Imperial College London

Department of Earth Science and Engineering

Simulation of Geothermal Reservoirs with Data Assimilation and Reduced Order Modelling

Zainab D. Titus

Submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy and the Diploma of Imperial College London
May 2024

Abstract

Generating realistic reservoir models that match observed well data is important for understanding and predicting subsurface behaviour. By combining static and dynamic data from wells with prior geological knowledge and fluid flow models, geoscientists and reservoir engineers can generate representative subsurface reservoir models.

Data assimilation integrates information from models and observations to generate improved estimates of systems for prediction and uncertainty analysis. In subsurface modelling, data assimilation has been applied for history matching reservoir models by assimilating spatio-temporal data such as downhole measurements of pressure, temperature, and flow rate. Integrating static and dynamic data from wells in a data assimilation workflow for history matching can aid the development of reliable reservoir models for forecasting reservoir performance and maximising fluid recovery. However, for large-scale, highly nonlinear dynamic systems, such as the geothermal reservoir models considered in this research, conventional data assimilation techniques can be computationally expensive, as the number of parameters to be estimated is orders of magnitude higher than the observed measurements.

This research develops models of geothermal reservoirs by applying artificial neural networks for data assimilation and predictive modelling. Using deep neural networks for well conditioning, history matching, and reduced order modelling, representative models of geothermal reservoirs that honour observed well data are generated and can be used to forecast reservoir behaviour at a fraction of the time required for direct numerical simulations. The approach presented here has many advantages over conventional methods: 1) the use of grid-free surface-based geological models (SBGMs) that preserve geological realism and model resolution, 2) neural networks for data assimilation and well conditioning to accelerate the history matching process, 3) space-filling curves for representing geothermal reservoir simulation data on unstructured meshes to aid data compression, and 4) transformer-based models for temporal evolution and replacing the geothermal reservoir simulator for rapid reservoir forecasting.

Copyright Declaration

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

Declaration

I hereby certify that the work presented in this thesis is the result of research during my Ph.D. programme and that the material in this thesis, which is not my own work, has been duely acknowledged.

Zainab D. Titus

Acknowledgements

Pursuing a Ph.D. in Earth Science and Engineering at Imperial College London has been an enjoyable, challenging, and invaluably rewarding experience. I have been enriched in every facet of life, and I have many to thank for this.

I am grateful to the Petroleum Technology Development Fund (PTDF) of Nigeria for the financial support and academic freedom to conduct research in a field that deeply interests me. Being a recipient of this scholarship is a distinct honour and I am grateful for this opportunity of a lifetime.

My immense gratitude goes to my supervisors and co-supervisors. I greatly appreciate my primary supervisor, Prof. Christopher Pain, for giving me the opportunity to work on this research, providing valuable guidance and feedback, and challenging me to grow as a researcher. I am equally grateful to Dr. Claire Heaney for her tutelage, counsel, and incredible support throughout my Ph.D., positively influencing my professional and personal growth. I would like to thank Prof. Matthew Jackson for his support towards my progress and guidance through his wealth of knowledge in subsurface modelling and development that has shaped my approach and contributions to this field. I'm grateful to Dr. Pablo Salinas for his technical assistance and willingness to guide me in working with the required tools, concepts, and resources for my research, particularly during the initial stages of my Ph.D.

I express my sincere gratitude to Imperial College London and specifically to the Department of Earth Science and Engineering (ESE) for an enriching Ph.D. programme. I have had the opportunity to participate in a wide range of projects and events that have made my doctoral programme worthwhile. Special thanks to Dr. Rebecca Bell, Sophia Quazi, and Victoria Murphy for supporting me in developing and implementing projects that have impacted young girls in my home community. My heartfelt appreciation to the Head of Department, Prof. Tina van De Flierdt, whose leadership and dedication to impactful research and community engagement have been a source of inspiration for my own efforts. Furthermore, I'm thankful to all the academic staff with whom I have worked as a graduate teaching assistant. I have learnt as much as I have taught others through the process. I'm grateful to the Director of Postgraduate Studies, Prof. Saskia Goes and the Postgraduate Administrator, Amanda Allotey, for their support throughout my programme and especially during the later stages of my Ph.D. as I navigated the challenges of being a new parent. I also appreciate my exam committee members, Prof. Ann Muggeridge and Prof. Ahmed Elsheikh, for their time, expertise, and useful feedback that will impact my future contributions to the field of subsurface modelling and machine learning.

It has been intellectually stimulating working within the NOvel Reservoir Modelling and Simulation (NORMS) group. Collaborating with innovative minds and creating memorable experiences together has been a cherished part of my Ph.D. journey. Thanks to Vinicius Santos Silva, Geraldine Regnier, Toby Phillips, and especially Jumanah Al Kubaisy for the fun times, support, and insightful discussions on reservoir simulation, machine learning, and life as a Ph.D. student. I'm grateful to colleagues and friends in ESE, Gabriela Salomao Martins, Deborah Pelacani-Cruz, Kenechukwu Nwegbu, Harriet Dawson, Idris Bukar, and Saira Baharuddin for their friendship and support.

My profound gratitude goes to all who stood by me during these years, offering motivation, encouragement, prayers, and assistance in every way possible especially as I faced the challenges presented by the Covid-19 pandemic. I'm indebted to my sisters and brothers, whose sacrifices have been instrumental to the successful completion of my studies. Despite being away from home for so long, you never stopped looking out for me. Thank you Afiniki for always being there to listen and encourage me at times when it was hard to find the strength to continue. To my parents, I'm eternally grateful for your love and sacrifices. Thank you Mummy for the timely words of advice and encouragement, and thank you for travelling all the way to take care of my family when we needed help and support as I worked to finish-up my programme. Daddy, I'm grateful you are alive to witness this, and I'm glad this achievement is as much a source of pride for you as it is for me.

To my husband, Emmanuel, thank you for believing in what I can achieve and for growing with me since I started this journey. Your presence and support every step of the way are more than words can express. Finally, to my daughter, Zikora, I'm grateful you came into my life during this journey. You have been a wellspring of joy and a huge source of inspiration, providing me with the powerful motivation needed to persevere and complete this chapter successfully.

Thank you, Jesus, for walking this path with me. I look forward to what you have prepared next for me.

Dedication

To the One who is my glory, my shield, and the lifter of my head.

Either write things worth reading or do things worth writing.

- Benjamin Franklin

Contents

1 Introduction **1**

1.1 Motivation and Objectives 1

1.2 Contributions and Significance of the Research 3

1.3 Statement of Originality 3

1.4 Thesis Outline 4

2 Literature Review **5**

2.1 Geothermal Systems 5

2.1.1 Opportunities for Geothermal Energy Production in Sedimentary Basins 7

2.1.2 Geothermal Reservoir Simulation 10

2.2 Data Assimilation 11

2.3 Data Assimilation for Reservoir History Matching 12

2.3.1 Ensemble-Based Methods for History Matching: Applications, Improvements and Limitations 14

2.3.2 Iterative Ensemble Kalman Smoother: An Efficient History Matching Technique 16

2.4 Generating Reservoir Models with Neural Networks 18

2.4.1 Geological Modelling 18

2.4.2 History Matching 20

2.5 Reduced Order Modelling for Rapid Reservoir Simulation 22

2.5.1 Data-Driven and Deep Learning Low Order Models for Subsurface Reservoirs 23
2.5.2 Simulation of Geothermal Reservoirs with Reduced Order Modelling 24
2.6 Research Questions 25

3 Conditioning Surface-Based Geological Models (SBGMs) to Well Facies Data using Artificial Neural Networks 27

3.1 Abstract 27
3.2 Introduction 28
3.3 Surface-Based Geological Model Description 31
3.3.1 NURBS for Generating Computationally Efficient SBGMs 31
3.3.2 Shallow Marine Example Case 32
3.4 Modelling SBGMs with ANNs 34
3.4.1 Forward Modelling 35
3.5 Conditioning SBGMs with ANNs 38
3.5.1 Inverse Modelling 38
3.5.2 Optimisation Methodology 38
3.6 Results and Discussions 40
3.6.1 Forward Modelling: Well Log Prediction 40
3.6.2 Inverse Modelling: Generating the SBGM Inputs 43
3.6.3 Uncertainty Assessment: Evaluating Spatial Variations 50
3.7 Conclusions 57

4 A Grid-Free Approach to Conditioning Surface-Based Geological Models to Well Data 59

4.1 Abstract 59
4.2 Introduction 60
4.3 Neural Networks for Surface-Based Modelling 61
4.3.1 Approximating SBGMs with Artificial Neural Networks 62

4.3.2 Conditioning SBGMs with Neural Networks . . . 67
4.4 Results and Discussions . . . 69
4.4.1 Predicting Well Data Using Neural Networks . . . 71
4.4.2 Conditioning SBGMs with Neural Networks . . . 72
4.4.3 Considerations for Application to Field Cases . . . 78
4.5 Conclusions . . . 83

5 Data Assimilation with Machine Learning for History Matching 85
5.1 Abstract . . . 85
5.2 Introduction . . . 86
5.2.1 Data Assimilation for History Matching Subsurface Reservoir Models . . . 86
5.2.2 Machine Learning in Subsurface Reservoirs for History Matching . . . 87
5.3 Methodology . . . 90
5.3.1 Conditioning SBGMs to Well Data . . . 90
5.3.2 Fluid Flow Modelling . . . 92
5.3.3 Data Assimilation With Machine Learning . . . 94
5.3.4 Uncertainty Analysis . . . 103
5.4 Results and Discussions . . . 103
5.4.1 Forward Modelling: Predicting the Production Data . . . 103
5.4.2 Inverse modelling: Optimising the Geological Model Inputs . . . 104
5.4.3 Uncertainty Assessments: Evaluating Spatial Variations and Well Correlations . . . 108
5.4.4 Discussions . . . 112
5.5 Conclusions . . . 115

6 Transformer-Based Reduced Order Models of Geothermal Reservoirs with Unstructured Meshes 116
6.1 Abstract . . . 116

6.2 Introduction . . . 117
6.2.1 Deep Learning and Low-Order Models . . . 118
6.3 Methodology . . . 121
6.3.1 Data Generation - Geothermal Reservoir Simulation . . . 121
6.3.2 Data Transformation - Space-Filling Curves . . . 123
6.3.3 Data Compression - Convolutional Autoencoders . . . 125
6.3.4 Temporal Evolution of the Latent State - Transformer-Based Model . . . 126
6.4 Results and Discussions . . . 130
6.4.1 Data Compression . . . 130
6.4.2 State Prediction . . . 131
6.4.3 Discussions and Future Work . . . 134
6.5 Conclusions . . . 137

7 A Proposed Workflow for the Simulation of Geothermal Reservoirs with Data Assimilation and Reduced Order Modelling 139

7.1 Abstract . . . 139
7.2 Introduction . . . 140
7.3 Methods . . . 143
7.3.1 Geological Reservoir Modelling . . . 144
7.3.2 Reservoir Simulation . . . 145
7.3.3 History Matching . . . 147
7.3.4 Reduced Order Modelling . . . 150
7.4 Application to Geothermal Reservoirs . . . 155
7.4.1 History Matching with Well Conditioning and Data Assimilation . . . 155
7.4.2 Reservoir Forecasting with the History-Matched Models . . . 159
7.5 Discussions . . . 162

7.6 Conclusions . . . 165

8 Conclusions and Future Work **166**

8.1 Conclusions . . . 166

8.2 Future Work . . . 168

8.2.1 Conditioning SBGMs to Stratigraphic Surfaces . . . 168

8.2.2 Physics-Informed Neural Networks (PINNs) . . . 169

8.2.3 Data Assimilation in the Latent Space . . . 170

8.2.4 Application to Real Field Cases . . . 170

Bibliography **172**

List of Figures

2.1 Geothermal play systems dominated by convection 7

2.2 Geothermal play systems dominated by conduction 7

2.3 The smoothing and filtering problems in data assimilation 17

3.1 The approach for conditioning surface-based geological models applied in this study. 31

3.2 Clinoform and clinothem templates for the SBGMs. 32

3.3 A surface-based geological model and corresponding well logs at three well locations. 33

3.4 The output of the SBGM for training the network is the facies type typically derived from well logs. 37

3.5 The neural network maps a vector representation of the 240 variables onto a discrete vector with n_y by n entries. 38

3.6 The workflow for conditioning the SBGMs using the trained neural network. 40

3.7 The learning curve for SurfNet2 with 100 epochs. 42

3.8 The loss and accuracy plots for SurfNet2 (left) and SurfNet3 (right). 42

3.9 A comparison of the well logs from the SBGM with the neural network prediction. 44

3.10 A comparison of the well logs from the SBGM with the SurfNet3 prediction. 45

3.11 The initialised, optimised and reference SBGMs results for an example using SurfNet2 - Example 1 46

3.12 The initialised, optimised and reference SBGMs results for an example using SurfNet2 - Example 2 47

3.13 The initialised, optimised and reference SBGMs results for an example using SurfNet3 - Example 1 48

3.14 The initialised, optimised and reference SBGMs results for an example using SurfNet3 - Example 2 49

3.15 The loss plot for the inverse modelling for (a) Example 1, and (b) Example 2 using SurfNet2 and SurfNet3. 50

3.16 The well data at different optimisation steps from the initial guess to the final model obtained after 10000 iterations for SurfNet2. 51

3.17 The well data at different optimisation steps from the initial guess to the final model obtained after 10000 iterations for SurfNet3. 52

3.18 The frequency of predictions plot for the initialised SBGMs. The uncertainty varies in equal proportions laterally in the model including at the well locations. 53

3.19 The frequency of predictions plot for SurfNet2 and SurfNet3 depicting the difference between the maximum frequency, and the minimum frequency at each discretised location in the domain. . . 55

3.20 The facies type with the maximum frequency or number of occurrences is shown for 100 plausible realisations that have been conditioned to the well data using (a) SurfNet2 and (b) SurfNet3. . . 56

3.21 The facies type with the maximum frequency or number of occurrences is shown for 100 models used as initial inputs to the inverse modelling. 57

4.1 The workflow for conditioning SBGMs to well data. 62

4.2 The randomly sampled points from the SBGMs have a uniform distribution of the coordinates. . 63

4.3 The output of the SBGM for training the network is the facies type typically derived from well logs. 65

4.4 The architecture of the multilayer perceptron showing the input, hidden, and output layers along with the activation functions used for each layer. 67

4.5 The workflow for conditioning SBGMs to data from multiple wells having an arbitrary number of data points. 70

4.6 The loss and accuracy plots for the network presented in Figure 4.4 and hyperparameters given in Table 4.2. 71

4.7 Comparisons of ground truth SBGMs with the neural network predictions. 73

4.8 The reference SBGM showing sampled vertical wells (top) and a comparison of the predicted neural network facies type (SurfNet) and the actual facies type (GroundTruth). 74

4.9 The reference SBGM showing sampled deviated and multilateral wells (top) and a comparison of the predicted neural network facies type (SurfNet) and the actual facies type (GroundTruth). . . 75

4.10 The reference surface-based geological model used for validating the predicted and optimised SBGMs. 76

4.11 The correlation of each of 100 initialised models and the optimised models to the reference data at well locations, sorted in descending order of the optimised model correlations. 77

4.12 Frequency of predictions for the initialised and optimised models - vertical wells example 79

4.13 Frequency of predictions for the initialised and optimised models - deviated wells example 80

4.14 Maximum predictions for the initialised (top) and optimised (bottom) models with five vertical wells showing good match with reference model at well locations. 81

4.15 Maximum predictions for the initialised (top) and optimised (bottom) models with one deviated well and one multilateral well with two branches deviated wells showing good match with reference model at well locations. 82

5.1 The data assimilation framework for history matching geologically constrained models proposed in this work. 89

5.2 Examples of 2D shallow marine surface-based geological models with one parasequence and 60 clinoforms generated using the method developed by Jacquemyn et al. (2019) 91

5.3 The reference surface-based geological model used for history matching with three wells located at x=100m, 500m and 900m. 92

5.4 The neural network, SurfNet, $\mathcal{S}$ used in conditioning SBGMs to well data. 93

5.5 The workflow for generating the dataset used for training and testing GeoNet, $\mathcal{G}$. 96

5.6 The architecture of the network showing GeoNet, the Discriminator and the adversarial layer. . . 98

5.7 The framework for generating SBGM inputs that generate geologically plausible models that match static well data and are history-matched to production data. 101

5.8 The loss plot of the adversarial network, GeoNet, as well as the discriminator, $\mathcal{D}$. 105

5.9 Sample predictions of the geothermal well simulations against the ground truth for unseen samples by the network. 106

5.10 Plots showing the mismatch between the predicted and observed data using the trained networks $\mathcal{S}$ and $\mathcal{G}$ for the well conditioning and data assimilation respectively - Example 1 107

5.11 Plots showing the mismatch between the predicted and observed data using the trained networks $\mathcal{S}$ and $\mathcal{G}$ for the well conditioning and data assimilation respectively - Example 2 108

5.12 Uncertainties from the predictions with the optimised inputs. 109

5.13 Temperature data at well1 (points 1 to 60), well2 (points 61 to 120) and well3 (points 121 to 180) from the reservoir simulation with the initialised, reference and optimised models. 110

5.14 The mean absolute temperature difference of the optimised and initialised forward models with the reference forward models showing that the optimised models have a better match to the reference model in comparison to the initialised models. 111

5.15 Correlation of the facies types at the wells with the reference well facies type for the initialised and optimised. 111

5.16 Frequency of Predictions (FOPs) from the predictions with the optimised inputs. 113

5.17 Maximum Predictions (MaxPreds) of the geological models generated with the optimised inputs. 114

5.18 Geology distribution plot showing the number of models with same facies type as the reference model at each grid point before (left) and after (right) conditioning and data assimilation. 115

6.1 The proposed method for generating Reduced order models using space-filling curves with convolutional autoencoder (SFC-CAE) and a transformer-based predictive network. 119

6.2 The process of creating the dataset for training and testing the compression and prediction neural networks. 122

6.3 The space-filling curve convolutional autoencoder (SFC-CAE) applied for compressing the high-dimensional geothermal reservoir simulation snapshots. 126

6.4 The Generative pre-trained transformer (GPT) applied for temporal evolution. 127

6.5 The loss curve showing the mean squared error over 400 training epochs of the SFC-CAE 132

6.6 The results of the SFC-CAE for two examples not used in training the model at five different time steps. 133

6.7 The loss curve showing the mean squared error over 25 training epochs of the transformer-based model. 134

6.8 The predictions from the GPT model for two examples not used in training the model at five different time steps. 135

6.9 The root mean squared error from the reconstruction and compression step for each time step for the two examples presented in the compression and state prediction. 136

7.1 The proposed method for rapid generation of history-matched geothermal reservoir models using neural networks for well conditioning, data assimilation and reduced-order modelling. 143

7.2 An example of a 2D surface-based geological model. This is the reference surface-based geological model used for history matching with three wells located at x=100m, 500m and 900m. 144

7.3 The neural network, SurfNet, $\mathcal{S}$ used in conditioning SBGMs to well data. 147

7.4 The process of generating the dataset for training and testing the neural network for the data assimilation. 148

7.5 The architecture of the network showing GeoNet, the Discriminator and the adversarial layer. . . 149

7.6 The process of generating the data for training and testing the compression and prediction neural networks. 152

7.7 The space-filling curve convolutional autoencoder (SFC-CAE) applied for compressing the high-dimensional geothermal reservoir simulation snapshots 153

7.8 The neural network applied in this study based on the Generative Pre-trained Transformer (GPT) architecture. 154

7.9 The loss plot of the adversarial network, GeoNet, $\mathcal{G}$ as well as the discriminator, $\mathcal{D}$. 156

7.10 Neural network predictions with the initialised, optimised and reference SBGM inputs for points in Well1 as a measure of uncertainties among models. 157

7.11 Temperature data at well1 (points 1 to 60), well2 (points 61 to 120) and well3 (points 121 to 180) from the reservoir simulation with the initialised, reference and optimised models. 158

7.12 Frequency of predictions for the initialised (top) and optimised (bottom) models for 100 conditioned models. 160

7.13 Maximum predictions for the 100 initialised (top) and optimised (bottom) models. 161

7.14 Correlation of the facies types at the wells with the reference well facies type for the initialised and optimised models. 162

7.15 The loss curves of the SFC-CAE (top) and GPT model (bottom) showing the mean squared error over 400 and 380 training epochs respectively. 163

7.16 The results of the SFC-CAE for one example not used in training the model at five different time steps. 164

7.17 The mean squared error for each time step in the reconstructed simulation for the history-matched (optimised) and initialised model. 164

List of Tables

3.1 Range of the clinoforms and facies zones for the parasequences. The parameters all sampled from a uniform distribution 34

3.2 Hyperparameters used for the training and the performance metrics of the model evaluated for the training set 43

3.3 Hyperparameters for optimising the SBGM inputs using the trained neural networks 44

4.1 Hyperparameters used for training the network 66

4.2 Performance metrics of the model evaluated for the training and test dataset 71

4.3 Parameters for optimizing the SBGM inputs using the trained neural networks 72

5.1 Reservoir properties and description of the fluid flow models 95

5.2 Architecture of the adversarial network applied for mapping the geological inputs, $\mathbf{p}$ to production data, $\mathbf{y}$. 99

5.3 Hyperparameters used for training GeoNet and the performance metrics of the model 104

5.4 Parameters for optimizing the SBGM inputs using the trained neural networks 107

6.1 Hyperparameters and performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$). . 130

6.2 Reservoir properties and description of the fluid flow models 136

6.3 Performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$). 137

7.1 Reservoir properties and description of the fluid flow models 146

7.2 Hyperparameters used for training GeoNet and the performance metrics of the model 150

7.3 Parameters for optimizing the SBGM inputs using the trained neural networks 150

7.4 Hyperparameters and performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$). . 155

7.5 Performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$). 159

Chapter 1

Introduction

1.1 Motivation and Objectives

Geothermal energy is a renewable source of baseload energy stored as heat in a variety of geological settings in the Earth's subsurface. Although this sustainable and environmentally friendly resource has been developed in various locations across the globe, geothermal energy has great potential to meet the growing global demand for energy as part of a diverse energy supply mix (Goldstein et al., 2011; Vargas et al., 2022).

Conventional geothermal reservoirs typically occur in igneous or metamorphic formations in volcanic settings. However, deep sedimentary basins in geological settings where reservoir rocks have sufficient temperature, thickness, porosity, and permeability, can extend the generation of geothermal energy to previously untapped areas (Anderson, 2013; Vivas et al., 2023). This suggests that sedimentary formations, including some containing oil or gas, are potential commercial geothermal reservoirs and present new opportunities for the generation of geothermal energy from reservoirs in sedimentary basins.

In these areas, data, knowledge, skills, and technology from the oil and gas sector can be applied to expand the development of geothermal resources (Clegg and Krase, 2022; Hardman et al., 2020; Salehi et al., 2022). For example, high risks are associated with determining the best strategies for reservoir development that ensure optimal fluid recovery and reservoir performance. Reliable subsurface models of geothermal reservoirs that honour observed data and can be rapidly generated can significantly improve the efficiency of modelling the heat extraction system and reduce uncertainties, thus ensuring sustainable heat production for decades to come. Techniques such as conventional data assimilation, adopted from numerical weather forecasting, have been applied as alternatives to traditional history matching in petroleum reservoirs (Emerick, 2016, 2018) to generate more accurate reservoir models that can be used for reservoir forecasting. Data assimilation involves the integration of numerical models with observed production data to condition reservoir models (Asch et al.,

2016). However, the application of traditional data assimilation techniques presents significant challenges where data is sparse and the physical system is characterised by high dimensionality (Apte et al., 2008).

Digitisation and machine learning have seen growing use in subsurface applications over the past two decades. Advancements such as automatic differentiation for high-dimensional parameter optimisation with backpropagation (LeCun et al., 2012), graphics processing units (GPUs) for parallelisation (Krizhevsky et al., 2017), large training datasets for improving model accuracy, and attention mechanisms for sequential modelling (Vaswani et al., 2017) have presented opportunities for developing more robust techniques that enable the application of machine learning and artificial intelligence for data assimilation and predictive modelling. Neural networks and other machine learning techniques have been applied with significant success to represent complex physical processes in geophysical systems characterised by sparse data, spatio-temporal heterogeneity, multiresolution data, high dimensionality, noise, and uncertainty (Karpatne et al., 2019). These include lithofacies classification (Brcković et al., 2017; Ma et al., 2009; Maiti et al., 2007), seismic interpretation and inversion (Fu et al., 2019; Liu et al., 2015; Mosser et al., 2020; Wang, Di, Shafiq, Alaudah and AlRegib, 2018), prediction of stratigraphic units (Zhou et al., 2019), generation of geological models (Chan and Elsheikh, 2017; Laloy et al., 2018; Zhang et al., 2019), performance prediction of coalbed methane and CO_2 sequestration (Gorucu et al., 2005), pressure transient analysis in fracture networks (Alajmi and Ertekin, 2007), shale gas reservoir modelling (Klie and Florez, 2018) and surrogate reservoir modelling (Jin et al., 2020; Kani and Elsheikh, 2019; Shahkarami et al., 2015; Temirchev et al., 2020).

Neural networks have been applied to solve ill-posed inverse problems such as generating geological reservoir models constrained by physical measurements (Dupont et al., 2018; Laloy et al., 2018; Mosser et al., 2018; Pan et al., 2021) and solving the history matching problem in reservoir simulation using a data assimilation framework (Brantson et al., 2018; Canchumuni et al., 2020; Jo et al., 2021). Researchers in the oil and gas industry continue to investigate the use of machine learning and other artificial intelligence techniques for developing reduced order models of subsurface reservoirs, also known as surrogate or proxy models (Bahrami et al., 2022; Samnioti and Gaganis, 2023). This research applies neural networks to develop methods for generating geothermal reservoir models that honour observed data and simulate reservoir behaviour at reduced computational costs. In summary, the objectives of the study are as follows:

1. Apply neural networks to generate reservoir models that are consistent with geological processes and geophysical information through the use of surface-based geological models (SBGMs) that better preserve geological features in subsurface reservoirs without posing constraints on model resolution.

2. Develop a history matching framework using neural networks for the assimilation of fluid flow data, using SBGMs that have been conditioned to well data to aid the history matching process.

3. Use neural networks and advanced deep learning techniques to develop robust reduced order models of

geothermal reservoirs that will aid the temporal evolution of compressed variables achieved through the use of space-filling curves to represent high-dimensional data on unstructured meshes in one dimension.

4. Propose a workflow that combines the developed methods into a framework that can allow the rapid generation of representative reservoir models that honour observed well data.

1.2 Contributions and Significance of the Research

The techniques from this research can be applied to develop a computationally efficient method for combining observations and models to produce reliable estimates of a geothermal system. The proposed approach involves the use of deep learning models in an inverse modelling framework to generate geological models that are calibrated to geophysical information and fluid flow models that match historical production data from wells. Furthermore, neural networks are also applied to generate reduced order models that can significantly accelerate reservoir simulation. This contributes to our understanding and development of techniques for real-time reservoir monitoring. Although the product of this research is aimed at geothermal reservoirs, the methodologies and techniques developed can be extended to oil and gas reservoirs, as well as other subsurface reservoirs.

The papers from this Ph.D. and presented in this thesis are under the names of myself, Zainab Titus, my supervisors Christopher Pain, Claire Heaney, Matthew Jackson, and Pablo Salinas, and co-author Carl Jacquemyn. One paper has been published in Computational Geosciences (Titus et al. (2022), Chapter 3). Chapter 4 is in preparation for submission to Computational Geosciences. Chapter 5 will be submitted to Mathematical Geosciences, and Chapter 6 to Computers and Mathematics. All publications will be made under open licences.

During my Ph.D., I have also given oral presentations (virtual) of my work at a number of conferences: European Conference on the Mathematics of Oil Recovery (ECMOR 2020), Machine Learning and Data Assimilation for Dynamical Systems (MLDADS 2021), and International Meeting for Applied Geosicence and Engineering (IMAGE 2021). Finally, I have also had the opportunity to present my work to colleagues at the Department of Earth Sciences and Engineering, Data Science Institute at Imperial College London, and to industry colleagues at CGG Services (UK) Limited.

1.3 Statement of Originality

The work presented herein has not previously been submitted for a degree or diploma at any other institution of learning. This thesis contains no material previously written or published by any other person with the exception of appropriately referenced materials and resources.

1.4 Thesis Outline

In this report, a brief introduction of the research motivation, aims, contributions, and significance is presented in Chapter 1. Chapter 2 gives an overview of geothermal reservoirs and presents applications of data assimilation in reservoir history matching, and also discusses the primary limitations of conventional methods. Current advances in addressing these challenges, particularly the use of deep learning frameworks, are highlighted. Applications of neural networks for geological modelling and reservoir simulation are also presented. In Chapter 3, the use of neural networks for conditioning surface-based geological models to facies data is demonstrated. An extension of the approach for a grid-free application to handle wells with varied trajectories is given in Chapter 4. Chapter 5 adopts surface-based geological models that have been conditioned to well data for reservoir history matching using adversarial neural networks. Chapter 6 presents the use of convolutional autoencoders and space-filling curves for data compression and transformer-based models for predictive modelling. In Chapter 7, the developed methods using data assimilation for history matching and reduced-order modelling for reservoir forecasting are demonstrated. The final chapter of this thesis is the conclusion that summarises the results of this research, recommendations for future work, and the required modifications for applications to field cases.

