally varying production profile cases. This was expected as the seasonally varying production profile has 3 times more data than the triennially varying profile dataset. Also from Table 8.2 and Figure 8.14, it is shown that the full rank POD-RBF model has the slowest implementation time for the actual predictions of the unseen cases. Nevertheless, the shorter time taken to train the full rank POD-RBF model makes its entire implementation time significantly faster than that of the POD-AE models. In other words, it takes hours to train the POD-AE model as opposed to the minutes taken to implement the full rank POD-RBF model.

Table 8.2: Overview of the time taken to implement POD-RBF and POD-AE models for gas field simulations in comparison to ECLIPSE. POD-RBF models are faster than POD-AE models, but the optimal POD-RBF is the only model that is faster than ECLIPSE for the seasonally varying production profile cases.

Model (time unit is seconds)	Seasonally varied profile			Triennially varied profile		
	POD-AE	Full POD-RBF	Optimal POD-RBF	POD-AE	Full POD-RBF	Optimal POD-RBF
Time to train model	13523	235	12.8	1346	6.94	0.99
Time for predictions	25.8	220	0.11	3.89	5.63	0.02
Total model time	13549	455	12.9	1350	12.7	1.01
ECLIPSE runtime	303	303	303	303	303	303
Speed-up	0.02	0.67	23.5	0.22	24.1	300

The optimal rank POD-RBF model was the fastest model considered in this chapter, with speed-ups of 1 and 2 orders of magnitude for the seasonal and triennially varying production profiles respectively.

Given that the prediction examples in this chapter were based on 50 training runs, which is quite high from a POD-RBF NIROM perspective, the full rank POD-RBF NIROM did not yield any speed-up in the simulation of the seasonally varying production scenarios. However as shown in chapters 6 and 7, also in Samuel and Muggeridge (2020) ^[125], NIROM predictions with fewer training runs, and fewer report steps/ temporal resolutions, are significantly faster and give adequate predictions of gas reservoir performance. This is linked to the amount of time required for POD reduction since the matrix inversion of smaller datasets is less time-consuming. So as expected, predictions from the smaller dataset of the triennially varying production profiles are the fastest for both POD-AE and POD-RBF models.

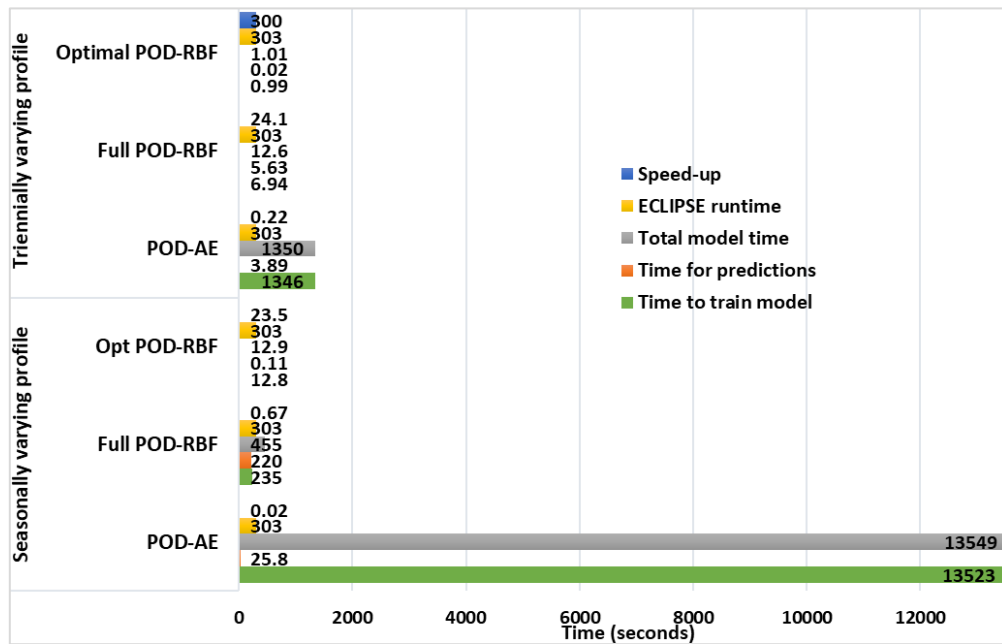


Figure 8.14: The breakdown of the time taken to train the POD-AE and POD RBF models that were implemented for predicting the pressure distribution in the modified Norne gas field, trained with 50 training runs. POD-AE is slower than ECLIPSE for both the seasonal and the triennially varying production profiles, while the full POD-RBF model is only faster than ECLIPSE for the triennially varying production profile cases.

SUMMARY

This is the first time that the POD-AE approaches have been applied to full field simulation, including pressure and saturation distribution prediction. In this chapter, POD-AE and POD-RBF models were analysed for modelling the performance of the heterogeneous Norne gas field for seasonal and triennially varying production profiles, with and without shut-in wells. Both models were applied to forecasting the pressure distribution in all the active grid cells of the Norne gas field. In both cases, the models were trained with 50 training runs. The POD-AE model was analysed for the prediction of 5 new cases with 1 to 5 dense encoders and the same number of decoder layers, and the POD-RBF

models were analysed for the same predictions but with a full and an optimal rank of POD basis functions.

For the POD-AE model, it was shown that models with different numbers of hidden layers are suited to making predictions for different production cases, as such no one autoencoder model suffices for forecasting different reservoir scenarios, i.e. production cases with or without shut-in or with seasonal or triennially varying production profiles or well controls. However, in this work, POD-AE models with 2 and 5 hidden layers were the most consistent models for obtaining satisfactory predictions of the Norne field's spatial pressure distribution for the cases with and without shut-in wells. Hence POD-AE is a potential NIROM for gas field modelling if its training time can be reduced.

In comparison, the POD-RBF models give better prediction results than the POD-AE models, with lower root mean square error values as well as stable predictions throughout the simulation lifecycle. The POD-RBF models were also significantly faster than the POD-AE models. It took over 3 hours to train the POD-AE models for the seasonally varying production profile, and less than ten minutes for the same cases with the POD-RBF models. In most cases, the optimal rank POD-RBF models (from data that represents 99.9% of the full POD basis functions) gave results with similar accuracy as the predictions with the full rank POD-RBF model but with a speed-up of an extra order of magnitude. POD-RBF predictions with the full rank POD basis functions were slightly more accurate than those from the optimal POD-RBF model, but this is only noticeable when the pressure in each grid cell is plotted side by side. Overall the training time for the POD-AE and the full rank POD-RBF models were prolonged because of the size of the training data, up to 4 orders of magnitude seconds for POD-AE. This again confirms the benefit of implementing NIROMs with the least sufficient amount of data (fewer numbers of training runs and/ or report steps), to obtain fast and equally accurate predictions. Table 8.3 is an overview of the key findings in this chapter.

From the cases considered in this work, POD-RBF NIROM is easier to set up, computationally faster and more accurate for gas field modelling than POD-AE and other machine learning based NIROMs. Machine learning based methods require lengthier training time and the process of coming up with a model with the most suitable combination of hyper-parameters that suit different production scenarios relies on several trials and errors. For a reservoir engineer without ML expertise or who is very busy, developing a satisfactory POD-AE model for field development planning may not be worth the effort and time. The POD-RBF NIROM would give more accurate and consistent results in a shorter time.

Table 8.3: A summary of the prediction cases and key findings in this chapter.

Cases Considered	Summary and Findings
• 2 NIROM methods were evaluated: ○ POD-RBF with full and optimal number of POD basis functions ○ POD-AE with 1 to 5 hidden layers • Time-varying prediction examples: ○ Seasonal production profiles ○ Triennial production profiles ○ Cases with shut-in wells and ○ Cases without shut-in wells • Mixed production controls: ○ Production rate with BHP or THP limit • 1 real gas reservoir evaluated: ○ Modified heterogeneous Norne gas field with 50 training runs	Key Findings
	• Demonstrated the POD-AE's predictive capacity for gas field modelling, based on different time-varying production constraints • No single POD-AE model is sufficient for modelling different production constraints • The prolonged time required to train POD-AE models makes their speed inefficient for gas field modelling • POD-RBF models give more consistent predictions than POD-AE models, and they are easier to build
	General Findings
	• Compared the performance of POD-AE to POD-RBF • POD-RBF is faster and more accurate than POD-AE • Pressure distribution predictions with the optimal POD-RBF are similar to those with the full rank model, and they are at least one order of magnitude faster.

CHAPTER 9: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

The aim of this research was to come up with a fast, straightforward and accurate non-intrusive reduced order model for gas reservoir modelling, benchmarked with a commercial black oil simulator, and then to evaluate the usefulness of the non-intrusive reduced order model. The non-intrusive reduced order model was intended to estimate the changing pressure distribution (and saturation distribution if appropriate) over time as well as the production rates at the wells. Therefore, a proper orthogonal decomposition-radial basis function non-intrusive reduced order model was implemented based on Xiao et al. (2016) ^[38] and Kostorz et al.'s (2019 to 2021) ^[37, 99, 100] implementations but for ECLIPSE's simulation results of gas reservoirs. Given that the previous implementation of the non-intrusive reduced order model was for constant parameter cases, that model has been improved upon in this work to accommodate varying parameter scenarios. A proper orthogonal decomposition-autoencoder non-intrusive reduced order model was also designed and implemented for modelling gas reservoir performance in time-varying production scenarios.

The performance of the two non-intrusive reduced order models were subsequently evaluated with four (4) gas reservoir models (including 2 real field cases) and for various production dynamics, with data of up to eight orders of magnitude. The four gas reservoirs are a homogeneous synthetic single well model; a heterogeneous gas reservoir with similar properties to the Norne field, with 3 wells; a water coning gas reservoir with one producer; and a sector model of a real gas reservoir (Field X), provided by Woodside Petroleum. The gas production scenarios that were modelled with both non-intrusive reduced order models, included constant and varying production rates and pressure limits, with different aquifer types and mixed production controls. With different combinations of the listed gas fields and production scenarios or dynamics, both non-intrusive reduced order models were evaluated for forecasting the gas fields' pressure and gas saturation distribution as well as more than 20 different production variables such as gas and water production, aquifer influx rates and cumulative production, in addition to the fields' and wells' pressure at different time intervals. Each production scenario or dynamics was evaluated for at least 20 years production lifecycle.

The evaluation of the non-intrusive reduced order models was implemented in three stages of complexity: estimating specific simulation cases, predicting new unseen simulation cases that had constant underlying production dynamics and predicting new unseen simulation cases that had time-varying and mixed underlying production dynamics as shown in Figure 9.1 below. The first stage was a proof of concept that the non-intrusive reduced order model can be implemented and is applicable for gas reservoir modelling. This stage also involved determining the most suitable radial basis function(s) for gas reservoir modelling as well as understanding if and how the number of proper

orthogonal decomposition basis functions affect the non-intrusive reduced order model's performance. In the second stage of evaluating the non-intrusive reduced order model(s), the model's predictive capacity was evaluated for new unseen simulation cases that had constant underlying production. The ideal number of simulation runs for training the proper orthogonal decomposition-radial basis function model was also analysed. In the last phase of this work, the non-intrusive reduced order model's predictive capacity was evaluated for practical time-varying production rates with mixed production controls and shut-in behaviours.

The findings from this work are discussed in the following sections.

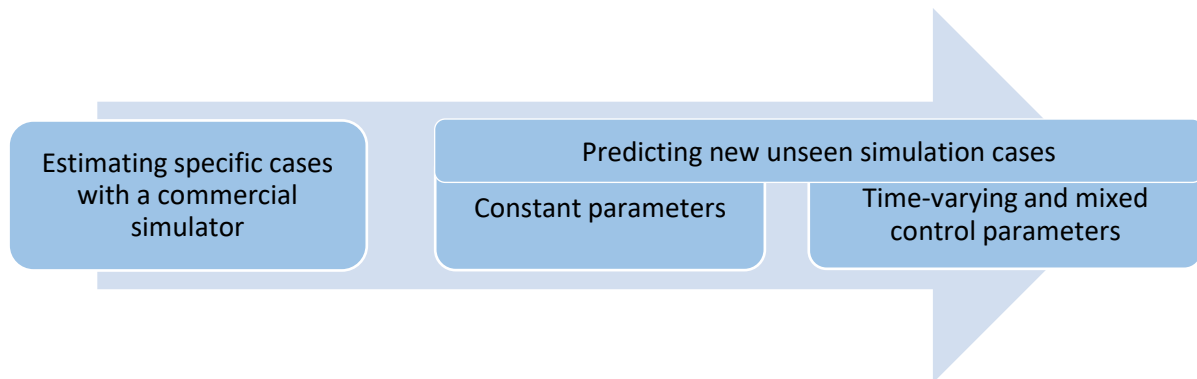


Figure 9.1: Summary of the 3 stages that the non-intrusive reduced order model was evaluated for, a) estimating specific simulation cases, and predicting new simulation cases with b) constant parameters c) time-varying and mixed control parameters.

9.1 SPECIFIC FINDINGS

In the first stage of this work, it was shown that the non-intrusive reduced order model can estimate the evolution of grid distributed properties (pressure and saturation) using 99.9% of the full set of proper orthogonal decomposition basis functions, determined with the energy criterion proposed by Xiao et al. (2016) ^[38]. The estimation of time series production variables require more proper orthogonal decomposition basis functions. This was initially shown in chapter 5 but was also indicated in the results in chapters 6 and 7. It was shown that the non-intrusive reduced order model required the most number of basis functions to satisfactorily model production variables with negligible change in magnitude over time and time-varying production profiles.

Furthermore, the linear radial basis function was found to be the most consistent and suitable interpolation function for modelling gas reservoirs. This was also highlighted in chapter 5, after analysing 7 different interpolation functions. Hence, subsequent applications of the proper orthogonal decomposition – radial basis function method in the other two phases of this work were based on the linear radial basis function.

From the results of the constant parameter prediction cases in chapter 6, it was highlighted that 10 training runs is a good balance between accuracy and speed of the non-intrusive reduced order

model's performance. It was demonstrated that the non-intrusive reduced order model can be adequately trained with as few as 5 training runs, although increasing the number of training runs could give better predictions that were more time consuming to implement. In addition, it was shown that as the number of training runs used to implement the non-intrusive reduced order model increases, only those predictions that were within similar range of parameters as the training runs were improved. Predictions outside the range of training runs (cases of extrapolation) only got worse as the number of training runs increased.

Furthermore, it was established that the accuracy of the non-intrusive reduced order model predictions was generally better during interpolation and worse during extrapolation. When the proper orthogonal decomposition – radial basis function model was applied to predicting reservoir performance for cases with parameters or simulation duration outside the range that it was trained on, it gave worse predictions. This performance of the non-intrusive reduced order model was consistent in all the chapters where the NIROM was applied to predicting reservoir performance, including during the time –varying production cases.

In chapter 8, the satisfactory predictive capacity of the non-intrusive reduced order models for practical time-varying production rates cases, with mixed production controls and shut-in well behaviours, was demonstrated as achieved with the new renormalisation method that is proposed in this work. This renormalisation approach aided the predictive capacity of both the proper orthogonal decomposition-radial basis function model and the proper orthogonal decomposition-autoencoder model for gas reservoir performance modelling. These improved non-intrusive reduced order models were therefore evaluated for modelling gas reservoir performance with two types of time-varying production profiles including the seasonally varying profile that is more typical of industry patterns.

In comparison to the proper orthogonal decomposition-autoencoder model, the proper orthogonal decomposition-radial basis function model gives more accurate and consistent results, and it is easier to implement. The proper orthogonal decomposition-autoencoder can be applied to modelling the performance of gas fields with acceptable accuracy, but its training time was more than 40 times the time taken to run the simulation with the commercial simulator. Having tested the autoencoder model with one to five encoder and decoder layers, it was also found that no one autoencoder model was sufficient to model the different production dynamics that were considered, even though the proper orthogonal decomposition-autoencoder with 2 and 5 hidden layers were the more consistent models.

The autoencoder, like other machine learning techniques, relies on several hyperparameters that are tuned to achieve the best model set-up. These hyperparameters include the activation functions for each hidden layer, number of hidden layers, epoch, batch size, error metrics, loss function, optimiser,

optimiser learning rate as well as provision of validation data. Each hyperparameter value can vary when training the autoencoder for different production scenarios. However, for the proper orthogonal decomposition radial basis function method, the only parameters that need to be tuned/defined in advance are the number of proper orthogonal decomposition basis functions and the type of interpolation function, for only one hidden layer. Once set up, the proper orthogonal decomposition-radial basis function model can be used for modelling different scenarios of the same or different reservoirs (or dynamic systems) as a standalone program that needs little or no re-tuning every time.

9.2 GENERAL FINDINGS

For the time varying parameter cases, which constitute the main novelty of this work, it was shown that the range of parameters for training versus prediction simulation runs was more important than the number of training runs used to implement the non-intrusive reduced order model. The proper orthogonal decomposition-radial basis function model predictions were more accurate when it was trained with a similar range of production parameters/ scenarios as the training runs, i.e. interpolation. During the prediction of mixed production profile cases, the non-intrusive reduced order model may not accurately show the onset of changing production controls, but overall it gives satisfactory predictions with over 97% accuracy.

A comparison of the full and optimal rank proper orthogonal decomposition-radial basis function models showed that the full rank model is more accurate than the optimal rank model, although the difference between both models is almost negligible. In addition, the optimal rank proper orthogonal decomposition-radial basis function model is an order of magnitude faster than the full rank model. Therefore the optimal rank proper orthogonal decomposition-radial basis function model is recommended as a suitable non-intrusive reduced order model for swiftly modelling gas reservoirs, since they give similar accuracy as the full rank non-intrusive reduced order model but with increased speed-up. The non-intrusive reduced order model was up to 4.4 times faster than ECLIPSE.

Overall, this work highlighted that the extent of speed-up that can be achieved with the non-intrusive reduced order model is dependent on the size and properties of the dynamic reservoir model. The reservoir size is reflected in the number of grid cells that the reservoir model is built with, as well as the frequency of report steps that the dynamic reservoir model is required to output simulation solutions, and the number of training runs that are used to implement the non-intrusive reduced order model. The other reservoir properties that affect the speed of the non-intrusive reduced order model include the complexity of the reservoir model (in terms of its degree of homogeneity or heterogeneity)

and the presence or absence of grid refinement, aquifer, shut-in wells or mixed production controls. In chapters 5 to 7 it is shown that less than 1% of the time taken to implement the non-intrusive reduced order models was spent on the actual non-intrusive reduced order model forecast. The majority of the remaining time is spent on extracting simulation solutions from the commercial simulator, and for proper orthogonal decomposition reduction of the training data or back projection of the model's solutions. This further reiterates the earlier recommendation that training the model with 10 runs is a reasonable balance between accuracy and speed of forecasts.

Having evaluated two non-intrusive reduced order models for different gas production scenarios, this work shows that the proper orthogonal decomposition-radial basis function model is suitable for gas reservoir modelling and can support engineers in implementing quick field performance estimates during field development planning and asset management. This work shows that proper orthogonal decomposition-radial basis function non-intrusive reduced order model can support engineers in predicting the following aspects of gas reservoir performance:

- pressure and gas saturation distribution in gas reservoirs;
- production variables, such as gas, water production and aquifer influx rates, cumulative production and pressures;
- production by gas rate and well control limits such as BHP and THP;
- constant or time-varying production controls;
- mixed production controls, with and without shut-in wells;
- in the presence of numerical and analytical aquifer influx;
- in the presence of water coning ;
- low porosity reservoirs (as low as 0.12 porosity);
- reservoir models built with Cartesian and radial grid as well as with local grid refinement or tartan grids; and
- initialisation by equilibration or enumeration.

9.3 NON-INTRUSIVE REDUCED ORDER MODEL WORKFLOW

The workflow for implementing non-intrusive reduced order model in this work is summarized below:

Generating and processing training data using proper orthogonal decomposition

- Step 1: Define the key simulation study parameters and their range of values of interest for the study. Examples include gas production rate, well bottom hole pressure, tubing head pressure, reservoir porosity and permeability.