Chapter 2

Literature Review

2.1 Geothermal Systems

Within the subsurface lies an immense amount of heat, with the core of the Earth reportedly reaching temperatures of approximately 6000°C (Anzellini et al., 2013). This heat is produced by various internal mechanisms and processes that can be cateapproaches, primordial or radioactive heat (Lowrie and Fichtner, 2020). Primordial heat is the energy produced during the initial formation of the Earth and its subsequent cooling, released in the form of kinetic energy. It also includes the adiabatic energy from compression during the Earth's growth and the crystallisation of the inner core, particularly the settling of iron crystals (Schubert, 1997; Vocadlo, 2015). Radioactive heat is produced from the decay of radioisotopes (uranium, thorium, and potassium) and is predominantly transferred within the structure of the Earth through conduction and convection (Lowrie and Fichtner, 2020). The distribution, movement, and surface expression of the produced heat is observed around the globe in various forms as geothermal energy (Hirst, 2017). Geothermal energy is a renewable resource that can be harnessed through various technologies and concepts (Sharmin et al., 2023), and exists as heat stored in different geological settings in the subsurface.

A geothermal system is a geological environment where thermal energy can be obtained from the Earth by circulating fluids, either naturally or artificially induced, and transported for later use (Moeck, 2014). Geothermal systems are characterised by the presence of heat, permeability, and fluid (water). The classification of geothermal systems differs widely due to the diverse range of settings in which geothermal resources are located, as well as the evolving methods for exploration and exploitation (Zarrouk and McLean, 2019). Various categorisations of geothermal resources have been employed based on:

- Accessibility and discovery status (Muffler and Cataldi, 1978),

- Total heat in place and limits of current technology—theoretical, technical, economic, sustainable, and developable (Rybach, 2015),
- Exergy for assessing the quality of the energy contained within recovered geothermal fluids (Lee, 2001),
- Geothermal resources—inferred, indicated, and measured, and geological reserves—probable and proven (Williams et al., 2010).

However, temperature remains one of the most critical factors for categorising geothermal resources. The classification based on temperature and geological settings is widely used and descriptive, although researchers have used various terminologies and factors such as economics, production mechanisms, and end use (direct heat use or indirect electrical use) to determine classifications. Using temperature (enthalpy) as the primary criterion, geothermal systems can be classified into high, medium, and low enthalpy systems (Lee, 1996; Williams et al., 2011).

High enthalpy systems are found in regions with increased surface heat flow and geothermal gradient, such as volcanic areas and plate margins in New Zealand, Iceland, Japan, Kenya, USA, Italy, Philippines, Turkey, and Kenya (Sharmin et al., 2023; Shi, Rop, Wang, Jiang, Wang and Hu, 2021). These systems contain fluids at high pressure and temperatures (usually exceeding 150°C) that can be harnessed for electricity generation, direct use, or reinjection of the condensed fluid containing residual heat. Medium enthalpy systems are characterised by temperatures below 150°C but above 100°C (Muffler and Cataldi, 1978). These systems have the potential to be used for direct heating or electricity generation. This can be achieved through the use of the organic rankine cycle (ORC), a type of binary cycle that employs a working fluid such as iso-pentane or iso-butane with a boiling point lower than that of water (Cano et al., 2022). The working fluid is flashed to steam in a heat exchanger to produce energy or circulated in a closed-loop network to absorb heat from the geothermal reservoir (Scherer et al., 2022). Low enthalpy systems are characterised by temperatures ranging from 40°C to 100°C. These systems are commonly used for direct heating purposes in various industries such as aquaculture, agriculture, paper and textile manufacturing, or for domestic purposes. Low enthalpy systems are predominantly found within permeable sedimentary units and are more prevalent compared to high enthalpy systems.

Moeck (2014) provided a classification of geothermal systems based on temperature, which is heavily influenced by geological and tectonic settings. The categorisation of geothermal plays takes into account tectonic controls and geological characteristics and their impact on hydrogeology, lithology, and geological stresses. The classification system considers the various types of geothermal plays that are influenced by convection and conduction (see Figures 2.1 and 2.2). The methods developed in this thesis primarily apply to low and medium enthalpy resources associated with conduction-dominated systems and sedimentary basins. These resources are located in geological settings similar to those explored and exploited for hydrocarbons.

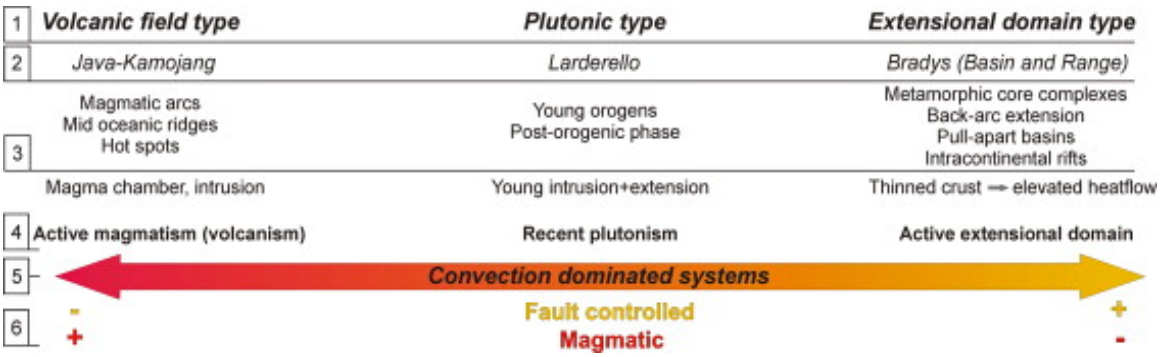


Figure 2.1: Geothermal play systems dominated by convection, categorised based on the geological controls of igneous activity. The three main categories are volcanic, plutonic, and extensional. The catalogue scheme includes the following information for each play: 1. Play type 2. Example 3. Plate tectonic setting 4. Geological habitat of potential geothermal reservoirs 5. Heat transfer type 6. Geological controls. Image from Moeck (2014).

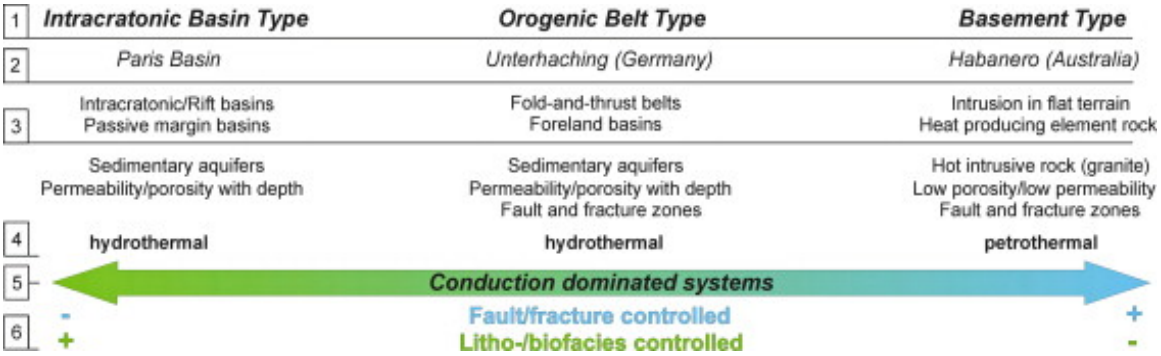


Figure 2.2: Conduction-dominated geothermal play types encompass a variety of geological settings, such as intracratonic basins and foreland basins within orogenic belts. The geological controls in these play types are primarily determined by the litho- or biofacies of the sedimentary rock, as well as faults and fractures. The catalogue scheme includes the following information for each play: 1. Play type 2. Example 3. Plate tectonic setting 4. Geological habitat of potential geothermal reservoirs 5. Heat transfer type 6. Geological controls. Image from Moeck (2014).

2.1.1 Opportunities for Geothermal Energy Production in Sedimentary Basins

Globally, geothermal projects have been restricted to locations where the heat flow and geothermal gradients are elevated near the surface, such as high-enthalpy volcanic environments (Sharmin et al., 2023). Conventional geothermal resources are typically produced from deep permeable formations having high temperatures. In some cases, the fracture permeability of faults could significantly enhance the local geothermal reserve (Busby, 2014). On the other hand, technological interventions are employed to enhance the natural permeability where natural flow rates are not economically feasible. These are known as enhanced geothermal systems (EGS) (Rybach, 2010). Although estimates of probable geothermal reserve rely on various assumptions, the presence and knowledge of local permeability and transmissivity are crucial for successful exploitation (Busby, 2014). Reservoir rocks in sedimentary basins, having sufficient temperature, thickness, porosity, and permeability, can extend the production of geothermal energy to previously untapped areas (Anderson, 2013; Vivas et al., 2023).

The geothermal potential of deep sedimentary basins has been extensively studied in different regions of the

world, both in the past and recently. Sedimentary basins in geological settings similar to those encountered in oil and gas production are becoming increasingly attractive for the development of geothermal resources (Sanyal and Butler, 2010). This presents new opportunities for the production of geothermal energy in various regions across the globe. Heat conducted from the lower crust of the Earth during the process of lithification of sedimentary rocks allows sedimentary basins to store significant amounts of thermal energy (Bethke et al., 1988). These resources, known as hot sedimentary aquifers (HSA), consist of geothermal resources at depths where temperatures and pressures are sufficiently high for direct and indirect use. The temperature field in hot sedimentary aquifers is predominantly driven by natural heat conduction (Moeck, 2014; Rybach, 1981). The hydraulics of groundwater in these systems is assumed to be single-phase, liquid-dominated, laminar, and normally considered in confined and saturated reservoir conditions, with permeability considered the dominant control on fluid flow and, therefore, on heat and mass transport (Blank et al., 2021). Although many sedimentary basins lack vertical permeability due to the existence of impermeable shale layers between permeable sandstone layers, a well in such a basin can still penetrate multiple layers of sand to reach a reservoir flow capacity that allows commercially viable flow rates (Sanyal and Butler, 2010).

The exploitation of hot sedimentary aquifers for geothermal energy has been considered in various locations such as Australia (Goldstein et al., 2009), Canada (Ferguson and Grasby, 2014; Jessop et al., 1991), the United States (Vivas et al., 2023), the United Kingdom (Busby, 2014; Gluyas et al., 2018; Hirst, 2017) and Morocco (Barkaoui et al., 2014). The Dogger aquifer in France serves as an example of a geothermal reservoir in a sedimentary basin that has been utilised for several decades (Lopez et al., 2010), suggesting that similar developments could be possible elsewhere. Sedimentary formations in the Imperial Valley of California are known to posses high heat flow (Rex, 1973), having installed geothermal power capacity of over 700MW for electricity generation (Quintero-Núñez et al., 2015). The Arkoma and Anadarko basins in Oklahoma are known to have regions with elevated geothermal gradients, which makes them attractive for the development of medium enthalpy geothermal systems (Vivas et al., 2023). Sedimentary basins in other regions, such as the Karoo Basin in South Africa (Campbell et al., 2016), and those in oil and gas fields, such as Quebec in Canada (Bdard et al., 2018), the Niger Delta in Nigeria (Odumodu and Mode, 2016), and the Wessex Basin in UK (Hirst, 2017; Watson et al., 2020), have been reported to have bottomhole temperatures sufficient for the development of geothermal resources.

Compared to traditional geothermal exploration, the development of geothermal energy in deep sedimentary basins offers several advantages. Many reservoirs, particularly those found in oil and gas sedimentary basins, already have a wealth of exploration data, such as core samples, seismic data, and logs. These data, along with historical records of reservoir performance can greatly aid in accurately characterising these reservoirs (Ferguson and Ufondu, 2017). Furthermore, the expertise gained from drilling and constructing wells in oil and gas sedimentary basins can be applied to the development of geothermal projects, making the drilling process

less risky, potentially faster, and more cost-effective compared to drilling in igneous and metamorphic rocks.

An additional benefit of exploring sedimentary basins is the opportunity to utilise existing, decommissioned, or abandoned oil and gas wells. This can make it easier to access geothermal energy by decreasing the high initial investments that limit the growth of this sustainable energy resource, particularly the costs associated with the construction of new wells for geothermal projects. Repurposing existing and inactive oil and gas infrastructure presents potential avenues for harnessing this renewable energy source in innovative ways (Davis and Michaelides, 2009). Recently, the use of existing oil and gas wells has been studied for geothermal development projects for example in Italy (Alimonti et al., 2021), Colombia (Pinto et al., 2021), United Kingdom (Watson et al., 2020), United States (Caulk and Tomac, 2017), Canada (Ferguson and Ufondu, 2017), Iran (Gharibi et al., 2018) and China (Wang et al., 2016). The co-production of geothermal energy from produced water of sufficient quantity and temperature to generate electricity using binary systems alongside oil operations was demonstrated at Teapot Dome in the United States (Johnson et al., 2010). Other studies have also explored the possibility of combining geothermal energy production with carbon capture and storage (CCS), where the captured CO_2 could be used to extract heat from geothermal resources in sedimentary basins (Bielicki et al., 2023).

The examples aforementioned demonstrate that the competencies, methodologies, technologies, approaches, and knowledge needed for geothermal energy exploration, well construction, and production are largely similar to those utilised in oil and gas exploitation. The extensive expertise gained by oil and gas companies and other industry participants through the exploration of hydrocarbon reserves in past and recent years can be advantageous for the advancement of geothermal projects (Clegg and Krase, 2022). Geoscientists and reservoir engineers can leverage their knowledge and understanding of seismic surveys, deep wellbore, and drilling sedimentary basins to evaluate and exploit geothermal resources while minimising subsurface risks and uncertainties (Hardman et al., 2020). Furthermore, the engineering, safety, and project management knowledge acquired in the oil and gas sector can be effectively re-applied in geothermal energy projects (Salehi et al., 2022). This presents fresh opportunities for individuals and companies in the oil and gas industry to transition to the field of geothermal energy. Using their expertise in exploration, drilling, and production, oil and gas workers can take advantage of these emerging developments.

The future of geothermal energy is promising as it has significant potential to meet the growing global energy demand as part of a diverse energy supply mix (Goldstein et al., 2011; Vargas et al., 2022). With recent technological advancements and continuous improvements in equipment, prospecting, and drilling techniques, the geothermal industry now has the capability to generate electricity from geothermal resources with significantly lower temperatures and in geological environments different from those required years ago.

This research focuses on the simulation of geothermal reservoirs in geological settings similar to oil and gas reservoirs by applying data assimilation for history matching, and reduced order modelling for enhanced reservoir simulation and management using deep learning techniques. The following section presents an overview of

geothermal reservoir simulation with emphasis on heat flow in conduction-dominated systems in sedimentary geothermal systems.

2.1.2 Geothermal Reservoir Simulation

The amount of thermal energy that can be extracted from a geothermal reservoir depends on the amount of stored energy, as well as the thermal conductivity and permeability of the rock. These govern heat and momentum transport (Blank et al., 2021). Geothermal reservoir simulation enables the estimation of the energy stored and that which can be recovered given the reservoir rock and fluid properties (such as porosity, permeability, temperature, pressure), under different production mechanisms and scenarios.

The temperature field in sedimentary settings (such as in intracratonic or foreland basin geothermal play types) is driven by natural heat conduction (Moeck, 2014; Rybach, 1981). An additional means of heat transport in geothermal systems is provided by advective forces resulting in heat flow becoming a combination of advection and conduction (Blank et al., 2021; Salinas et al., 2021). Permeability is the main control on fluid flow (and consequently on heat and mass transport) and is key to characterising and assessing geothermal reservoir performance. The mathematical model applied in this research consists of a fluid flow model described by Darcy's Law (Darcy, 1856) derived from the momentum balance equation for Newtonian fluids, and an advection diffusion model for the temperature distribution described by the conservation of energy equation. These two equations are coupled such that the velocity solution of Darcy's equation determines the advective field describing the temperature evolution. The model is applied to a confined and saturated reservoir, predominantly composed of consolidated sedimentary material. Geomechanical and chemical effects are neglected and thermal equilibrium between the rock and fluid is assumed. Flow of a single-phase, Newtonian fluid in an isotropic, non-deformable reservoir is modelled. Assuming conservation of mass and momentum, the mass balance equation is given as:

$$\frac{\partial(\phi\rho_f)}{\partial t} + \nabla \cdot \boldsymbol{u}\rho_f = s_c, \tag{2.1}$$

where ϕ is the porosity, ρ_f is the pressure and temperature-dependent fluid density, $\boldsymbol{u}$ is the Darcy velocity, s_c is a source term. The momentum equation is:

$$\mu(\boldsymbol{K})^{-1}\boldsymbol{u} = -\nabla p + \rho_f \boldsymbol{g} + s_m, \tag{2.2}$$

where $\boldsymbol{g}$ is the vector of the gravitational acceleration, $\boldsymbol{K}$ is the permeability tensor, p is the pressure, μ is the fluid viscosity,s_m is a source term in the momentum equation. The temperature field is described by the energy

balance equation:

$$\xi \frac{\partial T}{\partial t} = -\nabla \cdot (\kappa \nabla T) - \phi \rho_f C_{pf} \nabla \cdot (\boldsymbol{u} T) + s_t, \tag{2.3}$$

$$\xi = (1-\phi)\rho_p C_{Pp} + \phi \rho_f C_{P_f}, \tag{2.4}$$

$$\kappa = (1-\phi)\kappa_p + \phi \kappa_f, \tag{2.5}$$

where C_{P_p} is the porous rock heat capacity, C_f is the fluid porosity, ρ_f is the fluid density, κ_p is the rock thermal conductivity, κ_f is the fluid thermal conductivity, ξ is the effective heat capacity, T is the temperature field, $\boldsymbol{u}$ is the velocity of the fluid obtained using Darcy's Law and κ is the effective thermal conductivity. The resulting heat transport and governing porous-media fluid flow equations are solved using the control volume finite element method (CV-FEM) implemented in the open-source Imperial College Finite Element Reservoir Simulator (IC-FERST) (Jackson et al., 2015; Salinas et al., 2018, 2021).

However, although numerical and computational techniques can be employed to implement mathematical models and simulate physical phenomena, observed data must be used to validate a quantitative model. The capacity of a model to mimic observed behaviour is key to evaluating reservoir performance, optimising recovery, and forecasting production. Data assimilation provides a method for integrating observations into models to enable models to match observed historical data (Asch et al., 2016; Carrassi et al., 2018). This research applies data assimilation for history matching, and reduced-order modelling to develop representative models of geothermal reservoirs. The geothermal systems considered are located in sedimentary environments similar to oil and gas reservoirs. Drawing on extensive research and applications of data assimilation and reduced-order modelling in the oil and gas sector, these advanced techniques can be effectively re-applied to geothermal energy development using an artificial intelligence-driven approach. The following sections provide an overview of data assimilation and its application in the generation of representative models of subsurface reservoirs.

2.2 Data Assimilation

Data assimilation as the interface between observations and models enables the identification of the structure of a geophysical system from a set of discrete space-time data (Tarantola, 1986). The goal of data assimilation is to provide the best estimate of the system in the past, present, and future by combining information from two complementary sources—models and observations—while incorporating the limitations of the data, such as errors and uncertainties (Carrassi et al., 2018).

Data assimilation has been applied in problems where state estimation of a dynamic system from limited information is required; for example, the initial climate conditions from precipitation and temperature in a region as demonstrated by classical numerical weather prediction models (Daley, 1993; Evensen, 2003; Kalnay,

2012), or obtaining the subsurface structure of a reservoir using 4D seismic data to update models (Emerick, 2016; Yin et al., 2019). The various formulations of data assimilation methods involve matching model data with observed data. The methods for assimilating data into models are derived from control theory or estimation methods (Asch et al., 2016), although an exception to this approach is the nudging method (Auroux et al., 2016). Based on the application and the kind of information required, data assimilation methods can be broadly classified into variational, statistical, and hybrid methods.

Variational Methods employ a least-squares error estimation approach by minimising an objective function defined as the mismatch between the model and observed data. The assimilated state is defined as the solution that minimises the cost function obtained using iterative gradient-based algorithms. Commonly applied techniques include 3D and 4DVar analysis. While 3DVar computes the global solution for a particular time interval, 4DVar computes the solutions at several times within an assimilation interval (Asch et al., 2016; Bouttier and Courtier, 2002).

Statistical Methods apply probabilistic theorems for linear combinations of the model and observational data to produce the analysis. The data are assumed to be unbiased, the model linear, and the uncertainties are given in the form of covariance matrices. Statistical methods such as the Best Linear Unbiased Estimate and Kalman Filter (Kalman, 1960) provide information such as the variance and likelihood of the best estimates.

Hybrid Methods implement statistical and variational schemes to different degrees, such as iterative ensemble Kalman smoother (IEnKS) and the ensemble data assimilation (EDA). These more advanced assimilation methods combine the robustness of variational methods and the additional information provided by statistical methods to produce estimates of a system. A common hybrid approach is the combination of 4D variational method with ensemble Kalman filter in the 4DEnVar, where ensembles are used to determine the error structure and information for the covariance matrix. Using a robust covariance matrix derived from ensemble theory, a variational approach is then applied to assimilate the data (Yang et al., 2015).

The following section outlines the applications of data assimilation in reservoir engineering for history matching. Furthermore, limitations of current ensemble-based methods are highlighted, and improvements using hybrid methods are outlined. To relate these methods to the technique proposed in this research, emphasis is given to the IEnKS, an advanced DA scheme that applies variational analysis and ensembles.

2.3 Data Assimilation for Reservoir History Matching

Determining the spatial distribution of petrophysical properties that are key to fluid flow within the Earth's subsurface, such as porosity and permeability, is important for developing reliable reservoir models. However, subsurface environments are characterised by complex geological features at various length scales, which can

only be partially described or known from sparse observational data, leading to significant uncertainty in model development and predictions. Therefore, since the true state of the reservoir is uncertain, data assimilation serves as a powerful tool for integrating dynamic and static information into numerical models to aid characterising the reservoir and quantifying uncertainty.

For many applications of data assimilation in Earth science such as atmospheric and oceanic modelling, the knowledge of the distribution and evolution in space and time of the characteristics of the physical system is required. Therefore, the objective of applying data assimilation is to estimate the initial conditions of the model (Asch et al., 2016; Reichle, 2008) and produce an updated model state that most accurately approximates the true system state. However, the data assimilation framework is much more general and can also be used to estimate the parameters of a system (Chen et al., 2024; Hansen and Penland, 2007).

In subsurface modelling, data assimilation is more commonly applied for parameter estimation, where the objective is to calibrate the model to the time-distributed observations and estimate the model parameters (typically porosity and permeability) that govern the observed dynamic response (Gu and Oliver, 2006). This involves calibrating a spatial model of the petrophysical parameters with available physical measurements. The calibration task, commonly referred to as history matching, is an inverse problem that requires finding the underlying spatial distribution of petrophysical properties that explain the observed data (Mosser et al., 2019).

Traditional history matching involves manual adjustment of static parameters and running of the reservoir simulator to ensure that the results align with the dynamic data observed. This involves iteratively modifying initial inputs and geological model assumptions and generating multiple equivalent solutions to achieve a reasonable match at various production stages and under different operational conditions. However, the time-consuming nature of this trial and error process led to the development of automatic history matching, which offers greater accuracy and efficiency (Chen and Oliver, 2011). In automatic history matching, an objective function quantifying the disparity between measurements and model data is minimised within an optimisation framework. This process employs techniques such as gradient-based optimisation, response surface methods, genetic algorithms, particle swarm optimisation, simulated annealing, and variogram-based sensitivity to solve the inverse problem (Brantson et al., 2018).

Due to the sparse and indirect nature of the available data, the history-matching problem is an ill-posed inverse problem. Furthermore, the number of model parameters to be estimated far exceeds the number of observed data, significantly increasing the dimensionality and computational complexity of solving the inverse problem (Oliver and Chen, 2011). Solving inverse problems with spatially dependent parameters has also been shown to be often ill-posed due to the non-uniqueness (multiple geological models having various combinations of reservoir parameters can often fit the same set of observations equally well) and instability (lack of continuous dependence) associated with the solutions (Kabanikhin, 2008; Kravaris and Seinfeld, 1985). Hence, most approaches to solving the history problem adopt a Bayesian approach, where model parameters are sampled from a prior

distribution to honour the observed data of the forward problem. The randomized maximum likelihood (RML) method (Evensen et al., 2022) has frequently been used to generate an approximate sampling of the conditional posterior probability density function (pdf). If the set of history-matched models is obtained by correctly sampling the pdf, then the plausible models are consistent with the given information and data, and the ensemble of these realisations provides a characterisation of uncertainty in the reservoir (Zafari and Reynolds, 2007).

Over the past two decades, data assimilation has attracted huge attention in geosciences and specifically in reservoir history matching for accumulating observed data into subsurface models for recursively updating the model state. The rationale for the application of data assimilation techniques in reservoir modelling is attributed to the characteristic limited data, large-scale nonlinearities, and high uncertainty common to other geophysical models that apply data assimilation. In reservoir history matching, static data, such as petrophysical properties (usually porosity and permeability), represent the model parameters, while dynamic data such as production data (well oil and water flow rate, wellbore pressure), reservoir pressure, and fluid saturation, are the observed data integrated into reservoir models.

2.3.1 Ensemble-Based Methods for History Matching: Applications, Improvements and Limitations

Ensemble-based methods have been employed to generate multiple realisations of reservoir models that approximately characterise the reservoir. Information from the ensembles provides an estimate of the sensitivities and uncertainties. This typically uses the ensemble Kalman filter (EnKF) data assimilation method developed by Evensen (Evensen, 1994, 2003; Evensen et al., 2007) as a Monte-Carlo-based approach to data assimilation where statistical values for correlating the model parameters and observations are generated from a suite of model possibilities known as an ensemble (Anderson, 1996). The EnKF is suitable for the assimilation of data in time series, such as downhole measurements of pressure, temperature, and flow rate, which serve as observation data to be assimilated (Shuai et al., 2016).

Several applications of ensemble-based data assimilation to synthetic and field cases, particularly the works of Lorenz (2005), Naevdal et al. (2005), Gu and Oliver (2006), and Evensen et al. (2007) set the pace for these methods in petroleum reservoir simulation. Zafari and Reynolds (2007) explored the application of EnKF for uncertainty quantification in reservoir characterisation. The authors applied the randomised maximum likelihood (RML) approach to generate reservoir models using adjoint models. The approach was, however, computationally expensive, as the derivation of adjoint models required forward and equivalent backward simulations, in addition to assimilation steps.

Haugen et al. (2008) presented one of the first applications of EnKF in history matching production data for

a real problem of a field in the North Sea. The vertical and horizontal permeability constituted the model parameters used for generating the initial ensembles with the EnKF assimilation performed twice every month for five years. In another field example, Skjervheim et al. (2007) extended the method to the assimilation of seismic data along with the production data used for history matching. The results of assimilating production and seismic data improved the results of the model, as the permeability field showed a reasonable match with the reference case.

Although these early synthetic and real field cases yielded more satisfactory results than the traditional history matching approach, some limitations of these methods were identified, and improvements have been made particularly to address issues relating to the Gaussian and linear assumptions inherent in the EnKF concept. Jung et al. (2018) have provided a review of improvements in the application of ensemble-based data assimilation for history matching heterogeneous reservoirs. The assumption of a Gaussian distribution derived from the application of Kalman filters (Kalman, 1960) for small-scale linear problems has been extensively investigated by past and recent researchers. Since subsurface reservoir systems are characterised by high dimensionality and nonlinearity of the model, and also of the relationship between the model data and observed data, the EnKF can result in forecasts which are significantly different from the observations, especially for complex geology. For example, in a 2D horizontal reservoir model under waterflooding scheme, EnKF can produce a poor analysis of uncertainty in reservoirs if the variables have a multimodal distribution, which is typical for channel reservoirs (Zafari and Reynolds, 2007).

Parameterisation has been applied to transform non-Gaussian variables to have a Gaussian distribution to enable the implementation of the EnKF using techniques such as the normal score transform (Chen et al., 2009), the discrete cosine transform (Jafarpour and McLaughlin, 2008; Lorentzen et al., 2012), and the use of Gaussian mixture models (Dovera and Della Rossa, 2011), where the probability distribution function of the non-Gaussian parameters is transformed into weighted sums of Gaussian probability distribution functions.

To handle nonlinearities, iterative methods have been applied to large-scale models. These techniques estimate model parameters and forward simulate dynamic variables using the estimated initial conditions recursively. Iterative ensemble methods are based on minimising the objective function expressed as the data mismatch between the model and the observations. The ensemble randomised maximum likelihood approach introduced by Gu and Oliver (2007) presents one of the early applications of iterative ensemble methods in reservoir simulation. The authors used EnKF recursively to minimise the data mismatch function, where singular value decomposition (SVD) was applied to compute the sensitivity matrix required for optimisation.

More recently, Emerick (2016) and Jo et al. (2017) applied the iterative form of the ensemble Kalman filter (EnKF) for reservoir history matching. The capacity of the model for inversion even with ensembles having initial model parameters that do not contain the reference case was enhanced by recursively updating the state vector containing the model parameters. However, the need to restart the simulation at every assimilation step

in EnKF increases computational costs when combined with iterative methods. A more efficient approach is the use of iterative ensemble Kalman smoothers (IEnKS). Section 2.3.2 provides an overview of the method, applications, and current limitations to demonstrate the improvements that the proposed method in this research aims to achieve.

2.3.2 Iterative Ensemble Kalman Smoother: An Efficient History Matching Technique

An advanced data assimilation method using smoothing—the iterative ensemble Kalman smoother (IEnKS), has been proposed as more accurate than standard filtering, particularly for parameter estimation problems such as history matching (Emerick, 2018). Although the filtering problem is the most common in geoscience applications, smoothing has been shown to have similar accuracy as filtering with about 90% reduction in computational cost (Skjervheim et al., 2011). While filtering gives an estimate at a specific time, $t_n (0 \leq n \leq N_t)$, where N_t is the total simulation time and incorporates all observations until t_n, smoothing at t_n incorporates observations at all times from t_n, and updates both the current state and the past states. By combining information from the past, present, and future, a reanalysis of the state is performed by recursive forward and backward propagation of information. Nonetheless, at the final time, both filtering and smoothing are expected to converge as they would have assimilated the same data (Kalnay and Yang, 2010). Figure 2.3 illustrates the filtering and smoothing problems. The ensemble Kalman smoother (EnKS) uses information from the ensembles to estimate the tangent linear model (TLM) and the adjoint of observation models, or makes use of available TLM or adjoints if they exist. The IEnKS is an iterative form of the EnKS where the model is updated recursively until a specified tolerance is reached. In each iteration, observations far ahead in time from the updated state are assimilated to stabilise the scheme in the direction of the unstable nodes (Asch et al., 2016).

Skjervheim et al. (2011) introduced the use of an ensemble smoother for history matching a hydrocarbon reservoir model. Chen and Oliver (2014) applied the IEnKS to history matching a model of the Norne field in the North Sea. The study addressed the complexities of data assimilation for real field cases highlighted by early research, such as nonphysical values from poor initial ensembles. To improve the convergence of IEnKS to a solution, Luo et al. (2015) adopted a regularised Levenberge Marquardt (RLM) approach to optimise the objective function defined as a minimum average cost (MAC) function. The RLM-MAC algorithm outperformed IEnKS in estimating initial conditions, facies distribution, and history matching of production data. The key improvement in applying regularisation to the algorithm improved the computational efficiency of the method and provided a framework for biasing the solution of the optimisation problem.

These studies demonstrate that the use of ensembles, a variational approach with regularisation, and smoothing in the model update process can provide reasonably accurate assimilated models. Other applications of

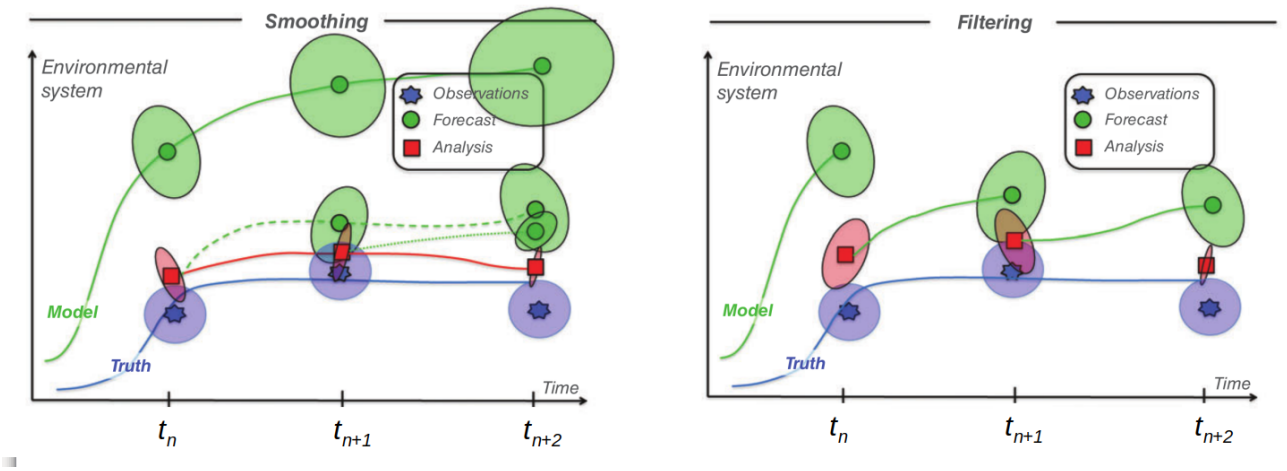


Figure 2.3: Smoothing (left) and Filtering (right), from (Carrassi et al., 2018). The true unknown measurement is represented by the blue line. The blue observations (blue stars), forecast (green squares), and analysis (red squares) are given as point estimators based on the corresponding probability density functions (pdfs), where a non-unique choice is taken as the mean of the pdfs. The associated ellipsoids display the pdfs of the measurements; the smaller the size, the smaller the estimated uncertainty. The observational error is assumed constant. In smoothing, observations from t_n to t_{n+2} are applied simultaneously to determine the analysis from t_n to t_{n+2} similar to the 4D variational analysis. The solutions at t_n and t_{n+1} provide initial conditions for the predictions until t_{n+2} as shown by the forecasts in green dash and dotted lines from t_n and t_{n+1} respectively , with predictions from the observation at t_{n+1} being the most accurate. Filtering shows that using the analysis obtained at t_n from the model and the observed data at that time, a prediction is made to t_{n+1}. A new analysis is obtained using the prediction and the observation at t_{n+1}, which is applied to generate the new forecast at t_{n+2}.

data assimilation, particularly ensemble techniques to generate multiple realisations that approximate the characteristics of the reservoir, have also been developed (Emerick, 2016, 2018; Evensen, 2003). However, the computational cost of implementing variational methods has limited their application, particularly in computing the adjoints for gradient-based optimisation for the iterative solver. Furthermore, the use of multiple data assimilation windows which are required for improving the accuracy for large-scale systems increases the complexity of the optimisation. The computational costs, time requirements, and quality of the results still present challenges in the use of conventional data assimilation methods in reservoir history matching. In this research, the use of neural networks is proposed within an integrated framework for well conditioning, history matching, and reservoir simulation to address the computational demands associated with other techniques, and to explore its potential for forward-looking reservoir management strategies such as real-time reservoir forecasting. Section 2.4 discusses the use of neural networks as an alternative approach to conventional data assimilation and for generating reliable models of the subsurface by integrating machine learning into geological modelling, history matching, and reservoir forecasting workflows.

2.4 Generating Reservoir Models with Neural Networks

Neural networks have been applied with significant success to solve inverse problems in geoscience characterised by sparse data, spatio-temporal heterogeneity, multiresolution data, high-dimensionality, noise, and uncertainty (Karpatne et al., 2019). Machine learning algorithms have been used in different phases of subsurface reservoir planning and development (Dramsch, 2020; Samnioti and Gaganis, 2023; Yu and Ma, 2021). In reservoir engineering, artificial neural networks have been used to predict performance indicators in CO_2 sequestration (Gorucu et al., 2005), pressure transient analysis in fracture networks(Alajmi and Ertekin, 2007), shale gas reservoir modelling (Klie and Florez, 2018), surrogate reservoir modelling (Shahkarami et al., 2015), and parameterisation and generation of geological models (Chan and Elsheikh, 2017; Laloy et al., 2018; Pan et al., 2021; Zhang et al., 2019).

Advancements in the field of machine learning, such as efficient automatic differentiation for backpropagation (Bishop, 2006; LeCun et al., 2012; Rumelhart et al., 1986*b*), high-dimensional parameter optimisation (Bottou, 2010), and graphics processing units (GPU) for parallelisation (Krizhevsky et al., 2017), have presented opportunities for developing techniques that allow the widespread application of machine learning and artificial intelligence in the development of more representative subsurface reservoir models. Furthermore, the emergence of state-of-the-art deep learning tools and models has enabled transfer learning, where models trained on other tasks can be used as foundation models for artificial intelligence applications in subsurface development to accelerate modelling and optimisation processes. Over the years, there have been various approaches to applying machine learning in geological modelling, reservoir simulation, and history matching.

2.4.1 Geological Modelling

Geological modelling aims to generate subsurface models that can describe the geological context— stratigraphic, structural, and depositional characteristics of the reservoir, given limited sample data. This involves applying techniques for interpolating and predicting geological features and architectures. In most applications, geological units are represented as discrete variables based on information about petrophysical properties and facies distributions or associations across different length scales from core plugs, well logs, to seismic data (Fitch et al., 2015; Miller and Castle, 2000). This allows for the formulation of the modelling as a supervised learning task, where a neural network can be trained to represent the underlying process of correlating stratigraphic and structural processes to rock types as a reflection of the depositional environment and lithology Li et al. (2023).

Numerous techniques have been applied for geological modelling of subsurface reservoirs using neural networks. Most applications have focused on conditioning geological models generated using pixel- or grid-based geostatistical techniques, such as multiple-point statistics and object-based modelling (Chan and Elsheikh, 2019; Dupont et al., 2018; Laloy et al., 2018; Mosser et al., 2018). However, these are less suitable for capturing

complex geological features, as such models do not always adhere to conceptual depositional rules governing interactions among sedimentary structures, for example, fluvial channels and interconnecting bodies (Osman et al., 2021). Forward modelling techniques such as grid-free process-, object-, and surface-based models produce more realistic reservoir architectures and property distributions than conventional modelling methods. Object-based modelling enables the representation of characteristic geological features by situating geobodies and sedimentary structures within a given domain (Deutsch and Wang, 1996). Process-based techniques leverage the underlying physical principles of sedimentological processes to generate geological models by simulating the temporal evolution of the physical processes responsible for shaping geological architectures (Miller et al., 2008). However, in process- and object-based methods, it is computationally expensive to depict geometry, stratigraphic sequences, variations in sediment, and other geological processes, since complex numerical simulations are typically required to generate these models. In addition, conditioning them to observed data to produce more representative geological models of the reservoir is challenging.

Surface-Based Geological Models

Surface-based modelling captures geometric complexities and effectively integrates heterogeneity across a broad spectrum of length scales. These models offer geological realism at a fraction of the computational cost associated with process- or object-based methods (Pyrcz et al., 2015). Surface-based geological models (SBGMs) use surfaces to control the temporal or spatial positioning of geological features (Jackson et al., 2014; Jacquemyn et al., 2019; Pyrcz et al., 2005; Ruiu et al., 2016; Sech et al., 2009; Zhang et al., 2009). Based on the reservoir modelled, hierarchical relationships are applied to establish the order in which surfaces are modelled to represent the interaction between surfaces where they truncate or conform.

Few studies have developed methods for generating SBGMs that are conditioned to well data. In one of the first studies, Pyrcz et al. (2005) produced SBGMs for deepwater lobe deposits that respect the geometry of the event deposit and the contact of the wells through localised adjustments of the lobe bounding. Other approaches include dual-spline interpolation (Zhang et al., 2009), inverse modelling to sequentially optimise SBGM input parameters (Bertoncello et al., 2013), reverse migration for channelised systems (Parquer et al., 2017), implicit filtering based on coordinate search (Wang, Pyrcz, Catuneanu and Boisvert, 2018), variational inference (Wingate et al., 2016). The computational complexity involved in generating and conditioning SBGMs, the time required for well correlation interpretation, the demand for numerous wells to effectively condition these models, and the potential to alter reservoir geometry, in some cases, have posed limitations for widespread application. Neural networks have been applied to address some of these issues. Using generative adversarial networks to extract the geological features of the models and generate realisations, Jo et al. (2021) conditioned rule-based model to well data. However, the use of grid-based approaches in some of the aforementioned studies poses constraints to model resolution and impacts the ease of use and manipulation of these models.

The approach presented here offers greater flexibility in achieving geological realism compared to conventional geostatistical or grid-based rule-based methods. The use of parametric non-uniform rational B-splines (NURBS) (Piegl and Tiller, 1997) to create models that better capture geological structures and require fewer parameters to represent complex geological features (Jacquemyn et al., 2019). The inputs to the SBGMs are derived by stochastically sampling variables from uniform distributions. These variables characterise the shape, orientation, and relative positioning of surfaces, including depositional and erosional characteristics such as angle of inclination and thickness of the parasequence. With few parameters, a realistic geological model can be accurately developed using this approach. More details on the surface-based modelling technique applied in this research are provided in Jacquemyn et al. (2019).

The use of such SBGMs with few parameters opens up the opportunity for applying neural networks to generate multiple realisations at the required level of detail and with reduced dimensionality for optimisation. This allows for conditioning geological models to well data, preserving the underlying geological processes, and avoiding constraints associated with grid-based approaches such as limitations on model resolution. A robust static model that represents the geological features and architectures of a reservoir, and is calibrated with well data forms the basis for simulation of fluid flow and integration of historical production data.

2.4.2 History Matching

History matching, a fundamental step in obtaining representative models of the reservoir, is key for reservoir forecasting and optimisation. Generating reliable models of the subsurface is crucial for reservoir prediction, optimisation, and sensitivity analysis. In one of the earliest applications of neural networks in reservoir history matching, Zangl et al. (2006) applied artificial neural networks (ANNs) to develop a proxy function for a gas storage reservoir to speed up the optimisation routine for history matching using a genetic algorithm. Although the authors reported that the ANN approach resulted in low CPU consumption, the method requires assessment for specific reservoir and production systems and may not be practical for high-dimensional reservoir models. Esmaili and Mohaghegh (2016) trained an ensemble of neural networks to model fluid flow in a shale gas reservoir and applied the model to match the monthly gas rate and cumulative production. A key benefit of this method is its ability to incorporate various field measurements to identify functional relationships and patterns in reservoir models. Similar applications have been developed using ANNs to directly model the relationship between uncertain reservoir parameters such as permeability and porosity, and outputs such as pressure and production rate, followed by an optimisation to match history data (Cullick et al., 2006; Denney, 2006; Mohmad et al., 2020; Reis, 2006).

History matching based on Bayesian methods has also been performed using neural networks to enable reservoir modelling and uncertainty quantification and reduction. This generally involves estimating the model parame-

ters using a probabilistic approach by sampling reservoir parameters from an initial distribution (prior), running several simulations with these parameters, and evaluating the likelihood of the posterior (Maschio and Schiozer, 2014). Models with the best fit are assigned higher probabilities, and the initial parameters are updated using Bayes' rule to obtain the posterior probability by applying sampling techniques to the posterior distribution such as Markov Chain Monte Carlo algorithms (Chai et al., 2018; Maschio and Schiozer, 2014), radial basis function neural networks (RBFNN) (Eltahan et al., 2020) or genetic algorithms (Christie et al., 2006).

Recurrent neural networks (RNNs) mainly used for modelling time series data or other sequential data and convolutional neural networks (CNNs) for modelling the spatio-temporal evolution of subsurface reservoirs have also been proposed to improve the accuracy and reliability of the history matching process (Ma et al., 2022; Pal et al., 2022). Other state-of-the-art deep learning algorithms and methods, such as generative networks, have also been applied to model subsurface reservoirs by exploring the low-dimensional Gaussian latent space to improve the speed of the history matching process (Canchumuni et al., 2020, 2019*a*; Jo et al., 2020; Mosser et al., 2019).

To speed up the reservoir modelling and history matching process, particularly for more complex large-scale reservoirs, dimensionality reduction of the reservoir parameters and/or production data have been performed. Liu et al. (2019) used principal component analysis (PCA) with CNNs to produce models that match the reference model, while Alguliyev et al. (2022) combined a convolutional variational autoencoder (VAE) with an ensemble smoother (ES) to perform history matching. In another study, Brantson et al. (2018) used a generalised regression neural network (GRNN), a one-pass learning algorithm based on nonparametric regression, for history matching of tight gas carbonate reservoir production profiles. However, the GRNN proved unsuccessful in matching the test data. Srinivasan et al. (2021) employed machine learning techniques for history matching of unconventional reservoirs, using Physics-Informed Machine Learning to capture and preserve the underlying physical phenomena and mechanisms of fluid flow in complex reservoirs, such as fractured shale gas reservoirs. The authors also demonstrated the use of transfer learning, where fast, low-fidelity models that offer speed advantages are combined with slower, high-fidelity models to enhance forecasting accuracy.

In another approach, Canchumuni et al. (2020) used generative models to construct parameterizations of the reservoir model, which were updated using ensemble smoother with multiple data assimilation (ES-MDA). The authors noted that the training time of the networks posed limitations, particularly considering the relatively small size of the model used. Similarly, Mosser et al. (2019) explored the low-dimensional Gaussian latent space in generative adversarial networks (GANs) to generate adjoints for variational data assimilation to improve the speed of history matching. Li and Misra (2021) applied reinforcement learning, using deep Q network (DQN) and deep deterministic policy gradients (DDPG), to replace the process of exploring the parameter space and performing inversion to predict reservoir permeability for history matching. This approach involves iteratively adjusting reservoir parameters and simulating production history, which can be time-consuming

for more complex and realistic reservoir models compared to the simple grid-based model presented in the study. Pal (2021) employed long short-term memory (LSTM) Networks (Hochreiter and Schmidhuber, 1997) to train a model on historical data from injection and production wells to forecast oil and water production from the wells. However, a notable limitation of this method is the substantial amount of continuous production and injection data required to achieve high accuracy. Jo et al. (2021) proposed the use of GAN-ENKF to assimilate dynamic fluid data into reservoir models, where the generative adversarial network (GAN) facilitated the parameterisation of high-dimensional features to update ensembles while preserving geological features of the subsurface models. History matching using neural networks which have been shown to produce reservoir models that are reasonably calibrated to observed data can benefit greatly from the use of more realistic and computationally convenient surface-based geological models that have been conditioned to well data.

2.5 Reduced Order Modelling for Rapid Reservoir Simulation

Reduced order models (ROMs) are simplified mathematical or computational representations of complex, high-dimensional systems or models. ROMs are used to approximate the behaviour of these systems with a significantly reduced number of variables, making them computationally more efficient while maintaining a reasonable level of accuracy (Lucia et al., 2004). Often referred to as surrogate or proxy models, ROMs estimate the state of a system using low-order representations of the degrees of freedom in the high-fidelity model. ROMs are used in various scientific and engineering fields where detailed, high-fidelity simulations are computationally expensive or time-consuming. These models offer a significant advantage for the exploration of model controls and parameters, analysis and quantification of uncertainties, data assimilation, multiphysics and multiscale modelling, real-time simulations, and optimisation. To generate reduced-order models, snapshots or solutions from the full-physics model are projected onto a low-dimensional space to approximate the evolution of the full-order model. Various techniques exist for dimensionality reduction to identify the most significant parameters that capture the system's variability. In one of the early studies, Bai (2002) presented Krylov subspace techniques for reduced order modelling of linear, second-order and nonlinear dynamical systems. Gratton and Willcox (2012) developed reduced order models using trajectory piecewise linearisation (TPWL) techniques for flow through an active controlled supersonic diffuser with significant shock motion.

In geological reservoir simulation, the widely utilised approach involves using the Galerkin method to project the governing equations onto the reduced basis space generated by applying proper orthogonal decomposition (POD) (Chatterjee, 2000; Liang et al., 2002) on snapshots of the full order simulation where singular value decomposition (SVD) (Henry and Hofrichter, 1992) is performed on the state vectors obtained from snapshots of simulation runs at different time steps. The objective is to extract the singular vectors associated with the largest singular values, which represent the reduced-space projection of the model. Several researchers have applied POD to

develop ROMs of subsurface systems. Cardoso and Durlofsky (2010) applied POD and piecewise linearisation in TPWL for the development of reduced-order models for simulating the thermal behavior of unconventional reservoirs undergoing in-situ heating. In this approach, a reduced representation of states and Jacobian matrices from a small set of training runs, achieved through POD, is used to derive new solutions. However, a notable limitation of projection-based methods using POD is their inability to adequately represent highly nonlinear physical systems, such as advection-dominated flows and problems with sharp gradients, which are frequently encountered in subsurface reservoir simulations. Although various techniques, including the discrete empirical interpolation method (DEIM) (Xiao et al., 2014), Petrov-Galerkin (Carlberg et al., 2017; Choi et al., 2020), adaptive algorithms (Carlberg, 2015), and shifted POD (Reiss et al., 2018), have been developed and applied to address these limitations, their success has often been specific to individual cases or situations where prior knowledge of the system is used to devise tailored solutions that enhance the subspace representation of the full-order simulations.

2.5.1 Data-Driven and Deep Learning Low Order Models for Subsurface Reservoirs

These approaches are particularly valuable in fields where it is difficult or impossible to derive governing equations to describe the real-world processes observed in systems, for example, when considering human behaviour and social interactions. Data-driven methods, such as non-intrusive reduced order models (NIROMs), have found application in creating simplified models for gas reservoirs (Samuel and Muggeridge, 2022). These NIROMs can run at least ten times faster than their high-fidelity counterparts, and they achieve this without necessitating modifications to the high-fidelity model's underlying code (Xiao et al., 2017). NIROMs are developed through a data-driven process, employing techniques such as proper orthogonal decomposition (POD) for dimensionality reduction and radial basis functions (RBF) for interpolating the reduced training data snapshots (Kostorz et al., 2020). NIROMs have been instrumental in generating low-order models that can operate orders of magnitude faster than their high-fidelity counterparts. This speed-up is achieved without the need for any alterations to the code of the high-fidelity model, although Kostorz et al. (2021) reported that the application of NIROMs is problem dependent and non-smooth behaviours characterised by a propagating shock such as immiscible displacements in two phase porous media are difficult to model accurately. A comprehensive review of NIROMs is available in Yu et al. (2019), and a brief overview of reduced order modeling in reservoir simulation is provided in Tang et al. (2021).

The development of ROM with neural networks have found widespread application in subsurface reservoirs, particularly in the context of hydrocarbon reservoir modelling. Canchumuni et al. (2017) developed a parameterisation of geological models by training a deep learning network using an ensemble of facies models to identify the essential features that are utilised to construct a reduced order model that follows a Gaussian distribution

which is applied in an ensemble smoother with multiple data assimilation for history matching. Applications of ROMs in reservoir simulation include: variational autoencoders (VAEs) and gated recurrent units (GRUs) for constructing ROMs of 3D reservoir models (Temirchev et al., 2020), convolutional and recurrent neural networks for predicting dynamic pressure and saturation maps and well rates in a history matching framework (Tang et al., 2020), image-to-image regression with a convolutional encoder-decoder network for ROMs of carbon storage process-based multiphase flow (Mo et al., 2018), deep residual recurrent neural networks (DR-RNN) with POD-Galerkin and DEIM for reduced model orders for uncertainty quantification (UQ) tasks (Kani and Elsheikh, 2019), conditional deep convolutional generative adversarial networks (cDC-GAN) for modelling fluid saturation in waterflooding (Zhong et al., 2020). A review of the application of ROMs in exploration, drilling and reservoir simulation is presented in Bahrami et al. (2022).

2.5.2 Simulation of Geothermal Reservoirs with Reduced Order Modelling

Direct numerical simulations provide high-fidelity models that closely emulate these physical behaviours and real-world processes but are time consuming. Simulators estimate reservoir behaviour by using numerical methods to solve differential and algebraic equations that describe the flow of fluids in porous media, derived from the integration of mass, momentum and energy conservation with thermodynamic equilibrium that describe the flow of fluids in porous media such as geothermal reservoirs (Garcia-Gutierrez et al., 2002). The output of a reservoir simulator is typically the spatio-temporal distribution, within the model domain, of variables with space- and time-dependent characteristics such as pressure, temperature, flow rate, saturation, or other rock and fluid properties.

Reservoir simulation combines principles from physics, mathematics, reservoir engineering, geoscience, and computer programming to model reservoir performance under various operating strategies, and in accordance with reservoir properties and conditions. This improves our understanding and analysis of the subsurface to aid in the design and optimisation of fluid recovery. Simulating the behaviour of geothermal reservoirs under different scenarios and conditions requires solving a computationally demanding numerical problem. This is due to the large-scale, nonlinear and high-dimensional behaviours associated with modelling fluid flow in the subsurface.

Despite ongoing research in geothermal reservoir simulation, the application of reduced order models (ROMs) has primarily relied on conventional methods like proper orthogonal decomposition (POD) (Li et al., 2020), TPWL (Cardoso and Durlofsky, 2010), and regression-based model reduction techniques (Mudunuru et al., 2017). While neural networks have found many applications in various aspects of geothermal systems, such as exploration, drilling, well engineering, and reservoir characterisation (Okoroafor et al., 2022), their application in the context of simulating geothermal reservoirs has been primarily limited to the simulation of wellbore

and reservoir pressures and temperatures (Ariturk, 2018; Aydin et al., 2020; Porkhial et al., 2015; Shi, Song and Song, 2021). Since geothermal resources in sedimentary basins exist in geological settings similar to oil and gas reservoirs, this research adapts methods developed for oil and gas reservoirs to geothermal resource development.

2.6 Research Questions

I have presented an overview of data assimilation in subsurface modelling mainly in oil and gas reservoirs for history matching at the beginning of this chapter. I have also highlighted some limitations of conventional data assimilation methods, particularly the time and computational requirements, which limit applications for rapid reservoir management. Recent techniques for data assimilation with neural networks have shown promising results. This inspires the adoption of a neural network-based approach to the key phases of subsurface reservoir modelling and development. On the basis of these, the following research questions and contributions arise:

Question 1: *Can we use machine learning to condition surface-based geological models to well data and generate geologically plausible models?*

Limited or sparse subsurface data and high uncertainty about the behaviour of subsurface reservoirs motivate the use of geological models that capture geological realism (Jackson et al., 2015; Jacquemyn et al., 2019) and can be conditioned on well data at reduced computational cost. In Chapter 3, I propose the conditioning of surface-based geological models using neural networks for inverting for the inputs to the geological mode. Surface based geological models can be rapidly generated, easily manipulated, and require fewer parameters to model subsurface reservoirs. This motivates the use of neural networks for well conditioning by leveraging the efficient techniques in deep learning for parameter optimisation. In Chapter 4, I extend this method to condition SBGMs using data from wells with a range of varied trajectories such as deviated, horizontal, and multilateral wells. This demonstrates the capabilities of the method for real field applications and the advantage of the technique over grid-based modelling approaches.

Question 2: *Given a well-calibrated static model, can neural networks be applied for assimilating dynamic flow data into reservoir simulation models for history matching?*

In Chapter 5, I adopt a data-driven approach to solving the history matching problem where data assimilation is performed using machine learning to iteratively calibrate the prediction from reservoir models to observed fluid flow data. This is similar to the iterative ensemble smoothing methods of data assimilation (ES-MDA) (Emerick and Reynolds, 2013). The static models for reservoir simulation are derived from the conditioning process developed in Chapter 3. The use of data-driven methods and, particularly, machine learning, has been thoroughly explored in reservoir history matching over the past 20 years. Integrating machine learning in data

assimilation provides capabilities for more robust data assimilation (Buizza et al., 2022), and, by extension, history matching in reservoir simulation.

Question 3: *Can state-of-the-art deep learning models be applied for generating reduced order models of geothermal reservoirs for reservoir forecasting?*

In Chapter 6, I propose an approach to generate reduced-order models of geothermal reservoirs based on neural networks. Advances in machine learning have enabled the application of neural networks for learning and predicting complex dynamics at reduced computational costs. The use of SBGMs, which provide geological realism and capture complex reservoir geometries, motivates the use of unstructured meshes to capture key flow characteristics in reservoir simulation models. Convolutional autoencoders with space filling curves (SFC-CAEs) (Heaney et al., 2024) are applied to compress data on unstructured meshes by exploring structural and spatial correlations, effectively reducing dimensionality. The transformer (Radford and Narasimhan, 2018), a state-of-the-art deep learning model in natural language processing (NLP), is applied for state prediction of geothermal reservoirs to generate reduced-order models of geothermal reservoirs in geological settings similar to oil and gas reservoirs.

The developed methods are integrated into a framework for geological modelling, history matching, and reservoir forecasting with neural networks, which is the final contribution of this research in Chapter 7.

Chapter 3

Conditioning Surface-Based Geological Models (SBGMs) to Well Facies Data using Artificial Neural Networks

3.1 Abstract

Surface-based modelling provides a computationally efficient approach for generating geometrically realistic repare theentations of heterogeneity in reservoir models. However, conditioning surface-based geological models (SBGMs) to well data can be challenging because it is an ill-posed inverse problem with spatially distributed parameters. To aid fast and efficient conditioning, we use here SBGMs that model geometries using parametric, grid-free surfaces that require few parameters to represent even realistic geological architectures. A neural network is trained to learn the underlying process of generating SBGMs by learning the relationship between the parametrised SBGM inputs and the resulting facies identified at well locations. To condition the SBGM to these observed data, inverse modelling of the SBGM inputs is achieved by replacing the forward model with the trained neural network and optimising the network inputs using the back-propagation technique applied in training the neural network. An analysis of the uncertainties associated with the conditioned realisations demonstrates the applicability of the approach for evaluating spatial variations in geological heterogeneity away from control data in reservoir modelling. This approach for generating geologically plausible models that are calibrated with observed well data could also be extended to other geological modelling techniques such as object- and process-based modelling.