- Step 2: Define output variables to be monitored and predicted in the simulation study such as pressure and saturation distribution or production variables like gas and water rates, gas-water ratio and well pressures.
- Step 3: Define time intervals for the output of simulation results as well as the number of training runs for training the non-intrusive reduced order model. We have shown that 10 runs are a good balance between speed and accuracy for constant parameter cases.
- Step 4: Build the base case reservoir model for the study or use an existing reservoir model.
- Step 5: Run simulation model based on all the parameters defined in step 1 and extract output results defined in step 2, within the programming interface. These are the training data.
- Step 6: Perform mean extraction on the training data, as in Equations 3.2 and 3.3
- Step 7: Perform proper orthogonal decomposition reduction on the data from step 6 above to obtain the left singular vectors matrix (Equation 3.4).
- Step 8: Implement energy criterion on left singular vectors matrix using Equations 3.5 and 3.6, to estimate the optimum number of proper orthogonal decomposition basis functions for implementing non-intrusive reduced order model predictions of the spatial distribution of the reservoir properties of interest.
- Step 9: Translate individual training runs data into hyperspace using the optimal left singular vector matrix and recombine/ concatenate them into a single matrix, i.e. the reduced training data matrix.
- Step 10: Re-normalise the reduced training data using the scaling factor in Equation 3.9 (for constant parameter scenarios) or Equation 3.10 (for varying parameter scenarios). Also, renormalize training parameters (parameterisation) with their corresponding scaling factors as in Equation 3.12 or 3.14.
- Step 11: Concatenate individual reduced and re-normalised training data with their corresponding re-normalised parameters as in Equation 3.15
- Step 12: Generate sample and response arrays from the reduced training data in a cause and effect manner i.e. in a similar manner as in the last paragraph of section 3.4.
- Step 13: Generate initial condition data for prediction cases to be evaluated with non-intrusive reduced order model, reduce and normalise the data via the same steps used to reduce and normalise training data (steps 6 to 11).

Interpolate using a) Radial Basis Functions or b) Autoencoder

a) Radial basis function interpolation

- Step 14: Generate pairwise distance matrix between successive columns of normalised and reduced sample training data from step 12.
- Step 15: Generate weighting coefficient matrix using the pairwise distance data from step 14 and a suitable radial basis interpolation function (Equation 3.19). The linear radial basis interpolation function is recommended for gas reservoir modelling.

Step 16: Execute non-intrusive reduced order model simulation with weighting coefficients from step 15 as well as initial condition data for the prediction cases from step 13.

b) Autoencoder interpolation

* use full rank left singular values matrix from step 7 rather than optimal rank data from step 8, to generate their training data.

Step 14: Generate additional simulation run(s) for validation data, reduce and normalise the validation data via the same steps used to reduce and normalise training data (steps 6 to 11).

Step 15: Define hyperparameters for autoencoder, train autoencoder with the defined hyperparameters and training data from steps 12 and 14. The data from step 12 is for training the autoencoder and the data from step 14 is for validating the autoencoder.

Step 16: Execute autoencoder simulation, to obtain prediction results for the data in step 13.

Back-project results from interpolation and compare results

Step 17: Un-normalise radial basis function and/ or autoencoder prediction results and translate them into full space using Equation 3.17.

Step 18: Compare results from step 17 to the full simulation outputs of the same production cases with the conventional simulator, using plots and/ or metrics.

9.4 LIMITATIONS AND RECOMMENDATIONS FOR FUTURE WORK

9.4.1 LIMITATIONS

Having highlighted the strengths and usefulness of the proper orthogonal decomposition-radial basis function non-intrusive reduced order model in gas field modelling, it is also important to recap its other realities. Nonetheless, the views mentioned below are more cautionary than constraints for the application of non-intrusive reduced order model for fast modelling of practical aspects of gas field development planning such as during history matching and sensitivity analysis.

First, the non-intrusive reduced order model is at best a complementary tool to standard reservoir simulation, it depends on conventional simulation results for the data with which it learns a system's dynamics. Likewise, the proper orthogonal decomposition-radial basis function model does not give the exact same results as ECLIPSE. So for production scenarios that rely on high modelling accuracy/ precision (such as during enhanced oil recovery studies – which have not been evaluated in this work), an approved commercial simulator remains the best simulation tool.

In this work also, it has been shown that for simple simulation cases like the base case of the synthetic gas reservoir, the proper orthogonal decomposition-radial basis function model did not give as significant speedup as for the heterogeneous reservoirs or more complex production scenarios such

as grid refinement, water coning or presence of aquifer influx. So for sensitivity analysis of simple simulations cases, it is probably more beneficial to use the direct results of a conventional simulator, rather than build a non-intrusive reduced order model.

Lastly, proper orthogonal decomposition-based non-intrusive reduced order models are not ideal for simulation cases with large amounts of data because of the time requirement for back projection. For instance in chapter 8 it was shown that in the seasonally varying production profile case with 50 training runs and the full rank proper orthogonal decomposition basis functions, the non-intrusive reduced order model was slower than ECLIPSE. Back projection constitutes one of the different sources of uncertainty in this work. Although it has not been evaluated, it is expected to have a small consequence on the results that were discussed.

9.4.2 RECOMMENDATIONS

Although this work has considered and evaluated several aspects of gas reservoir modelling, several other production scenarios exist that could be evaluated with proper orthogonal decomposition-radial basis function non-intrusive reduced order model or other modelling techniques. For instance, this work only considered reservoirs with vertical and deviated (in Field X) producers, however further work could focus on applying the proper orthogonal decomposition-radial basis function to reservoir systems with injectors and more complicated well configurations. The non-intrusive reduced order model could also be applied to modelling more complex reservoir scenarios such as naturally fractured carbonates or sandstone reservoirs. This work was also focussed on conventional gas reservoir modelling but proper orthogonal decomposition-radial basis function could be applied to modelling the performance of unconventional shale gas reservoirs, condensate gas reservoirs or CO₂ sequestration in depleted hydrocarbon reservoirs or aquifers, especially based on the renormalisation approach for time-varying parameters that is presented in this work. This should be feasible, since the non-intrusive reduced order model approaches are data driven and independent of the physics of the reservoir systems. The non-intrusive reduced order model could also be evaluated for modelling the performance of enhanced recovery strategies, for instance chemical, thermal and immiscible displacement. It could also be expanded to simulate other aspects of the energy and other industries that rely on big data and simulations such as healthcare, telecommunications and studies on natural events and disasters.

In addition to the above recommendations, other machine learning techniques can be investigated for gas reservoir modelling, especially those with faster training time than the proper orthogonal decomposition-autoencoder method presented in this work. These could include other machine

learning data reduction techniques like convolutional neural networks with graphics processing units and hybrid modelling that incorporates machine learning techniques with optimisation techniques. The hyperparameters and architectures applied to these machine learning techniques can also be further investigated to determine suitable combinations that could improve the fast modelling of gas reservoirs.

From an industry perspective, it would be good to collaborate with software/ programming experts to implement the proper orthogonal decomposition-radial basis function non-intrusive reduced order model in an efficient compiled programming language. Possibly, a software or commercial product can be developed that incorporates all the aspects of gas reservoir modelling that have been evaluated in this work, as well as other aspects of field development planning that may be considered in future. A proper orthogonal decomposition-radial basis function non-intrusive reduced order model tool would benefit even more from being applied in the daily activities of practising engineers who can highlight areas for development and improvement of the tool. This would further assist in fine-tuning the non-intrusive reduced order model's usefulness and application to suit specific industry purposes.

NOMENCLATURE

A	$(d+1 \times nt)$ array of reduced training data and parameters
AE	Autoencoder
AI	Artificial Intelligence
ANN	Artificial neural network
B	formation volume factor
BCF	billion cubic feet
BHP	Bottom Hole Pressure
bp	British Petroleum
c	compressibility
CNN	convolutional neural network
cP	centipoise
d	number of POD basis functions that represent 99.9% energy criterion
DEIM	discrete empirical interpolation method
EIA	US Energy Information Administration
ELU	Exponential Linear Unit
f	a function
FFNN	feedforward neural network
ft	feet
g	gram
G	volume of Gas
GIP	Gas in Place
GNAT	Gauss-Newton with approximated tensors
GPU	graphics processing unit
GRU	gated recurrent units
GWC	Gas Water Contact
HL	Hidden Layer
I	energy criterion
K	permeability
kg	kilogram
lb	pound
LGR	Local Grid Refinement
LSTM	Long Short Term Memory
m	meter

mD	milliDarcy
Mscf	thousand standard cubic feet
ML	Machine Learning
MMscf	million standard cubic feet
NIROM	Non-Intrusive Reduced Order Model
NN	Neural Networks
np	total number of parameters
ns	total number of simulation snapshots
nt	total number of dimensions in the temporal resolution
nx	total number of dimensions in the spatial resolution
NX	number of grid cells/blocks in the x-direction
NY	number of grid cells/blocks in the y-direction
NZ	number of grid cells/blocks in the z-direction
p	parameter
PCA	principal components analysis
PDE	Partial Differential Equation
POD	Proper Orthogonal Decomposition
psi	pound square inch
p/Z	pressure (P) divided by gas deviation factor (Z)
Δp	total pressure drop
R	rank
RBF	Radial Basis Function
ReLU	rectified linear unit
RMSE	Root Mean Square Error
RNN	recurrent neural network
ROM	Reduced Order Model
s	index for the number of training runs
sm	standard meter
S_{wc}	connate water saturation
t	index for temporal resolution, t th time instance/snapshot
T	superscript denoting transpose
THP	tubing head pressure
TPWL	trajectory piecewise linearization
U	columns of the training data

V	left singular vectors' matrix of a POD
V^T	transpose of the left singular vectors or eigenvectors of a POD
w	weighting coefficient for linear RBF
W	right singular vectors' matrix of a POD
W_e	water influx

Greek symbols

α	vectors describing the projection of the training solution in hyperspace
λ	singular values
μ	viscosity
$\Sigma = \lambda^2$	eigenvalues
ϕ	porosity
φ	radial basis interpolation function
ω	weighting coefficient matrix

Subscripts

p	produced
g	gas
i	initial
w	water

REFERENCES

1. IEA, Key World Energy Statistics 2020. 2020, IEA: Paris.
2. Project, G.C., Supplemental data of Global Carbon Budget 2020 2020, Global Carbon Project.
3. (EIA), U.S.E.I.A., Annual Energy Outlook 2021. 2020.
4. bp, Energy Outlook. 2020.
5. Bansal, A.K. Understanding Natural Gas Sales & Purchase Contracts and Principal Contractual Terms. 2017 [cited 2021 25/02/2021]; Available from: <https://www.linkedin.com/pulse/understanding-natural-gas-sales-purchase-contracts-principal-bansal/>.
6. bp, Statistical Review of World Energy. 2020.
7. Harrison, M., Campbell, P., Coeiho, F., Grecco, C., Ikeocha, X., Wang, X., Economides, M. J., Natural Gas: Overcoming Limitations toward Energy Supremacy, in SPE Annual Technical Conference and Exhibition. 2006, SPE-102061-MS. Society of Petroleum Engineers: San Antonio, Texas, USA.
8. BP, Statistical Review of World Energy. 2018.
9. SPE, Comparison of Selected Reserves and Resource Classifications and Associated Definitions. 2005.
10. Guzman, J.D.D., J.A.A. Arevalo, and O. Espinola. Reserves Evaluation of Dry Gas Reservoirs through Flowing Pressure Material Balance Method. SPE Energy Resources Conference. 2014, SPE-169989-MS.
11. Dake, L.P., Chapter 6 Gas Reservoir Engineering, in Developments in Petroleum Science, L.P. Dake, Editor. 2001, Elsevier. p. 473-536.
12. Bernard, W.J., Reserves estimation and performance prediction for geopressured gas reservoirs. Journal of Petroleum Science and Engineering, 1987. 1(1): p. 15-21.
13. Temirchev, P., Simonov, M., Kostoev, R., Burnaev, E., Oseledets, I., Akhmetov, A., Margarit, A., Sitnikov, A., Koroteev, D., Deep neural networks predicting oil movement in a development unit. Journal of Petroleum Science and Engineering, 2020. 184: Article. 106513.
14. Yue, J., Water-drive gas reservoir: sensitivity analysis and simplified prediction, in Department of Petroleum Engineering. 2002, Louisiana State University, Agricultural and Mechanical College.
15. Al-Marhoon, N. and S. Al-Shidhani. The Challenges of Predicting Commingled Production of Rich and Dry Gas Reservoirs of the Saih Rawl Field, Oman. Middle East Oil Show. 2003, SPE-81440-MS.
16. Nwachukwu, A., Jeong, H., Pyrcz, M., Lake, L. W., Fast evaluation of well placements in heterogeneous reservoir models using machine learning. Journal of Petroleum Science and Engineering, 2018. 163: p. 463-475.
17. Rietz, D.C. and A.H. Usmani. Reservoir Simulation and Reserves Classifications-Guidelines for Reviewing Model History Matches to Help Bridge the Gap between Evaluators and Simulation Specialists. SPE Annual Technical Conference and Exhibition. 2005, SPE-96410-MS.
18. Mustafiz, S. and M.R. Islam, State-of-the-art Petroleum Reservoir Simulation. Petroleum Science and Technology, 2008. 26(10-11): p. 1303-1329.
19. Hemmati-Sarapardeh, A., Larestani, A., Nait Amar, M., Hajirezaie, S., Chapter 7 - Weaknesses and strengths of intelligent models in petroleum industry, in Applications of Artificial Intelligence Techniques in the Petroleum Industry, A. Hemmati-Sarapardeh, et al., Editors. 2020, Gulf Professional Publishing. p. 295-301.
20. Ersahin, A., Gul, S., Ertekin, T., Temizel, C. Artificial Neural Network Modeling of Cyclic Steam Injection Process in Naturally Fractured Reservoirs. SPE Western Regional Meeting. 2019, SPE-195307-PA.

21. Gonzalez, F.J. and M. Balajewicz, Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems. arXiv preprint, 2018. 1808.01346.
22. Fang, F., Zhang, T., Pavlidis, D., Pain, C.C., Buchan, A.G., Navon, I.M., Reduced order modelling of an unstructured mesh air pollution model and application in 2D/3D urban street canyons. *Atmospheric Environment*, 2014. 96: p. 96-106.
23. Cortial, J., Farhat, C., Guibas, L. J., Rajashekhar, M., Compressed Sensing and Time-Parallel Reduced-Order Modeling for Structural Health Monitoring Using a DDDAS. 2007. Berlin, Heidelberg: Springer Berlin Heidelberg.
24. Maurya, M.R., Bornheimer, S. J., Venkatasubramanian, V., Subramaniam, S., Reduced-order modelling of biochemical networks: application to the GTPase-cycle signalling module. *Systems biology*, 2005. 152(4): p. 229-242.
25. Kaiser, E., Morzyński, M., Daviller, G., Kutz, J. N., Brunton, B. W., Brunton, S. L., Sparsity enabled cluster reduced-order models for control. *Journal of Computational Physics*, 2018. 352: p. 388-409.
26. Griffiths, L.M., Gaitonde, A., Jones, D., Friswell, M.I., Updating of aerodynamic reduced order models generated using computational fluids dynamics. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 2018. 232(9): p. 1739-1763.
27. Fang, F., Pain, C. C., Navon, I. M., Piggott, M. D., Gorman, G. J., Allison, P. A., Goddard, A. J. H., Reduced-order modelling of an adaptive mesh ocean model. *International journal for numerical methods in fluids*, 2009. 59(8): p. 827-851.
28. Dowell, E., Hall, K., Thomas, J., Florea, R., Epureanu, B., Heeg, J., Reduced order models in unsteady aerodynamics. 40th Structures, Structural Dynamics, and Materials Conference and Exhibit. 1999.
29. Ripepi, M., Verveld, M.J., Karcher, N.W., Franz, T., Abu-Zurayk, M., Görtz, S., Kier, T.M., Reduced-order models for aerodynamic applications, loads and MDO. *CEAS Aeronautical Journal*, 2018. 9(1): p. 171-193.
30. Casas, C.Q., R. Arcucci, and Y. Guo, Urban air pollution forecasts generated from latent space representation. *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*, 2020.
31. Feder, J., Parallel Simulation and Cloud Computing Can Optimize Large-Scale Field Development. *Journal of Petroleum Technology*, 2019. 71(09): p. 81-82.
32. He, J. and L.J. Durlofsky, Reduced-Order Modeling for Compositional Simulation by Use of Trajectory Piecewise Linearization. *SPE Journal*, 2014, SPE-163634-PA. 19(05): p. 858-872.
33. Ghasemi, M., Yang, Y., Gildin, E., Efendiev, Y., Calo, V., Fast Multiscale Reservoir Simulations using POD-DEIM Model Reduction, in *SPE Reservoir Simulation Symposium*. 2015, SPE-173271-MS, Society of Petroleum Engineers: Houston, Texas, USA.
34. Yoon, S., Alghareeb, Z.M., and J.R. Williams, Hyper-Reduced-Order Models for Subsurface Flow Simulation. *SPE Journal*, 2016, SPE-181740-PA. 21(06): p. 2128-2140.
35. Jiang, R. and Durlofsky, L.J., Implementation and detailed assessment of a GNAT reduced-order model for subsurface flow simulation. *Journal of Computational Physics*, 2019. 379: p. 192-213.
36. Klie, H., Unlocking Fast Reservoir Predictions via Nonintrusive Reduced-Order Models, *SPE Reservoir Simulation Symposium*. 2013, SPE-163584-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
37. Kostorz, W., Muggeridge, A. H., Jackson, M. D., Non-Intrusive Reduced Order Modelling for Reconstruction of Saturation Distributions, in *SPE Reservoir Simulation Conference*. 2019, SPE-193831-MS, Society of Petroleum Engineers: Galveston, Texas, USA. p. 13.
38. Xiao, D., Non-intrusive reduced order models and their applications, PhD Thesis, in Department of Earth Science and Engineering. 2016, Imperial College London.

39. Cardoso, M.A. and Durlofsky, L.J., Use of Reduced-Order Modeling Procedures for Production Optimization. SPE J. 15 (2010): 426–435, SPE-119057-PA.
40. Ghasemi, M., A. Ibrahim, and E. Gildin, Reduced Order Modeling In Reservoir Simulation Using the Bilinear Approximation Techniques, in SPE Latin America and Caribbean Petroleum Engineering Conference. 2014, SPE-169357-MS, Society of Petroleum Engineers: Maracaibo, Venezuela.
41. He, J., Xie, J., Wen, X.-H., Chen, W., Improved Proxy For History Matching Using Proxy-for-data Approach And Reduced Order Modeling, in SPE Western Regional Meeting. 2015, SPE-174055-MS, Society of Petroleum Engineers: Garden Grove, California, USA.
42. Temizel, C., Nabizadeh, M., Kadkhodaei, N., Ranjith, R., Suhag, A., Balaji, K., Dhannoon, D., Data-Driven Optimization of Injection/Production in Waterflood Operations. SPE Intelligent Oil and Gas Symposium. 2017. SPE-187468-MS.
43. van Doren, J.F.M., Markovinović, R., and Jansen, J.-D., Reduced-order optimal control of water flooding using proper orthogonal decomposition. Computational Geosciences, 2006. 10(1): p. 137-158.
44. Artun, E., Performance assessment and forecasting of cyclic gas injection into a hydraulically fractured well using data analytics and machine learning. Journal of Petroleum Science and Engineering, 2020. 195: Article. 107768.
45. Hemmati-Sarapardeh, A., Larestani, A., Nait Amar, M., Hajirezaie, S., Chapter 4 - Application of intelligent models in reservoir and production engineering, in Applications of Artificial Intelligence Techniques in the Petroleum Industry, A. Hemmati-Sarapardeh, et al., Editors. 2020, Gulf Professional Publishing. p. 79-227.
46. Alvarado, V. and E. Manrique, Chapter 3 - Simulations and Simulation Options, in Enhanced Oil Recovery, V. Alvarado and E. Manrique, Editors. 2010, Gulf Professional Publishing: Boston. p. 17-42.
47. PetroWiki. Petroleum Engineering Handbook: Gas Reservoirs - Forecasting. 2017 26 April 2017 [cited 2021 26 February 2021]; Available from: https://petrowiki.org/PEH:Gas_Reservoirs#Forecasting.
48. Temizel, C., Canbaz, C. H., Gok, I. M., Roshankhah, S., Palabiyik, Y., Deniz-Paker, M., Hosgor, F. B., Ozyurtkan, H., Aksahan, F., Gormez, E., Kaya, S., Kaya, O.A., A Thorough Review and Latest Advances in Shale Reservoirs: Seismic to Surveillance. SPE Latin American and Caribbean Petroleum Engineering Conference. 2020, SPE- 198994-MS.
49. Odeh, A.S., Reservoir Simulation ...What Is It. Journal of Petroleum Technology, 1969. 21(11): p. 1383-1388.
50. Clarke, A.L., Imber, J., Davies, R. J., van Hunen, J., Daniels, S. E., Yielding, G., Application of material balance methods to CO₂ storage capacity estimation within selected depleted gas reservoirs. Petroleum Geoscience, 2017. 23(3): p. 339-352.
51. Walsh, M.P., J. Ansah, and R. Raghavan. The New, Generalized Material Balance as an Equation of a Straight Line: Part 1 - Applications to Undersaturated, Volumetric Reservoirs. Permian Basin Oil and Gas Recovery Conference. 1994, SPE 27684. Society of Petroleum Engineers: Midland, Texas, USA.
52. Satter, A. and G.M. Iqbal, 13 - Decline curve analysis for conventional and unconventional reservoirs, in Reservoir Engineering, A. Satter and G.M. Iqbal, Editors. 2016, Gulf Professional Publishing: Boston. p. 211-232.
53. Satter, A. and G.M. Iqbal, 15 - Petroleum reservoir simulation: a primer, in Reservoir Engineering, A. Satter and G.M. Iqbal, Editors. 2016, Gulf Professional Publishing: Boston. p. 247-287.