3.2 Introduction

Generating realistic geological models that match observed well data is important for understanding and predicting subsurface reservoir behaviour. By combining prior geological knowledge of sedimentary and structural processes, stratigraphy, and facies distributions and associations, with data from wells, geoscientists can generate representative subsurface geological models that honour the sparse data available.

Surface-based geological models (SBGMs) generate representative geologic heterogeneities by applying hierarchical or rule-based relationships to control temporal or spatial positioning of surfaces that represent geological features. Structural, stratigraphic, and diagenetic heterogeneities are modelled as discrete geological domains bounded by these surfaces. Within these domains, petrophysical properties such as porosity and permeability are homogeneous, and further property modelling is not required (Jackson et al., 2015). Previous work has shown that capturing the geometry, spatial distribution, and connectivity of these domains is key to capturing subsurface flow (Osman et al., 2021). The use of parametrised surfaces allows modelling of heterogeneity free from any underlying grid, providing greater flexibility for the creation and manipulation of complex, realistic geological geometries at low computational cost. For flow simulation, unstructured meshes are used to discretise space, which more accurately preserve the modelled geometries and key flow features (Jackson et al., 2015; Salinas et al., 2018). However, although there are significant benefits to the use of SBGMs, conditioning these models to observed data at wells remains a challenging ill-posed inverse problem and a barrier to widespread practical application.

Few studies on conditioning SBGMs to well data have been carried-out. In one of the earliest approaches, Karssenberg et al. (2001) conditioned 3D alluvial-architecture process-imitating models using Monte-Carlo simulation in a trial-and-error approach. Pyrcz et al. (2005) generated SBGMs of deepwater lobe deposits that honour event deposit geometry and well contacts by local adjustments of the lobe bounding surfaces. The study presented an approach in which a stochastic residual accounted for fluctuations in the bounding surfaces and is conditioned to well data. Using detailed interpretations of well logs to identify and order depositional events, Michael et al. (2010) conditioned SBGMs generated using geologic process models to well data. Zhang et al. (2009) used a dual-spline error surface interpolation and a surface acceptance criteria to constrain SBGMs to data in a large number of wells. The computational expense of generating and conditioning the SBGMs, time requirements for interpreting well correlations, and the large number of wells required for efficiently conditioning these models in some cases have limited their application. Using an inverse modelling method of the channel migration (reverse migration), Parquer et al. (2017) conditioned channelised systems represented with parametric surfaces to observed channel path and belt structures from satellite and seismic data. To extend the method to subsurface reservoirs deposited in fluvial or deepwater environments, the authors suggest conditioning to well data.

Conditioning SBGMs to well data has been explored using optimisation techniques and probabilistic methods. Using inverse modelling to sequentially optimise input parameters controlling global and local spatial variations, Bertoncello et al. (2013) conditioned SBGMs to well log and thickness data. The speed of convergence and quality of fit to the data were improved by significantly reducing the dimensionality of the problem (from 103 to 37 parameters to be optimised for the case considered). However, since the method can significantly modify the model geometry, which results in high degrees of uncertainty, its applicability is limited for cases with sparse data. Although the uncertainty on input parameters was quantified and a sensitivity analysis on leading parameter uncertainty was performed, there is a need for uncertainty assessment of the conditioned realisations as suggested by the authors to evaluate the accuracy and applicability of the method. Wang, Pyrcz, Catuneanu and Boisvert (2018) also applied optimisation techniques to condition parameterised object-based models. The optimisation used an implicit filtering algorithm based on coordinate search and gradient approximation to minimise the cost function that measures the level of conditioning. The difficulty in parameterising the objects presents a limitation of the method, as this could impact the computational expense and time cost of solving the inverse problem, due to the non-uniqueness and high-dimensionality of the problem. To generate multiple lobe models that fit observed data at well locations, Wingate et al. (2016) applied a variational inference method that used a stochastic gradient optimiser to determine the optimal parameters of a surrogate probability density function (PDF) that approximates the posterior PDF. The method allowed for the rapid generation of multiple conditioned samples when compared to a Markov Chain Monte Carlo (MCMC) method, although the MCMC yielded models that better fit the data.

The use of neural networks has been explored with remarkable success in the forward and inverse modelling of complex systems and physical processes in geoscience. Dramsch (2020) gives a review of the historic application of neural networks in geoscience. In subsurface reservoir applications, neural networks have become increasingly popular for representing complex nonlinear functions, particularly for lithofacies classification (Brcković et al., 2017; Ma et al., 2009; Maiti et al., 2007), seismic interpretation and inversion(Fu et al., 2019; Liu et al., 2015; Wang, Di, Shafiq, Alaudah and AlRegib, 2018), and generating realisations that honour spatial observations (Chan and Elsheikh, 2017; Dupont et al., 2018; Laloy et al., 2018; Mosser et al., 2020). In geological modelling, neural networks have also been used to predict stratigraphic units by interpreting borehole data as oriented spatial sequences or series of data that are processed using Recurrent Neural Networks (RNNs) (Zhou et al., 2019). By providing a parameterisation for the direct generation of conditional realisations obtained using generative adversarial networks (GANs), Chan and Elsheikh (2019) produced plausible realisations of binary channelised images while achieving dimensionality reduction of two orders of magnitude. Jo et al. (2020) explored the use of neural networks for conditioning rule-based models by training a GAN to extract the geological features of the models and applying semantic image inpainting to generate realisations of the rule-based models.

We present here a method based on neural networks to condition SBGMs to well data. The use of parametrised,

surface-based models to describe even complex geological features significantly decreases the computational expense of training the network compared to grid-based models as there are fewer parameters to optimise. The use of neural networks for generating conditioned, grid-free, surface-based models could also be extended to grid-based, object- and process-based modelling. Compared to existing conventional geostatistical or other rule-based methods, our approach has more flexibility for capturing geological realism without reference to a predefined grid. We use parametric non-uniform rational B-splines (NURBS) which enable the generation of models that better capture geological architectures and preserve geological features such as continuity of channelised bodies compared to grid-based approaches. Furthermore, we require few parameters to model geometries and describe heterogeneities as even complex geological features/bodies can be modelled using surfaces that define volumes within which petrophysical properties are homogeneous. Our workflow allows for the conditioning of these models to well data, specifically well log information such as facies type described in this study, whilst ensuring that the underlying process for generating these models such as the stacking patterns is preserved and constraints posed by using gridding or other resolution-dependent approach are avoided. Thus, the proposed method ensures the preservation and representation of subsurface heterogeneity while conditioning to well data for subsurface models. The use of surface-based modelling to describe geometrically realistic geological features compared to grid-based modelling could significantly decrease the computational expense of training the network for more complex geologies or 3D models as there are fewer parameters to optimise. Figure 3.1 illustrates the framework for our proposed methodology for generating realistic geological realisations that match observed well data. Using this approach, we 1) train a neural network to learn the geological patterns and underlying phenomenon for generating the SBGMs, 2) generate plausible geological models conditioned to data at well locations by using inverse modelling to optimise the inputs to the neural network, and 3) perform uncertainty analysis of the conditioned realisations.

Our methodology is illustrated with synthetic data as a proof-of-concept. The example case represents a reservoir that was deposited in a shallow marine environment. We use the neural network to generate inputs for surface-based models that fit the well data at multiple well locations. The optimised inputs describe the locations, orientation, and distribution of the depositional bodies within the domain. In the following section, a description of the surface-based model considered is given, along with an illustration of the process for generating SBGMs. Section 3.4 describes the procedure for training the neural network to represent the SBGM modelling process. Section 3.5 explains the framework for the optimisation methodology using the trained neural network. The results of the neural network prediction are presented in Section 3.6, along with examples of conditioned realisations using the proposed methodology. An assessment of the uncertainty associated with the results and the proposed methodology is also provided in Section 3.6.3 and the concluding remarks are presented in the final section.

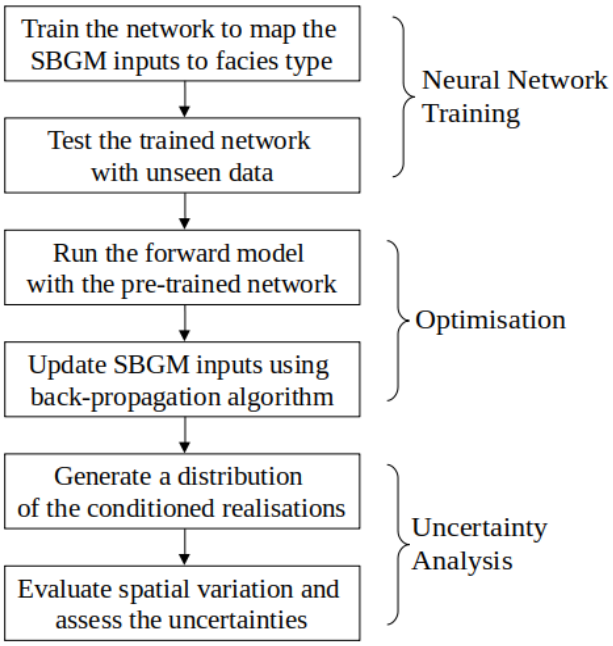


Figure 3.1: The approach for conditioning surface-based geological models applied in this study. The workflow involves training the neural network to represent the forward model, optimising the network inputs to generate conditioned SBGMs and assessing the uncertainty associated with these realisations.

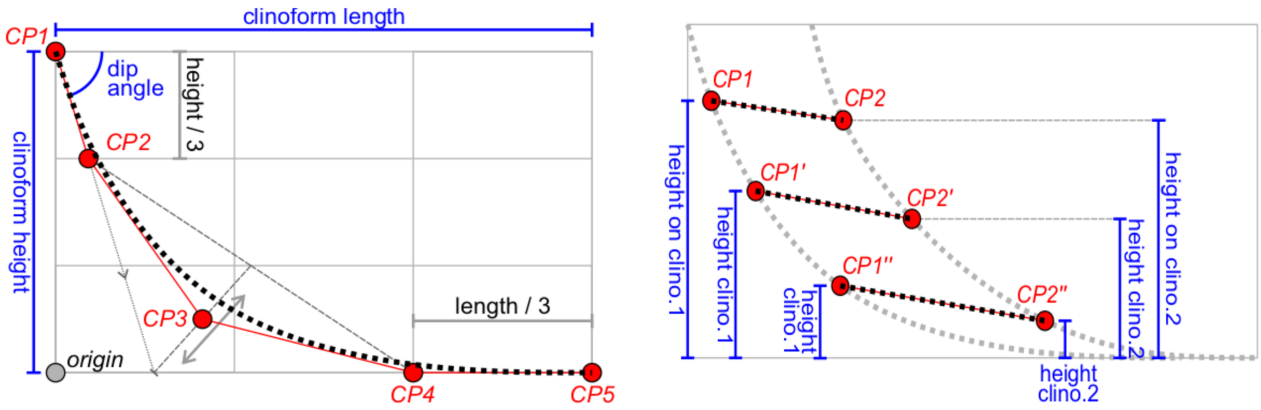
3.3 Surface-Based Geological Model Description

Surface-based geological models are forward models in the sense that these models predict depositional sequences or generate geological architectures and features given a set of input parameters. SBGMs can be deterministic and/or stochastic, empirical and/or theoretical (Karssenberg et al., 2001), and provide realistic representations of subsurface structures (Jackson et al., 2014). We use SBGMs that model geometries with parametric, grid-free surfaces using non-uniform rational B-splines (NURBS) interpolation (Piegl and Tiller, 1997) based on the work of Jacquemyn et al. (2019).

3.3.1 NURBS for Generating Computationally Efficient SBGMs

In this approach, geological architectures are described using parameterised inputs that define geological geometries at multiple levels of surface hierarchy. The resulting surfaces represent realistic geometries and determine the distribution of heterogeneity as the surfaces define discrete rock volumes within which petrophysical properties are assumed to be uniform. If heterogeneity exists within these volumes, they are further subdivided by additional surfaces. In these geologically meaningful domains, no further property modelling is required (Jackson et al., 2015; Osman et al., 2021). The application of NURBS provides a computationally efficient representation, as few control points are required to represent even complex curves and surfaces. The efficiency of generating and manipulating SBGMs using this approach allows for the rapid generation of multiple realisations

of realistic reservoir models for training the neural network. Jacquemyn et al. (2019) gives details of generating SBGMs using the NURBS approach depicted in Figures 3.2a and 3.2b, where clinoforms and facies boundaries are constructed using the defined control points (CP).



(a) Clinoform template for the SBGMs

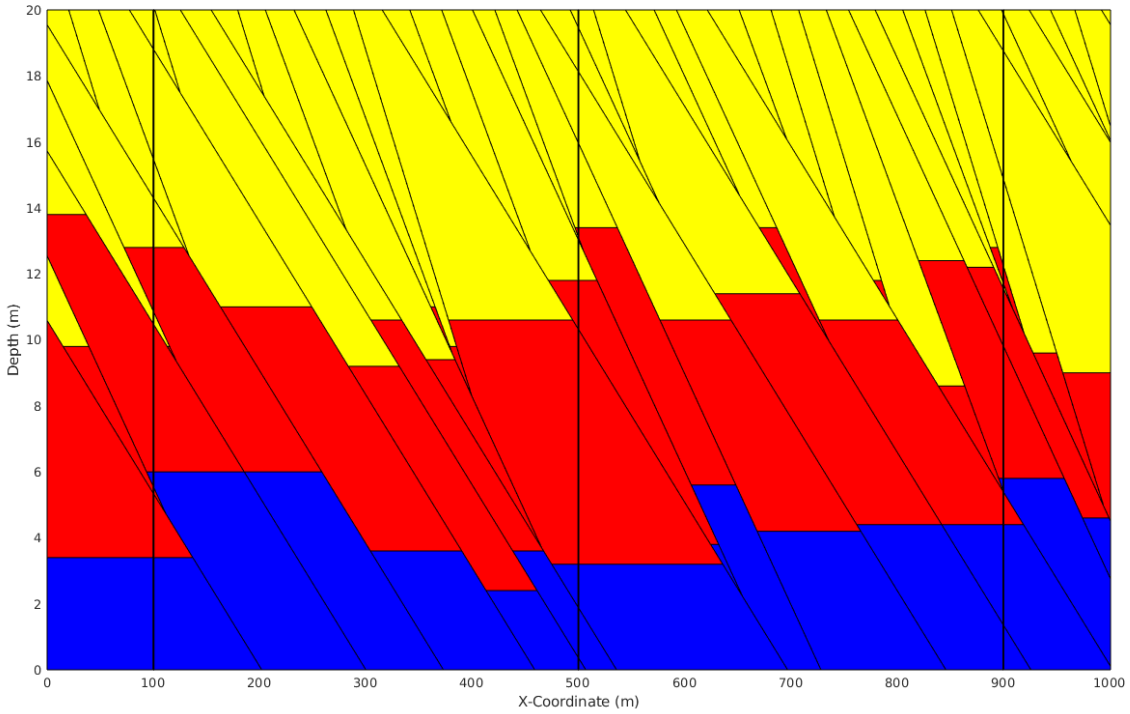
(b) Clinothem template for the SBGMs

Figure 3.2: a) The dipping clinoform template constructed using the height, length and dip angle, and b) facies boundaries inside a clinothem are constructed using variables relative to their bounding clinoforms, both from Jacquemyn et al. (2019).

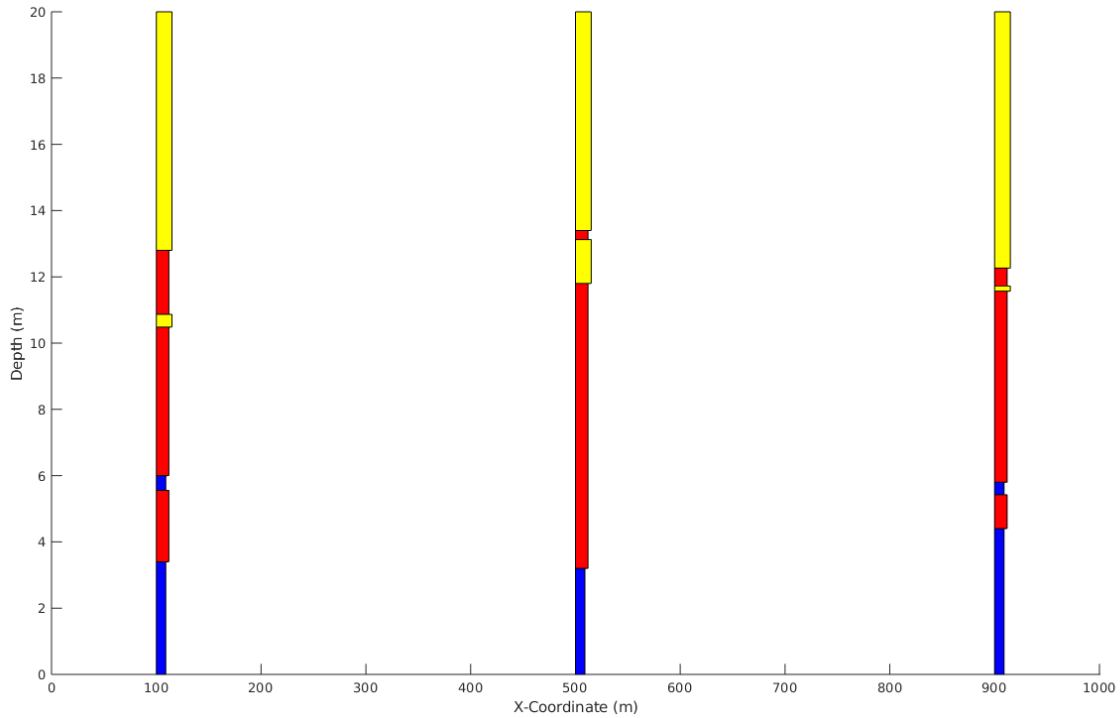
3.3.2 Shallow Marine Example Case

We demonstrate our approach here using a 2D model representing a shallow marine sedimentary succession, similar to what would be encountered in many subsurface reservoirs (Rønning and Steel, 1987; Sech et al., 2009). This is simplified to a single 20m thick parasequence containing dipping clinoform surfaces and horizontal facies boundaries (see Figure 11 of Jacquemyn et al. (2019)) with 1000m lateral extent. Three facies are considered: upper shoreface (USF; yellow), lower shoreface (LSF) and offshore mudstone (OM). The inputs to the SBGMs are derived from stochastically sampled variables from uniform distributions that parameterize the shape, orientation, and relative location of the surfaces (depositional or erosional), which are sequentially placed to represent clinoform surfaces and facies boundaries. Figure 3.3a shows the resulting 2D model, and Figure 3.3b illustrates the corresponding facies type (wells logs) at three well locations in the domain.

The clinoforms surfaces are parametrised using the dip angle of the surfaces and the parasequence thickness. The toe of each clinoform truncates at the parasequence base or underlying clinoform, depending on the dip. The spacing between clinoforms is determined at the top of the parasequence. Within each clinothem volume, two horizontal facies boundaries are placed to differentiate between different facies. The vertical relative position of the lower facies boundary (LSF-OM boundary) is measured from the parasequence base. The relative vertical position of the upper facies boundary (USF-LSF boundary) is measured from the lower facies boundary. In total 60 clinoforms are constructed to fill the parasequence; each clinothem is subdivided by two facies boundaries and a spacing so the model requires 240 input values (i.e. 60x clinoform dip angle, clinoform



(a) An example of the resulting 2D shallow marine model with one parasequence and three wells with 50x vertical exaggeration. The parameters for the SBGM were randomly sampled from Table 3.1



(b) The well log for three wells at x=100m, x=500m and x=900m from the SBGM given in Figure 3.3a. The three facies considered are the upper shoreface (yellow), the lower shoreface (red), and the offshore mudstone (blue).

Figure 3.3: Surface-based geological model and corresponding well logs at three well locations.

spacing, lower facies boundary position, upper facies boundary position). The clinoform surfaces are defined by length and dip angle (Figure 3.2a) while facies boundaries inside a clinothem are positioned relative to the parasequence boundaries (flooding surfaces) or underlying facies boundary (Figure 3.2b). Table 3.1 gives the range of parameter distributions sampled to generate the stochastic SBGMs.

Table 3.1: Range of the clinoforms and facies zones for the parasequences. The parameters all sampled from a uniform distribution

	Lower Limit (L_b)	Upper Limit (U_b)
Clinoform Shape		
Dip (°)	3.0	6.0
Clinoform Spacing (m)	20.0	30.0
Facies Zones		
Distal Lower Shoreface Base (%)*	10.0	30.0
Upper Shoreface Base (%)**	30.0	40.0

*% of parasequence thickness
**% of thickness above LSF base

3.4 Modelling SBGMs with ANNs

To condition SBGMs to well log data (facies classification), we use neural networks in two ways: 1) to map SBGM inputs to the facies type at well locations in the forward modelling step and 2) to generate optimised SBGM inputs using a backpropagation algorithm to update randomly selected SBGM inputs so that the resulting SBGM honours the observed facies type along the well. This conditions the SBGM to the observed data at the well location. A key benefit of our method is the adoption of a grid-free approach in the forward and inverse modelling steps. In contrast to other methods such as the use of convolutional neural networks for image generation, which have been applied in GANs to condition rule-based geological models (Jo et al., 2020), our method does not impose constraints or impact on the resolution of the models.

The flexibility of the approach we present allows the generation of the SBGM inputs which are consistent with the geological modelling process, where the underlying phenomenon for generating parametric grid-free SBGMs is preserved. With this approach, we are not artificially or mechanically changing the underlying process for generating the SBGMs but generating the inputs to the models that take into account the geological relations between surfaces. The proposed method ensures that the spatial continuity of the geological model is preserved, the features are well connected, and patterns such as the hierarchical arrangement of surfaces in the SBGM generation process remain unchanged. The method described here guides the conditioning of SBGMs generated using a parametric, grid-free approach while ensuring that any constraint posed by the model resolution does not impact the conditioning. Using this approach for more complex or realistic 3D models should also deliver these advantages. The forward modelling and inverse modelling steps are described in more detail in the following

sections.

3.4.1 Forward Modelling

We use a multilayer perceptron (MLP) with fully-connected layers to obtain a parameterisation of the surface-based modelling process in the forward modelling step. In this approach, we are given a training set comprising examples of the inputs, $\mathbf{p}$, and the outputs $\mathbf{y}$, with which to train the network. By training the neural network to represent the underlying process for generating SBGMs, we aim to find a set of parameters, $\boldsymbol{\theta}$, that allow the mapping of the SBGM inputs to the observations. Thus, the MLP is applied in a supervised learning mode where the goal is to determine $\boldsymbol{\theta}$ so that $\mathbf{y} = \mathcal{S}_{\boldsymbol{\theta}}(\mathbf{p})$, where $\mathcal{S}$ represents the neural network so that $\mathcal{S}_{\boldsymbol{\theta}} : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_y}$ and $\mathbf{y} \in \mathbb{R}^{n_y}$ is the property of interest or the outputs, i.e. the facies type at n_y discretised locations along the well.

The problem is formulated as a classification problem, where the facies type at locations along the well is predicted. The input is represented as a feature vector $\mathbf{p} = [p^{(1)}, \ldots, p^{(n_p)}] \in \mathbb{R}^{n_p}$, where n_p is the number of inputs to the surface-based model. The aim is to find the classification mapping from the input feature vector, $\mathbf{p}$ to the output vector $\mathbf{y}$ whose values are taken from a discrete set of n distinct classes, $\mathcal{C} = \{c^{(1)}, \ldots, c^{(n)}\}$, for all discretised locations along the well. The input feature vector is continuous and can take on non-discrete values, while the set of classes is discrete and non-continuous. For a given collection of training examples, $\mathcal{D}$, of M input-output pairs: $\mathcal{D} = \{(\mathbf{p}^m, \mathbf{y}^m) | \mathcal{S}_{\boldsymbol{\theta}}(\mathbf{p}^m) = \mathbf{y}^m\}_{m=1}^{M}$, we seek to find the parameters $\boldsymbol{\theta}$, which map the inputs of the surface-based model to the facies class or type at discretised locations along the well. Therefore, the inputs to the feed-forward network, $\mathbf{p}$, are the 240 variables that define the control points of the SBGM, while the outputs are the facies types, $\mathbf{y}$, at discretised locations along the well. In the following sections, we discuss a data normalisation applied to the data and the training procedure for deriving the parameters of the network.

Data normalisation

Since the input feature vectors span varying orders of magnitude, we perform a normalisation of the data set. This is necessary to remove any bias to feature vectors with large magnitudes. We standardise the input data to a normal distribution by scaling the input feature vector to $\tilde{\mathbf{p}} \sim \mathcal{N}(0, 1)$, by computing

$$\tilde{\mathbf{p}} = \frac{\mathbf{p} - \boldsymbol{\mu}}{\boldsymbol{\sigma}}, \tag{3.1}$$

where $\boldsymbol{\mu} = [\mu_1, ..., \mu_{n_p}]^T$ such that,

$$\mu_j = \frac{\sum_{m=1}^{M} p_j^m}{M}, \quad \forall j \in \{1, 2, \ldots, n_p\}, \tag{3.2}$$

where p_j^m represents the jth SBGM input parameter in the mth example and $\boldsymbol{\sigma} = [\sigma_1, ..., \sigma_{n_p}]^T$ such that

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{M} (p_j^m - \mu_j)^2}{M}}, \quad \forall j \in \{1, 2, \ldots, n_p\}. \tag{3.3}$$

One-hot encoding

A commonly applied approach for representing non-ordinal or categorical data for classification problems is one-hot encoding. Since a natural ordered relationship among facies types only exists within clinothems (volumes between two clinoforms), and not necessarily between neighbouring clinothems, one-hot encoding is required to represent the output data rather than label encoding where integer labels are used to encode variables, as this may result in poor performance or unrealistic results. Using the one-hot encoding, binary variables are used to represent facies types such that the output data at each discretised location, i, is given as a vector, $\mathbf{z}_i$ of equal length to the number of classes, n and whose entries are zero except where the class type corresponds to the facies type at that location:

$$z_i^f = \begin{cases} 0 & \text{for} f \in \{1, \ldots, n\}, f \neq l \\ 1 & \text{for} f = l, \end{cases} \tag{3.4}$$

where $c^{(l)}$ corresponds to the facies type or class at the discretised location. Figure 3.4 illustrates the well representation and one-hot encoding of the facies type, which can be interpreted as the probability of each facies type at the discretised locations. The facies type at the centre of each discretised location was selected to represent the facies type of the discretized location.

Training

For the multiclass classification problem, the network is trained to predict the probability of each facies type at a discrete location, $\hat{\mathbf{z}}_i$ using the softmax operator. The final output of the network is the vector of facies type with the highest predicted probability at all discretised locations along the well, $\hat{\mathbf{y}}$. To learn the parameters of the network, $\boldsymbol{\theta}$, we minimise the objective function, $\mathcal{J}_t$ which defines the mismatch, $\mathcal{L}$ between the one-hot encoded vector of facies type $\mathbf{z}_i$ and the predicted probabilities by the network $\hat{\mathbf{z}}_i$ at each discretised location along the well, given as:

$$\mathcal{J}_t(\boldsymbol{\theta}) = \mathcal{L}(\mathbf{z}_{i,m}, \hat{\mathbf{z}}_{i,m}), \tag{3.5}$$

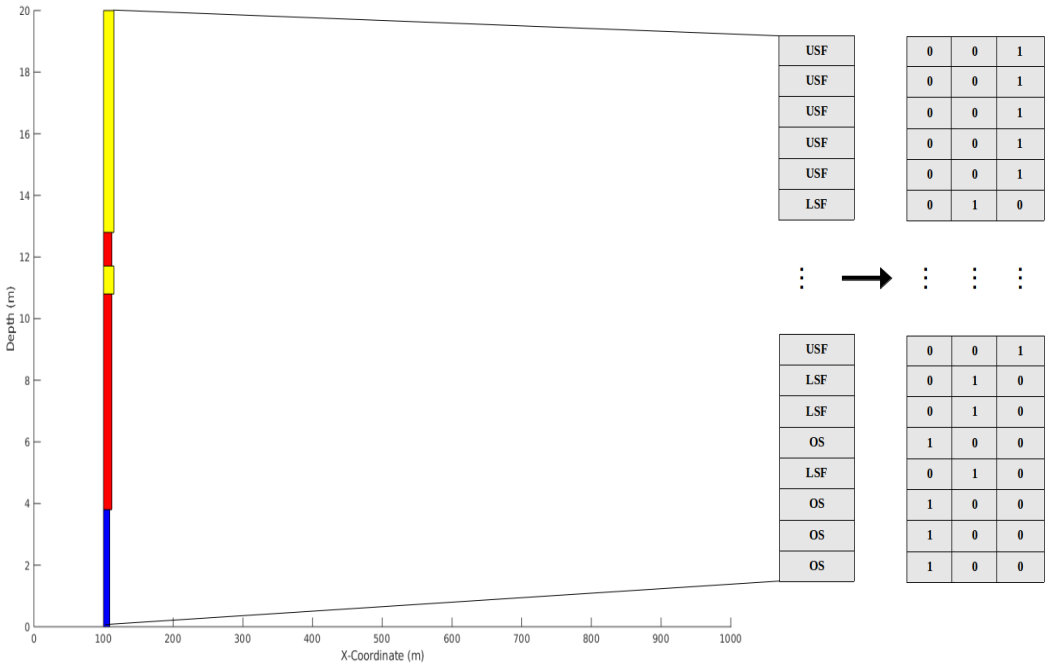


Figure 3.4: The output of the SBGM for training the network is the facies type typically derived from well logs. To represent this information in the network, one-hot encoding of the facies type for each grid along the well was used. Facies types: Upper Shoreface (USF), Lower Shoreface (LSF) and Offshore Mudstone (OM) deposits.

where $\boldsymbol{\theta}$ represents the trainable parameters of the network, i.e the weights and biases, and $\hat{\mathbf{z}}_i$ depends on $\boldsymbol{\theta}$ and the loss function, $\mathcal{L}$ is the multi-class cross-entropy loss function (Bishop, 2006) defined as:

$$\mathcal{L}(\mathbf{z}_{i,m}, \hat{\mathbf{z}}_{i,m}) := -\frac{1}{n_y M} \sum_{m=1}^{M} \sum_{i=1}^{n_y} \left(z_{i,m}^{l} \log(\hat{z}_{i,m}^{l}) \right) , \tag{3.6}$$

where l is such that $z_{i,m}^{l} \neq 0$. The trainable parameters are iteratively updated using the Adam optimiser (Kingma and Ba, 2014) and the final output of the model, $\hat{\mathbf{y}}$ consists of the facies with the highest predicted probability at each discretised location, where each entry in $\hat{\mathbf{y}}$ is

$$y_i = \underset{z_i}{\operatorname{argmax}}\, \hat{\mathbf{z}}_i. \tag{3.7}$$

Figure 3.5 illustrates the model training procedure and shows the architecture of the multilayer perceptron having 3 hidden layers.

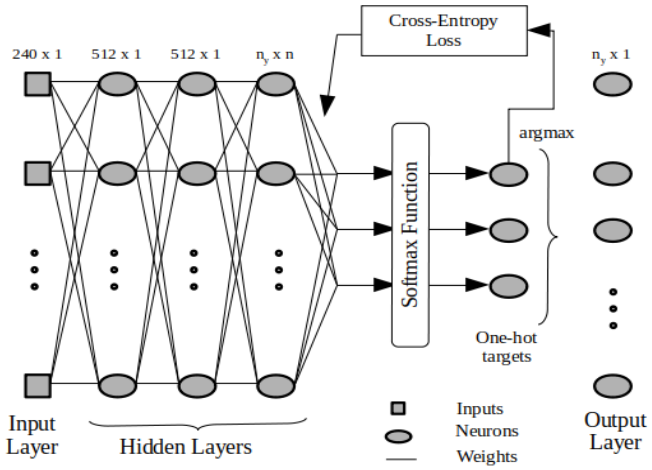


Figure 3.5: The neural network maps a vector representation of the 240 variables onto a discrete vector with n_y by n entries. The softmax function gives the probabilities of the facies type at each discretised location. To calculate the parameters of the network, a backpropagation of the mismatch between the predicted probabilities $\hat{\mathbf{z}}_i$ and the one-hot targets, $\mathbf{z}_i$ at each location is performed to update $\boldsymbol{\theta}$. The loss function is the cross-entropy loss used for the multiclass classification given in equation (3.6). The vector of facies type with the maximum probability at all discretised location, $\hat{\mathbf{y}}$ is returned as the output of the network.

3.5 Conditioning SBGMs with ANNs

3.5.1 Inverse Modelling

An efficient and robust approach to condition SBGMs is the use of an iterative method in an optimisation framework with neural networks where the network prediction, $\tilde{\mathbf{y}}$ is compared with the observed data, $\mathbf{y}_{obs}$ and the inputs, $\mathbf{p}$ are updated at each iteration in accordance with the data mismatch. This is achieved using a backpropagation algorithm (Rumelhart et al., 1986*a*) based on gradient descent schemes similar to those used in the training process, with the network trainable parameters now set to the inputs. Using this approach, we can learn the inputs to the neural network, $\mathbf{p}$, which based on the given parameters of the neural network, $\boldsymbol{\theta}$ will give the forward model having observations $\mathbf{y}_{obs}$ at the well location. The method is described in more detail in Section 3.5.2.

3.5.2 Optimisation Methodology

Given the trained network $\mathcal{S}$, we aim to find the set of input feature vectors $\mathbf{p}^*$ whose outputs from the network, $\tilde{\mathbf{y}}$, match a reference data set or observations, $\mathbf{y}_{obs}$. In other words, to condition the SBGM model to observations, we are to find the SBGM inputs whose predictions from the neural network, match the observed well data, i.e. find $\mathbf{p}^*$, such that $\mathcal{S}(\mathbf{p}^*) = \mathbf{y}_{obs}$. We formulate the problem as finding a $\mathbf{p}^*$ that minimises the

mismatch, $\mathcal{L}$, between the network prediction and the given observations,

$$\mathbf{p}^* = \underset{\mathbf{p}}{\operatorname{argmin}}\, \mathcal{J}_c(\mathbf{p}), \tag{3.8}$$

where $\mathcal{J}_c$ is the objective function to be minimized defined as:

$$\mathcal{J}_c(\mathbf{p}) = \mathcal{L}(\mathbf{z}_{i,obs}, \tilde{\mathbf{z}}_\mathbf{i}) := -\frac{1}{n_p}\sum_{i=1}^{n_p}\left(z^l_{i,obs}\log(\hat{z}^l_i)\right), \tag{3.9}$$

where $\mathbf{z}_{i,obs}$ is the one-hot encoding of the observed facies type at each discretized location, l is such that $z^l_{i,obs} \neq 0$, the mismatch or loss $\mathcal{L}$ remains the multi-class cross-entropy used in training the model and $\tilde{\mathbf{z}}_\mathbf{i}$ is the predicted facies type by the trained network at that location. The vector of facies with the highest predicted probability in all discretised locations, $\tilde{\mathbf{y}}$ is given as the network output, where each entry in $\tilde{\mathbf{y}}$ is

$$\tilde{y}_i = \underset{z_i}{\operatorname{argmax}}\, \tilde{\mathbf{z}}_\mathbf{i}\,. \tag{3.10}$$

To generate realisations conditioned to the observations, we solve equation (3.9) repeatedly starting with different random initial guesses. The mismatch of the facies type predicted at the well locations is compared with the reference well data and a back-propagation algorithm is applied to minimise the objective function, $\mathcal{J}(\mathbf{p})$ and update the input parameters, $\mathbf{p}$ based on the mismatch. A further evaluation of the generated SBGMs with the optimised inputs is performed by calculating a well correlation, $\text{Corr}_{\text{well}}$ between the observations at the well, $\mathbf{y}_{obs}$ and facies type of the SBGMs generated with the optimised inputs, defined as

$$\text{Corr}_{\text{well}} = \frac{n^*_y}{n_y}100\%, \tag{3.11}$$

where n^*_y is the number of discretized locations where the facies type at the well locations of the SBGM generated with the optimised outputs matches the observed data at the well. The correlations associated with each well are then averaged to obtain the mean well correlation for the model, $\overline{\text{Corr}}_{\text{well}}$ given as

$$\overline{\text{Corr}}_{\text{well}} = \frac{1}{n_w}\sum_{w=1}^{n_w}\text{Corr}_{\text{well},w}, \tag{3.12}$$

where n_w is the number of wells in the model.

The well correlation provides the proportion of facies type predicted by the conditioned realisations that exactly match the observed data, which can be interpreted as an accuracy metric. Thus, the network accuracy from the training process is used as a cut-off for determining conditioned realisations. The approach presented here (see Figure 3.6) provides a robust methodology for identifying the set of input parameters that generate the forward model, i.e. plausible SBGMs which honour the well data, or models that are conditioned to observations.

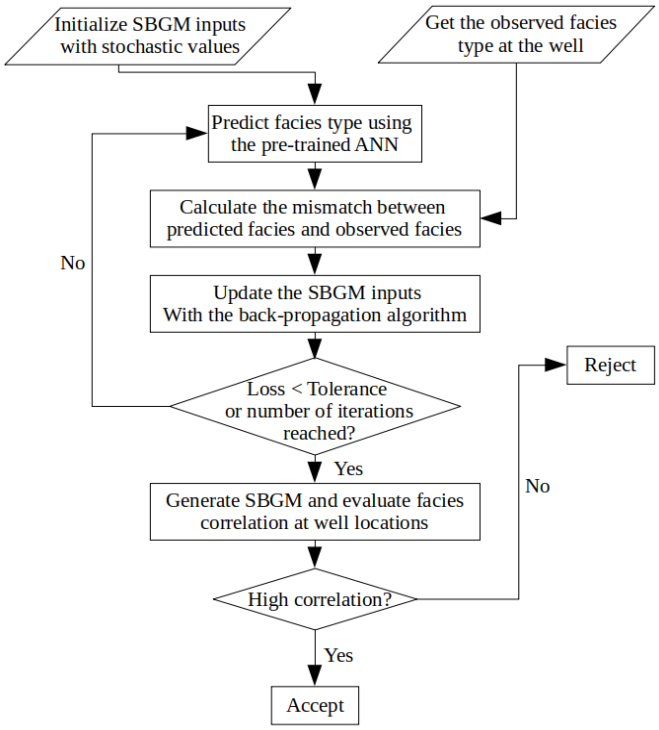


Figure 3.6: The workflow for conditioning the SBGMs using the trained neural network.

Since the process of generating SBGMs is not modified to match the observed data at the wells, the geological consistency of the model is preserved. In other words, the model only takes a stochastic set of SBGM inputs that produce a realistic geological model and applies it to produce an SBGM that matches the observed facies type along the wells. The workflow is integrated within the neural network framework (see Figure 3.6) and the algorithm for the methodology is presented in algorithm 3.1.

3.6 Results and Discussions

3.6.1 Forward Modelling: Well Log Prediction

The results are given here for the forward modelling using the neural networks. The number of examples in the dataset is 10^6 SBGMs each having 240 SBGM input parameters and 60 facies definitions for each well corresponding to 60 locations along each well. The number of discretised locations was selected based on the vertical resolution of gamma well logging tools of 0.3048m $\pm 5\%$ of leading industry wireline service providers (Schlumberger, 2015). Consideration was also given to the number of clinoforms used in generating the model, as this can be interpreted as the maximum number of geological volumes that intersect the well. A vertical resolution of 0.33m was used to obtain the 60 discretized locations along each well location, corresponding to a grid size of 1×0.3m.

Two networks with a similar architecture to Figure 3.5 were trained to predict the facies type along the well for

Algorithm 3.1 Back-propagation algorithm for conditioning SBGMs

initialization: get random input parameters, $\mathbf{p}^0 \in \mathbb{R}^{n_p}$; set counter, $k = 0$
initialization: define the optimizer; set network trainable parameters to $\mathbf{p}^0$
read $\mathbf{y}_{obs}$ and encode output data at each discretized location as $\mathbf{z}_{obs,i}$
scale $\mathbf{p}^0$ to a normal distribution, $\tilde{\mathbf{p}}^0 \sim \mathcal{N}(0, 1)$
while $k \leqslant K$ **do** ▷ K is the number of iterations
    calculate $\tilde{\mathbf{z}}_\mathbf{i}$ $\forall i \in \{1, \ldots, n_p\}$
    $\tilde{\mathbf{y}} = \mathcal{S}_{\boldsymbol{\theta}}(\tilde{\mathbf{p}}^k)$ ▷ $\tilde{\mathbf{y}} \in \mathbb{R}^{n_p}$
    $\mathcal{L}(\mathbf{z}_{i,obs}, \tilde{\mathbf{z}}_\mathbf{i})$
    $\frac{\delta \mathcal{L}}{\delta \tilde{p}_j^k} \forall j \in \{1, \ldots, n_p\}$
    $\tilde{p}_j^{k+1} := \tilde{p}_j^k - \eta \frac{\hat{m}}{\sqrt{\hat{v}} + \epsilon}$ ▷ update inputs with the Adam optimizer
    $\tilde{p}_j^{k+1} = \begin{cases} \tilde{p}_j^{k+1} & \text{if } L_b \leqslant \tilde{p}_j^{k+1} \leqslant U_b \\ L_b & \text{if } \tilde{p}_j^{k+1} < L_b \\ U_b & \text{if } \tilde{p}_j^{k+1} > U_b \end{cases}$
    ▷ $\sim \mathcal{U}(L_b, U_b)$, $\mathcal{U}$ is a uniform distribution
    $\tilde{\mathbf{p}}^* = \tilde{\mathbf{p}}$
    $k = k + 1$
return $\tilde{\mathbf{p}}^*$

SBGMs having two and three wells, hereafter referred to as SurfNet2 and SurfNet3. For SurfNet2, there are 120 outputs (2 wells each having 60 discretised locations); for SurfNet3, there are 180 outputs (3 wells each having 60 discretized locations). The inputs to both networks are the 240 SBGM parameters, and the outputs are the facies types at the discretised locations along each well in the model. The same dataset and hyperparameters were applied to train both networks. One-third of the dataset was used as the test set, the remaining two-thirds was used for training the network. Figure 3.7 shows the loss and accuracy learning curves for SurfNet2 for 100 epochs. The loss is the mismatch calculated in equation (3.6) and the accuracy metric gives the average percent of facies types at the well correctly predicted for all examples. The performance on the training and test set with increasing number of examples was applied to determine the number of examples to use, and also to ascertain that the training and test datasets are representative. It can be seen from Figure 3.7 that the dataset containing 10^6 examples shows significant improvements in the test accuracy and loss, whilst maintaining a high accuracy for the training set after only 100 epochs. This indicates that the network generalises well and is not under- or over-fitting the data. Therefore, a model trained with 10^6 examples is capable of predicting unseen data with high accuracy. Table 3.2 shows the hyperparameters and performance metrics of both networks. From Figure 3.8, both networks show comparable accuracy. Although SurfNet2 is slightly more accurate, it can be argued that from the loss plots for SurfNet3, the network is capable of further improvements with more epochs.

In SurfNet2, the network is able to predict the facies along the well with $\sim 99\%$ accuracy for samples in the training set and $\sim 98\%$ accuracy for samples in the test set, which are unseen data not used in the training. The trained MLP is able to represent the SBGM forward model as it is capable of emulating the patterns of the model and identifying facies along the well reasonably. The predictions of the facies type at well locations for unseen samples (SBGM examples not included in the training) is shown in Figures 3.9 and 3.10 for SurfNet2 and SurfNet3 respectively. Although the neural network inherently does not make assumptions about the underlying

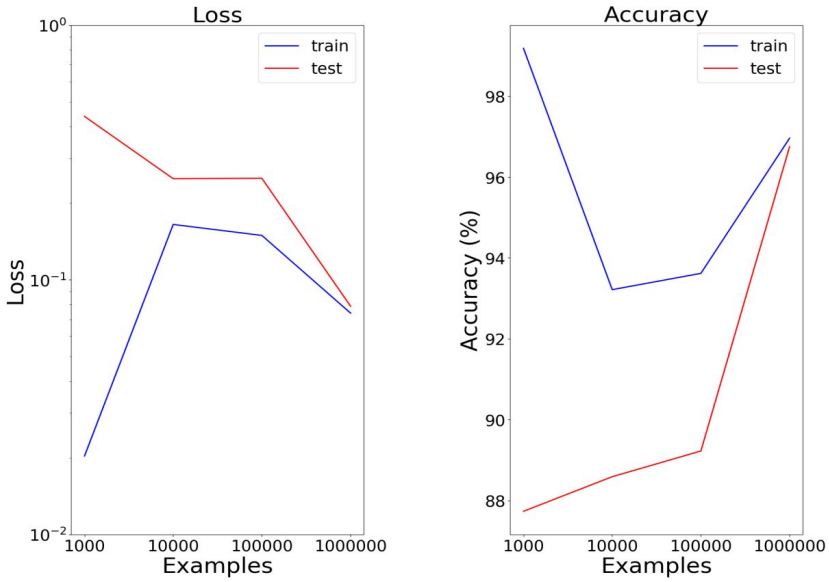


Figure 3.7: The learning curve for SurfNet2 with 100 epochs. The learning curves of the loss and accuracy provide information on the model performance on the training and test datasets, and are used to diagnose an underfit, overfit, or well-fit model, and also the optimum number of examples. The comparable performance of the training and test examples for 10^6 SBGMs shows the model is well-fit with that dataset.

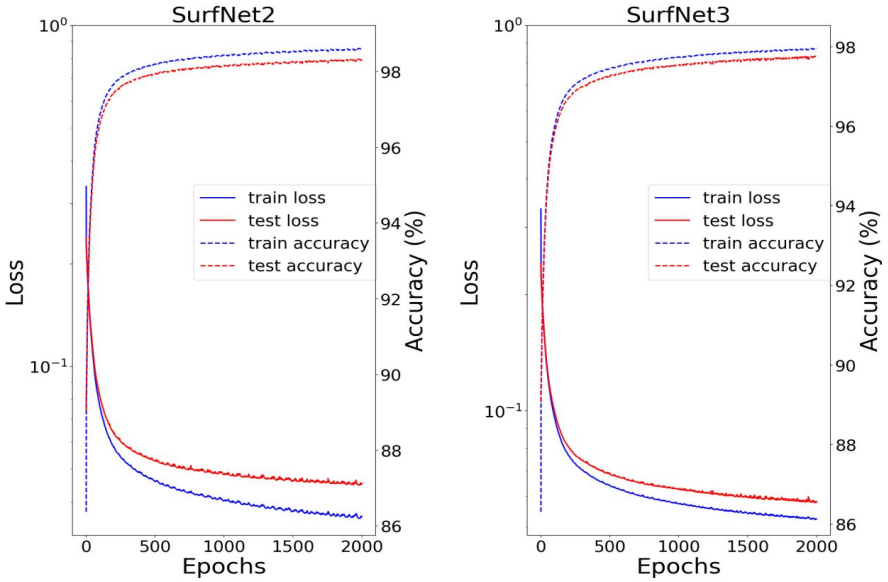


Figure 3.8: The loss and accuracy plots for SurfNet2 (left) and SurfNet3 (right). The loss and accuracy in both networks, SurfNet2 and SurfNet3, show comparable performance.

Table 3.2: Hyperparameters used for the training and the performance metrics of the model evaluated for the training set

	SurfNet2	SurfNet3
Hyperparameters		
Learning rate	0.001	0.001
Weight decay	0.0	0.0
Train-test split	0.67	0.67
Number of epochs	2000	2000
Weight initialisation	Kaiming	Kaiming
Optimizer	Adam	Adam
Metrics		
Final Training Loss	0.03617	0.05242
Final Training Accuracy (%)	98.588	97.944
Final Test Loss	0.04514	0.05822
Final Test Accuracy (%)	98.312	97.735
Training time on GPUs (hrs)	5.05	5.25

physical phenomena of modelling SBGMs, but rather selects parameters according to the characteristics of the data, it can be seen from the results that the model is able to generalise well as it predicts well logs for unseen data with high accuracy in both networks.

In the case of this 2D model, we have demonstrated that our approach for generating SBGMs using NURBS allowed for the rapid generation of 10^6 examples to train the network to a very high degree of accuracy as each SBGM can be generated in less than 2s, and this can be done in parallel. Although this might not be practical for more complex and realistic models, Figure 7 illustrates that it could be possible to reduce the number of models by a factor of two at a reduced accuracy of only about 2%. In addition, training the network with 10^5 examples gives accuracy of 90%, which is sufficient for obtaining practicable results in the prediction step. However, to demonstrate and ascertain the validity of the approach for the conditioning step, it was necessary to ensure a high degree of accuracy is obtained in the forward modelling step, justifying the use of 10^6 examples. Nonetheless, it is also possible to explore other network architectures and hyperparameter combinations for 10^5 training examples, that would not sacrifice the accuracy of the network as this is also required to aid good performance in the conditioning step.

3.6.2 Inverse Modelling: Generating the SBGM Inputs

Using the trained networks, SurfNet2 and SurfNet3, we performed the inverse modelling. A stochastically determined set of SBGM inputs was generated to produce observed data at the wells, which we refer to as the reference. The set of SBGM inputs that generated the reference data was not used in training or testing the network. The workflow presented in Section 3.5.2 was used to generate the optimised SBGMs. We use a maximum of 10^4 iterations for the optimisation or a tolerance of 10^{-4}, which is over two orders of magnitude lower than the neural network training loss. Table 3.3 gives a summary of the hyperparameters used for

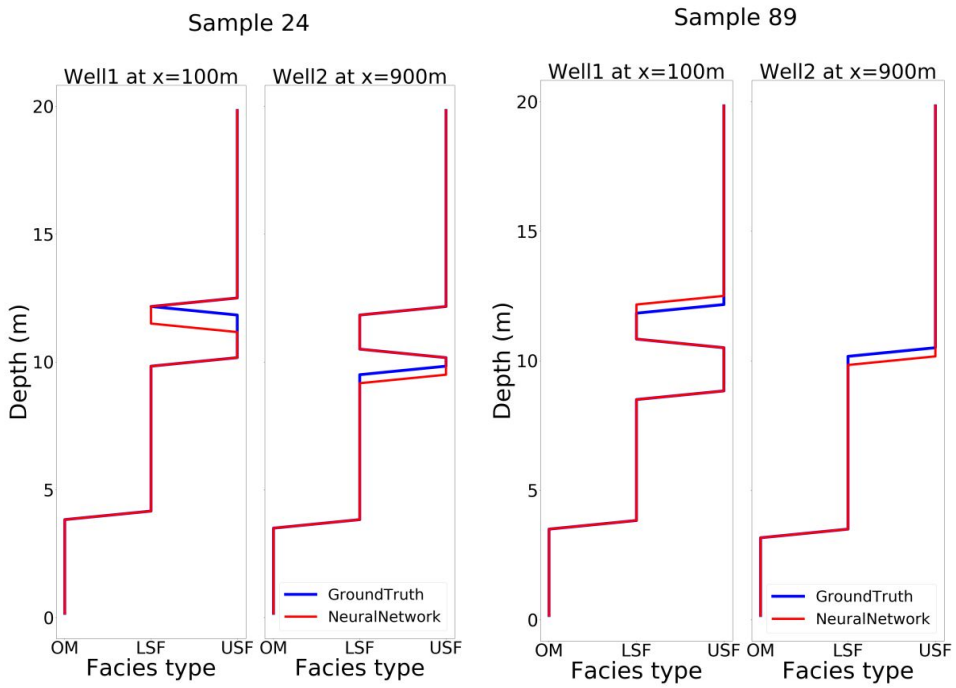


Figure 3.9: A comparison of the well logs from the SBGM (referred to as the GroundTruth in the figure) with the neural network prediction for 2 samples not included in the training examples. The results are shown here for SurfNet2. The neural network prediction of the facies type along the well shows a reasonable match with the SBGM.

optimising the SBGM inputs using the trained networks.

Table 3.3: Hyperparameters for optimising the SBGM inputs using the trained neural networks

	SurfNet2	SurfNet3
Learning rate	0.001	0.001
Weight decay	0.00	0.00
Maximum iterations	10^4	10^4
Tolerance	10^{-4}	10^{-4}
Optimizer	Adam	Adam
Well correlation cut-off (%)	98.0	97.5

Figures 3.11a to 3.14b compare the SBGM of the initialised, reference, and optimised neural network inputs from the inversion using SurfNet2 and SurfNet3 for two examples. It can be observed that the conditioned neural network geological models do not match the reference SBGM at all locations. However, a high correlation between the facies type from the SBGM derived from the neural network and that from the reference model is observed at the well locations (see Figures 3.11b, 3.12b, 3.13b and 3.14b). In Examples 1 and 2, the value of $\overline{\text{Corr}}_{\text{well}}$ is 98.3% for SurfNet2 and 97.8% with SurfNet3.

Figures 3.16a to 3.17b compare the well data of the SBGMs at different stages of the optimisation (0, 100, 500, 2000 and 10000 iterations) with the reference (observed) data. It can be seen that the initial guess presents a poor match with the reference data. The quality of the match improves as the number of iterations increases.

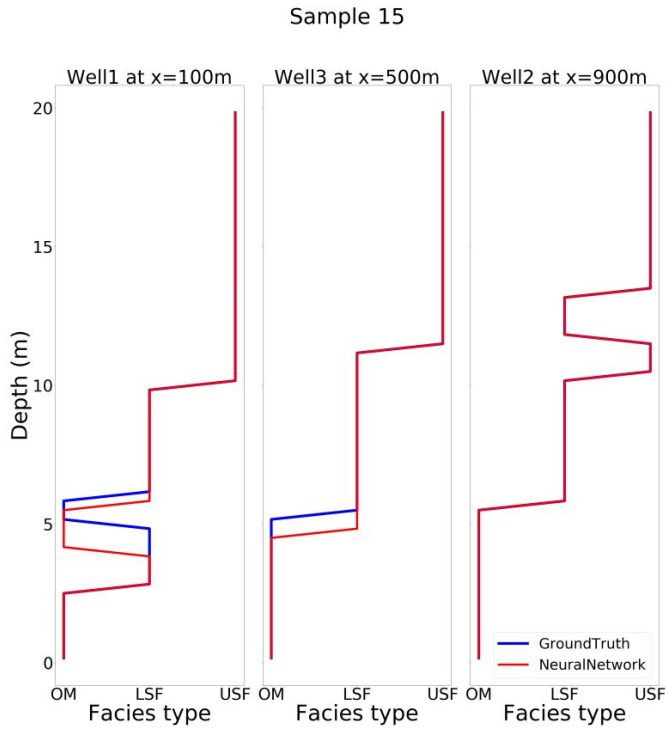
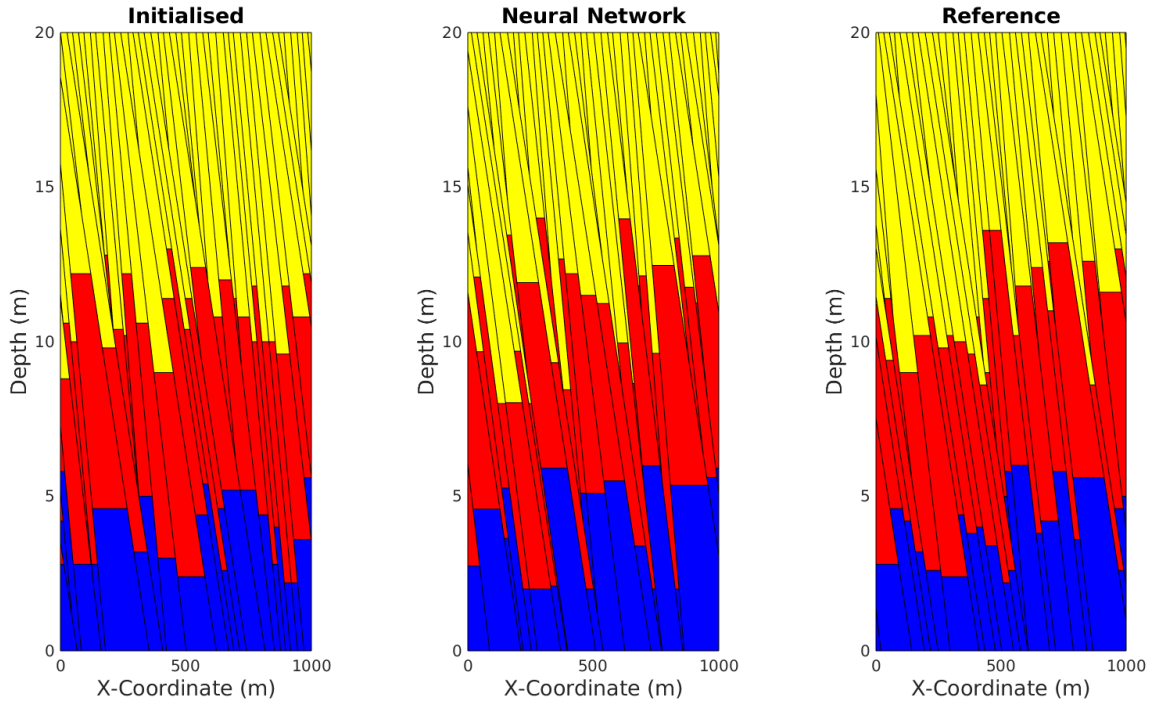


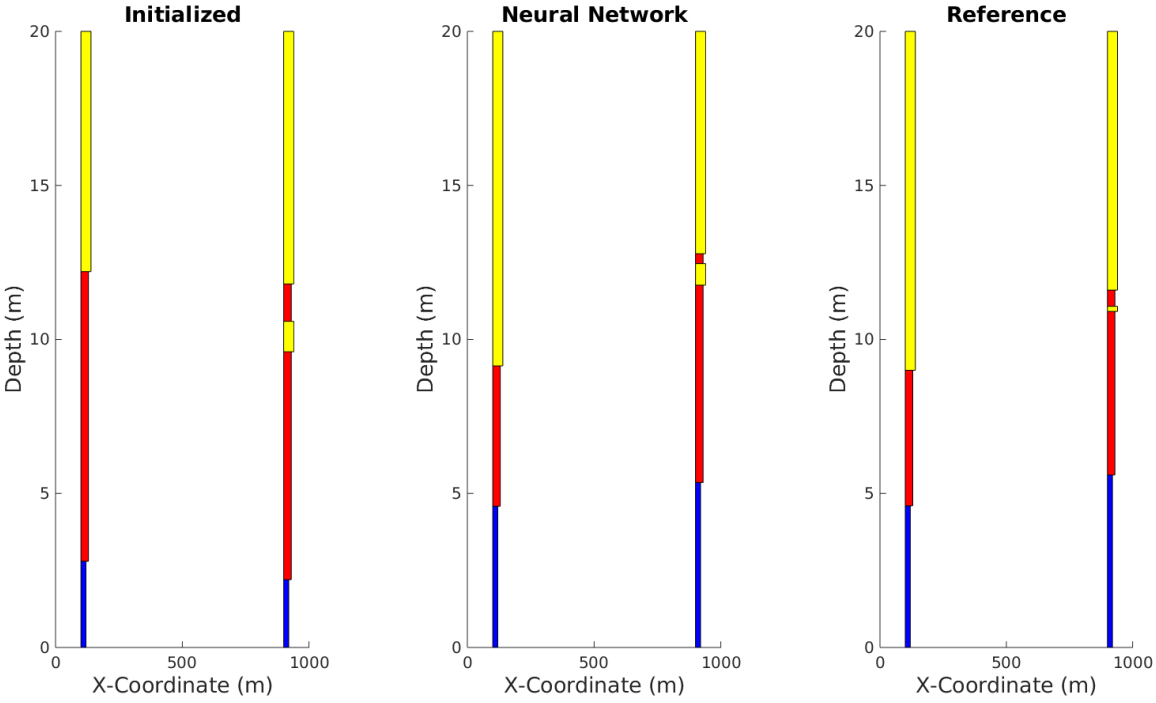
Figure 3.10: A comparison of the well logs from the SBGM with the SurfNet3 prediction for 1 sample not included in the training examples. The neural network prediction of the facies type along the well shows a reasonable match with the SBGM.