54. Shahkarami, A., G. Wang, and H. Belyadi. Performance Forecast and Optimization Workflow for Horizontal Wells: A Case Study on Utica-Point Pleasant Shale Play. SPE Eastern Regional Meeting. 2017, SPE-187511-MS.
55. Sun, S. and T. Zhang, Chapter six - Recent progress in machine learning applications in reservoir simulation, in *Reservoir Simulations*, S. Sun and T. Zhang, Editors. 2020, Gulf Professional Publishing. p. 259-288.
56. Chen, H., H. Klie, and Q. Wang, A Black-Box Interpolation Method to Accelerate Reservoir Simulation Solutions, in SPE Reservoir Simulation Symposium. 2013, SPE-163614-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
57. Bayat, M. and J.E. Killough, An Experimental Study of GPU Acceleration for Reservoir Simulation, in SPE Reservoir Simulation Symposium. 2013, SPE-163628-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
58. Hemmati-Sarapardeh, A., Larestani, A., Nait Amar, M., Hajirezaie, S., Chapter 1 - Introduction, in *Applications of Artificial Intelligence Techniques in the Petroleum Industry*, A. Hemmati-Sarapardeh, et al., Editors. 2020, Gulf Professional Publishing. p. 1-22.
59. Du, S., L.S. Fung, and A.H. Dogru. Accelerating Parallel Field-Scale Reservoir Simulation with Peripheral and/or Bottom Aquifers. SPE Middle East Oil & Gas Show and Conference. 2017, SPE-183902-MS.
60. Christie, M.A., Upscaling for Reservoir Simulation. *Journal of Petroleum Technology*, 1996. 48(11): p. 1004-1010.
61. Christie, M.A. and P.J. Clifford, Fast Procedure for Upscaling Compositional Simulation. 1998. SPE J. 3 (03): 272–278
62. Portella, R.C.M. and T.A. Hewett, Fast 3-D Reservoir Simulation and Applications Using Streamlines, in *Latin American and Caribbean Petroleum Engineering Conference*. 1997, SPE-39061-MS. Society of Petroleum Engineers: Rio de Janeiro, Brazil.
63. Ma, X., Al-Harbi, M., Datta-Gupta, A., Efendiev, Y., An Efficient Two-Stage Sampling Method for Uncertainty Quantification in History Matching Geological Models. *SPE Journal*, 2008, SPE-102476-PA. 13(01): p. 77-87.
64. Gries, S. and H.-J. Plum, Status of System-AMG for Reservoir Simulation Applications, in SPE Reservoir Simulation Symposium. 2015, SPE-173241-MS, Society of Petroleum Engineers: Houston, Texas, USA.
65. Lu, P. and B.L. Beckner, An Adaptive Newton's Method for Reservoir Simulation, in SPE Reservoir Simulation Symposium. 2011, SPE-141935-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
66. Appleyard, J., Appleyard, J., Wakefield, M., Desitter, A., Accelerating Reservoir Simulators using GPU Technology, in SPE Reservoir Simulation Symposium. 2011, SPE-141402-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
67. Klie, H.M., Sudan, H. H., Li, R., Saad, Y., Exploiting Capabilities of Many Core Platforms in Reservoir Simulation, in SPE Reservoir Simulation Symposium. 2011, SPE-141265-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
68. Yang, Y., Ghasemi, M., Gildin, E., Efendiev, Y., Calo, V., Fast Multiscale Reservoir Simulations with POD-DEIM Model Reduction. *SPE J.* 21 (2016): 2141–2154. SPE-173271-PA.
69. Chaturantabut, S. and D.C. Sorensen, Application of POD and DEIM on dimension reduction of non-linear miscible viscous fingering in porous media. *Mathematical and Computer Modelling of Dynamical Systems*, 2011. 17(4): p. 337-353.
70. Schilders, W., Introduction to Model Order Reduction, in *Model Order Reduction: Theory, Research Aspects and Applications*, W.H.A. Schilders, H.A. van der Vorst, and J. Rommes, Editors. 2008, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 3-32.

71. Haykin, S., *Neural Networks and Learning Machines*. 3rd ed. 2009: Prentice Hall PTR. 938.
72. Du, K.-L. and M.N.S. Swamy, *Neural Networks in a Softcomputing Framework*. Springer London, 2006: p. 251-294.
73. Tanaka, S., Wang, Z., Dehghani, K., He, J., Velusamy, B., Wen, X.-H., Large Scale Field Development Optimization Using High Performance Parallel Simulation and Cloud Computing Technology. SPE Annual Technical Conference and Exhibition. 2018, SPE-191728-MS.
74. Kumar, A., Dutt, A., Nahar, S., Batshas, S., Majumdar, C., Saraiya, R., Sarkar, S., Chatterjee, C., It's Raining Barrels: Cloud Computing in the O&G Industry. SPE Middle East Artificial Lift Conference and Exhibition. 2018. SPE-192515-MS.
75. Darman, N.H., K.S. Sorbie, and G.E. Pickup, The Development of an Optimal Grid Coarsening Scheme: Interplay of Fluid Forces and Higher Moments of Fine-scale Flow Data, in SPE Asia Pacific Oil and Gas Conference and Exhibition. 2005, SPE-93348-MS, Society of Petroleum Engineers: Jakarta, Indonesia. p. 8.
76. Cardoso, M.A. and L.J. Durlofsky, Use of Reduced-Order Modeling Procedures for Production Optimization, in SPE Reservoir Simulation Symposium. 2009, SPE-119057-MS, Society of Petroleum Engineers: The Woodlands, Texas.
77. He, J., P. Sarma, and L.J. Durlofsky, Use of Reduced-order Models for Improved Data Assimilation within an EnKF Context, in SPE Reservoir Simulation Symposium. 2011, SPE-141967-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA.
78. Dowell, E.H. and K.C. Hall, Modeling of fluid-structure interaction. *Annual review of fluid mechanics*, 2001. 33(1): p. 445-490.
79. Zimmermann R, V. A, and G. S., Reduced-order modeling of steady flows subject to aerodynamic constraints. *American Institute of Aeronautics and Astronautics*, 2014. 52(2): p. 255-266.
80. Bai, Z., P.M. Dewilde, and R.W. Freund, *Reduced-Order Modeling Handbook of Numerical Analysis*, 2005. 13: p. 825-895.
81. Amsallem, D. and C. Farhat, Stabilization of projection-based reduced-order models. *International Journal for Numerical Methods in Engineering*, 2012. 91(4): p. 358-377.
82. Amsallem, D., R. Tezaur, and C. Farhat, Real-time solution of linear computational problems using databases of parametric reduced-order models with arbitrary underlying meshes. *Journal of Computational Physics*, 2016. 326: p. 373-397.
83. Oliver, J., Caicedo, M., Huespe, A. E., Hernández, J. A., Roubin, E., Reduced order modeling strategies for computational multiscale fracture. *Computer Methods in Applied Mechanics and Engineering*, 2017. 313: p. 560-595.
84. Lucia, D.J., P.S. Beran, and W.A. Silva, Reduced-order modeling: new approaches for computational physics. *Progress in Aerospace Sciences*, 2004. 40(1): p. 51-117.
85. Xiao, D., Fang, F., Du, J., Pain, C.C., Navon, I.M., Buchan, A. G., ElSheikh, A.H., Hu, G., Non-Linear Petrov-Galerkin Methods for Reduced Order Modelling of the Navier-Stokes Equations using a Mixed Finite Element Pair. *Computer Methods in Applied Mechanics and Engineering*, 2013. 255: p. 147-157.
86. Ly, H.V. and H.T. Tran, Modeling and control of physical processes using proper orthogonal decomposition. *Mathematical and Computer Modelling*, 2001. 33(1-3): p. 223-236.
87. Yu, G., Yu, B., Han, D., Wang, L., Unsteady-state thermal calculation of buried oil pipeline using a proper orthogonal decomposition reduced-order model. *Applied Thermal Engineering*, 2013. 51(1): p. 177-189.
88. Capellari, G., S. Eftekhari Azam, and S. Mariani. Hybrid Reduced-Order Modeling and Particle-Kalman Filtering for the Health Monitoring of Flexible Structures. *International Electronic Conference on Sensors and Applications*. 2014.

89. Alawneh, S., J.W. Thijssen, and M. Richard, Accelerating Numerical Ice Engineering Tools Using GPGPU, in Arctic Technology Conference. 2016, OTC-27386-MS. Offshore Technology Conference: St. John's, Newfoundland and Labrador, Canada.
90. Sagar, A., Wei D., Mason M., Jeffrey D V., LeCover R., Reduced order modeling and analysis of the human complement system. PLoS ONE 12(11): e0187373. BioRxiv, 2017.
91. Xiao, D., Lin, Z., Fang, F., Pain, Christopher C., Navon, L. M., Salinas, P., Muggeridge, A., Non-intrusive reduced-order modeling for multiphase porous media flows using Smolyak sparse grids. International Journal for Numerical Methods in Fluids, 2017. 83(2): p. 205-219.
92. Xiao, D., Fang, F., Pain, C. C., Navon, I. M., Salinas, P., Muggeridge, A., Non-intrusive reduced order modeling of multi-phase flow in porous media using the POD-RBF method. Journal of Computational Physics, 2015;1:1-25.
93. Cardoso, M.A., Reduced-Order Models for Reservoir Simulation, in SPE Annual Technical Conference and Exhibition. 2009, SPE-129636-STU, Society of Petroleum Engineers: New Orleans, Louisiana.
94. Ghasemi, M. and E. Gildin, Model order reduction in porous media flow simulation using quadratic bilinear formulation. Computational Geosciences, 2016. 20(3): p. 723-735.
95. Xiao, D., Fang, F., Buchan, A. G., Pain, C. C., Navon, I. M., Muggeridge, A., Non-intrusive reduced order modelling of the Navier–Stokes equations. Computer Methods in Applied Mechanics and Engineering, 2015. 293: p. 522-541.
96. Xiao, D., Fang, F., Pain, C., Hu, G., Non-intrusive reduced-order modelling of the Navier–Stokes equations based on RBF interpolation. International Journal for Numerical Methods in Fluids, 2015. 79(11): p. 580-595.
97. Jackson, M.D.D., Percival, J.R., Mostaghimi, P., Tollit, B.S., Pavlidis, D., Pain, C.C., Gomes, J.L.M.A., El-Sheikh, A.H., Salinas, P., Muggeridge, A.H., Blunt, M.J., Reservoir Modeling for Flow Simulation by Use of Surfaces, Adaptive Unstructured Meshes, and an Overlapping-Control-Volume Finite-Element Method. SPE Reservoir Evaluation & Engineering, 2015. SPE-163633-PA. 18(02): p. 115-132.
98. Christie, M.A. and D.J. Bond, Detailed Simulation of Unstable Processes in Miscible Flooding. SPE Reservoir Engineering, 1987. SPE-14896-PA. 2(04): p. 514-522.
99. Kotorz, W.J., A.H. Muggeridge, and M.D. Jackson, An efficient and robust method for parameterized nonintrusive reduced-order modeling. International Journal for Numerical Methods in Engineering, 2020. SPE-193831-MS. 121(20): p. 4674-4688.
100. Kotorz, W.J., A.H. Muggeridge, and M.D. Jackson, Non-intrusive reduced order modeling: Geometrical framework, high-order models, and a priori analysis of applicability. International Journal for Numerical Methods in Engineering, 2021 ; 122: 2545– 2565.
101. Marr, B., Deep Learning Vs Neural Networks - What's The Difference? Available from: <https://bernardmarr.com/default.asp?contentID=1789#:~:text=What%20is%20deep%20learning%3F&text=Well%20an%20ANN%20that%20is,is%20what%20underpins%20deep%20learning.> Accessed: April 2021.
102. Sharma, S., Artificial Neural Network in Machine Learning. 2017, Data Science Central. Available: [https://www.datasciencecentral.com/profiles/blogs/artificial-neural-network-ann-in-machine-learning#:~:text=Artificial%20Neural%20networks%20\(ANN\)%20or,systems%20composed%20of%20%E2%80%9Cneurons%E2%80%9D.&text=A%20neural%20network%20is%20a,model%20of%20a%20human%20neuron.](https://www.datasciencecentral.com/profiles/blogs/artificial-neural-network-ann-in-machine-learning#:~:text=Artificial%20Neural%20networks%20(ANN)%20or,systems%20composed%20of%20%E2%80%9Cneurons%E2%80%9D.&text=A%20neural%20network%20is%20a,model%20of%20a%20human%20neuron.) Accessed: April 2021.
103. Hajizadeh, Y., Machine learning in oil and gas; a SWOT analysis approach. Journal of Petroleum Science and Engineering, 2019. 176: p. 661-663.

104. Jin, Z.L., Y. Liu, and L.J. Durlofsky, Deep-learning-based surrogate model for reservoir simulation with time-varying well controls. *Journal of Petroleum Science and Engineering*, 2020. 192: Article. 107273.
105. Otchere, D.A., Arbi Ganat, T. O., Gholami, R., Ridha, S., Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering*, 2020: Article. 108182.
106. David, F., *Generative Deep Learning*. 2019: O'Reilly Media, Inc.
107. Hemmati-Sarapardeh, A., Larestani, A., Nait Amar, M., Hajirezaie, S., Chapter 2 - Intelligent models, in *Applications of Artificial Intelligence Techniques in the Petroleum Industry*, A. Hemmati-Sarapardeh, et al., Editors. 2020, Gulf Professional Publishing. p. 23-50.
108. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A. and Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., *Scikit-learn: Machine Learning in Python*. 2011 [cited 2020; Available from: <https://scikit-learn.org/stable/about.html#citing-scikit-learn>.
109. Kanin, E.A. Osiptsov, A. A., Vainshtein, A. L., Burnaev, E. V., A predictive model for steady-state multiphase pipe flow: Machine learning on lab data. *Journal of Petroleum Science and Engineering*, 2019. 180: p. 727-746.
110. Chollet, F. and others, *Keras*. 2015. Available at: <https://github.com/fchollet/keras>.
111. Basheer, I.A. and M. Hajmeer, *Artificial neural networks: fundamentals, computing, design, and application*. *J Microbiol Methods*, 2000. 43(1): p. 3-31.
112. Suhag, A., R. Ranjith, and F. Aminzadeh. Comparison of Shale Oil Production Forecasting using Empirical Methods and Artificial Neural Networks. *SPE Annual Technical Conference and Exhibition*. 2017. SPE-187112-MS
113. Hemmati-Sarapardeh, A., Larestani, A., Nait Amar, M., Hajirezaie, S., Chapter 6 - Application of intelligent models in exploration engineering, in *Applications of Artificial Intelligence Techniques in the Petroleum Industry*, A. Hemmati-Sarapardeh, et al., Editors. 2020, Gulf Professional Publishing. p. 279-294.
114. Kalam, S., Khan, M. R., Tariq, Z., Siddique, F. A., Abdulraheem, A., Khan, R. A., A Novel Correlation to Predict Gas Flow Rates Utilizing Artificial Intelligence: An Industrial 4.0 Approach. *SPE/PAPG Pakistan Section Annual Technical Symposium and Exhibition*. 2019. SPE-201170-MS.
115. Taware, S.V., T. Friedel, and A. Datta-Gupta, A Practical Approach For Assisted History Matching Using Grid Coarsening And Streamline-Based Inversion: Experiences in a Giant Carbonate Reservoir, in *SPE Reservoir Simulation Symposium*. 2011, SPE-141606-MS, Society of Petroleum Engineers: The Woodlands, Texas, USA. p. 16.
116. Al-Dhamen, M.A. Application of Neural Network to Predict Gas Well Performance during Two-Phase Flow in Gas Condensate Reservoirs. *SPE Saudi Arabia Section Technical Symposium and Exhibition*. 2014. SPE-172200-MS.
117. Wood, D.A., Predicting porosity, permeability and water saturation applying an optimized nearest-neighbour, machine-learning and data-mining network of well-log data. *Journal of Petroleum Science and Engineering*, 2020. 184: Article. 106587.
118. MATLAB, version 9.6.0 (R2019a). 2019, Natick, Massachusetts: The MathWorks Inc.
119. Thomas, K., Ragan-Kelley B., Perez F., Granger B., Bussonnier M., Frederic J., Kelley K., Hamrick J., Grout J., Corlay S., Ivanov P., Avila D., Abdalla S., Willing C., *Jupyter Notebooks - a publishing format for reproducible computational workflows. Positioning and Power in Academic Publishing: Players, Agents and Agendas*. 2016: IOS Press.
120. Van Rossum, G.A.D.J., Fred L., *Python 3 Reference Manual (version 3.7.6)*. 2019, Scotts Valley, CA: CreateSpace.
121. Schlumberger, *Eclipse Industry-Reference Reservoir Simulator Manual*. 2017.

122. Sævik, P.N., read_ecl. 2014: https://uk.mathworks.com/matlabcentral/fileexchange/47607-read_ecl-m. Accessed: July 2018.
123. Recktenwald, G., readColData. 1995: <https://web.cecs.pdx.edu/~gerry/MATLAB/programming/intermediate.html>. Accessed: May 2018.
124. Sirovich, L., Turbulence and the Dynamics of Coherent Structures Part III: Dynamics and Scaling. Quarterly of Applied Mathematics, 1987. 45(3): p. 583-590.
125. Samuel, J.-S. and A.H. Muggeridge. Fast Modelling of Gas Reservoirs Using POD-RBF Non-Intrusive Reduced Order Modelling. SPE Asia Pacific Oil & Gas Conference and Exhibition. 2020. SPE-202436-MS.
126. Wu, Y., Wang, H., Zhang, B., Du, K. L., Using Radial Basis Function Networks for Function Approximation and Classification. ISRN Applied Mathematics. 2012. 10.5402/2012/324194.
127. Martín, A., Ashish, A., Paul, B., Eugene, B., Zhifeng, C., Craig, C., Greg, S. C., Andy, D., Jeffrey, D., Matthieu, D., Sanjay, G., Ian, G., Andrew, H., Geoffrey, I., , Michael, I., Rafal, J., Yangqing, J., Lukasz, K., Manjunath, K., Josh, L., Dan, M., Mike, S., Rajat, M., .Sherry, M., Derek, M., Chris, O., Jonathon, S., Benoit, S., Ilya, S., Kunal, T., Paul, T., Vincent, V., Vijay, V., Fernanda, V., Oriol, V., Pete, W., Martin, W., Martin, W., Yuan, Y., Xiaoqiang, Z., TensorFlow: Large-scale machine learning on heterogeneous systems. 2015.
128. NTNU, I.C. Norne Field Case Description. 2011 01/08/2018]; Available from: http://www.ipt.ntnu.no/~norne/wiki/data/media/english/norne_field_case_desc.pdf.
129. Rasmussen, A.F., Norne. 2014, Github. Available from: <https://github.com/OPM/opm-data/tree/master/norne>
130. Armenta, M., Mechanisms and control of water inflow to wells in gas reservoirs with bottom water drive. 2003, Louisiana State University.
131. The Cleeton Field, Block 42/29, UK North Sea. Geological Society, London, Memoirs, 1991. 14(1): p. 409-415.
132. Bailey, R.J. and J.E. Clever, The Windermere Gas Field, Blocks 49/9b, 49/4a, UK Southern North Sea. Geological Society, London, Memoirs, 2003. 20(1): p. 893-901.
133. Heinrich, R.D., Ravenspurn South Field, Blocks 42/29, 42/30, 43/26, UK North Sea. Geological Society, London, Memoirs, 1991. 14(1): p. 469-475.
134. Hillier, A.P., The Sean North, Sean South and Sean East Fields, Block 49/25a, UK North Sea. Geological Society, London, Memoirs, 2003. 20(1): p. 825-833.
135. McCrone, C.W., The Davy, Bessemer, Beaufort and Brown Fields, Blocks 49/23, 49/30a, 49/30c, 53/5a, UK North Sea. Geological Society, London, Memoirs, 2003. 20(1): p. 705-712.
136. Sarginson, M.J., The Barque Field, Blocks 48/13a, 48/14, UK North Sea. Geological Society, London, Memoirs, 2003. 20(1): p. 663-670.
137. Smith, B. and V. Starcher, The Mercury and Neptune Fields, Blocks 47/9b, 47/4b, 47/5a, 42/29, UK North Sea. Geological Society, London, Memoirs, 2003. 20(1): p. 777-787.
138. Schlumberger, Petrel E&P Software Platform. 2017.
139. Cai G., Fang Z., and Chen Y., (2019). 'Estimating the Number of Hidden Nodes of the Single-Hidden-Layer Feedforward Neural Networks.' 15th International Conference on Computational Intelligence and Security (CIS), pp. 172-176, doi: 10.1109/CIS.2019.00044.
140. Teoh E. J., Tan K. C., and Xiang C., (2006) 'Estimating the Number of Hidden Neurons in a Feedforward Network Using the Singular Value Decomposition' IEEE Transactions on Neural Networks, vol. 17, no. 6, pp. 1623-1629, doi: 10.1109/TNN.2006.880582.