At only 2000 iterations, there is a reasonable match with the reference data. Figures 3.15a and and 3.15b show the loss plots for the two examples using SurfNet2 and SurfNet3. The decrease in the mismatch as shown in the loss plots does not significantly increase between 2000 iterations (obtained after ≈3mins) and 10000 iterations (obtained after ≈11mins), as observed in Figures 3.16a to 3.17b. Although the learning rate is annealed by a factor of 0.5 after every 2000 iterations, a sufficient match of the model with the reference data is achieved even with 2000 iterations. The results also show that the fit with the reference data is especially good at the upper and lower sections of the model and at boundaries where the thickness of the layers are significant. However, the model does not produce an exact match at well locations where the thickness of the layer is less than the resolution of the encoding of the output data presented in Section 3.4.1 as the facies type is given as the value at the centre of the discretised location.

It is useful to point out that both SBGM and ANN can be used in the conditioning step. For current 2D models the speed difference is minimal, but looking forward to more general and 3D cases, this will become more significant, and ANN will likely be faster. For the 2D models presented in this research, the wall time for running 10000 iterations for conditioning the model is approximately 7 to 11mins. In some cases, this time can be reduced by a factor of 0.2 since after 2000 iterations, as shown in Figures 3.16a to 3.17b, the neural network optimised inputs achieve a reasonable match with the observed (reference) model at the well location. An assessment of the uncertainties is presented in the following section to validate the applicability of the method

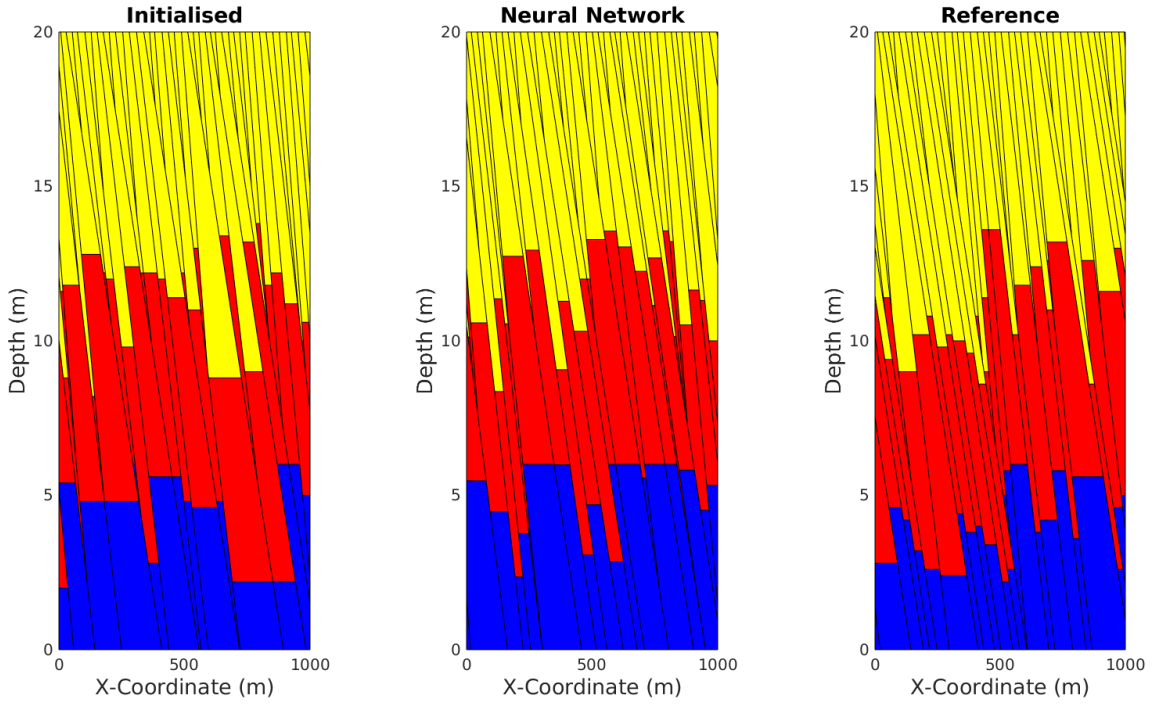


(a) SBGMs using SurfNet2 - Example 1

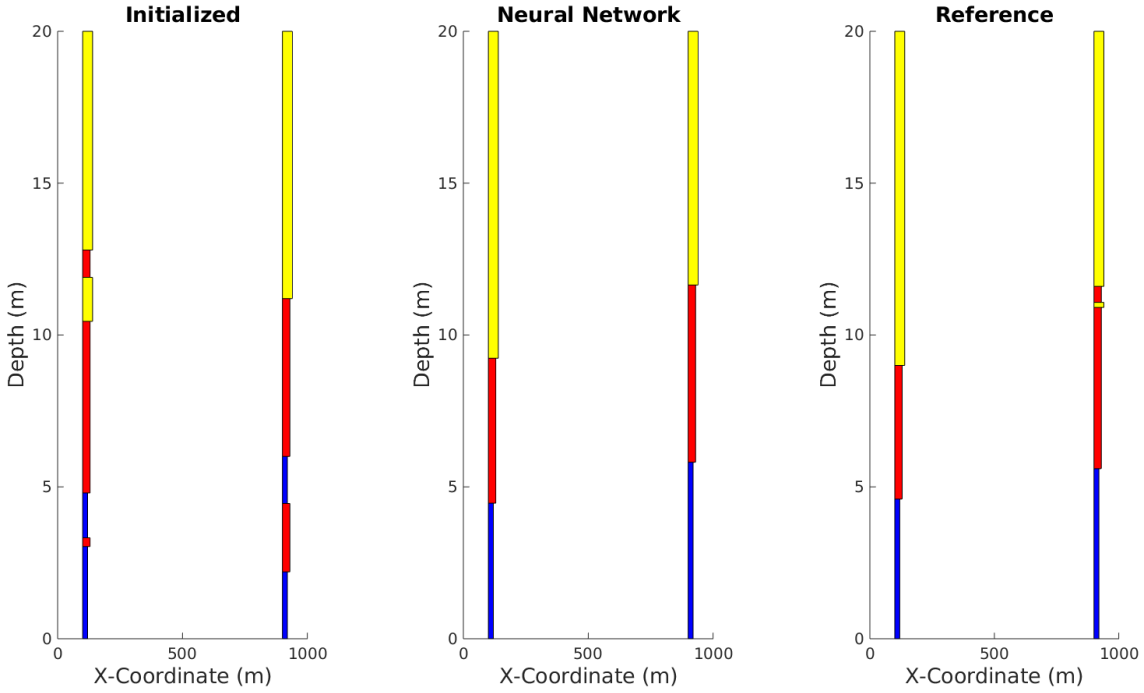


(b) Well data for SBGMs using SurfNet2 - Example 1

Figure 3.11: The initialised, optimised and reference SBGMs results for an example using SurfNet2 (Top). The inputs to the initialised model are optimised in the back-propagation algorithm to produce the inputs that generate the neural network SBGM to match reference SBGM at the well location in the conditioning step. The facies types at the well for models with the initialised SBGM inputs, the optimised SBGM inputs from SurfNet2 and the reference SBGM are shown in the bottom. The optimised inputs from the neural network yield a reasonable match of the facies type at the well with the reference model (observed data).

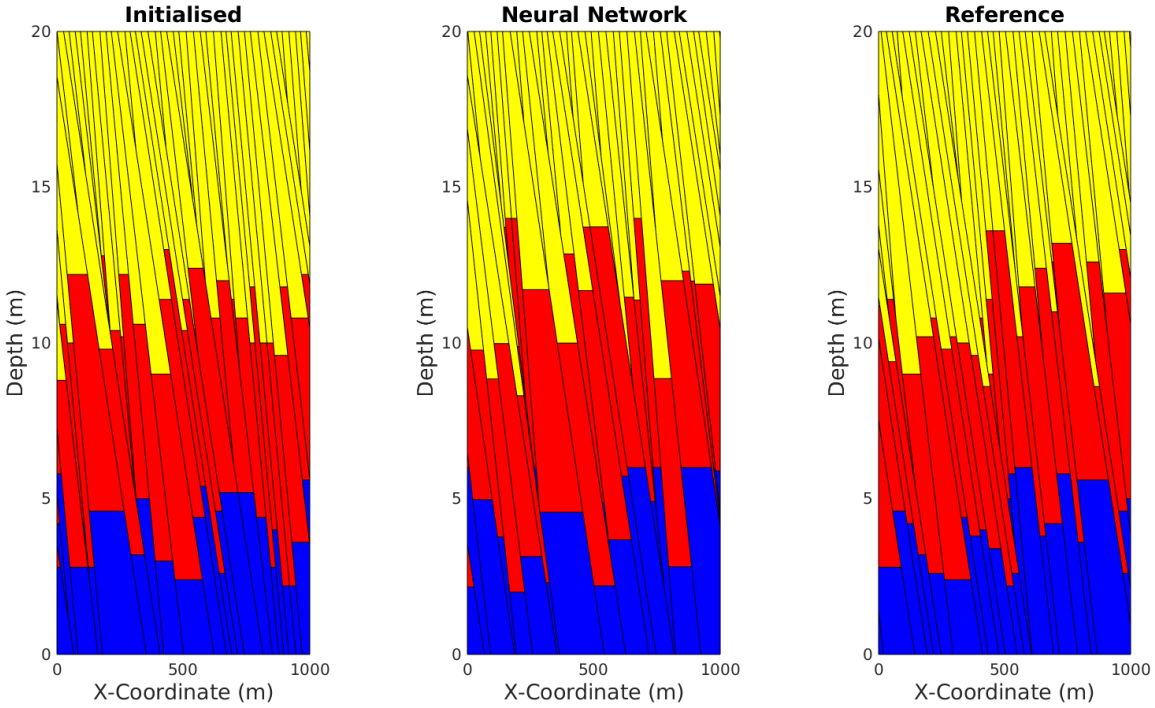


(a) SBGMs using SurfNet2 - Example 2

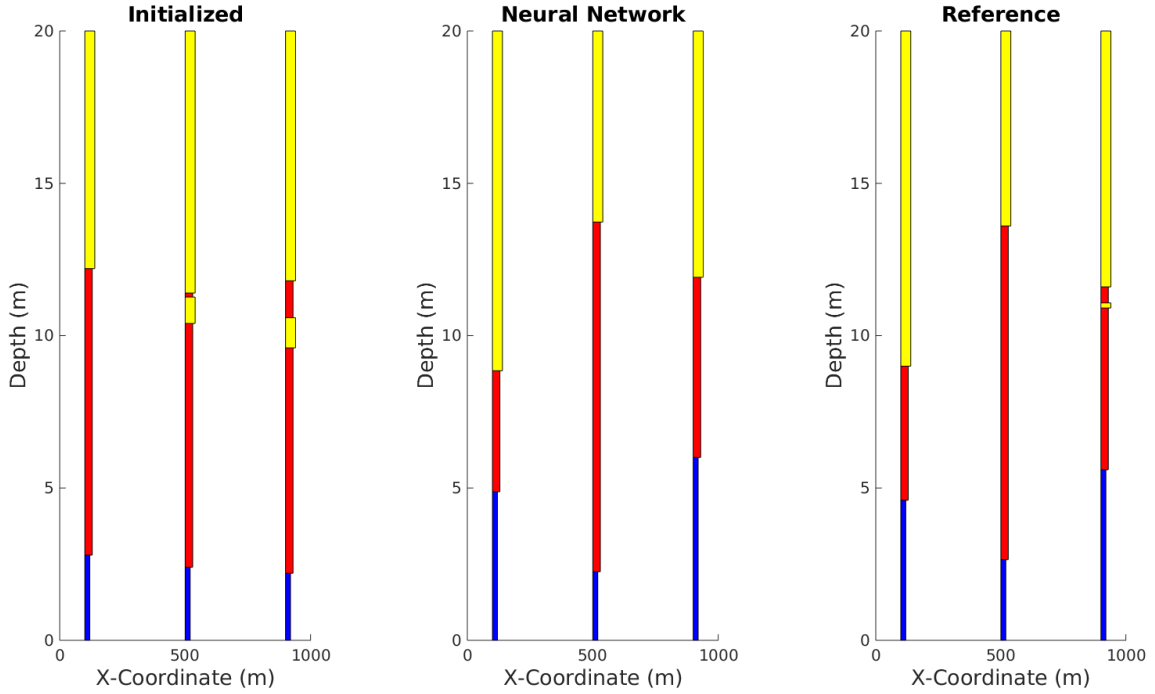


(b) Well data for SBGMs using SurfNet2 - Example 2

Figure 3.12: The initialised, optimised and reference SBGMs results for an example using SurfNet2 (Top). The inputs to the initialised model are optimised in the back-propagation algorithm to produce the inputs that generate the neural network SBGM to match reference SBGM at the well location in the conditioning step. The facies types at the well for models with the initialised SBGM inputs, the optimised SBGM inputs from SurfNet2 and the reference SBGM are shown in the bottom. The optimised inputs from the neural network yield a reasonable match of the facies type at the well with the reference model (observed data).

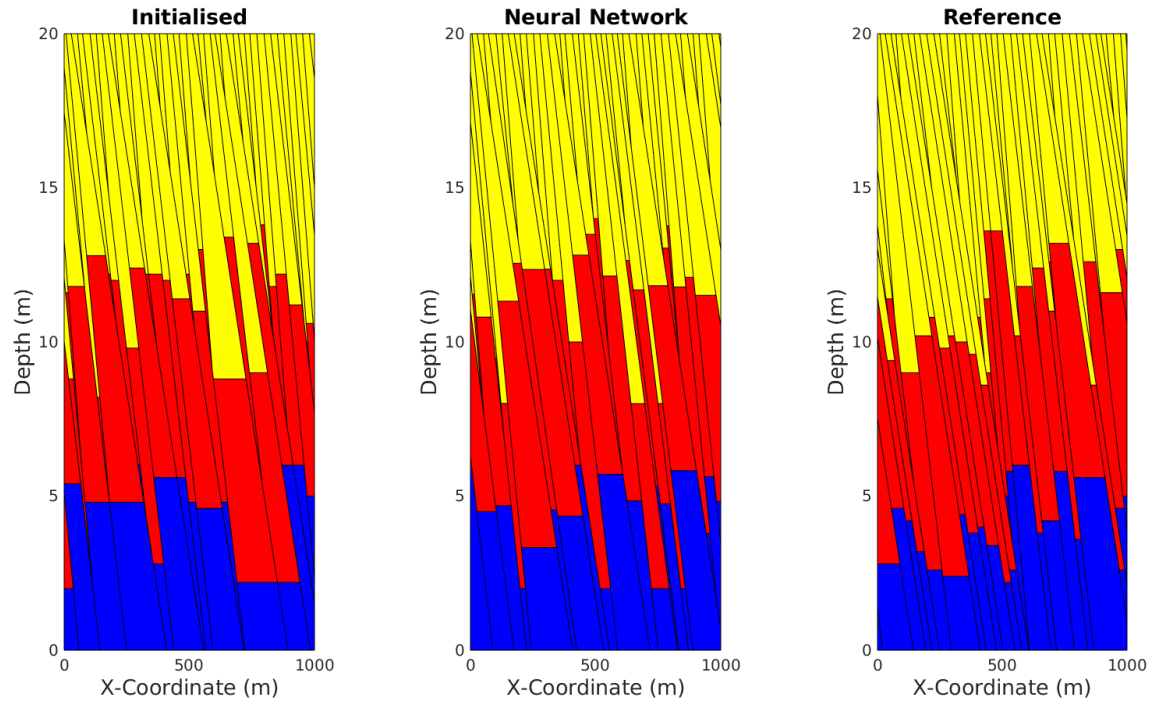


(a) SBGMs using SurfNet3 - Example 1

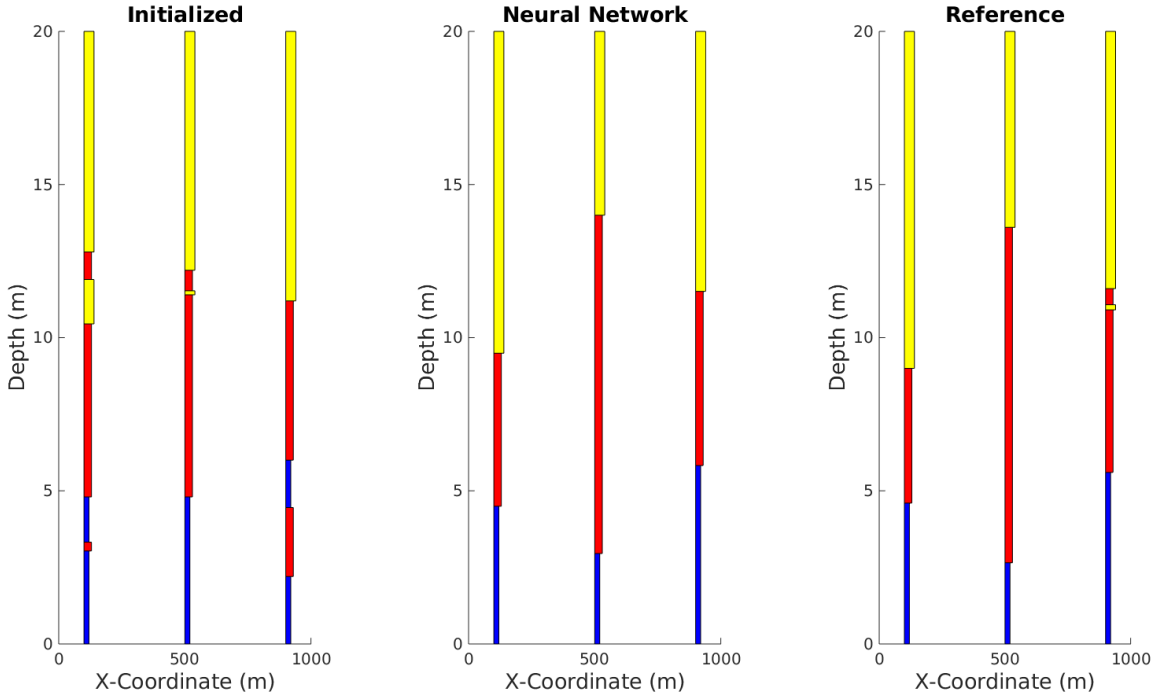


(b) Well data for SBGMs using SurfNet3 - Example 1

Figure 3.13: The initialised, optimised and reference SBGMs results for an example using SurfNet3 are shown in the top. The inputs to the initialised model are optimised in the back-propagation algorithm to produce the inputs that generate the neural network SBGM to match reference SBGM at the well location in the conditioning step. The facies types at the well for models with the initialised SBGM inputs, the optimised SBGM inputs from SurfNet3 and the reference SBGM are shown in the bottom. The optimised inputs from the neural network yield a reasonable match of the facies type at the well with the reference model (observed data).



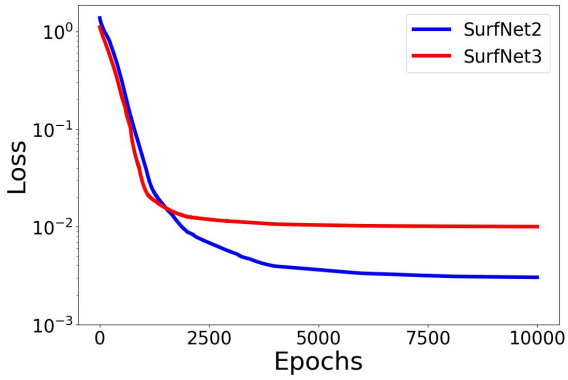
(a) SBGMs using SurfNet3 - Example 2



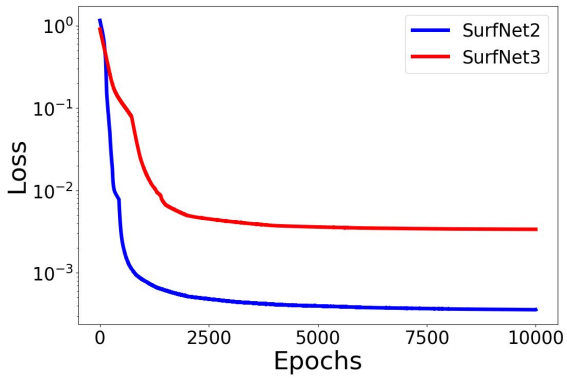
(b) Well data for SBGMs using SurfNet3 - Example 2

Figure 3.14: The initialised, optimised and reference SBGMs results for an example using SurfNet3 are shown in (a). The inputs to the initialised model are optimised in the back-propagation algorithm to produce the inputs that generate the neural network SBGM to match reference SBGM at the well location in the conditioning step. The facies types at the well for models with the initialised SBGM inputs, the optimised SBGM inputs from SurfNet3 and the reference SBGM are shown in the bottom. The optimised inputs from the neural network yield a reasonable match of the facies type at the well with the reference model (observed data).

for evaluating spatial variations.



(a) Example 1 back-propagation loss plot using SurfNet2 and SurfNet3



(b) Example 2 back-propagation loss plot using SurfNet2 and SurfNet3

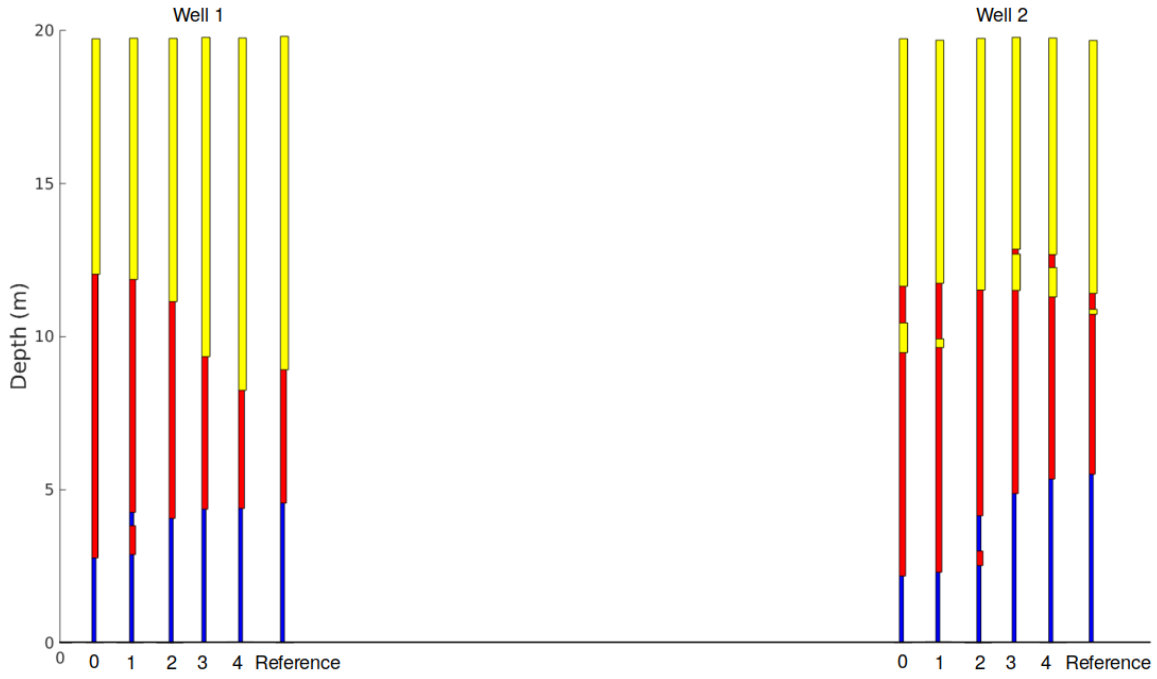
Figure 3.15: The loss plot for the inverse modelling for (a) Example 1, and (b) Example 2 using SurfNet2 and SurfNet3.

3.6.3 Uncertainty Assessment: Evaluating Spatial Variations

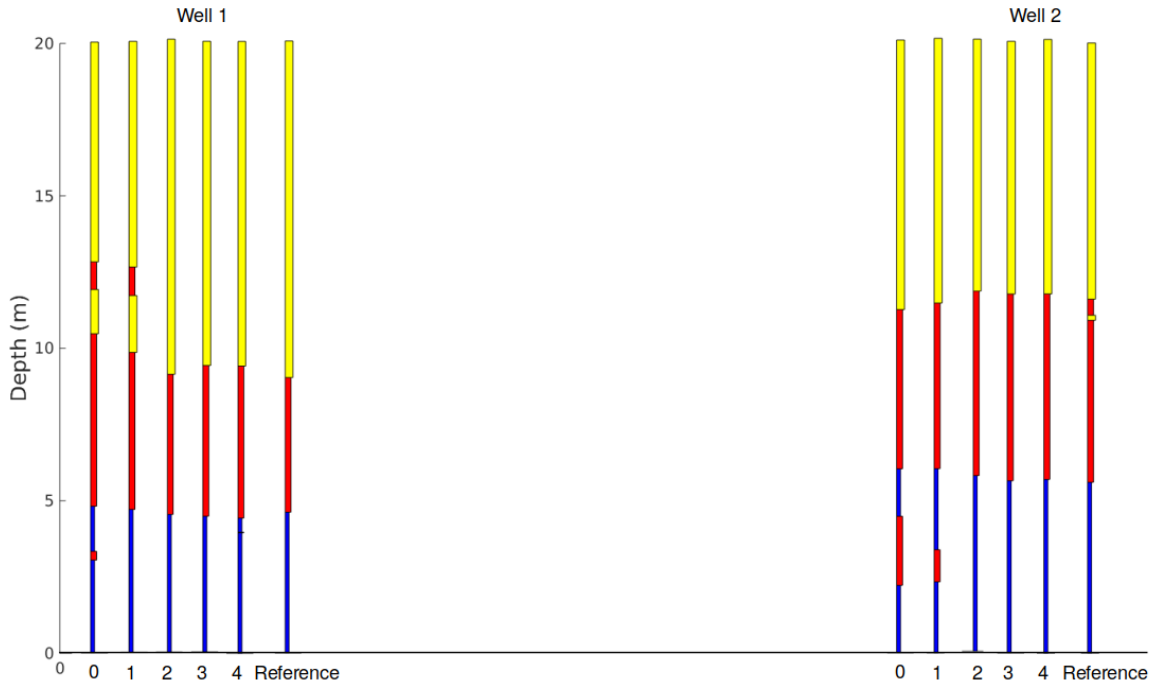
Due to the sparsity of the data and non-uniqueness of the solution, numerous plausible realisations of the SBGM exist. Although conditioning SBGMs to well data using local optimisation methods such as applied in this study is complicated by the stochastic nature of generating SBGMs (Pyrcz et al., 2015), this allows the generation of multiple realisations for the assessment of uncertainty. Using our optimisation methodology presented in Figure 3.6, we run the back-propagation algorithm multiple times with different initialised SBGM inputs to rapidly obtain an ensemble of conditioned realisations and perform an analysis of the uncertainties associated with the conditioned models.