APPENDICES

APPENDIX A: OTHER RENORMALISATION METHODS

In chapter 3, two methods for parameterisation/ renormalising the data for NIROM predictions were described, i.e. the Euclidean distance approach developed by Kostorz et al (2020) ^[99] which is suitable for constant parameter cases; and the maximum reduced data approach that was proposed by this work, which is suitable for varying parameter cases. Recall that the renormalisation approach of Kostorz et al (2020) [99] entails computation of equally spaced Euclidean distances between the different parameters of each dataset. But for varying/ forced parameter cases where parameters can increase, reduce or stay the same haphazardly, this is not feasible. Hence the 3 renormalisation methods discussed in this section are the other techniques that were contemplated and evaluated for varying parameter cases, before arriving at the maximum reduced data method that was described in chapter 3.

These 3 renormalisation methods are based on Kostorz et al. (2020) ^[99] Euclidean distance renormalisation method for constant parameter cases. They include a modified renormalisation method, using proxy parameters (which are related to the actual parameters but have an ordered sequence of progression with time) and constraining the NIROM's temporal resolution based on the production recovery efficiency rather than chronological time. Each method was initially examined for correctness by applying them to constant parameter cases and establishing their accuracy before they were evaluated for varying parameter examples.

A.1 MODIFIED NORMALIZATION METHOD

For this method, the reduced training data (α_s from Equation 3.7) and their corresponding parameters (p_s from Equation 3.11) are combined into a single dataset before they are normalised. In other words, after the training runs are reduced into $(d \times nt)$ arrays, the reduced training data are merged with their corresponding un-normalised varying parameters to obtain a $(d+np \times nt)$ array of data.

$$\hat{\alpha}_s = [\alpha_s ; p_s] \quad \text{Equation 10.1}$$

$$\text{where } \alpha_s = V_d^T \tilde{U}_s ; s = 1, 2, 3, \dots, ns$$

Then a second mean extraction and POD are implemented on the combined data. The corresponding Eigenvalues, Σ_2 , and vectors, W , from the second POD are then used as the renormalisation factors for the combined data as shown in Equation 10.2.

$$A_s = \frac{W^T (\hat{\alpha}_s - \bar{\alpha})}{\Sigma_2} \quad \text{Equation 10.2}$$

$$\text{where } \hat{\alpha} \hat{\alpha}^T = W \Sigma_2 W^T \text{ and } \bar{\alpha} = \frac{1}{nt \times ns} \sum_{t=1}^{nt \times ns} \hat{\alpha}_t$$

Figure 10.1 is a summary of how this method differs from the typical renormalisation approach of Kotorz et al (2020) ^[99].

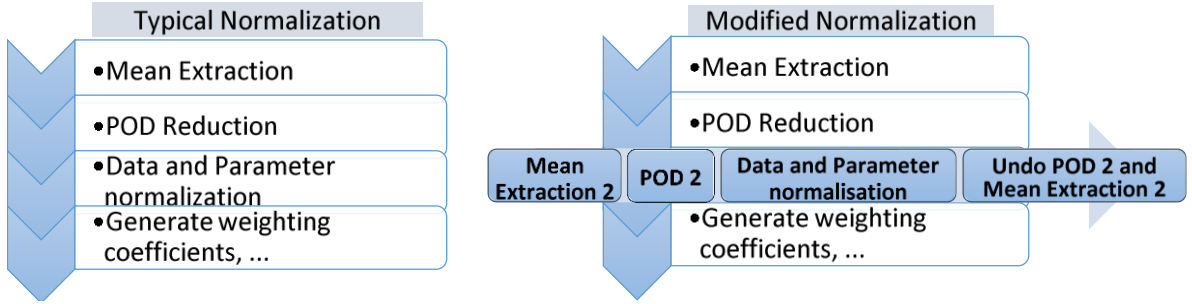


Figure 10.1: Summary of procedure for the modified normalization approach of forced parameter scenarios. Parameterization is executed in hyperspace to nullify the need for sorting parameters in any order of magnitude.

A.2 USING A PROXY AS PARAMETER

In this approach, Kotorz et al.'s (2020) ^[99] renormalisation approach was implemented exactly for the varying parameter cases, but with a different production variable serving as the simulation parameters. Instead of the actual simulation parameter which would be the varying production rates, another production variable that is more stable (continuously increasing or decreasing during the production lifecycle) is used as a substitute parameter. For example the cumulative gas production can be a proxy for the production rate since it is directly related to the production rate and it would typically not fluctuate haphazardly during the period of production. This method seemed practical since Kotorz et al.'s (2020) ^[99] method requires that simulation parameters should increase (or decrease) steadily throughout the production time in consideration.

A.3 CONSTRAINING THE TEMPORAL RESOLUTION WITH A PROXY PARAMETER

This approach involved using a different type of variable as the temporal resolution of the NIROM, in addition to the last approach of using proxy parameters instead of varying rates. The typical NIROM is implemented based on the chronological time in report steps serving as the temporal resolution for the NIROM. However in this approach, a non-time temporal resolution such as the recovery efficiency of production was used to compute the regular intervals that correspond to the temporal resolution of the NIROM. For example, the temporal resolution was pegged on successive increments of cumulative recovery from the reservoir, for instance at 3, 6, 9, 12%, etc. cumulative recovery.

A.4 RESULTS AND DISCUSSION

A.4.1 MODIFIED NORMALIZATION APPROACH AND CONSTRAINED TEMPORAL RESOLUTION

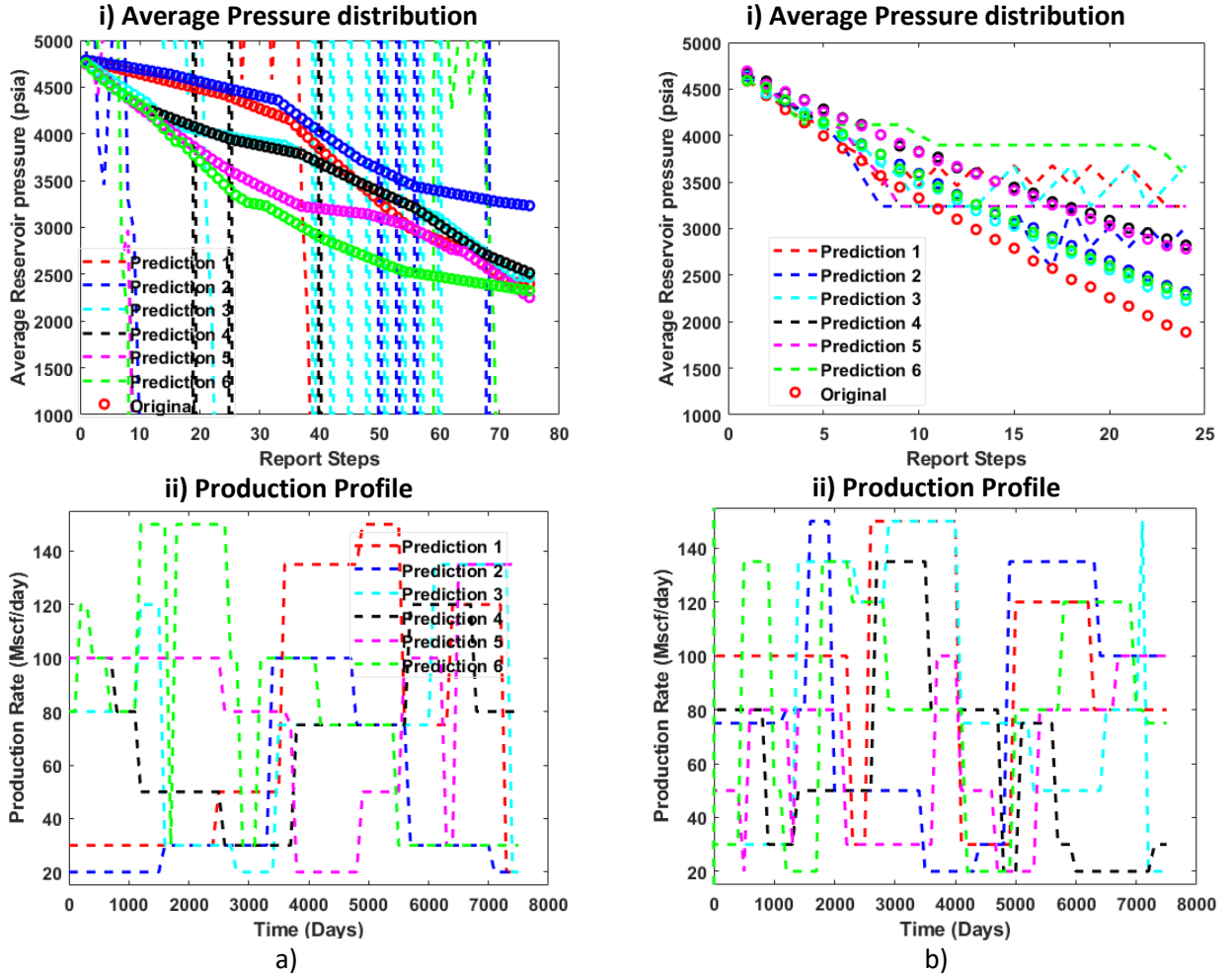


Figure 10.2: Examples of NIROM's attempt at predicting the average pressure distribution in 2 varying rate cases of the synthetic gas reservoir model. In both cases, a) modified normalization approach and b) temporal resolution constrained at equal intervals of cumulative production, NIROM's predictions were not representative. The corresponding production profiles for both cases are the (ii) plots.

None of the above 3 methods of renormalisation was suitable for the varying parameter case. Figure 10.2 shows some results of applying the first two renormalisation methods described above. The erratic behaviour of NIROM predictions with these renormalisation methods was similar to those of transient systems described by Kostorz et al. (2021) ^[100]. Kostorz et al. (2021) proposed that it is more challenging for NIROM to describe transient systems in hyperspace and more training runs' data are required for their dynamics to be adequately learnt. Besides, interpolation errors accumulate as NIROM predictions progress, leading to even poorer interpolations at the time intervals when new forced parameters are imposed on the system.

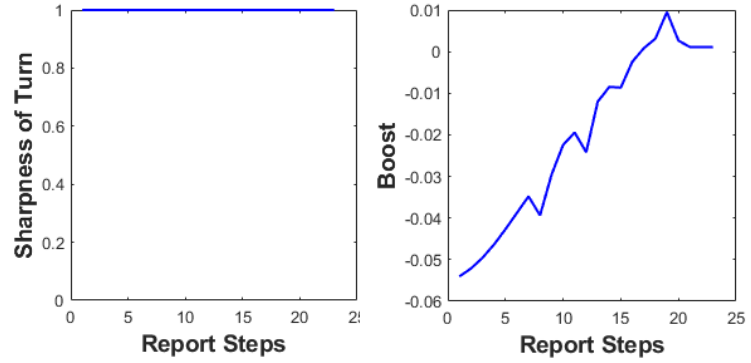
A.4.2 BOOST AND SHARPNESS OF TURN CORRELATION ANALYSIS

To confirm if the NIROM's inability to give representative predictions of the varying parameter cases was a fundamental issue with the gas reservoir systems or the renormalisation methods, the simulation data for the constant and varying parameter cases were tested with Kotorz et al.'s (2021) ^[100] validation framework. Kotorz et al.'s (2021) ^[100] validation framework checks the suitability of a dataset for POD-RBF modelling. The framework was tested on the spatial pressure distribution data of the synthetic gas reservoir model (Figure 10.3).

This framework analyses the phase behaviour of simulation data based on 2 criteria: the extent to which the simulation data experiences change in direction between successive temporal resolutions (sharpness of turn or *mth*-order correlation) and the amount of steadiness or explosion that the data experiences at every interval when a new forced parameter is introduced into the system (boost). The sharpness of turn criteria shows how close a dataset is to being a transient system, so an ideal sharpness of turn plot should have a value of 1, lower values indicate that the case will be more difficult for NIROM to predict. The boost plot should evolve smoothly with values that are close to zero, it tells whether or not the effect of the varying parameters is too sudden for the NIROM to cope with during interpolation.

Figure 10.3 shows a comparison of sharpness of turn and boost plots for constant and varying parameter examples of the synthetic gas model. In Figure 10.3 the sharpness of turn plots for constant and varying rate simulation data are similar, and they are close to the ideal value of 1. This implies that the synthetic gas model's data does not experience a significant change in direction in the phase space and NIROM should be able to model its performance. This can be attributed to the pressure depletion mechanism of gas reservoirs and this was a good sign to show that the gas reservoir system can be modelled with a POD-RBF NIROM. On the other hand, the boost plot for the forced parameter case is much more 'spiky' than for the constant parameter case, with maximum values ~ 0.2 as opposed to ~ 0.06 for the constant rate case. This confirms that when modelling the time-varying/ forced parameter cases, the NIROM will find it more difficult to capture the dynamics of the reservoir's pressure distribution at the time intervals of each sudden change in production rate, than in the constant parameter case.

a) Constant rate production



b) Varying rate production

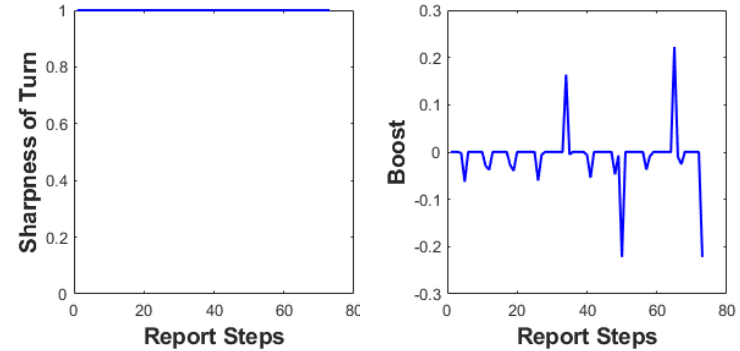


Figure 10.3: Sharpness of turn and boost plots for the spatial pressure data of the synthetic gas reservoir. Both plots are compared for a) constant rate production and b) varying rate production, to evaluate the extent of instability inherent in the reservoir's pressure data and to evaluate if the NIROM can cope with the data for interpolation. The sharpen of turn plots in both cases are satisfactory, but the spikiness of the boost plot for the varying rate case suggests that it would be more challenging for the NIROM to accurately interpolate.

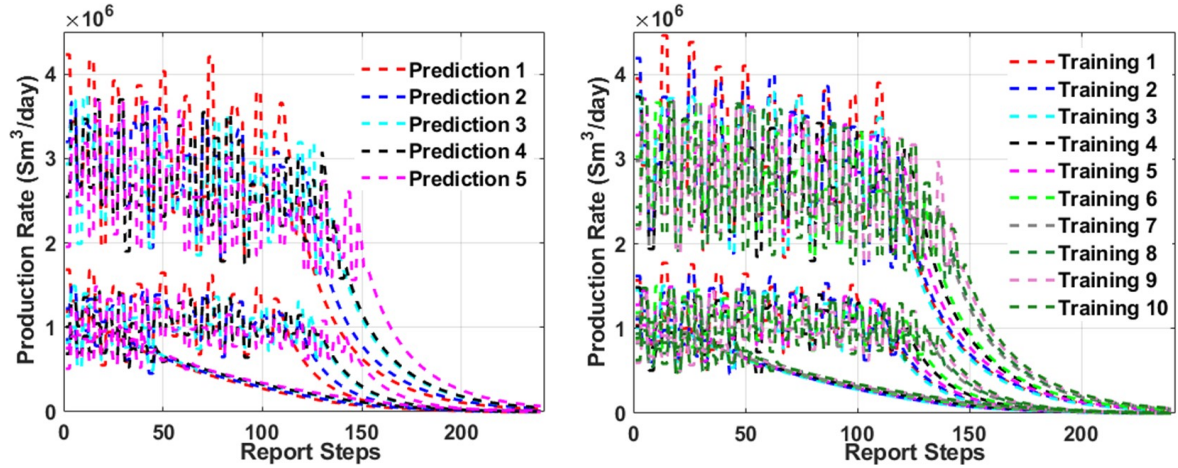
APPENDIX B: MORE POD-AE RESULTS AND OTHER MACHINE LEARNING METHODS

This section considers other aspects of the POD-AE model that were evaluated for this work. It was earlier mentioned that the autoencoder model that was discussed in chapters 3 and 8 was evaluated with fewer than 50 training runs, and apart from the autoencoder, other machine learning approaches were evaluated and analysed. These are discussed in the following sections. A similar optimiser and activation function was used in these examples as for the POD-AE framework that was described in chapter 3, i.e. Adam optimiser and ELU activation function.

B.1 POD-AE RESULTS WITH 10 AND 20 TRAINING RUNS

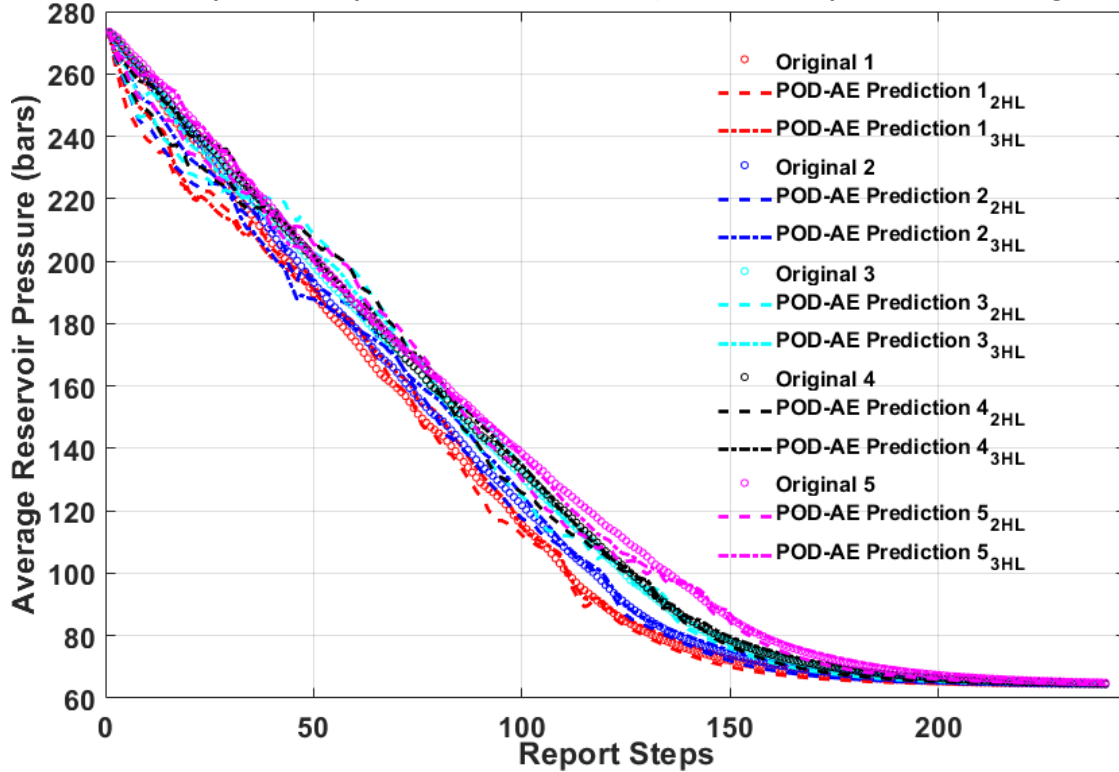
The results of POD-AE modelling that were showed in chapter 8 were based on training the POD-AE with 50 training runs. This section provides some results on the performance of the POD-AE when it was trained with 10 and 20 runs. The results in Figure 11.1 and Figure 11.2 show the instability that was experienced while making predictions of gas reservoir performance with POD-AE. The predictions with 10 training runs (Figure 11.1) were more stable and accurate than those with 20 training runs (Figure 11.2), whereas ML techniques are expected to be more accurate as more data is provided for them to learn a system's dynamics. In the case with 10 training runs (Figure 11.1), the best POD-AE predictions were obtained with the model with 2HL and 3HL, while in Figure 11.2 the best POD-AE

results were obtained from the model with 2, 4 and 5 HLs. Besides, Figure 11.2 showed the undulating under and over prediction of the AE model for the prediction examples of the modified Norne gas field with varying production profiles and shut-in well(s). The AE model was overall less stable and inconsistent for modelling other cases of the gas field.

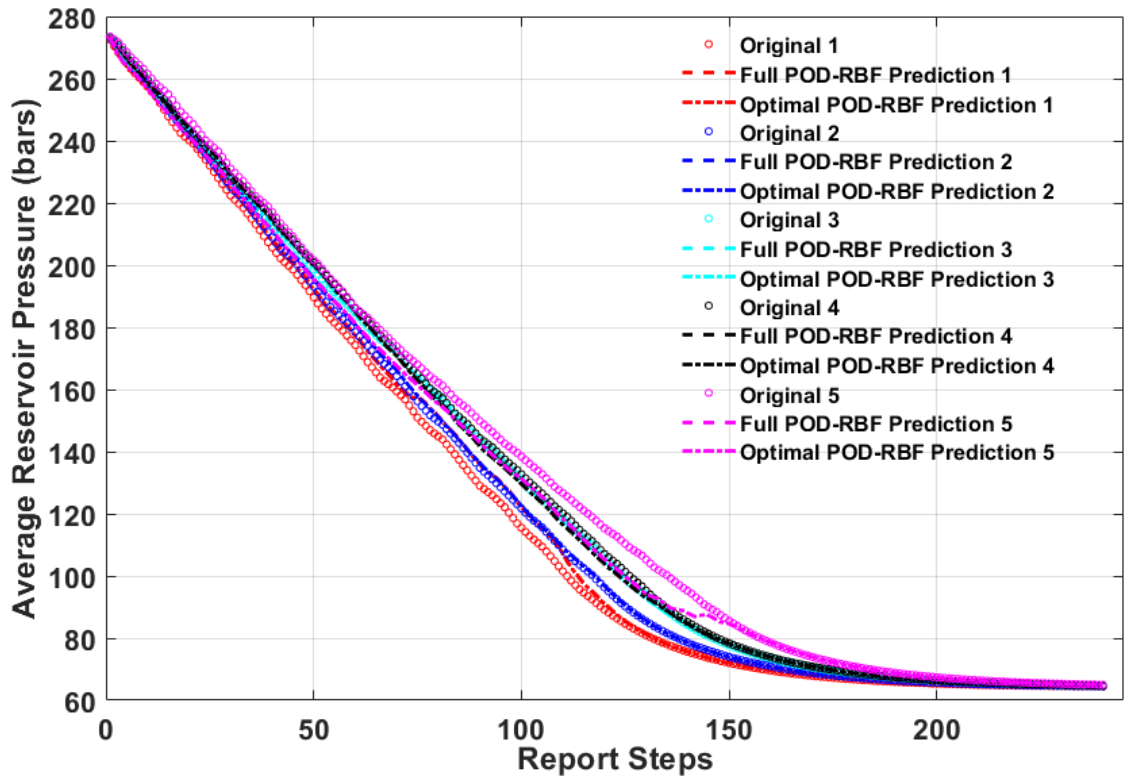


a) Production profiles for prediction runs

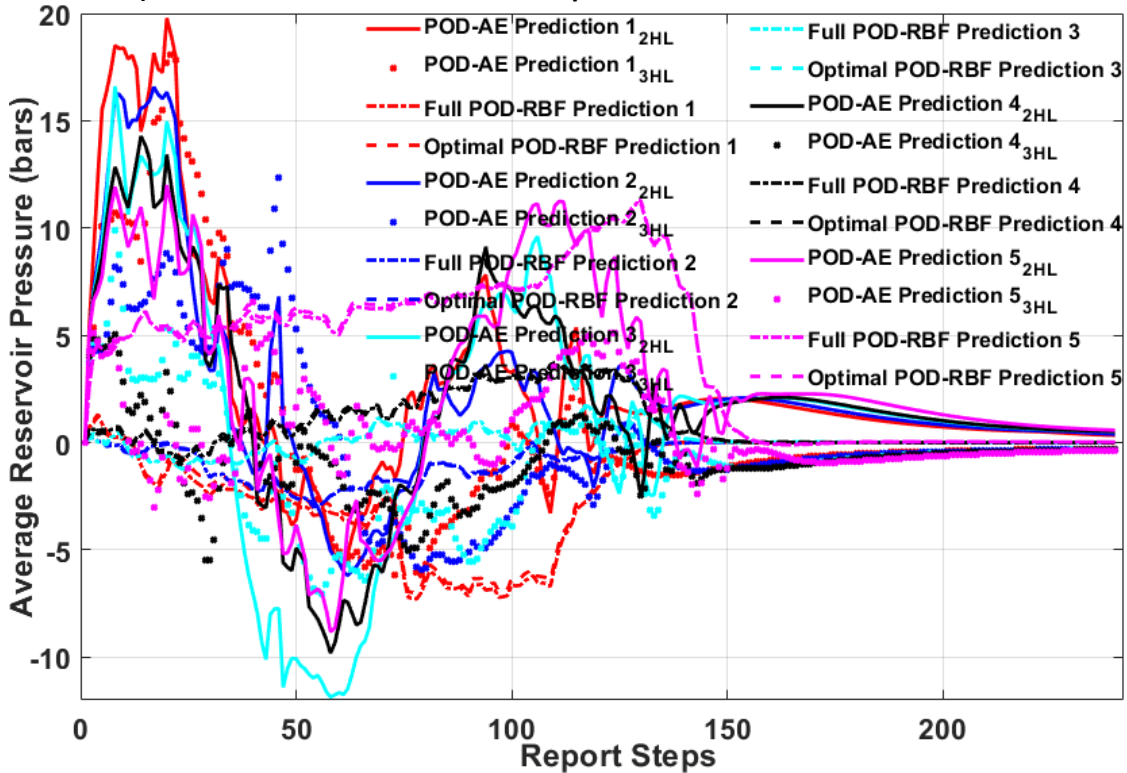
b) Production profiles for training runs



c) Predictions with POD-AE with 2 and 3 HLs

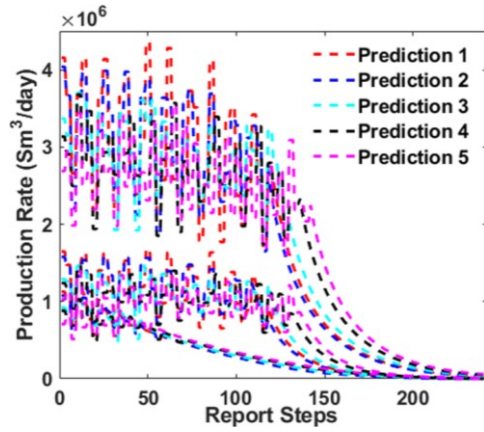


d) Predictions with POD-RBF with optimal and full rank POD basis functions

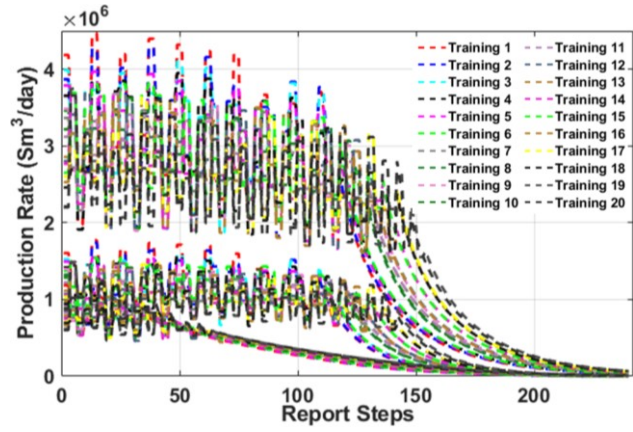


e) Average errors in the predictions with POD-AE and POD-RBF

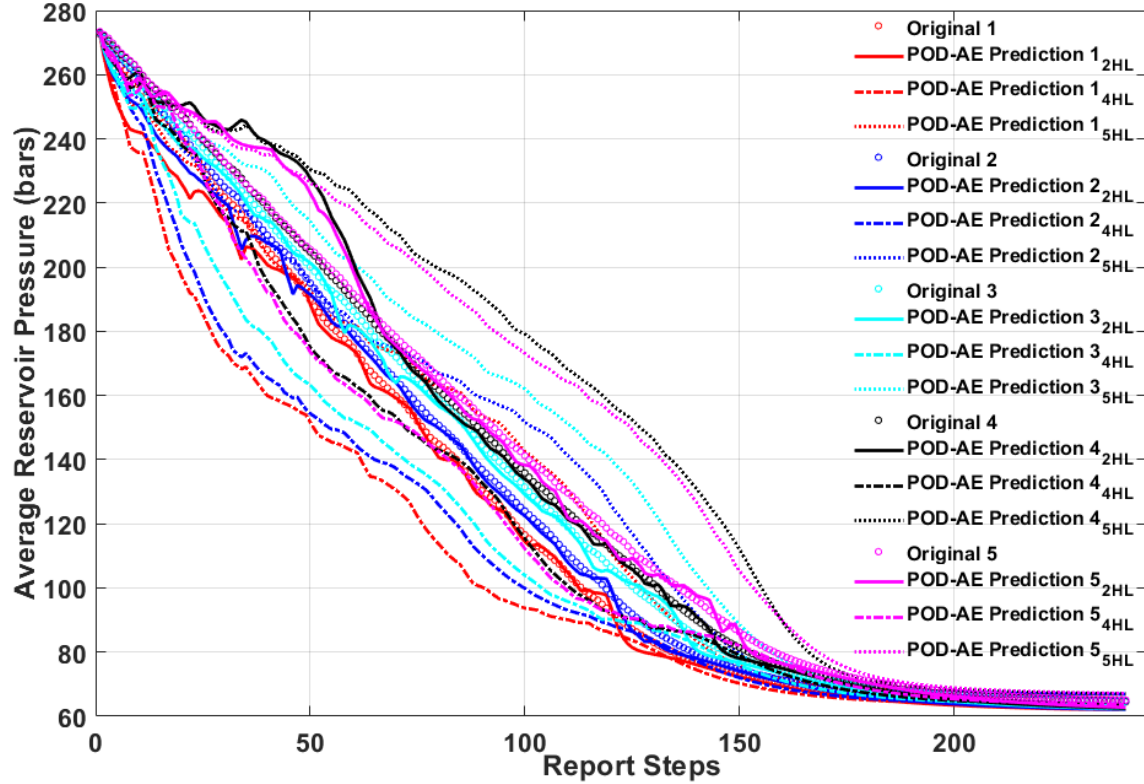
Figure 11.1: Results of modelling the modified Norne gas field with 10 training runs when at least one of the wells was shut-in during the production lifecycle. a) The production profiles for the prediction cases tested with POD-AE and POD-RBF b) The production profiles for the 10 training runs used in this example c) Predictions of the average pressure distribution in the field using POD-AE with 2 and 3 HLs d) Predictions of the average pressure distribution in the field using the optimal and full rank POD-RBF and e) the average error in the prediction of pressure distribution from the 4 models in c and d.



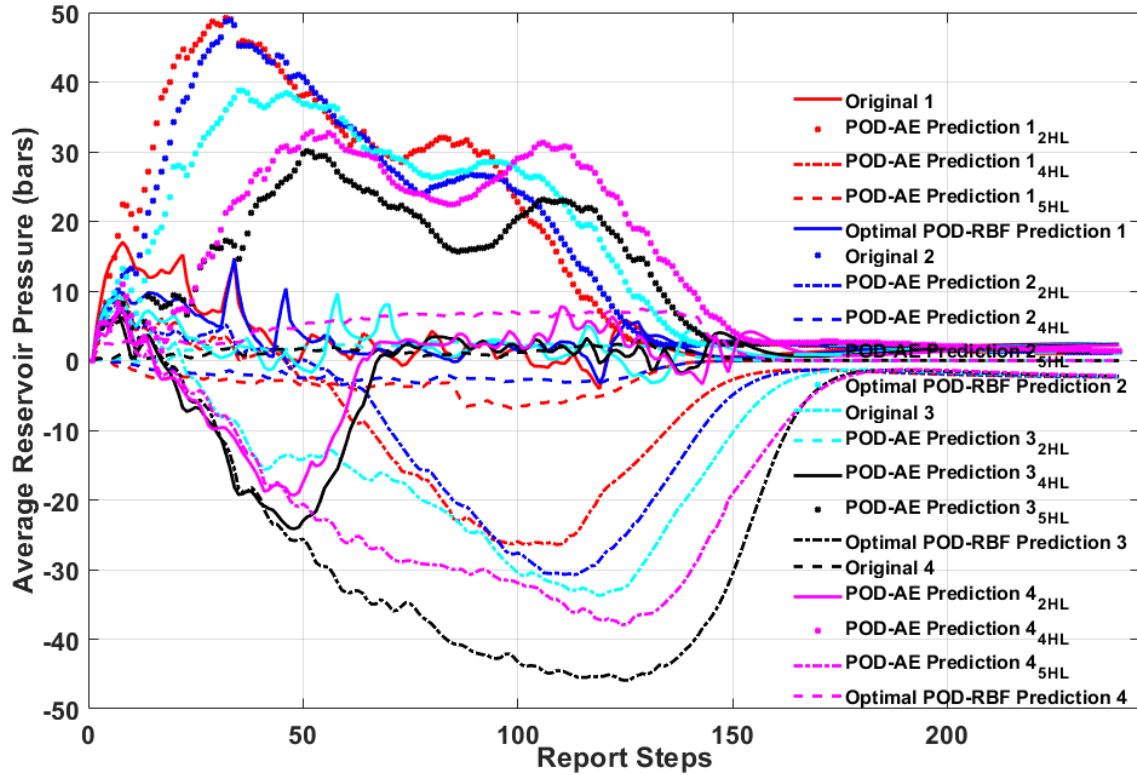
a) Production profiles for prediction runs



b) Production profiles for training runs



c) Predictions with POD-AE with 2, 4 and 5 HLs



d) Average errors in the predictions with POD-AE and POD-RBF

Figure 11.2: Results of modelling the modified Norne gas field with 20 training runs when at least one of the wells was shut-in during the production lifecycle. a) The production profiles for the prediction cases tested with POD-AE and POD-RBF b) The production profiles for the 20 training runs used in this example c) Predictions of the average pressure distribution in the field using POD-AE with 2, 4 and 5 HLs d) The average error in the prediction of pressure distribution from the 3 models in (c) as well as the optimal POD-RBF model.

B.2 POD-AE WITH DROPOUT LAYERS

In section 3.5.2 the POD-AE structure that was used in this work was described as having fully connected/ dense layers. A neural network with dense layers has all its layers in a given layer connected to all the layers in the preceding layer. In some cases, to improve the generalisation/ predictive ability of NNs, not all the preceding layers of a NN are connected to their next layer. Such models are said to have dropout layers. Dropout layers randomly removing some HL connections during the training of ML models to minimise overfitting^[106]. The decision to use dense layers in this work was taken because dropping out some layers of the autoencoder network yielded less satisfactory results. Figure 11.3 shows an example of where dropout layers have been added to the autoencoder for modelling a triennially varying production profile case of the modified Norne gas field. Figure 11.3 shows that as more layers were dropped out of the autoencoder model, its predictive capacity became worse.

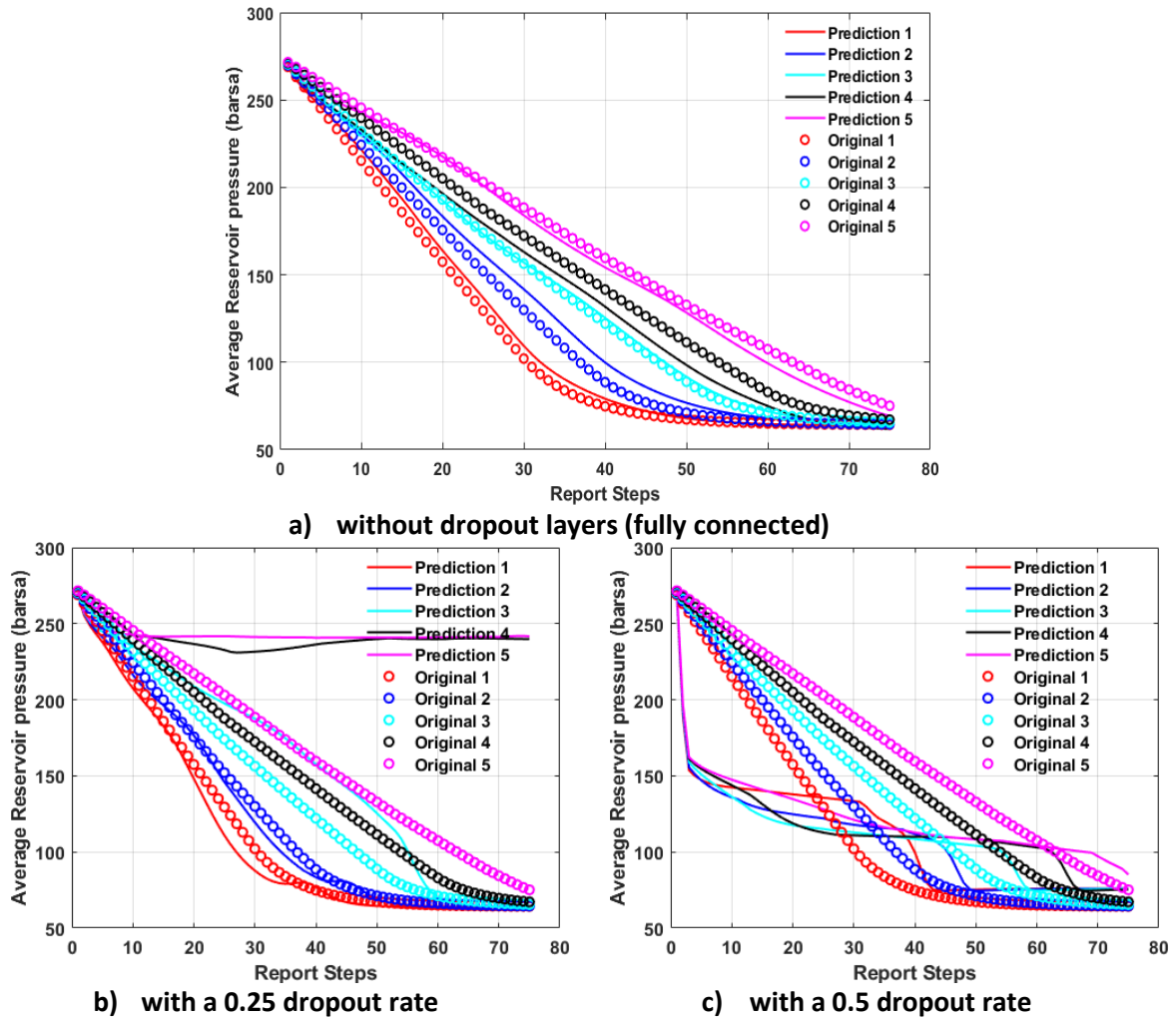


Figure 11.3: Modelling a triennially varying production case of the modified Norne field case with shut-in well(s). This example is from the 4HLs model a) without dropout layers b) with 0.25 dropout rate and c) with 0.5 dropout rate. The best prediction was obtained with the fully connected model.