Many models which fit the observed well data were obtained by continuously sampling and optimising SBGM inputs with different initial guesses. Using our approach, fast generation of plausible geological models that honour observed data at the wells was achieved, and the uncertainties associated with these models were evaluated. We optimised the inputs to 1000 SBGMs and obtained the models (here referred to as the conditioned realisations) with high well correlation using equation (3.12). The well-correlation cut-offs used are 98.0% and 97.5% for models generated using SurfNet2 and SurfNet3. The values of the well correlation used correspond to the test accuracy of the models (within a range of +0.5%). The method lends itself to parallelism and this massively improves the speed of computation and the performance as realisations can be conditioned independently. For SurfNet2, 36.4% of the optimised SBGM inputs showed high correlation with the observed well data, while 38.8% of the optimised inputs with SurfNet3 showed high correlation with the well data using the well-correlation cut-offs.

To assess the uncertainties associated with the conditioned realisations, we use a frequency of predictions. The frequency of predictions provides a measure of the agreement among the conditioned realisations and the

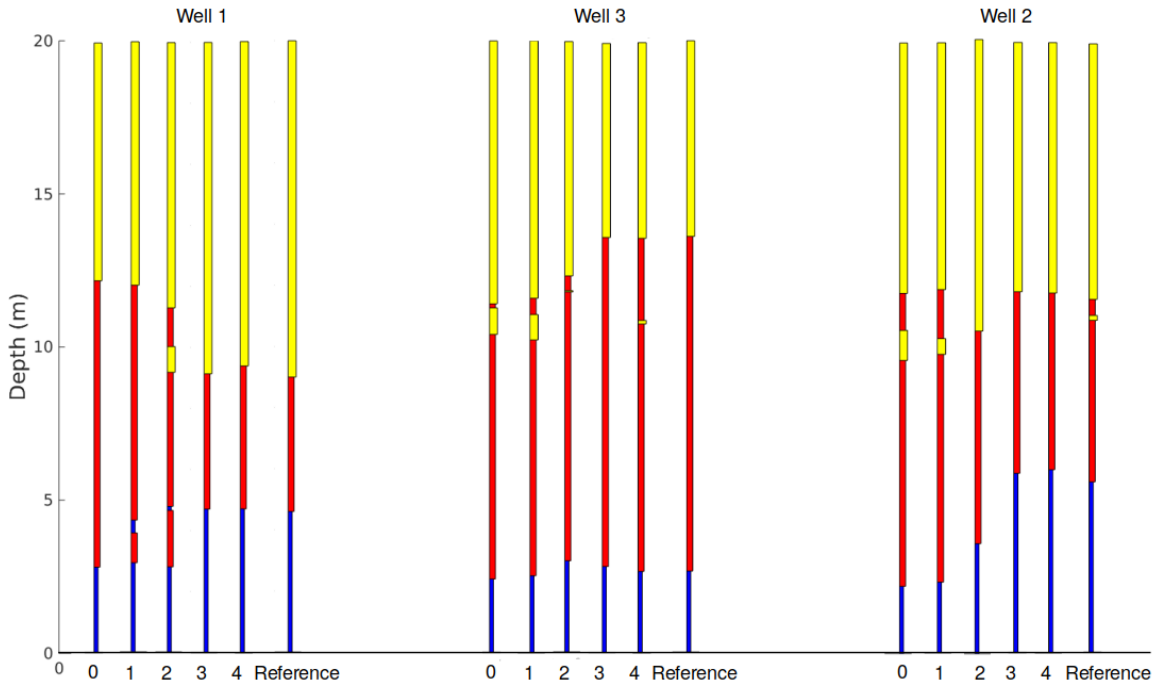


(a) Well data for conditioned SBGMs using SurfNet2 for Example 1

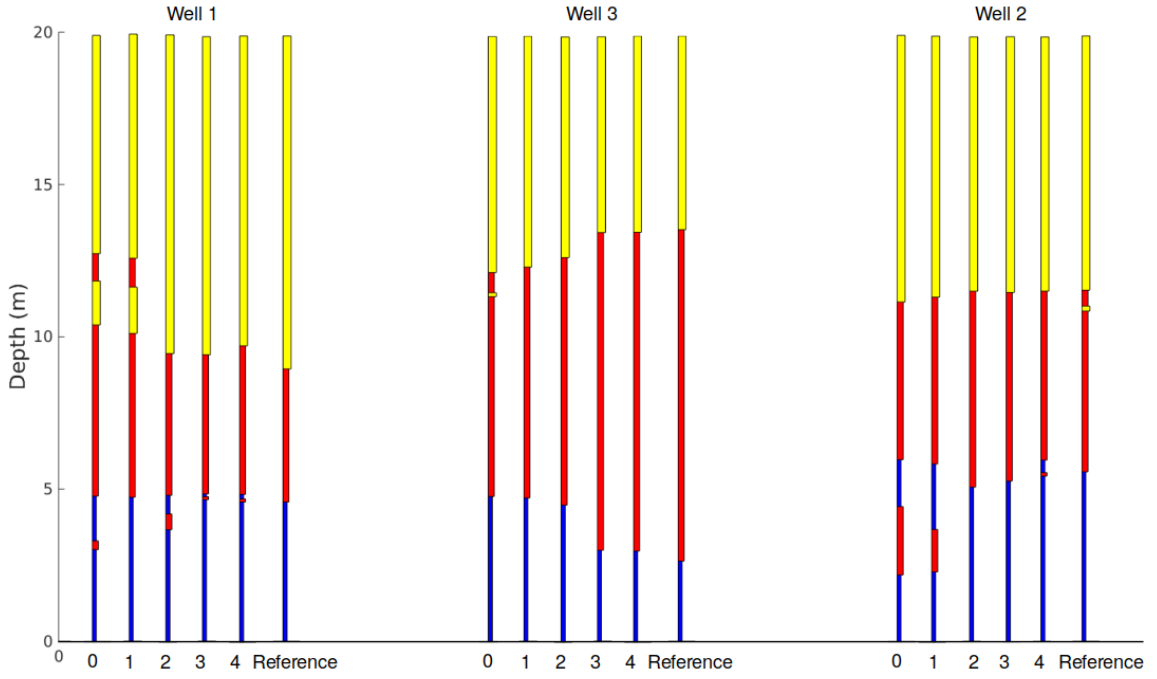


(b) Well data for conditioned SBGMs using SurfNet2 for Example 2

Figure 3.16: The well data at different optimisation steps from the initial guess to the final model obtained after 10000 iterations for SurfNet2. The results are shown above for the initial guess, and after 100, 500, 2000 and 1000 iterations given as 1,2,3 and 4 respectively. The observed data is also indicated as the reference. The optimised inputs from the neural network yield a reasonable match of the facies type at the well with the reference model (observed data) except at locations where the resolution of the grid does not allow for the representation of the data at that point.



(a) Well data for conditioned SBGMs using SurfNet3- Example 1



(b) Well data for conditioned SBGMs using SurfNet3- Example 2

Figure 3.17: The well data at different optimisation steps from the initial guess to the final model obtained after 10000 iterations for SurfNet3. The results are shown above for the initial guess, and after 100, 500, 2000 and 1000 iterations given as 1,2,3 and 4 respectively. The observed data is also indicated as the reference. The optimised inputs from the neural network yield a reasonable match of the facies type at the well with the reference model (observed data) except at locations where the resolution of the grid does not allow for the representation of the data at that point.

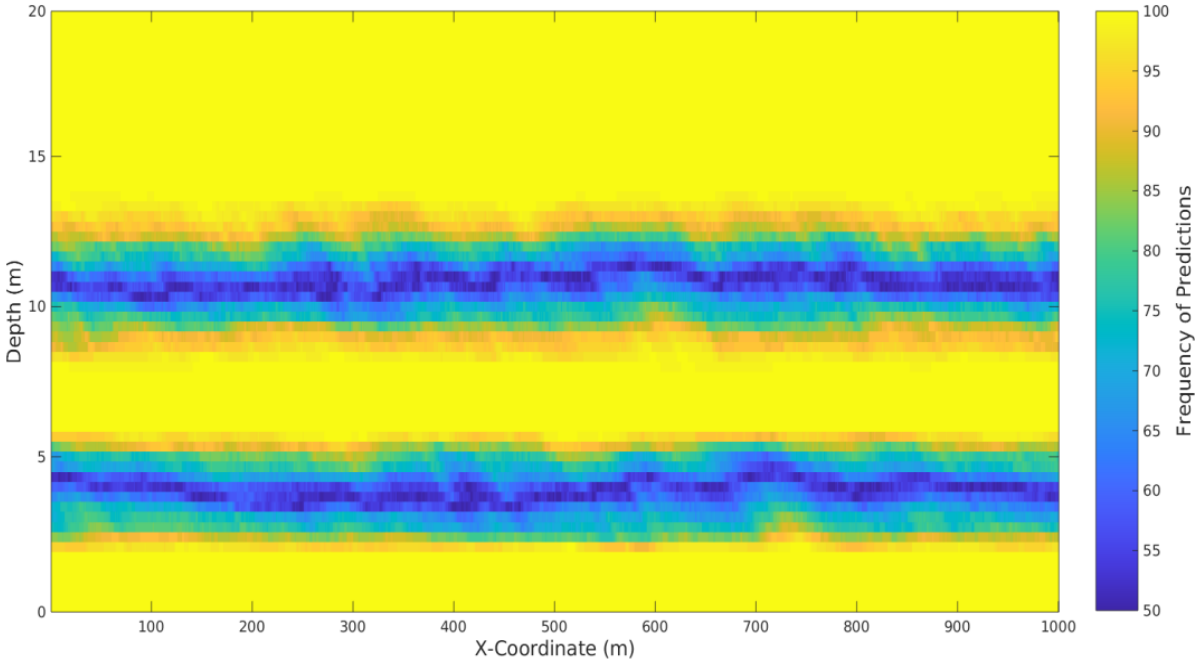


Figure 3.18: The frequency of predictions plot for the initialised SBGMs. The uncertainty varies in equal proportions laterally in the model including at the well locations.

reference data. This is useful for understanding the spatial variations and uncertainties in the conditioning methodology. The frequency of predictions, f^{pred} is given as the difference between the maximum frequency and the minimum frequency defined as

$$f_{i,j}^{pred} = f_{i,j}^{max} - f_{i,j}^{min}, \tag{3.13}$$

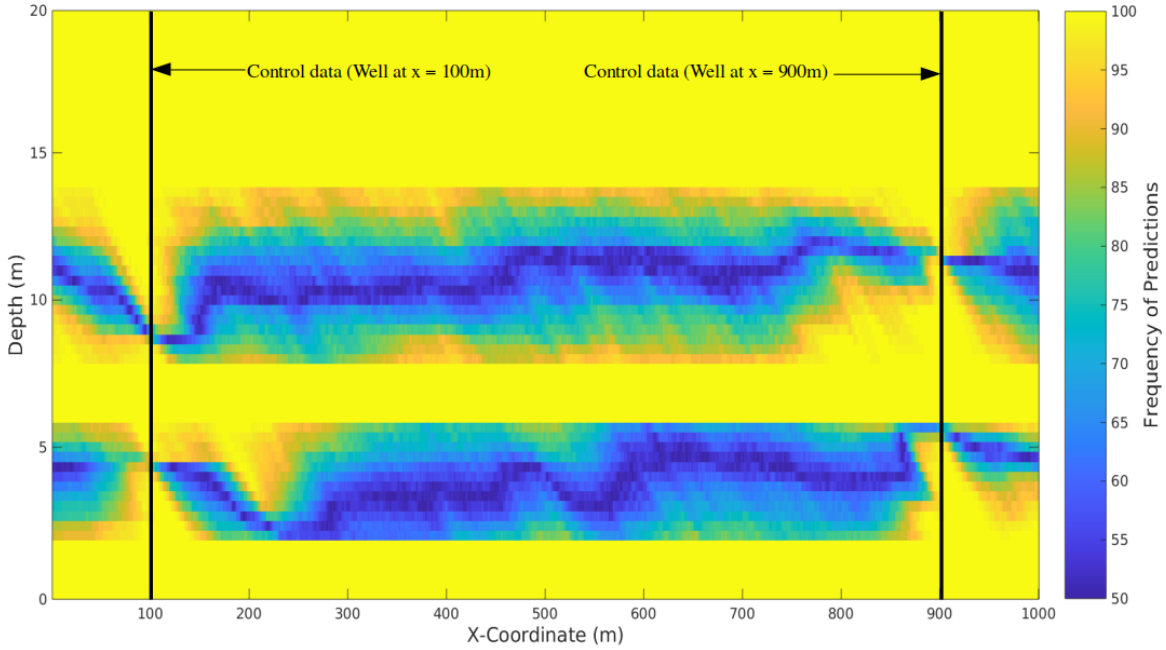
where $f_{i,j}^{max}$ is the frequency of the facies with the highest occurrence (also described as the occurrences of the facies type with the most predictions), and $f_{i,j}^{min}$ is the frequency of the facies with the least occurrence at each discretized location within the domain. For example, if the facies type at a particular discretised location is predicted to be USF 80 times, LSF 15 times and OM 5 times, then $f_{i,j}^{max} = 80$ and $f_{i,j}^{min} = 5$, and therefore $f_{i,j}^{pred} = 75$.

Figure 3.18 shows the frequency of predictions for the initial guesses of the SBGMs. It can be observed that there is high uncertainty in the model as the uncertainty varies in equal proportions laterally in the model including at the well locations. Figure 3.19a and Figure 3.19b depict the frequency of predictions and the variations among 100 conditioned realisations for models conditioned using SurfNet2 and SurfNet3 respectively. The 100 realisations used have the same initial guesses for the optimisation in both networks, and are above the cut-offs used for both networks. In both figures, spatial variations and uncertainties among the conditioned realisations are considerably lower at locations close to the wells compared to locations away from the control data, where we observe a smaller number of the conditioned realisations predicting the same facies types. In Figure 3.19b, the addition of the well at x=500m constrains the model at that location and its surrounding, resulting in a more global reduction in the uncertainties.

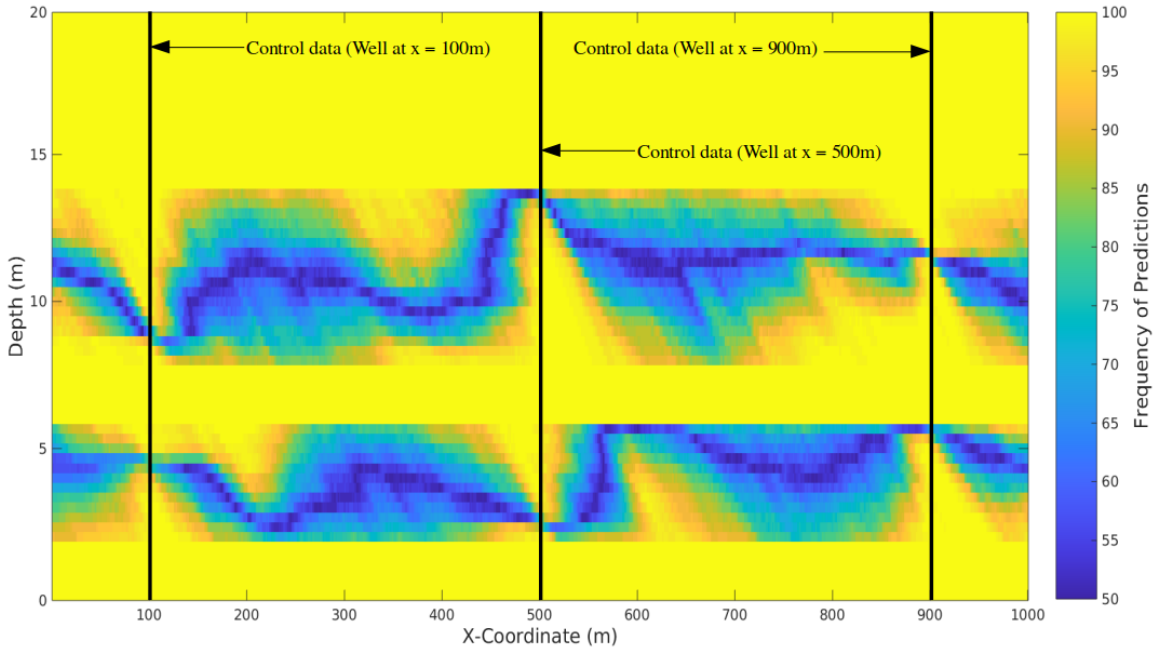
Figures 3.20a and 3.20b show the facies with the highest predictions at discretised grids over the entire domain of the SBGM for 100 conditioned realisations. Qualitatively, it can be observed that salient features are captured and patterns in the reference wells are preserved as the facies predicted in most of the models from the neural network follow similar trends with the reference model. In contrast, Figure 3.21 which shows the maximum predicted facies for the initialised SBGM models does not depict the significant variations at well location that correspond to the observed data. An assessment of Figures 3.20a and 3.20b show similar patterns at x=100m and x=900m, while at x=500m, variations between the two maximum predictions are observed. By constraining the model with more information at x=500m, there is a better agreement with the reference model.

We found good results using the approach proposed here for conditioning surface-based models, particularly for initialised input values that largely agree with the reference data, where the conditioning is very effective and the method produces plausible realisations. However, if initialised model inputs significantly disagree with the spatial statistics of the reference dataset, conditioning may be more challenging. Since geological and petrophysical properties are often spatially correlated, exploring the spatial variations among variables through the use of convolutional neural networks (CNNs) (Lecun et al., 1998), which are capable of learning representative features of spatially correlated data should be very interesting. Furthermore, the hierarchical arrangement of surfaces and stacking patterns of SBGMs present opportunities for conditioning individual surfaces using recurrent neural networks (RNNs) to model temporal sequences (Grana et al., 2020; Hochreiter and Schmidhuber, 1997; Zhang et al., 2018). Because perturbing previous events of the forward model changes subsequent events, data conditioning on an event-by-event basis may significantly improve the results, as variations among the conditioned realisations along the well locations can still be observed at points that correspond to the boundaries of facies associations in the reference data. The addition of information regarding variations in sequence numbers and thickness of layers using RNNs may significantly improve the results. This will allow for the extension of the method to conditioning SBGMs to stratigraphic surfaces in addition to the facies type investigated in this research. The proposed methodology may also be reapplied to object-based models where parameterized objects are used to describe complex reservoir geometry. Future work will have far-reaching practical application for more accurately describing the spatial distribution of geological features in 3D reservoir modelling and characterisation.

Although the application of the method is presented on a 2D synthetic example, the method described in this work can also be applied to real cases, where SBGMs can be generated using the method presented in Jacquemyn et al. (2019). For field applications, the approach outlined in Figure 3.6 can be utilised for conditioning geological models to well data although more attention will be focused on generating SBGMs that are 3D and comprise a broader range of geometries that are representative of subsurface geology. Furthermore, dealing with more than three facies codes and a larger number of wells are steps to take for a more general application. We also note that, in contrast to the example demonstrated here, the following challenges will apply to real cases:

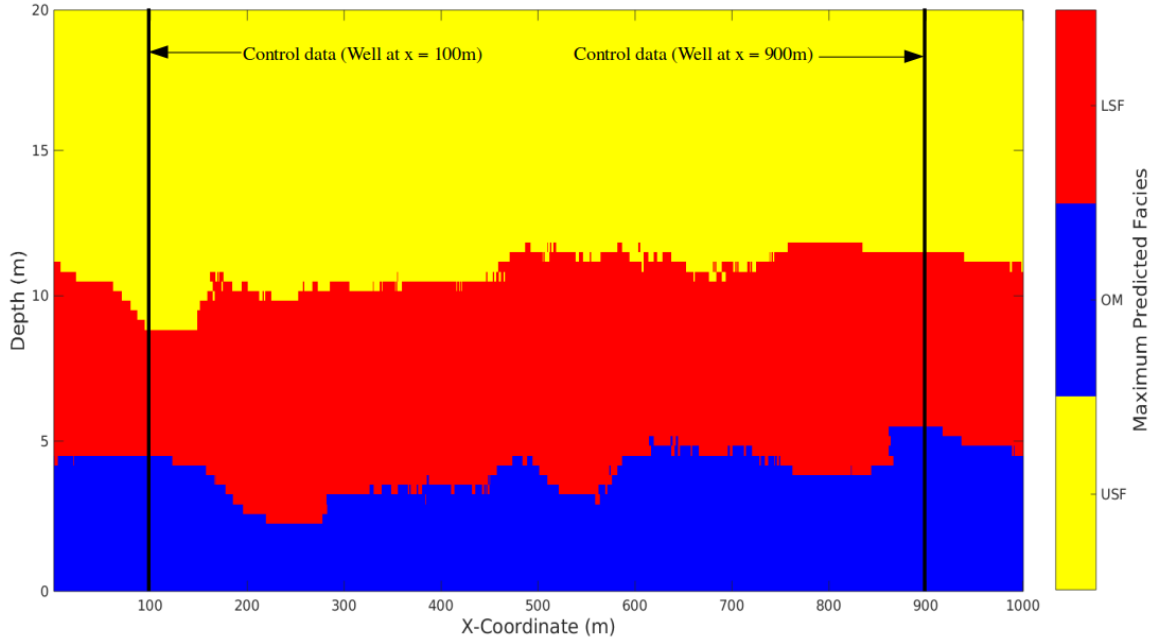


(a) Frequency of Predictions for Conditioned realisations with SurfNet2

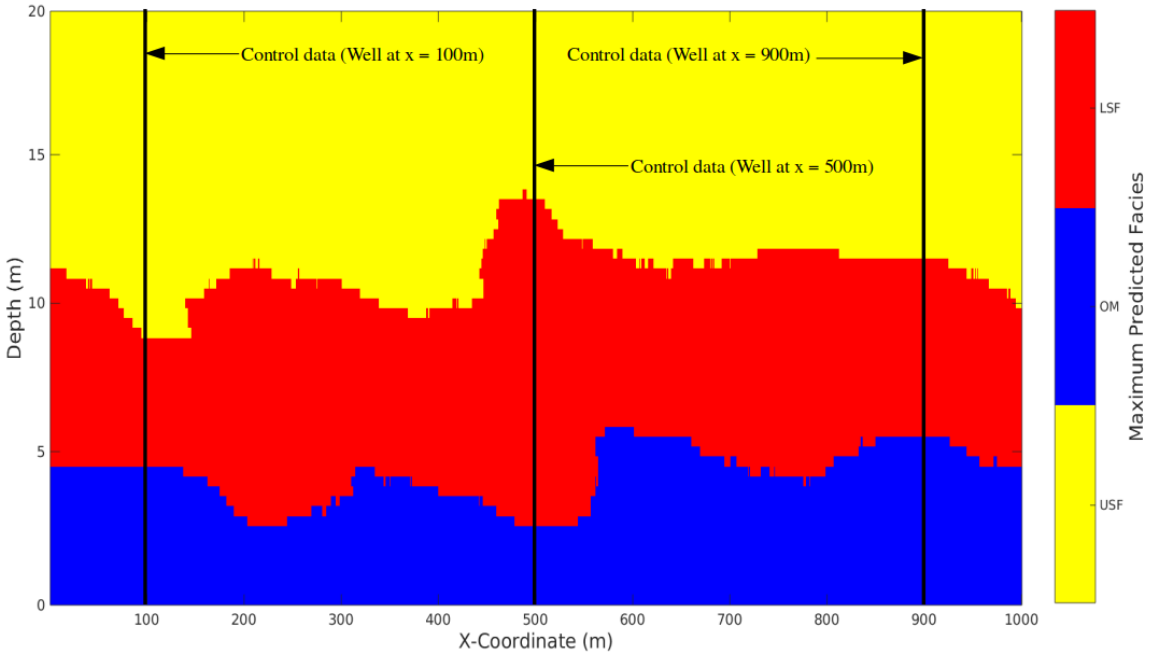


(b) Frequency of Predictions for Conditioned realisations with SurfNet2

Figure 3.19: The frequency of predictions plot for SurfNet2 and SurfNet3 depicting the difference between the maximum frequency (number of facies with the highest occurrence), and the minimum frequency (number of facies with the least occurrence) at each discretized location in the domain. At regions close to the control data (wells at x=100m, x=500m and x=900m), there is a higher number of predictions of the same facies type among the 100 conditioned realisations, compared to regions away from the control data.



(a) Maximum predicted facies for the conditioned realisations with SurfNet2



(b) Maximum predicted facies for the conditioned realisations with SurfNet2

Figure 3.20: The facies type with the maximum frequency or number of occurrences is shown for 100 plausible realisations that have been conditioned to the well data using (a) SurfNet2 and (b) SurfNet3. Patterns associated with the reference well are preserved as the facies predicted in most of the models from the neural network follow similar trends with the reference model.

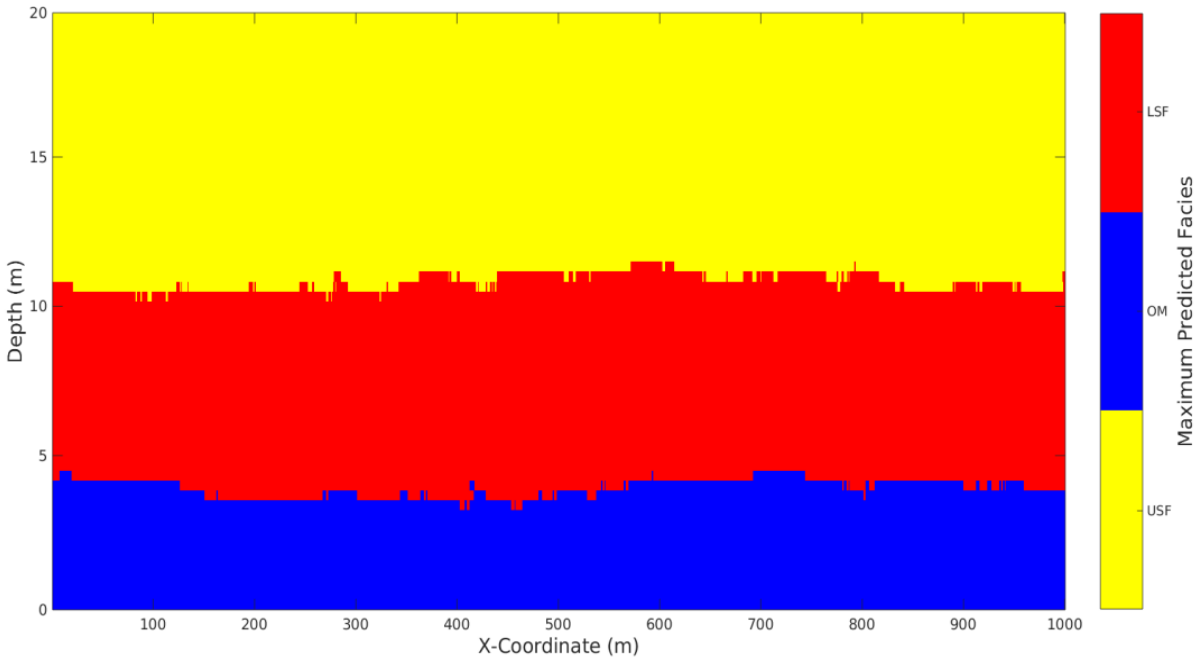


Figure 3.21: The facies type with the maximum frequency or number of occurrences is shown for 100 models used as initial inputs to the inverse modelling. General trends and similarities with the reference data cannot be observed at well locations.

1. Longer time requirements for generating the SBGMs to be used as training examples for the neural network. However, compared to other methods, the use of parametric, grid-free surface-based models for representing geological heterogeneity of interest such as features which influence the flow conditions has many advantages which outweigh any anticipated computational limitations. Furthermore, the number of surfaces required to model realistic reservoirs using surfaces has been shown by Osman et al. (2021) to be low to generate representative models.

2. Increased model training time/computational requirements due to the increase in the number of input parameters for the model, which might require more complex network architectures. However, for 3D model equivalents of the 2D example shown, the number of parameters only increases by a factor of two.

3.7 Conclusions

We have demonstrated how surface-based geological models (SBGMs) can be conditioned to well data using neural networks. A multilayer perceptron was trained to learn the features and underlying relationship between the SBGM inputs and facies types along a well. The ease of representing realistic surface-based geological models enabled the rapid generation of 10^6 SBGM samples with which the network could be trained to a high degree of accuracy. Using a backpropagation algorithm, the inputs of the SBGM were generated by optimising the inputs of the trained neural network based on the forward model. The inverse modelling provided an

ensemble of realisations that fit the observed well data; these multiple realisations can be used to understand the impact of uncertainty on predictions of fluid flow and other important reservoir metrics. An assessment of the uncertainties revealed low variations between models close to the control data as expected, and high variations away from control data, further informing the spatial uncertainty of the sequences in the geological model. The use of an additional well constrains the resulting SBGMs further and validates the applicability of the method as a robust framework for conditioning SBGMs to well data.

This proof-of-concept work is essential to underpin future research to improve the methodology for application to real cases, particularly 3D reservoir models of more complex heterogeneities and a larger number of wells. The use of CNNs, which apply sparse connections to describe spatial correlations, and RNNs, which allow for variability in building series models, are proposed in future work to improve the methodology presented here.