B.3 OTHER ML MODELS CONSIDERED IN THIS WORK

This section gives a brief overview of the other ML techniques that were considered during this work. For a more detailed insight into these concepts, the reader is referred to respective references.

B.3.1 FEED-FORWARD NEURAL NETWORK (FFNN),

These are the most basic and straightforward ANNs. They consist of several interconnected layers between their input and output layers, they are unidirectional and have no looped or feedback connections ^[105]. In this work FFNNs with one to five hidden layers were evaluated for their suitability for modelling gas reservoirs. Figure 11.4 is a schematic of a three-layer FFNN.

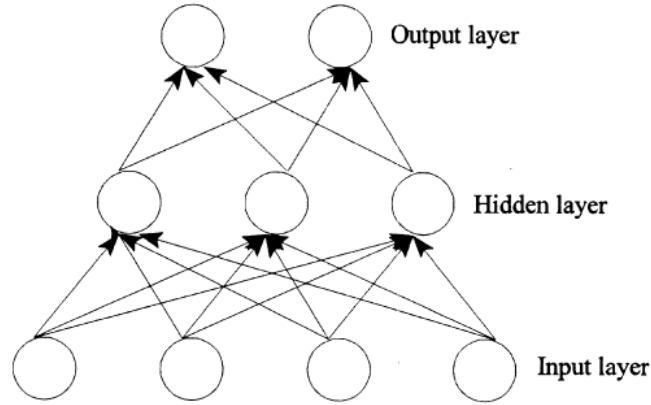


Figure 11.4: An example of a three-layer feedforward neural network ^[111].

Figure 11.5 shows an example of predictions of the average pressure distribution in the modified Norne gas field using FFNNs. Figure 11.5 shows unsatisfactory predictions from all the FFNNs that were tested, 1 to 5 HLs. The FFNNs gave the same average prediction for all 5 predictions, they were unable to distinguish between individual predictions.

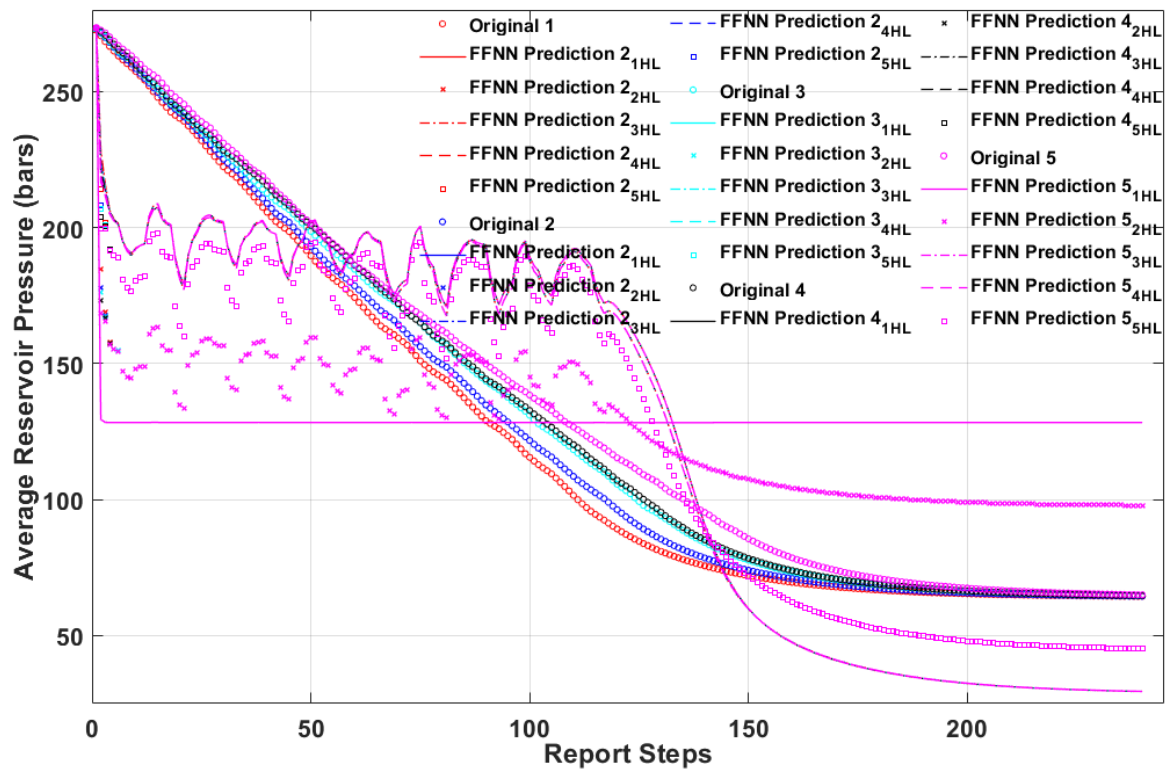


Figure 11.5: Predictions of the average pressure distribution in the modified Norne gas field with shut-in wells. 5 FFNNs were tested and they all gave averaged values for all 5 predictions, being unable to distinguish the dynamics of the different prediction examples.

B.3.2 LONG SHORT TERM MEMORY (LSTM) MODEL

LSTMs are a type of recurrent neural network that has a looped feedback mechanism to improve their generalisation/ predictive ability. For LSTMs, the feedback loops help the model to learn from both recent layers and other layers farther backwards in the network, thereby fine-tuning their predictions and minimising their prediction errors ^[105, 106]. This makes LSTM models more time consuming to run than other networks without feedback loops e.g. FFNN. LSTMs also require a large computational capacity for storage of the extra parameters and more data that are saved for the LSTMs to look back to in future steps ^[13]. For this work, an autoencoder-LSTM was applied to modelling the gas reservoir system. The AE-LSTM was of a similar structure as the AE model, except that the innermost layer of the AE was replaced with an LSTM layer as shown in Figure 11.6. Just like in the FFNN case, the AE-LSTM model was unable to identify the difference in individual prediction examples, for all of the AE-LSTM models with 1 to 5 HLs Figure 11.7.

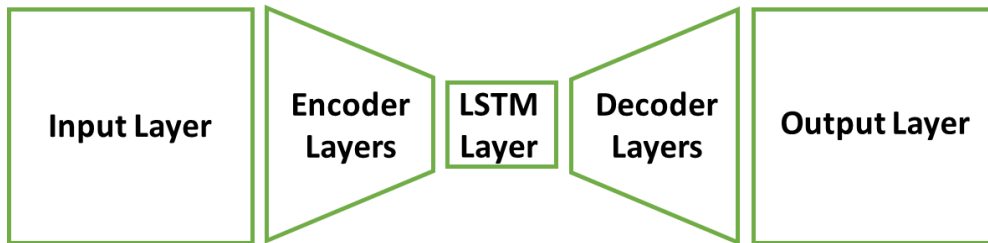


Figure 11.6: A schematic of the AE-LSTM model that was tested for this work. The model consists of an autoencoder whose innermost layer is an LSTM model.

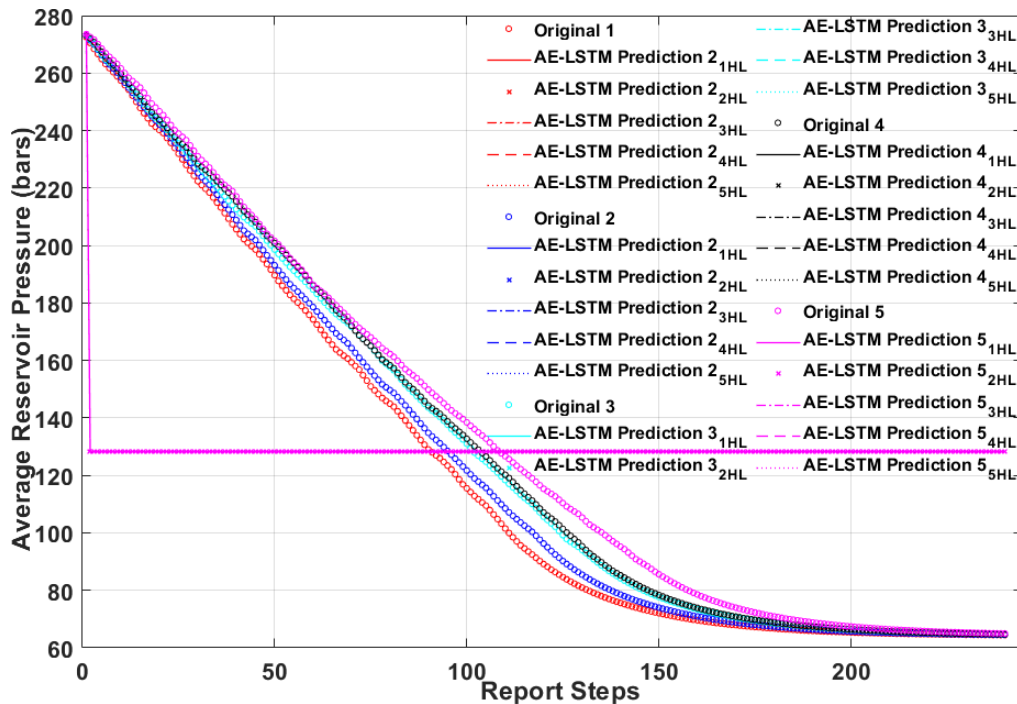


Figure 11.7: Results of using an AE-LSTM for modelling the average pressure distribution in a modified Norne field example with shut-in well(s). All the AE-LSTM models gave identical predictions for all the cases (HLs) that were tested.

APPENDIX C: RESERVOIR DATA FILES AND NIROM MODELS

Below are some of the data files and models for the gas reservoirs that were considered in this work.

C.1 RESERVOIR MODELS

C.1.1 SYNTHETIC GAS MODEL WITH FETKOVICH AQUIFER

Figure 12.1 is the relative permeability curve for the synthetic gas reservoir model that was built using Schlumberger's Petrel ^[138], showing the relative permeability to gas and water (K_{rg} and K_{rw} respectively). Next is ECLIPSE data file for the base case synthetic gas reservoir, including the PVT data and relative permeability curves data under the "PROPS" section.

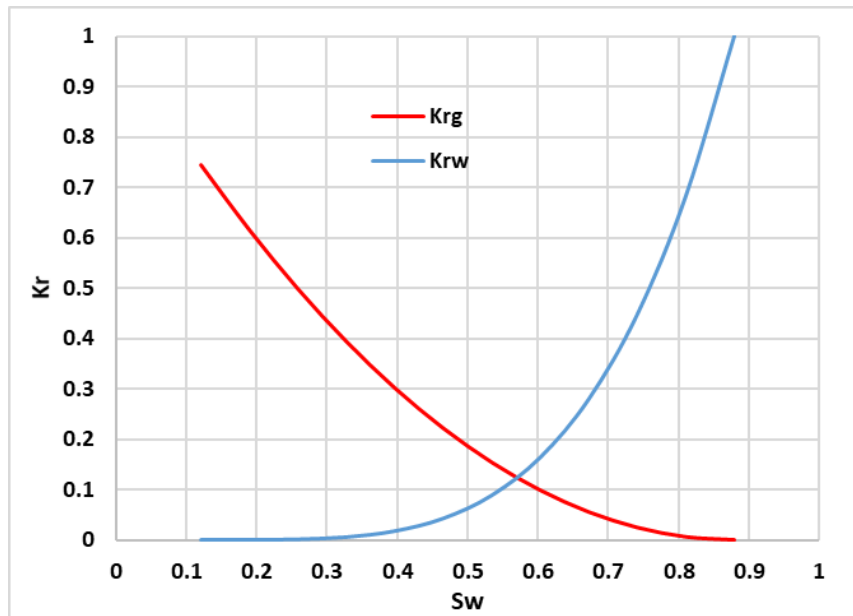


Figure 12.1: Relative permeability curves for the synthetic gas reservoir.

```
-- SYNTHETIC GAS RESERVOIR MODEL
```

```
RUNSPEC =====
TITLE
GAS-WATER CASE
```

```
DIMENS
-- Nx  Ny  Nz
19  21  6/
```

```
WATER
```

```
GAS
```

```
FIELD
```

```
EQLDIMS
1      100  10  1  1 /
```

```

TABDIMS
  1      1      16      20      5      12 /

WELLDIMS
  1      100      1      1 /

START
  1 'JAN' 1995 /

AQUDIMS
  30      3*      4      3000      2 /

FMTIN

UNIFOUT

UNIFIN

GRID      =====
DXV
19*105/

DYV
21*105/

DZ
2394*55/

TOPS
399*8095/

PORO
2394*0.2/

PERMX
2394*50/

PERMZ
2394*5 /

COPY
PERMX PERMY/
/

PROPS      =====
SGFN      -- Generated: Petrel      Corey 2
           0.12      0      0
0.1890909091      0.006164196435      0
0.2581818182      0.02465678574      0
0.3272727273      0.05547776791      0
0.3963636364      0.09862714295      0
0.4654545455      0.1541049109      0
0.5345454545      0.2219110716      0
0.6036363636      0.3020456253      0
0.6727272727      0.3945085718      0
0.7418181818      0.4992999112      0
0.8109090909      0.6164196435      0
           0.88      0.7458677686      0
/

SWFN      -- Generated: Petrel      Corey 4
           0.12      0      0
0.1890909091      6.830134554E-05      0
0.2581818182      0.001092821529      0
0.3272727273      0.005532408988      0
0.3963636364      0.01748514446      0
0.4654545455      0.04268834096      0
0.5345454545      0.08851854382      0

```

Appendix C: Reservoir Data files and NIROM Models

```

0.6036363636      0.1639915306      0
0.6727272727      0.2797623113      0
0.7418181818      0.4481251281      0
0.8109090909      0.6830134554      0
0.88              1              0
/

PVTW      4200      1.0205      2.1317E-06      -- Generated: Petrel
0.56127      0 /

RVCONSTT      -- Generated: Petrel
0      1150 /

PVDG
--      PGAS      BGAS      VISGAS      -- Generated: Petrel
      1150      2.609789961      0.01464603392
1360.526316      2.186577631      0.01505977657
1571.052632      1.879732362      0.01550970561
1781.578947      1.648163466      0.0159938051
1992.105263      1.468116226      0.01650958703
2202.631579      1.324878837      0.01705405537
2413.157895      1.208833088      0.01762376274
2623.684211      1.113409707      0.01821493658
2834.210526      1.033958126      0.01882364638
3044.736842      0.967087626      0.01944598108
3255.263158      0.9102656529      0.02007820927
3465.789474      0.8615630748      0.02071690365
3676.315789      0.8194867376      0.02135902178
3886.842105      0.7828657371      0.02200194429
4097.368421      0.7507718551      0.02264347793
4307.894737      0.7224624388      0.02328183313
4518.421053      0.6973384853      0.02391558595
4728.947368      0.6749133276      0.02454363225
4939.473684      0.654788901      0.02516514021
      5150      0.6366375455      0.02577950483
/

DENSITY      -- Generated: Petrel
37.457      71.397      0.046568 /

RTEMP
180 /

ROCK
4000      .3E-5      /

RPTPROPS
/

SOLUTION =====
EQUIL
      8115      4200      8420      0      0      0      0      0      0 /

RPTSOL
'PRES' /

AQUFETP
1      8420 1*      100E8 5.6216E-6 500      3*      /
/

AQUANCON
1      1      19 1      21 1      6      K+ /
/

SUMMARY =====
RUNSUM

EXCEL

```

```

FWPR
FWPT
FPR
WBHP
/
WGPR
/
WGPT
/
WWGR
/
ANQR
/
ANQT
/
ANQP
/
AAQR
/
AAQT
/
AAQP
/
WPI
/
SCHEDULE =====

RPTSCHED
'PRES' 'SWAT' 'SGAS' 'RS' 'RESTART=2' 'FIP=1' 'WELLS=2' 'SUMMARY=2'
'CPU=2' 'NEWTON=2' /

WELSPERCS
'PROD' 'G' 10 11 8115 'GAS' /
/

COMPDAT
'PROD' 10 11 1 6 'OPEN' 0 -1 0.5 /
/

WCONPROD
'PROD' 'OPEN' 'BHP' 5* 1000/ 'GRAT' 2* 6772 2* 14.7 /
/

TSTEP
25*300 /

END =====

```

C.1.2 MODIFIED NORNE GAS FIELD MODEL FOR THE SEASONALLY VARYING PRODUCTION CASE

Below is the eclipse data file for the base case modified Norne gas field, including the PVT data and relative permeability curves under the “PROPS” section. The other INCLUDE files that defined the field’s structure and grid properties are available on Github by Rasmussen (2014) ^[128, 129].

```

-- This reservoir simulation deck is made available under the Open Database
-- License: http://opendatacommons.org/licenses/odbl/1.0/. Any rights in
-- individual contents of the database are licensed under the Database Contents
-- License: http://opendatacommons.org/licenses/dbcl/1.0/
-- Copyright (C) 2015 Statoil

-- Norne full field model for SPE ATW 2013
RUNSPEC

```

```

DIMENS
46 112 22 /

GRIDOPTS
'YES' 0 /

WATER

GAS

METRIC

START
06 'NOV' 1997 /

EQLDIMS
1 /

WELLDIMS
130 36 15 84 /

TABDIMS
1 1 33 60 16 60 /

CPR
/

NSTACK
25 /

UNIFIN
UNIFOUT
-----
--      Input of grid geometry
-----

GRID

GRIDFILE
2 /

GRIDUNIT
METRES /

MESSAGES
8*10000 20000 10000 1000 1* /

NOECHO
-----
--      Grid and faults
-----

INCLUDE
'IRAP_1005.GRDECL' /

INCLUDE
'ACTNUM_0704.prop' /
-----
--      Input of grid parameters
-----

INCLUDE
'PORO_0704.prop' /

INCLUDE
'NTG_0704.prop' /

INCLUDE
'PERM_0704.prop' /

```

```

COPY
  PERMX PERMY /
  PERMX PERMZ /
/

-- based on same kv/kh factor
MULTIPLY
  'PERMX' 0.5      1 46 1 112 1 22 /
  'PERMY' 0.5      1 46 1 112 1 22 /
  'PERMZ' 0.2      1 46 1 112 1 1 /   Garn
  'PERMZ' 0.04     1 46 1 112 2 2 /   Garn
  'PERMZ' 0.25     1 46 1 112 3 3 /   Garn
  'PERMZ' 0.0      1 46 1 112 4 4 /   Not      (inactive anyway)
  'PERMZ' 0.13     1 46 1 112 5 5 /   Ile 2.2
  'PERMZ' 0.13     1 46 1 112 6 6 /   Ile 2.1.3
  'PERMZ' 0.13     1 46 1 112 7 7 /   Ile 2.1.2
  'PERMZ' 0.13     1 46 1 112 8 8 /   Ile 2.1.1
  'PERMZ' 0.09     1 46 1 112 9 9 /   Ile 1.3
  'PERMZ' 0.07     1 46 1 112 10 10 /   Ile 1.2
  'PERMZ' 0.19     1 46 1 112 11 11 /   Ile 1.1
  'PERMZ' 0.13     1 46 1 112 12 12 /   Tofte 2.2
  'PERMZ' 0.64     1 46 1 112 13 13 /   Tofte 2.1.3
  'PERMZ' 0.64     1 46 1 112 14 14 /   Tofte 2.1.2
  'PERMZ' 0.64     1 46 1 112 15 15 /   Tofte 2.1.1
  'PERMZ' 0.64     1 46 1 112 16 16 /   Tofte 1.2.2
  'PERMZ' 0.64     1 46 1 112 17 17 /   Tofte 1.2.1
  'PERMZ' 0.016    1 46 1 112 18 18 /   Tofte 1.1
  'PERMZ' 0.004    1 46 1 112 19 19 /   Tilje 4
  'PERMZ' 0.004    1 46 1 112 20 20 /   Tilje 3
  'PERMZ' 01.0     1 46 1 112 21 21 /   Tilje 2
  'PERMZ' 01.0     1 46 1 112 22 22 /   Tilje 1
/

-----
--      Barriers
-----

-- MULTZ multiplies the transmissibility between blocks
-- (I, J, K) and (I, J, K+1), thus the barriers are at the
-- bottom of the given layer.

-- Region barriers
INCLUDE
  'MULTZ_HM_1.INC' /

-- Field-wide barriers
EQUALS
  'MULTZ'      1.0      1 46 1 112 1 1 / Garn3      - Garn 2
  'MULTZ'      0.05     1 46 1 112 15 15 / Tofte 2.1.1 - Tofte 1.2.2
  'MULTZ'      0.001    1 46 1 112 18 18 / Tofte 1.1   - Tilje 4
  'MULTZ'      0.00001  1 46 1 112 20 20 / Tilje 3     - Tilje 2
-- The Top Tilje 2 barrier is included as MULTREGT = 0.0
/

-----
PROPS
-----

--      Input of fluid properties and relative permeability
-----

SWFN
--Sw      Krw      Pcw
0.151090      0.0      400.0
0.151230      0.0      359.190
0.151740      0.0      257.920
0.152460      0.0      186.310
0.156470      0.0      79.060
0.165850      0.0      40.010
0.178350      0.0      27.930
0.203350      0.000010  20.400
0.253350      0.000030  15.550

```

0.350000	0.000280	11.655
0.352000	0.002292	8.720
0.354000	0.004304	5.947
0.356000	0.006316	3.317
0.358000	0.008328	1.165
0.360000	0.010340	0.463
0.364395	0.015548	-0.499
0.368790	0.020756	-1.139
0.370000	0.022190	-1.194
0.380000	0.035890	-1.547
0.400000	0.069530	-1.604
0.433450	0.087900	-1.710
0.461390	0.104910	-1.780
0.489320	0.123290	-1.860
0.517250	0.143030	-1.930
0.573120	0.186590	-2.070
0.601060	0.210380	-2.130
0.656930	0.261900	-2.260
0.712800	0.318650	-2.380
0.811110	0.430920	-2.600
0.881490	0.490000	-2.750 /

SGFN

--Sg	Krg	Pcg
.00	.0	0
.04	.0	.2
.1	.022	.5
.2	.1	1
.3	.24	1.5
.4	.34	2
.5	.42	2.5
.6	.5	3
.7	.8125	3.5
0.84891	1.	3.9 /

PVTW

306.1 1.03 0.000041 0.3 0 / (BAR RM3/SM3 1/BAR CP)

ROCK

306.1 0.000053 / (BAR 1/BAR)

DENSITY

800 1022 0.9907 / (KG/M3)

PVDG

30	0.04234	0.01344
530	0.003868	0.02935

/

RTEMP

1* /

NOECHO

SOLUTION

EQUIL

-- Datum	P	woc	Pc	goc	Pc	Rsvd	Rvvd	
2582.0	269.46	2618.0	0.0	2582.0	0.0	1*	0	0/ E: Garn

RPTSOL

FIP=3 / --SWAT PRES

SUMMARY

RUNSUM

```

EXCEL

RPTONLY

FGPR
FGPT
FWPR
FWPT
FPR
FPRGZ
FWGR
WBHP
/
WGPR
/
WWPR
/
WWGR
/
ANQR
/
ANQT
/
ANQP
/
WGPT
/
-----
SCHEDULE

RPTSCHED
FIP=1 RESTART=6    /PRES

RPTRST
3/

NOECHO

WEL SPECS
    'PROD01' 'P'      8 26 1* 'GAS' 7* /
    'PROD02' 'P'     17 31 1* 'GAS' 7* /
    'PROD03' 'P'     38 96 1* 'GAS' 7* /
/

COMP DAT
    'PROD01' 8    26    3    3    'OPEN' 1* 118.457 /
    'PROD02' 17   31    9    9    'OPEN' 1* 36.540 /
    'PROD03' 38   96    2    2    'OPEN' 1* 27.041 /
/

WPIMULT
PROD01 50 /(all connections in the well PROD1)
PROD02 50 /
PROD03 50 /
/

INCLUDE
ProdProfile.inc/ -- generated from NIROM interface based on "SeasonalProdProfile"
function

END

```

C.1.3 WATER CONING GAS RESERVOIR MODEL

This water coning gas reservoir model is based on the work of Armenta (2003) ^[130], although this model has 4 components in the theta direction (divided into 4 quadrants) and Armenta (2003) ^[130] has one component in the theta direction.

```

RUNSPEC
TITLE
--Comparison of Water Coning in Oil and Gas Wells After Water
--Breakthrough
--Gas-Water Model
MESSAGES
  6* 3*5000 /
RADIAL
DIMENS
-- NR Theta NZ
  26 4 127/
GAS
WATER
FIELD
REGDIMS
  2 /
WELLDIMS
  2 100 1 2 /
START
  1 'Jan' 2002 /
NSTACK
  25/
UNIFOUT
CPR
/
-----
GRID
TOPS
104*5000 /
INRAD
0.333 /
COORDSYS
-- K1 K2 completed --
  2* 'COMP' /
DRV
0.4170  0.3016  0.4229  0.5929  0.8313  1.166
1.634   2.292   3.213   4.505   6.317   8.857
12.42   17.41   24.41   34.23   48.00   67.30
94.36   132.3   185.5   260.1   364.7   511.4
717.0   2500 /
DZ
12480*0.5
624*5
104*550
/
EQUALS
'DTHETA' 90 /
'PERMR' 10 /
'PERMTHT' 100 /
'PERMZ' 5 /
'PORO' 0.25 /
'PORO' 0.00001  26 26 1 1 1 100 /
'PORO' 10      26 26 1 1 101 127/--8 /
/
INIT
-----
PROPS
DENSITY
45 64 0.046 /

```

```

ROCK
2500 10E-6 /
PVTW
2500 1 2.6E-6 0.68 0 /
PVZG
-- Temperature
120 /
-- Pres Z    Visc
100  0.989  0.0122
300  0.967  0.0124
500  0.947  0.0126
700  0.927  0.0129
900  0.908  0.0133
1100 0.891  0.0137
1300 0.876  0.0141
1500 0.863  0.0146
1700 0.853  0.0151
1900 0.845  0.0157
2100 0.840  0.0163
2300 0.837  0.0167
2500 0.837  0.0177
2700 0.839  0.0184
3200 0.844  0.0202
/
--Saturation Functions
--Sgc = 0.20
--Krg @ Swir = 0.9
--Swir = 0.3
--Sorg = 0.0
SGFN
--Using Honarpour Equation 71
-- Sg Krg Pc
0.00 0.000 0.0
0.10 0.000 0.0
0.20 0.020 0.0
0.30 0.030 0.0
0.40 0.081 0.0
0.50 0.183 0.0
0.60 0.325 0.0
0.70 0.900 0.0
/
SWFN
--Using Honarpour Equation 67
-- Sw Krw Pc
0.3 0.000 0.0
0.4 0.035 0.0
0.5 0.076 0.0
0.6 0.126 0.0
0.7 0.193 0.0
0.8 0.288 0.0
0.9 0.422 0.0
1.0 1.000 0.0
/
-----
REGIONS
FIPNUM
10400*1
2808*2 --2912*2
/
-----
SOLUTION
EQUIL
5000 1500 5050 0 5050 0 /
RPTSOL
6* 2 2 /
-----
SUMMARY
EXCEL

```

```

RPTONLY
FGPR
FGPT
FWPR
FWPT
FPR
FWGR
WBHP
/
WWCT
/
-----
SCHEDULE
RPTSCHED
  'PRES' 'SWAT' 'SGAS' 'RS' 'RESTART=2' 'FIP=1' 'WELLS=2' 'SUMMARY=2'
  'CPU=2' 'NEWTON=2' / --6* 2 /
RPTRST
  4 / restarts once a year
TUNING
  0.0007 30.4 0.0007 0.0007 / --1.2 /
  4* 12/-- 3* 0.00001 3* 0.0001 /
/
WELSPECS
-- WELL GROUP LOCATION BHP PI shut-in cross
-- NAME I J DEPTH DEFN instanc flow?
  'P' 'G' 1 1 5000 'GAS' 2* 'STOP' 'YES' /
/
COMPDAT
--WELL -LOCATION- OPEN/ SAT CONN well
--NAME I J K1 K2 SHUT TAB FACT dia
  'P' 1 1 1 100 'OPEN' 2* 0.666 /
/
INCLUDE
'WCONPROD.IN'/

TSTEP
  6*30.4 /
TSTEP
  35*30.4 /

TUNING
  0.0007 3 0.0007 0.0007 /
/
/
TSTEP
  2*30.4 /
TUNING
  0.0007 3 /
/
/
TSTEP
  12*30.4 /
TUNING
  0.0007 3 /
/
/
TSTEP
  1*30.4 /
TUNING
  0.0007 .5 /
/
/
TSTEP
  2*30.4 /
TUNING
  0.0007 3 /
/
/

```

```

TSTEP
1*30.4 /
TUNING
0.0007 3 /
/
/
TSTEP
8*30.4 /
TUNING
0.0007 .1 /
/
/
TSTEP
1*30.4 /
TUNING
0.0007 3 /
/
/
TSTEP
3*30.4 /
TUNING
1 30.4 /
/
/
TSTEP
48*30.4 /
TSTEP
48*30.4 /
TSTEP
48*30.4 /
TSTEP
25*30.4 /

END

```

C.2 NIROM MODELS

C.2.1 POD-RBF MODEL WITH SEASONALLY VARYING PRODUCTION PROFILE

Below is a POD-RBF NIROM for the seasonally varying production file case. This NIROM creates the time-varying production profiles (using the `SeasonalProdProfile` function) as well as runs ECLIPSE to generate training and prediction runs. The NIROM is also trained and used for predictions within the same interface.

```

close all
clear all
%% Setting up data for creating seasonal rate profile
ns = 20; % Number of training runs
header = 6; % number of header lines in ECLIPSE .RSM file
Parameter = 'GRAT'; % or 'BHP';% production parameter
prod_prof = {[18e5 17e5 16E5 15e5 14e5 13e5 12e5 11e5 10e5 9e5 8E5],...
[]}; % parameter for production profile (Table 8)
if strcmp(Parameter,'GRAT') == 1 % setting the main production control
    e = 1;
else
    e = 2;
end

%% Generating data for Training runs production profiles (combination of rates and
duration)
[Rate_Yr_training, Durtn_F_training, Durtn_Fsum_training] = SeasonalProdProfile(ns,
prod_prof,e);

```

```

Para_intended = Rate_Yr_training;

%% Simulation runs for Training NIROM
% Initializing NIROM for training runs data
ii=1;
Z = cell(1,ns); PRES = cell(1,ns); GSAT = cell(1,ns); PRES_ = cell(1,ns);
GSAT_ = cell(1,ns); ProdVar = cell(1,ns); ProdVar_ = cell(1,ns);
TIME_DAYS1 = cell(1,ns); Para_actual = cell(1,ns);
TEclipseTrainingData = zeros(1,ns); T_ExtractTrainingData = zeros(1,ns);

tic
for i=1:ns
    aaa=1;

    % creating production profile INCLUDE text file
    fid1 = fopen('ProdProfile.inc', 'wt');
    for a=1:length(Para_intended{i})

        fprintf(fid1, '\n');
        fprintf(fid1, 'WCONPROD');
        fprintf(fid1, '\n');
        fprintf(fid1, ['PROD01', ' ', 'OPEN', ' ', 'GRAT', ' 2* ']);
        fprintf(fid1, num2str(Para_intended{i}(1,a)));
        fprintf(fid1, [' 2* ', '60']);
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        fprintf(fid1, ['PROD02', ' ', 'OPEN', ' ', 'GRAT', ' 2* ']);
        fprintf(fid1, num2str(Para_intended{i}(2,a)));
        fprintf(fid1, [' 2* ', '60']);
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        fprintf(fid1, ['PROD03', ' ', 'OPEN', ' ', 'GRAT', ' 2* ']);
        fprintf(fid1, num2str(Para_intended{i}(3,a)));
        fprintf(fid1, [' 2* ', '60']);

        fprintf(fid1, '/'); fprintf(fid1, '\n');
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        fprintf(fid1, 'TSTEP');
        fprintf(fid1, '\n');
        fprintf(fid1, [num2str(Durtn_F_training{i}(a)), '*']);
        fprintf(fid1, num2str(30));
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        aaa=aaa+Durtn_F_training{i}(a);
    end
    fclose(fid1);

    % Running Simulations/ Generating training data
    disp(['Training run #', num2str(ii)]);
    tic
    [~,RESULTS] = system('$eclipse NORNE_field');
    Problems = strfind(RESULTS, 'UN-PHYSICAL VALUE');
    TEclipseTrainingData(ii) = toc;
    tic
    % reading simulation output files into NIROM
    out = read_ecl('NORNE_field.UNRST');

    if ~isempty(Problems)
        disp('UN-PHYSICAL VALUE somewhere')
        ii=ii+0;
        continue
    else
        [~,TIME_DAYS1{ii},PPTY1] = readColData('NORNE_field.RSM',10,header-1);
        [~,~,PPTY2] = readColData('NORNE_field.RSM', 10, length(TIME_DAYS1{ii}) + 2
*header-1);
        [~,~,PPTY3] = readColData('NORNE_field.RSM', 6, 2*length(TIME_DAYS1{ii}) + 3
*header+1);
    end
end

```

```

fclose('all');

FGPR = PPTY1(:,2); FGPT = PPTY1(:,3); FWPR = PPTY1(:,4);
FWPT = PPTY1(:,5); FPR = PPTY1(:,6); FWGR = PPTY1(:,8);
WBHP_PROD1 = PPTY1(:,9); WBHP_PROD2 = PPTY2(:,1);
WBHP_PROD3 = PPTY2(:,2); WGPR_PROD1 = PPTY2(:,3);
WGPR_PROD2 = PPTY2(:,4); WGPR_PROD3 = PPTY2(:,5);
WWPR_PROD1 = PPTY2(:,6); WWPR_PROD2 = PPTY2(:,7);
WWPR_PROD3 = PPTY2(:,8); WWGR_PROD1 = PPTY2(:,9);
WWGR_PROD2 = PPTY3(:,1); WWGR_PROD3 = PPTY3(:,2);

% saving Training data into NIROM
Z{ii} = [FGPR FGPT.*10^6 FWPR FWPT FPR FWGR WBHP_PROD1 ...
WBHP_PROD2 WBHP_PROD3 WGPR_PROD1 WGPR_PROD2 WGPR_PROD3 WWPR_PROD1 ...
WWPR_PROD2 WWPR_PROD3 WWGR_PROD1 WWGR_PROD2 WWGR_PROD3];
g = Z{ii}(2:end,:);
ProdVar{ii} = g'; % production variables training data

PRES{ii} = out.PRESSURE; % spatial pressure training data
GSAT{ii} = ones(size(out.SWAT))-out.SWAT; % spatial gas saturation training
data
Para_actual{ii} = (Z{ii}(2:end,10:12))'; % Actual production profile from
WGPRs

ii=ii+1;
end

T_ExtractTrainingData(ii) = toc;
end
TGeneratingTrainingData = toc;
%% Mean Extraction
% Initialising NIROM for mean extraction
nx = length(PRES{1}(:,1)); nt = length(PRES{1}(1,:));
yx = length(ProdVar{1}(:,1)); yt = length(ProdVar{1}(1,:));
pressure_mean = mean(cell2mat(PRES),2);
GSAT_mean = mean(cell2mat(GSAT),2);
ProdVar_mean = mean(cell2mat(ProdVar),2);
rates4Para = Para_actual;
tic

% Mean Extraction
for t = 1:ns
    PRES_{t}(1:nx,:) = PRES{t}(1:nx,:)-pressure_mean;
    GSAT_{t}(1:nx,:) = GSAT{t}(1:nx,:)-GSAT_mean;
    ProdVar_{t}(1:yx,:) = ProdVar{t}(1:yx,:)-ProdVar_mean;
end
Tnirom_meanExctn=toc;
%% POD
disp(' POD Reduction');
% computing singular values and vectors
tic
[V,D] = svd(cell2mat(PRES_), 'econ'); D = diag(D).^2;
[F,J] = svd(cell2mat(GSAT_), 'econ'); J = diag(J).^2;
[W,Q] = svd(cell2mat(ProdVar_), 'econ'); Q = diag(Q).^2;

% computing optimal number of POD basis functions
Hmax = 0.999; % Energy criteria
d = 0; Hn = 0; sD = sum(D);
dd = 0; HHn = 0; sJ = sum(J);
ddd = 0; yHn = 0; sQ = sum(Q);

for j = 1:min(nx,nt*ns) % for spatial pressure
    h = D(j); Hn = Hn+h; H = Hn/sD; d = d + 1;
    if H >= Hmax, break, end
end

for j = 1:min(nx,nt*ns) % for spatial pgsat

```

```

h = J(j); HHn = HHn+h; HH = HHn/sJ; dd = dd + 1;
if HH >= Hmax, break, end

end
%
for j = 1:min(yx, yt*ns) % for production variables
h = Q(j); yHn = yHn+h; yH = yHn/sQ; ddd = ddd + 1;
if yH >= Hmax, break, end

end

P = V(:,1:d); PP = F(:,1:dd); PPP = W(:,1:ddd);

% Training Data Reduction
PRES_Red_ = cell(1,ns); GSAT_Red_ = cell(1,ns); ProdVar_Red_ = cell(1,ns);

for tt = 1:ns
PRES_Red_{tt} = P'*PRES_{tt};
GSAT_Red_{tt} = PP'*GSAT_{tt};
ProdVar_Red_{tt} = PPP'*ProdVar_{tt};
end

Tnirom_POD_Red = toc;
%% Sorting reduced training data into sample and response arrays
disp(' Generating NIROM Response Array');
tic
% initialising NIROM for sample and response array data
a = nt-1; aa = length(ProdVar_{1}(1,:))-1;
b = 1; bb = 1; nPara = length(rates4Para{1}(:,1));
nP = length(P(1,:)); nPP = length(PP(1,:)); nPPP = length(PPP(1,:));

PRES_Samp = zeros(nP+1+nPara, length(cell2mat(PRES_Red_))-ns);
PRES_Resp = zeros(nP, length(cell2mat(PRES_Red_))-ns);

GSAT_Samp1 = zeros(nPP+1+nPara, length(cell2mat(GSAT_Red_))-ns);
GSAT_Resp = zeros(nPP, length(cell2mat(GSAT_Red_))-ns);

ProdVar_Samp = zeros(nPPP+1+nPara, length(cell2mat(ProdVar_Red_))-ns); %a);
ProdVar_Resp = zeros(nPPP, length(cell2mat(ProdVar_Red_))-ns); %a)
Yr = cell(1,ns);

% computing max reduced training data for renormalisation (Equation 21 and 25)
PRES_ScalFac = max(max(abs(cell2mat(PRES_Red_))));
GSAT_ScalFac = max(max(abs(cell2mat(GSAT_Red_))));
ProdVar_ScalFac1 = max(max(abs(cell2mat(ProdVar_Red_))));
maxPara = max(max(cell2mat(rates4Para)));

for t = 1:ns
% creating sample and response array with their corresponding
% paramaters (Equation 26)
PRES_Samp(:,b:b+a-1) = [PRES_Red_{t}(:,1:end-1)/PRES_ScalFac;
rates4Para{t}(:,1:end-1)/maxPara; ones(1,a)];
PRES_Resp(:,b:b+a-1) = PRES_Red_{t}(:,2:end)/PRES_ScalFac;
GSAT_Samp1(:,b:b+a-1) = [GSAT_Red_{t}(:,1:end-1)/GSAT_ScalFac;
rates4Para{t}(:,1:end-1)/maxPara; ones(1,a)];
GSAT_Resp(:,b:b+a-1) = GSAT_Red_{t}(:,2:end)/GSAT_ScalFac;
ProdVar_Samp(:,bb:bb+aa-1) = [ProdVar_Red_{t}(:,1:end-1)/ProdVar_ScalFac1;
rates4Para{t}(:,1:end-1)/maxPara; ones(1,aa)];
ProdVar_Resp(:,bb:bb+aa-1) = ProdVar_Red_{t}(:,2:end)/ProdVar_ScalFac1;

b = b+a;
bb = bb+aa;
end

%% Modelling Dynamics
% computing pairwise distances between training data
A = squareform(pdist(PRES_Samp'));
AA = squareform(pdist(GSAT_Samp1'));

```

```

AAA = squareform(pdist(ProdVar_Samp'));

% Radial Basis Functions (RBFs)
rbf = @(r)r.^1; % Linear RBF
Coef_PRES = rbf(A)\PRES_Resp'; cc = rbf(AA)\GSAT_Resp';
ccc = rbf(AAA)\ProdVar_Resp';

TimeGeneratingWeight = toc;
%% NIROM Prediction Section
%New Initial Condition
% Initialize storage for predictions
ntest = 6; ParaNew = cell(1,ntest); PRES_Exact = cell(1,ntest);
GSAT_Exact = cell(1,ntest); xExY = cell(1,ntest); ProdVar_Exact = cell(1,ntest);
PRES_NIROM = cell(1,ntest); GSAT_NIROM = cell(1,ntest);
ProdVar_NIROM = cell(1,ntest); ProdVarNIROM = cell(1,ntest);
PRES_IC_Red = cell(1,ntest); GSAT_IC_Red = cell(1,ntest);
ProdVar_IC_Red = cell(1,ntest); Tnirom_Interpolation = zeros(1,ntest);
TEclipsePredData = zeros(1,ntest); T_ExtractPredData = zeros(1,ntest);
RE_multiplier = linspace(1,.5,ntest); NT = linspace(1,nt,nt);
TIME_DAYS = cell(1,ntest); ParaNew_actual = cell(1,ntest);
rates4ParaNew = cell(1,ntest);
%% Generating data for Prediction runs production profiles
[Rate_Yr_prediction,Durtn_F_prediction, Durtn_Fsum_prediction] = SeasonalProdProfile
(ntest,prod_prof,e);
ParaNew_intended = Rate_Yr_prediction;

%% Simulation runs for NIROM Prediction
tic
for kk = 1:ntest
    aaa = 1;
    % creating production profile text file
    fid1 = fopen('ProdProfile.inc', 'wt');
    for a=1:length(ParaNew_intended{kk})

        fprintf(fid1, '\n');
        fprintf(fid1, 'WCONPROD');
        fprintf(fid1, '\n');
        fprintf(fid1, ['PROD01', ' ', 'OPEN', ' ', 'GRAT', ' 2* ']);
        fprintf(fid1, num2str(ParaNew_intended{kk}(1,a)));
        fprintf(fid1, [' 2* ', '60']);
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        fprintf(fid1, ['PROD02', ' ', 'OPEN', ' ', 'GRAT', ' 2* ']);
        fprintf(fid1, num2str(ParaNew_intended{kk}(2,a)));
        fprintf(fid1, [' 2* ', '60']);
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        fprintf(fid1, ['PROD03', ' ', 'OPEN', ' ', 'GRAT', ' 2* ']);
        fprintf(fid1, num2str(ParaNew_intended{kk}(3,a)));
        fprintf(fid1, [' 2* ', '60']);

        fprintf(fid1, '/'); fprintf(fid1, '\n');
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        fprintf(fid1, 'TSTEP');
        fprintf(fid1, '\n');
        fprintf(fid1, [num2str(Durtn_F_prediction{kk}(a)), '*']);
        fprintf(fid1, num2str(30));
        fprintf(fid1, '/'); fprintf(fid1, '\n');

        aaa=aaa+Durtn_F_prediction{kk}(a);
    end
    fclose(fid1);

    % Running ECLIPSE for prediction runs
    disp(['Prediction run #',num2str(kk)]);
tic
    [~,RESULTS] = system('$eclipse NORNE_field');

```

```

Problems = strfind(RESULTS, 'UN-PHYSICAL VALUE');
TEclipsePredData(kk) = toc;

% extracting simulation output data
tic
out = read_ecl('NORNE_field.UNRST');
if ~isempty(Problems)
    disp('Un-Physical Value in Simulation')
    %xExact{kk} = [];
    continue
else
    PRES_Exact{kk} = out.PRESSURE;%(:,1:7) ; % New simulation pressure field
    GSAT_Exact{kk} = ones(size(out.SWAT))-out.SWAT;

    [~,TIME_DAYS{kk},PPTY1] = readColData('NORNE_field.RSM',10,header-1);
    [~,~,PPTY2] = readColData('NORNE_field.RSM', 10, length(TIME_DAYS{kk}) + 2
*header-1);
    [~,~,PPTY3] = readColData('NORNE_field.RSM', 6, 2*length(TIME_DAYS{kk}) + 3
*header+1);

    fclose('all');

    FGPR = PPTY1(:,2); FGPT = PPTY1(:,3); FWPR = PPTY1(:,4);
    FWPT = PPTY1(:,5); FPR = PPTY1(:,6); FWGR = PPTY1(:,8);
    WBHP_PROD1 = PPTY1(:,9); WBHP_PROD2 = PPTY2(:,1);
    WBHP_PROD3 = PPTY2(:,2); WGPR_PROD1 = PPTY2(:,3);
    WGPR_PROD2 = PPTY2(:,4); WGPR_PROD3 = PPTY2(:,5);
    WWPR_PROD1 = PPTY2(:,6); WWPR_PROD2 = PPTY2(:,7);
    WWPR_PROD3 = PPTY2(:,8); WWGR_PROD1 = PPTY2(:,9);
    WWGR_PROD2 = PPTY3(:,1); WWGR_PROD3 = PPTY3(:,2);

    xExY{kk} = [FGPR FGPT.*10^6 FWPR FWPT FPR FWGR WBHP_PROD1 ...
    WBHP_PROD2 WBHP_PROD3 WGPR_PROD1 WGPR_PROD2 WGPR_PROD3 WWPR_PROD1 ...
    WWPR_PROD2 WWPR_PROD3 WWGR_PROD1 WWGR_PROD2 WWGR_PROD3];
    g = xExY{kk}(2:end,:);
    ProdVar_Exact{kk} = g';

end

ParaNew_actual{kk} = (xExY{kk}(2:end,10:12))';%9:10));%--%
rates4ParaNew{kk} = ParaNew_actual{kk};
T_ExtractPredData(kk) = toc;

% Initializing NIROM simulation from Initial condition data
tic
PRES_IC_Red{kk}(:,1) = [(P'*(PRES_Exact{kk}(:,1)-pressure_mean))/PRES_ScalFac;
rates4ParaNew{kk}(:,1)/maxPara; ones(1,1)];
GSAT_IC_Red{kk}(:,1) = [(PP'*(GSAT_Exact{kk}(:,1)-GSAT_mean))/GSAT_ScalFac;
rates4ParaNew{kk}(:,1)/maxPara; ones(1,1)];
ProdVar_IC_Red{kk}(:,1) = [(PPP'*(ProdVar_Exact{kk}(:,1)-
ProdVar_mean))/ProdVar_ScalFac; rates4ParaNew{kk}(:,1)/maxPara; ones(1,1)];

% Propagating New IC for NIROM nt-1 report steps - PRES_Exact and GSAT_Exact
disp(['Propagating New IC #',num2str(kk)]);
for n = 2:length(PRES_Exact{kk}(1,:))%nt
    r = rbf(sqrt(sum((PRES_Samp-[PRES_IC_Red{kk}(1:nP,n-1);
rates4ParaNew{kk}(:,n-1)/maxPara; ones(1,1)])^2,1)))';
    PRES_IC_Red{kk}(1:nP,n) = Coef_PRES'*r;
    rr = rbf(sqrt(sum((GSAT_Samp1-[GSAT_IC_Red{kk}(1:nPP,n-1);
rates4ParaNew{kk}(:,n-1)/maxPara; ones(1,1)])^2,1)))';
    GSAT_IC_Red{kk}(1:nPP,n) = cc'*rr;
end
% Propagating New IC for NIROM nt-1 report steps - ProdVar_Exact
for nn = 2:length(TIME_DAYS{1})-1%
    rrr = rbf(sqrt(sum((ProdVar_Samp-[ProdVar_IC_Red{kk}(1:nPPP,nn-1);
rates4ParaNew{kk}(:,nn-1)/maxPara; ones(1,1)])^2,1)))';
    ProdVar_IC_Red{kk}(1:nPPP,nn) = ccc'*rrr;
end

```

```

% Un-normalising and Back projection of NIROM predictions
PRES_NIROM{kk} = (P*(PRES_IC_Red{kk}(1:nP,:)*PRES_ScalFac))+pressure_mean;
GSAT_NIROM{kk} = (PP*(GSAT_IC_Red{kk}(1:nPP,:)*GSAT_ScalFac))+GSAT_mean;
ProdVarNIROM{kk} = (PPP * (ProdVar_IC_Red{kk}(1:nPPP,:) * ProdVar_ScalFac1)) +
ProdVar_mean;
ProdVar_NIROM{kk} = [xExY{kk}(1,:); ProdVarNIROM{kk}'];
ProdVar_Exact{kk} = [xExY{kk}(1,:); ProdVar_Exact{kk}'];
Tnirom_Interpolation(kk)=toc;
end
TNIROMPred = toc;
%% Comparing and Plotting Results ...

%% Error Analysis ...

%% Time/durations ...

%% Function to create seasonal production profiles
function [Rate_Yr, Durtn_F, Durtn_Fsum,len_Durtn_F] = SeasonalProdProfile(ns,
prod_prof,e)
    ii=1;
    durtn = cell(1,4); %4 seasons starting from november: autumn, winter, spring,
summer
    durtn_F = cell(1,4); Durtn_F = cell(1,ns); Durtn_Fsum = zeros(1,ns);
    Durtn = cell(1,4); % stores the varying rates within seasons
    rate_yr = cell(1,ns); rate_Yr = cell(1,ns); Rate_Yr = cell(1,ns);
    RE_multiplier = linspace(1,.5,ns);

    for i=1:ns
        disp(['ProdProfile #',num2str(ii)]);

        aa = 1;
        prod_prof_ = prod_prof;
        for jj= 1:20
            prod_prof_{e} = prod_prof_{e}.*aa;
            for j=1:4
                durtn{j} = randi([1 3],1,3);
                Durtn{j} = cumsum(durtn{j})';

                k=find(Durtn{j}>=3,1);
                if Durtn{j}(k)>3
                    durtn{j} = [durtn{j}(1:k-1) 3-sum(Durtn{j}(k-1))];
                elseif Durtn{j}(k)==3
                    durtn{j} = durtn{j}(1:k);
                elseif isempty(k)
                    durtn{j} = [durtn{j} 3-sum(durtn{j})];
                end

                for b=1:length(durtn{j})

% with shut-in
                    if e==1 %this affects if it's going to be BHP control (e=2) or
rate control (e=1)
                        if j == 1 % differentiates between seasons
                            if b <= length(durtn{j})
                                z = prod_prof_{e}(randi(round(length(prod_prof_{e}
(1,:)).*25))));%
                                else
                                    z = prod_prof_{e}(b-1);
                                end
                                rate1 = (z-randi(1e5))*RE_multiplier(ii);
                                rate2 = (2.5*z-randi(1e5))*RE_multiplier(ii);
                                rate3 = (.7*z-randi(1e5))*RE_multiplier(ii);
                                if rate3 < 0
                                    rate3 = .7*z ;
                                end
                                elseif j == 3

```

```

        if b <= length(durtn{j})
            z = prod_prof_{e}(randi([round(length(prod_prof_{e}
(1,:))*0.75) round(length(prod_prof_{e}(1,:)))]));%
        else
            z = prod_prof_{e}(b-1);
        end

        rate1 = z-randi(1e5);
        rate2 = 2.5*z-randi(1e5);
        rate3 = .7*z-randi(1e5);
        if rate3 < 0
            rate3 = .7*z ;
        end
    else
        if b <= length(durtn{j})
            z = prod_prof_{e}(randi([round(length(prod_prof_{e}
(1,:))*0.25)+1 round(length(prod_prof_{e}(1,:))*0.75)]));%
        else
            z = prod_prof_{e}(b-1);
        end

        rate1 = z-randi(1e5);
        rate2 = 2.5*z-randi(1e5);
        rate3 = .7*z-randi(1e5);
        if rate3 < 0
            rate3 = .7*z ;
        end
    end
end
else
end
%}
rate_yr{j} = ones(3,b).*[rate1; rate2; rate3];
end
end
rate_Yr{jj} = cell2mat(rate_yr);
durtn_F{jj} = cell2mat(durtn);
aa = aa-.003;
end
Durtn_F{ii} = cell2mat(durtn_F);
Rate_Yr{ii} = cell2mat(rate_Yr);

len_Durtn_F = length(Durtn_F{ii});
Durtn_Fsum(ii) = sum(Durtn_F{ii});
ii=ii+1;
end
end
end

```

C.2.2 AUTOENOCDER (AE) MODEL WITH 2HLS

An example of the autoencoder model that was used in this work is given below.

```

# In[ ]:
from numpy.random import seed
seed(1)
import tensorflow
tensorflow.random.set_random_seed(2)

# In[ ]:
from scipy.io import loadmat, savemat
import numpy as np

# load training data
PRES_Samp = loadmat('PRES_Samp.mat') # training data sample array
PRES_Samp.keys()
PRES_Samp = ( PRES_Samp['PRES_Samp'])

```

```

PRES_Resp = loadmat('PRES_Resp.mat') # training data response array
PRES_Resp.keys()
PRES_Resp = (PRES_Resp ['PRES_Resp'])

Test_Samp = loadmat('Test_Samp.mat') # validation data sample array
Test_Samp.keys()
Test_Samp = (SxT['Test_Samp'])

Test_Resp = loadmat('Test_Resp.mat') # validation data response array
Test_Resp.keys()
Test_Resp = (Test_Resp['Test_Resp'])

# In[ ]:
from keras.layers import Dense, Input
from keras.models import Model
from keras import optimizers

### PRE-TRAINING
# initialise for AE
n = len(Sx[0]) # spatial dimension
nt_train = len(Sx[:,0]) # number of report steps
nt_steps_out = len(SxT[:,0])
activation = 'elu'

# In[ ]:
#Assemble the AE model
#Encoder section compressing from n to 10
input_img = Input(shape=(n,))
encoded = Dense(100, activation=activation)(input_img)
encoded = Dense(10, activation=activation)(encoded)
#Decoder section expanding from 10 to n
decoded = Dense(100, activation=activation)(encoded)
decoded = Dense(n, activation=activation)(decoded)
#Full Model assembled
autoencoder=Model(input_img,decoded)

# In[ ]:
autoencoder.summary()

# In[ ]:
from keras import metrics
from keras.optimizers import Adam

###TRAINING
#Compile the model, choosing an optimizer and loss function
Adam_mee = Adam(lr=0.001)
autoencoder.compile(optimizer=Adam_mee,loss='mean_squared_error',metrics=[metrics.MeanAbsoluteError()])
history = autoencoder.fit(PRES_Samp,PRES_Resp,epochs= 1000,batch_size = 200,shuffle
= True,validation_data=(Test_Samp,Test_Resp))

# In[ ]:
#RECORD KEEPING (losses & metrics)
loss_history = history.history["loss"]
metrics_history = history.history["mean_absolute_error"]
val_loss_history = history.history["val_loss"]
val_metrics_history = history.history["val_mean_absolute_error"]

# to save history losses as a numpy array
numpy_loss_history = np.array(loss_history)
np.savetxt("loss_history2.txt", numpy_loss_history, delimiter=",")
numpy_val_loss_history = np.array(val_loss_history)
np.savetxt("val_loss_history2.txt",numpy_val_loss_history,delimiter=",")

# to save history metrics as a numpy array
numpy_metrics_history = np.array(metrics_history)
np.savetxt("metrics_history2.txt",numpy_metrics_history,delimiter=",")
numpy_val_metrics_history = np.array(val_metrics_history)

```

```

np.savetxt("val_metrics_history2.txt",numpy_val_metrics_history, delimiter=",")

# In[ ]:
#Save full AE model and the weights
autoencoder.save('autoencoder2.h5')
autoencoder.save_weights('autoencoder2wt.h5')

# In[ ]:
# Loading Initial Condition Prediction data
PRES_Exact1 = loadmat('PRES_Exact1.mat')
PRES_Exact1.keys()
PRES_Exact1=(PRES_Exact1['PRES_Exact1'])

PRES_Exact2 = loadmat('PRES_Exact2.mat')
PRES_Exact2.keys()
PRES_Exact 2=(PRES_Exact2['PRES_Exact2'])

PRES_Exact3 = loadmat('PRES_Exact3.mat')
PRES_Exact3.keys()
PRES_Exact3=(PRES_Exact3['PRES_Exact3'])

PRES_Exact4 = loadmat('PRES_Exact4.mat')
PRES_Exact4.keys()
PRES_Exact4=(PRES_Exact4['PRES_Exact4'])

PRES_Exact5 = loadmat('PRES_Exact5.mat')
PRES_Exact5.keys()
PRES_Exact5 =(PRES_Exact5['PRES_Exact5'])

# In[ ]:
# Loading prediction parameters
Para1 = loadmat('Para1.mat')
Para1.keys()
Para1=(Para1['Para1'])

Para2 = loadmat('Para2.mat')
Para2.keys()
Para2=(Para2['Para2'])

Para3 = loadmat('Para3.mat')
Para3.keys()
Para3=(Para3['Para3'])

Para4 = loadmat('Para4.mat')
Para4.keys()
Para4=(Para4['Para4'])

Para5 = loadmat('Para5.mat')
Para5.keys()
Para5=(Para5['Para5'])

# In[ ]:
# AE Predictions
para = len(Para1[:,0])
predictedPRES = []
predictedTest = []
predicted1 = []
predicted2 = []
predicted3 = []
predicted4 = []
predicted5 = []

predicted.append(PRES_Samp[0,:])
predictedT.append(Test_Samp[0,:])
predicted1.append(PRES_Exact1[0,:])
predicted2.append(PRES_Exact2[0,:])
predicted3.append(PRES_Exact3[0,:])
predicted4.append(PRES_Exact4[0,:])

```

```

predicted5.append(PRES_Exact5[0,:])

training = PRES_Samp[0,:].reshape(1,n)
test_run = U_test_in[0,:].reshape(1,n)

for i in range(nt_train-1):
    predtd = autoencoder.predict(training)
    predicted.append(predtd[0,:])
    training = predtd

for i in range(nt_steps_out):
    predtdT = autoencoder.predict(test_run)
    predictedT.append(predtdT[0,:])
    test_run = predtd

for i in range(nt_steps_out):
    predtd1 = autoencoder.predict(np.concatenate((PRES_Exact1[:, :-para], (Para1[:, i])
.reshape(1, para)), axis=1))
    predicted1.append(predtd1[0,:])
    PRES_Exact1 = predtd1

    predtd2 = autoencoder.predict(np.concatenate((PRES_Exact2[:, :-para], (Para2[:, i])
.reshape(1, para)), axis=1))
    predicted2.append(predtd2[0,:])
    uNew2 = predtd2

    predtd3 = autoencoder.predict(np.concatenate((PRES_Exact3[:, :-para], (Para3[:, i])
.reshape(1, para)), axis=1))
    predicted3.append(predtd3[0,:])
    PRES_Exact3 = predtd3

    predtd4 = autoencoder.predict(np.concatenate((PRES_Exact4[:, :-para], (Para4[:, i])
.reshape(1, para)), axis=1))
    predicted4.append(predtd4[0,:])
    PRES_Exact4 = predtd4

    predtd5 = autoencoder.predict(np.concatenate((PRES_Exact5[:, :-para], (Para5[:, i])
.reshape(1, para)), axis=1))
    predicted5.append(predtd5[0,:])
    PRES_Exact5 = predtd5

predicted=np.array(predicted)
predictedT=np.array(predictedT)
predicted1=np.array(predicted1)
predicted2=np.array(predicted2)
predicted3=np.array(predicted3)
predicted4=np.array(predicted4)
predicted5=np.array(predicted5)

predicted = {'HL2_predicted':predicted.transpose()}
savemat('HL2_predicted.mat',predicted)
predictedT = {'HL2_predictedT':predictedT.transpose()}
savemat('HL2_predictedT.mat',predictedT)
predicted1 = {'HL2_predicted1':predicted1.transpose()}
savemat('HL2_predicted1.mat',predicted1)
predicted2 = {'HL2_predicted2':predicted2.transpose()}
savemat('HL2_predicted2.mat',predicted2)
predicted3 = {'HL2_predicted3':predicted3.transpose()}
savemat('HL2_predicted3.mat',predicted3)
predicted4 = {'HL2_predicted4':predicted4.transpose()}
savemat('HL2_predicted4.mat',predicted4)
predicted5 = {'HL2_predicted5':predicted5.transpose()}
savemat('HL2_predicted5.mat',predicted5)

```