Chapter 4

A Grid-Free Approach to Conditioning Surface-Based Geological Models to Well Data

4.1 Abstract

Conditioning geological models to well data is important for predicting reservoir performance and improving fluid recovery. In this work, a method is presented for inferring the parameterised inputs to surface-based geological models (SBGMs) that match well log data for any well trajectory. The study builds on an approach that applies neural networks to approximate SBGMs and optimise the inputs to the SBGMs for a given well data in an inverse modelling framework. The developed technique aids in the generation of plausible geological models that are conditioned to data from an arbitrary number of wells that have simple or complex trajectories. The adoption of a grid-free approach in generating and conditioning the SBGMs allows for the flexibility of representing various geological architectures and wells. The method is demonstrated with realistic well paths, such as deviated wells and multi-laterals. An ensemble of conditioned models is evaluated to understand spatial variations and uncertainties associated with the generated models. The methods developed in this study provide a computationally convenient approach to generating realistic and reliable geological models that can aid the accuracy of reservoir modelling and simulation for efficient production or injection of subsurface fluids.

4.2 Introduction

Reservoir modelling is key to developing, managing, and optimising performance and production. Due to the sparsity of observed data such as logs from few exploratory and appraisal wells, determining the spatial distribution of petrophysical or geological properties and characterising reservoirs can be extremely difficult. Calibrating static geological models to well data such as facies association, enables the development of models that are reliable representations of a subsurface reservoir (Aminzadeh and Dasgupta, 2013). Several techniques have been developed for producing representative models of the subsurface (a review of methods for geological modelling is discussed in Pyrcz and Deutsch (2014)). Surface-based geological models (SBGMs) (Jackson et al., 2014, 2015; Jacquemyn et al., 2019; Pyrcz et al., 2005, 2015; Ruiu et al., 2016; Sech et al., 2009; Zhang et al., 2009), a type of rule-based models, provide the advantage of geological realism over geostatistical methods such as Kriging and multiple-point statistics (MPS), and at a fraction of the computational expense of other rule-based models such as object-and process-based models (Pyrcz et al., 2015).

In surface-based modelling, depositional facies are described using surfaces that honour the sequence and time scales of geological units and allow for the modelling of large- and small-scale heterogeneities (Jacquemyn et al., 2019). By modelling heterogeneities as geologically meaningful discrete volumes bounded by surfaces having constant petrophysical properties, surface-based modelling honours underlying geological process and is not restricted by the limitations imposed by the spatial resolution of the reservoir model (Jackson et al., 2015). Depending on the depositional environment, the type of available data, and the heterogeneity to be modelled, surfaces can be generated using a variety of techniques: infilling of surface-based volumes with pillar grids (Parquer et al., 2017; Ruiu et al., 2016), rasterising channel geometries onto locally refined grids (Hassanpour et al., 2013), constraining structures of stacking elements such as lobes using surfaces (Pyrcz et al., 2005; Zhang et al., 2009). Jacquemyn et al. (2019) use a grid-free framework that applies parametric functions to generate surfaces using non-uniform rational B-splines (NURBS) to build reservoir models that capture realistic geometries.

The use of parameterised surfaces to represent geological features and architectures in SBGMs enables the ease of generating and manipulating surfaces, as few input parameters are required for producing realistic geological models. However, the difficulty of conditioning these models to observed data has limited the widespread application of SBGMs for reservoir modelling. A variety of methods have been applied for conditioning rule-based models including SBGMs: sequential optimisation of input parameters (Bertoncello et al., 2013), implicit filtering (Wang, Pyrcz, Catuneanu and Boisvert, 2018), variational inference (Wingate et al., 2016), Markov-Chain Monte-Carlo methods (Karssenberg et al., 2001), local adjustment techniques (Pyrcz et al., 2005), dual-spline error surface interpolation (Zhang et al., 2009) and reverse migration (Parquer et al., 2017). Some recurring issues in the above-mentioned methods include: 1) the computational expense of generating and

conditioning the geological models, 2) the difficulty of conditioning the geological models to well data when the density of the control data increases, and 3) the inefficiency of conditioning the models when observations are sparse. Jo et al. (2020) presented a method using generative adversarial networks for conditioning rule-based models to local well data.

In a previous work (Titus et al., 2022), we have shown that neural networks can be applied to condition SBGMs to facies data at multiple well locations while preserving geological realism and at a reasonable computational cost. To condition SBGMs to well data with more complex trajectories and geometries, we present here an approach based on artificial neural networks (ANNs). The method is applied to SBGMs generated using parametric, grid-free surfaces described in Jacquemyn et al. (2019), which more accurately capture geological architectures. The adoption of a grid-free framework in this work allows for the flexibility of representing heterogeneities in subsurface reservoirs and a variety of well trajectories. The work presented here improves on the method demonstrated in Titus et al. (2022), for conditioning SBGMs to facies data at discretised locations along vertical well. A significant difference with previous studies is the applicability of our method for generating SBGMs that can honour observed data at any location within the model domain. The developed method allows SBGMs to be conditioned to data from an arbitrary number of wells having varying number of points and any trajectory. We illustrate our approach by conditioning 2D shallow marine SBGMs to facies data from multiple wells.

Figure 4.1 illustrates the methodology applied in this work for generating SBGMs that honour observed data at any location within the model, as well as the steps for evaluating spatial variations using the generated geologically plausible models. This paper is organised as follows: a brief description of the example case and data generation process of the forward models is given in the next section, followed by a description of the steps for training the ANN to learn the forward modelling process. A brief explanation/illustration of the algorithm for conditioning SBGMs to observed data using the trained neural network is subsequently presented. The results of the neural network prediction and the optimisation are shown, along with examples of conditioned realisations using the proposed methodology. The final section of the results shows the application of the conditioned realisations using the proposed methodology for assessing the uncertainties associated with the technique. A discussion of the results obtained and future work/applications are then presented. The concluding section is presented last.

4.3 Neural Networks for Surface-Based Modelling

We apply neural networks in a supervised learning mode, where the aim is to determine the set of parameters that map the SBGM inputs to the facies association along a well. The problem is defined as a classification task where the network is trained to map a normalised vector representation of SBGM input parameters and coordinates of a location within the model, to an output vector whose entries approximate the probabilities

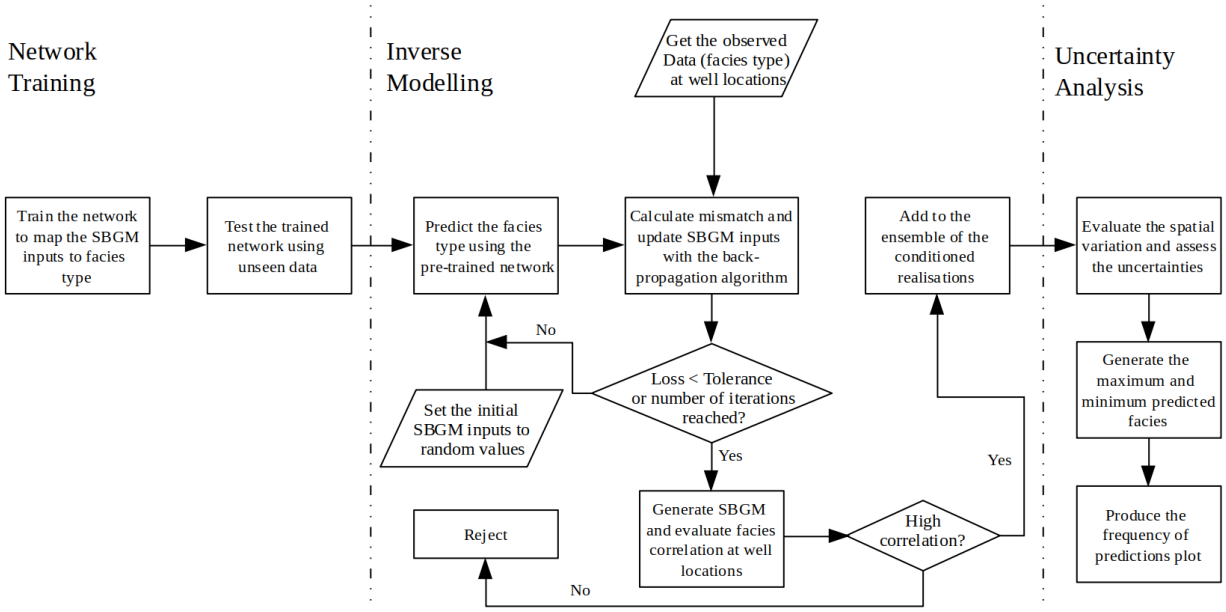


Figure 4.1: The workflow for conditioning SBGMs to well data. This involves training the neural network to represent the forward model, optimising the network inputs to generate conditioned SBGMs and assessing the uncertainties of the realisations.

of the distinct facies classes at that location. The loss function is the traditional cross-entropy loss used for multiclass classification. The inverse modelling leverages the optimisation capabilities of neural networks. Given the trained network, we find the set of input parameters whose network predictions match the observed data at the well by minimising the mismatch between the network prediction and observed data. The following sections describe the method in more detail.

4.3.1 Approximating SBGMs with Artificial Neural Networks

A multilayer perceptron (MLP) with fully-connected layers is trained to approximate the process of generating SBGMs. We use SBGMs generated based on the work of Jacquemyn et al. (2019). This allows the generation of geometrically realistic subsurface reservoir models that are not constrained to a pre-defined grid. The surfaces are built using a parametric non-uniform rational B-splines (NURBS) interpolation (Piegl and Tiller, 1997) for rapidly constructing surfaces that can describe even models with complex geometries and heterogeneities. The reservoir models generated also honour hierarchical relationships between surfaces and preserve geological realism.

Data generation

Our approach is demonstrated using 2D SBGMs similar to what was presented in Titus et al. (2022). The 2D SBGM is a simplified shallow marine depositional environment having a 20m thick parasequence of length

1000m. There are three facies in the model: upper shoreface (USF; yellow), lower shoreface (LSF; red), and offshore mudstone (OS; blue). The SBGM is constructed using stochastic input parameters sampled from a uniform distribution of variables describing the clinoform shape (dip and spacing) and the facies zones (upper and lower shoreface base). Table 1 in Titus et al. (2022) gives a summary of the distribution of the variables which represent the parameterised inputs to the model. Figure 4.2 shows an example of a 2D SBGM generated using parameters from Table 1 in Titus et al. (2022). We create the dataset for training the networks by extracting point data from the SBGMs. For each SBGM, we randomly sample a point and record the facies type at that location to provide one training example. To generate more data, multiple points can be randomly sampled from each SBGM to provide additional training examples. Figure 4.2 shows the distribution of the randomly sampled points, i, where for each point, $i_x \sim \mathcal{U}([a_x, b_x])$ and $i_y \sim \mathcal{U}([a_y, b_y])$, where a and b are the upper and lower limits of the coordinates or dimensions.

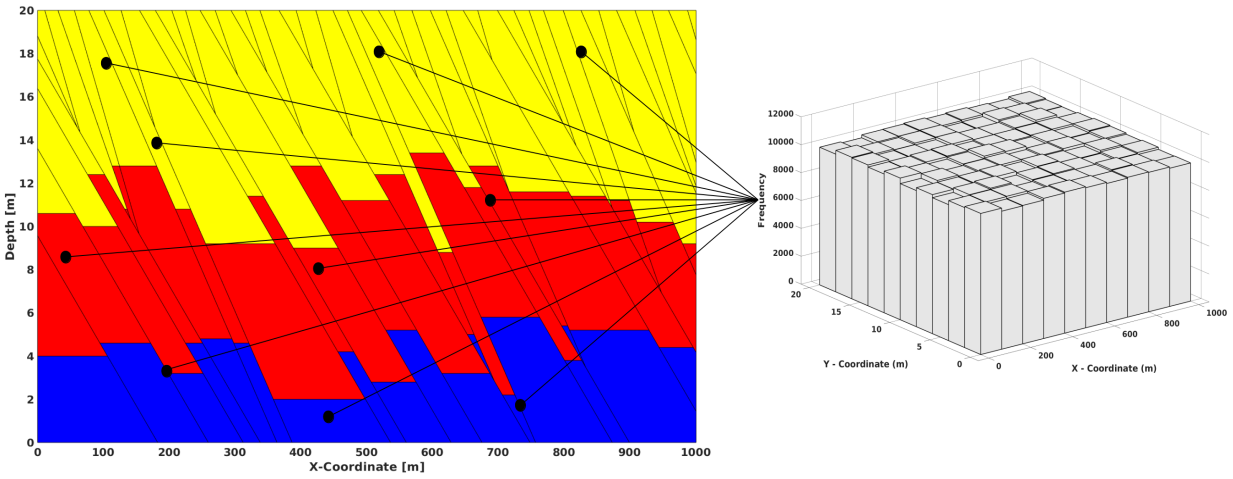


Figure 4.2: The randomly sampled points from the SBGMs have a uniform distribution of the coordinates.

Data preprocessing

Let the input feature vector, $\mathbf{f} = [f_1, f_2, \ldots, f_P] \in \mathbb{R}^P$, be an array of P real-valued parameters that comprise the inputs to the SBGM and the coordinates of a chosen point within the model domain. The entries in $\mathbf{f}$ can span multiple orders of magnitude, since these are independent and characterise different properties such as the shape, orientation, and relative location of the surfaces. We normalise the feature vectors of each example by standardising to a multivariate normal distribution, $\mathbf{f} \sim \mathcal{N}(0, 1)$. We compute the mean:

$$\overline{f}_p = \frac{\sum_{i=1}^{N} f_p^n}{N}, \quad \forall p \in \{1, 2, \ldots, P\}, \tag{4.1}$$

where N is the number of SBGM examples, and the standard deviation is:

$$\hat{f}_p = \sqrt{\frac{\sum_{i=1}^{N} f_p^n - \overline{f}_p^2}{N}}, \quad \forall p \in \{1, 2, \ldots, P\}. \tag{4.2}$$

We define the mean-shifted and standardised array for each example, n as:

$$\mathbf{f}^n = \frac{f_1^n - \overline{f}_1}{\hat{f}_1}, \frac{f_2^n - \overline{f}_2}{\hat{f}_2}, \ldots, \frac{f_P^n - \overline{f}_P}{\hat{f}_P}. \tag{4.3}$$

Since geological and petrophysical properties are often spatially correlated, the inputs in $\mathbf{f}$ were sequentially arranged to ensure depositional rules are respected. Thus, $f_1, \ldots, f_4$ represent the parameters of the first surface (clinoform), $f_5, \ldots, f_8$ represent the parameters of the second clinoform, ... and $f_{P-4}, \ldots, f_P$ are the parameters of the last clinoform. The concatenation of parameters according to temporal and spatial dynamics provides information that can significantly aid the neural network in learning fundamental relationships among stratigraphic structures as the hierarchical arrangement of surfaces and stacking patterns of SBGMs is important for capturing and preserving underlying geological trends.

To transform the non-continuous and non-ordinal facies type data into numeric data, we apply one-hot encoding (Bishop, 2006) to generate a binary encoding of the facies classes. We use binary digits to represent the output data as a definite natural ordering does not strictly exist between neighbouring clinothems, which are the volumes between any two clinoforms. Using this technique, the output for a given input vector $\mathbf{f}_n$ is represented as a vector of length equal to the number of classes, $\mathcal{C}_K$, whose entries are zero except for the element k, where the class correlates to the facies type at that location. This can be interpreted as the probability of each facies type at the discretised locations. Since the approach does not rely on discretised grids, the facies type at an exact point can be represented with the output vector. Figure 4.3 illustrates this.

Training

The problem is formulated as a classification problem where the aim is to parameterise the class probability. A multilayer perceptron, $\mathcal{S}_{\boldsymbol{\theta}}$, referred to as SurfNet, learns to map $\mathbf{f}$ to a discrete and non-continuous set $\mathcal{C}$ of K distinct classes by finding a set of parameters $\boldsymbol{\theta}$ that allow for this mapping. This is formally denoted as:

$$\mathcal{S}_{\boldsymbol{\theta}} : \mathbf{f} \rightarrow \mathcal{C} \tag{4.4}$$

where the distinct classes in $\mathcal{C}$ correspond to the three facies types identified: $\mathcal{C} = \{\text{USF, LSF and OM}\}$. The approach is similar to the use of indicator functions, where the output vector, $\mathbf{I}$ returns 1 for the index, l,

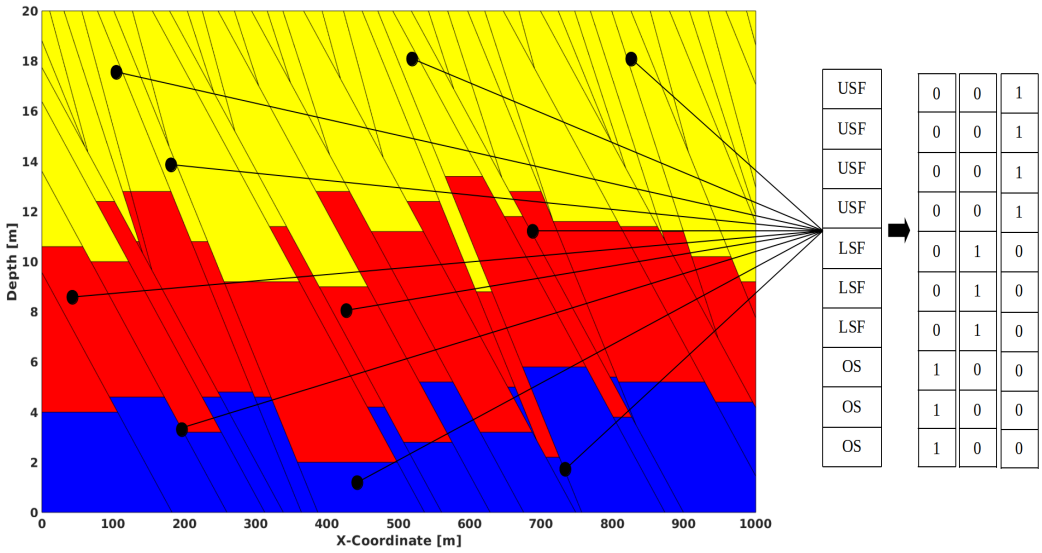


Figure 4.3: The output of the SBGM for training the network is the facies type typically derived from well logs. To represent this information in the network, one-hot encoding of the facies type for each grid along the well was used. Facies types: upper shoreface (USF), lower shoreface (LSF) and offshore mudstone (OS) deposits.

uniquely identifying the class of the facies at the point, and 0 for all other entries. This is given as

$$\mathbf{I}(\mathbf{f}) = \begin{cases} 0 & \text{if } l \neq k, \text{for } k \in \{1, \ldots, K\} \\ 1 & \text{for } l = k. \end{cases} \tag{4.5}$$

In the data-driven approach, we aim to construct a proxy for $\mathbf{I}$ from a set of samples. To obtain the facies type at any location within the SBGM, we map the feature vectors which consist of the SBGM inputs and coordinates of the point of interest to a discrete set of the facies classes, $\mathcal{C}$. The neural network learns to map a vector representation of $p + d$ variables (where d is the dimension of the samples that represent the number of coordinates in the model domain) onto a discrete vector with k entries. We apply a softmax operation, σ, to compute the probabilities of the facies type at the given point:

$$\sigma(\mathbf{I}_j^n) = \frac{e^{I_j}}{\sum_{k=1}^{K} e^{I_k}}. \tag{4.6}$$

Therefore, given a data set of N SBGM examples with randomly sampled points within the models, we solve the classification problem by finding the parameters of the network, $\boldsymbol{\theta}$, which minimise the mismatch (loss) between the neural network predictions, $\overline{\mathbf{y}}$ and the observed facies type $\mathbf{y}$. A number of optimisation techniques based on gradient schemes can be applied to iteratively optimise the network parameters (Dogo et al., 2018; Haji and Abdulazeez, 2021; Ruder, 2017). The mismatch between the predicted probabilities $\overline{I}_k^n$ and the one-hot encoded targets, I_k^n at each location is backpropagated to update/optimise $\boldsymbol{\theta}$. We define the objective function for the

optimisation, also known as the loss function, $\mathcal{L}$ as

$$\mathcal{L}(\mathbf{I}, \bar{\mathbf{I}}) = -\frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} I_k^n \log(\bar{I}_k^n) \tag{4.7}$$

where $\bar{I}_l^n \neq 0$. The entries in $\mathbf{I}$ give the probability of each class, and the final predicted facies type for each example, $\bar{y}$ is the class with the highest probability.

Network Architecture

The neural network used to learn the relationship between the SBGM parameters and facies classification at any location within the model domain is a fully connected MLP (see Figure 4.4). The size of the input layer corresponds to the number of SBGM inputs and dimensions in the models (two in this case). The inputs to the network are sequentially arranged with careful consideration of the order of generating the surfaces. This accounts for the depositional sequence in the models and serves to further constrain the network. The architecture of the network consists of an initial upsampling of the variables from 242 to 1024 and a subsequent downsampling to 3 neurons. The Rectified Linear Unit (ReLU) activation function is used in all layers except the last layer, where a softmax activation is used to transform the values of the 3 neurons (logits) into the probabilities of the three facies classes at the point of interest. The hyperparameters for optimising the weights of the network include: the weight decay, dropout probability, learning rate, momentum, and annealing. The network is trained using the ADAM (ADAptive Momentum) (Kingma and Ba, 2014) gradient-based optimiser, where each weight of the network, θ_j, is updated every k-th iteration as

$$\theta_j^{k+1} := \theta_j^k - \eta \frac{\hat{m}}{\sqrt{\hat{v}} + \epsilon}, \tag{4.8}$$

where η is the learning rate, $\hat{m}$ is the momentum, $\hat{v}$ is the velocity, and ϵ is the tolerance. Table 4.1 gives an overview of the neural network hyperparameters used for each training run.

Table 4.1: Hyperparameters used for training the network

Hyperparameter	**Value**
Learning rate	10^{-3}
Weight decay	0.00
Train-test split	0.8
Number of epochs	700
Batch size	10000
Weight initialisation	Kaiming (He et al., 2015)
Optimizer	Adam (Kingma and Ba, 2014)

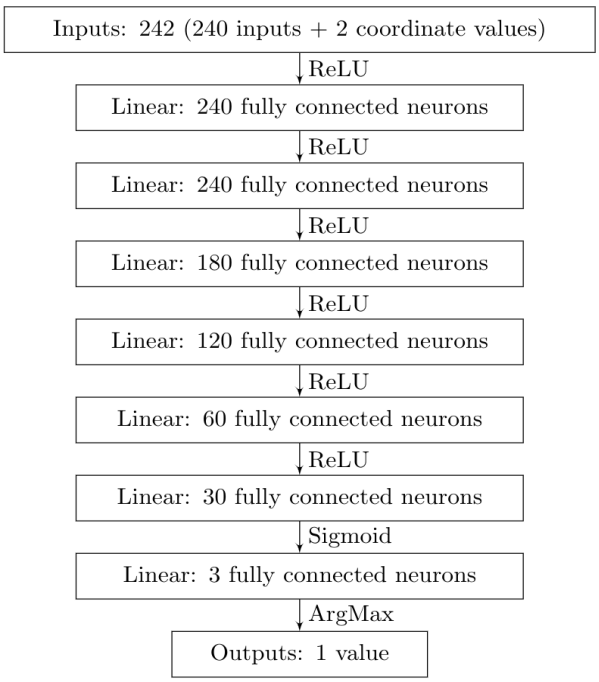


Figure 4.4: The architecture of the multilayer perceptron showing the input, hidden, and output layers along with the activation functions used for each layer.

4.3.2 Conditioning SBGMs with Neural Networks

Conditioning SBGMs involves inferring a set of parameters that generate a model whose properties of interest match observed data. This requires solving an ill-posed inverse problem where a solution may 1) not exist, 2) be non-unique, and/or 3) lack continuous dependence. We solve the inverse problem by optimising the inputs in two basic steps:

- The process of generating the SBGM is replaced with the trained network. For any point within the model domain, the trained network is capable of predicting the facies type given the SBGM inputs and the location information of the point.

- Given the observed data, the inputs of the trained network are optimised using a similar technique used to train the weights of the network. Backpropagation with stochastic gradient schemes allows the optimisation of any number of points (Rumelhart et al., 1986*a*).

With this approach, it is possible to generate surface-based geological models (SBGMs) that are calibrated to observed data. The parameters that are optimised are the inputs to the surface-based model, $\mathbf{f}$, which describe the location, orientation, and distribution of the depositional bodies within the domain. The following paragraphs describe the method in more detail.

Inverse Modelling

Given the trained network, $\mathcal{S}_\theta$, we can learn the inputs to the network $\mathbf{f}$ whose predictions match the observed data, $\mathbf{y}_{obs}$. To achieve this, the trainable parameters of the network are set to the model inputs in place of the weights as previously done. The network is then "trained" as performed in the forward modelling step with the multi-class cross-entropy loss function reformulated as

$$\mathcal{L}_c(\mathbf{I}_o, \hat{\mathbf{I}}) = -\sum_{k=1}^{K} I_o^k \log(\hat{I}^k), \tag{4.9}$$

where $\mathbf{I}_o$ is the one-hot encoded vector of the observed data (targets) and $\hat{\mathbf{I}}$ is the neural network predictions of the probabilities of each facies type from the softmax layer. By iteratively solving Equation (4.9), we can find a set of inputs $\hat{\mathbf{f}}$ that minimise the mismatch, $\mathcal{L}_c$ between the network predictions and the given observations at the points of interest, denoted as:

$$\hat{\mathbf{f}} = \underset{\mathbf{f}}{\operatorname{argmin}}\, \mathcal{L}_c(\mathbf{f}). \tag{4.10}$$

In practice, however, we condition the SBGM to data from multiple wells using several points along the well. The optimisation framework applies stochastic gradient descent (SGD) scheme (Robbins and Monro, 1951; Ruder, 2017) to condition the network inputs to randomly or sequentially selected points along the well. The use of SGD with annealing of the step size for optimising parameters of neural networks has been suggested as a robust method for accelerating convergence in SGD as noise can improve generalization (Smith et al., 2020). The modified loss function is expressed as:

$$\mathcal{L}_c(\mathbf{I}_o, \hat{\mathbf{I}}) = -\frac{1}{N_w N_v} \sum_{n_w=1}^{N_w} \sum_{n_v=1}^{N_v} \sum_{k=1}^{K} I_{o,n_w,n_v}^k \log(\hat{I}_{n_w,n_v}^k), \tag{4.11}$$

where N_w is the number of wells in the model and N_v is the number of points along the wells used in the optimisations, and the formula for updating each SBGM input, f_p is:

$$f_p^{k+1} := f_p^k - \eta \nabla_{f_p^k} \mathcal{L}_c(f_p^k). \tag{4.12}$$

Evaluation Criteria

A pre-defined criteria is required to evaluate the accuracy of conditioned models. This is important since the approach applies an iterative method in the optimisation and a stopping criteria or condition is necessary to determine well-calibrated models. To generate realisations conditioned to well data at multiple points, we run several iterations with Equation (4.11) starting with stochastic initial guesses of the inputs. The number of iterations is determined by a tolerance. The training loss in the forward modelling step is used as the tolerance

for the optimisation. A second evaluation criteria is applied to the optimised inputs from the inverse modelling. The training accuracy in the forward modelling step is used to determine if the SBGM generated from the optimised inputs is a valid conditioned realisation to be included in the ensemble of plausible models. Similar to the formula presented by Titus et al. (2022), a correlation between the facies from the conditioned SBGMs and the reference model at locations with observed data is computed. This is defined as:

$$\overline{Corr}_{model} = \frac{N_{\text{match}}}{N_{\text{model}}}, \tag{4.13}$$

where N_{match} is the number of points that match the facies types at locations along all wells in the model, and N_{model} is the number of data points from all wells in the reference model. The model correlation, $\text{Corr}_{\text{model}}$ can be interpreted as an accuracy metric. Figure 4.5 illustrates the workflow using neural networks to condition SBGMs to multiple points from wells.

Frequency of Predictions

To evaluate uncertainties, we run the algorithm presented in Figure 4.5 using different initial guesses of the SBGM inputs to obtain multiple conditioned realisations which fit the observed (reference) well data. We define a measure of the variations among the conditioned realisations at any location within the model as a frequency of prediction, f^{pred}, given as

$$f^{pred}_{x,y} = f^{max}_{x,y} - f^{min}_{x,y}, \tag{4.14}$$

where $f^{max}_{x,y}$ and $f^{min}_{x,y}$ are the maximum frequencies or number of occurrence of the facies with the highest predictions at a point, and the minimum frequencies or number of occurrence of the facies with the least predictions at a point, respectively at the given point.

4.4 Results and Discussions

The following sections present the results of the proposed method for conditioning SBGMs to data from wells with varying trajectory. Section 4.4.1 discusses the applications of the proposed grid-free approach to conditioning surface-based models, and Section 4.4.3 provides the steps for the application of the method to field cases. As explained in the methodology (see Section 4.3), we use neural networks to generate SBGMs that are conditioned to well data in two ways: 1) we train a neural network to represent the forward modelling process of generating geological models by learning the parameters that describe the relationship between the inputs to the SBGM and the facies type at any location within the model, 2) using the trained network, we optimise a randomly selected set of stochastic inputs to the SBGM to predict a given reference well data as described in Section 4.3.2.

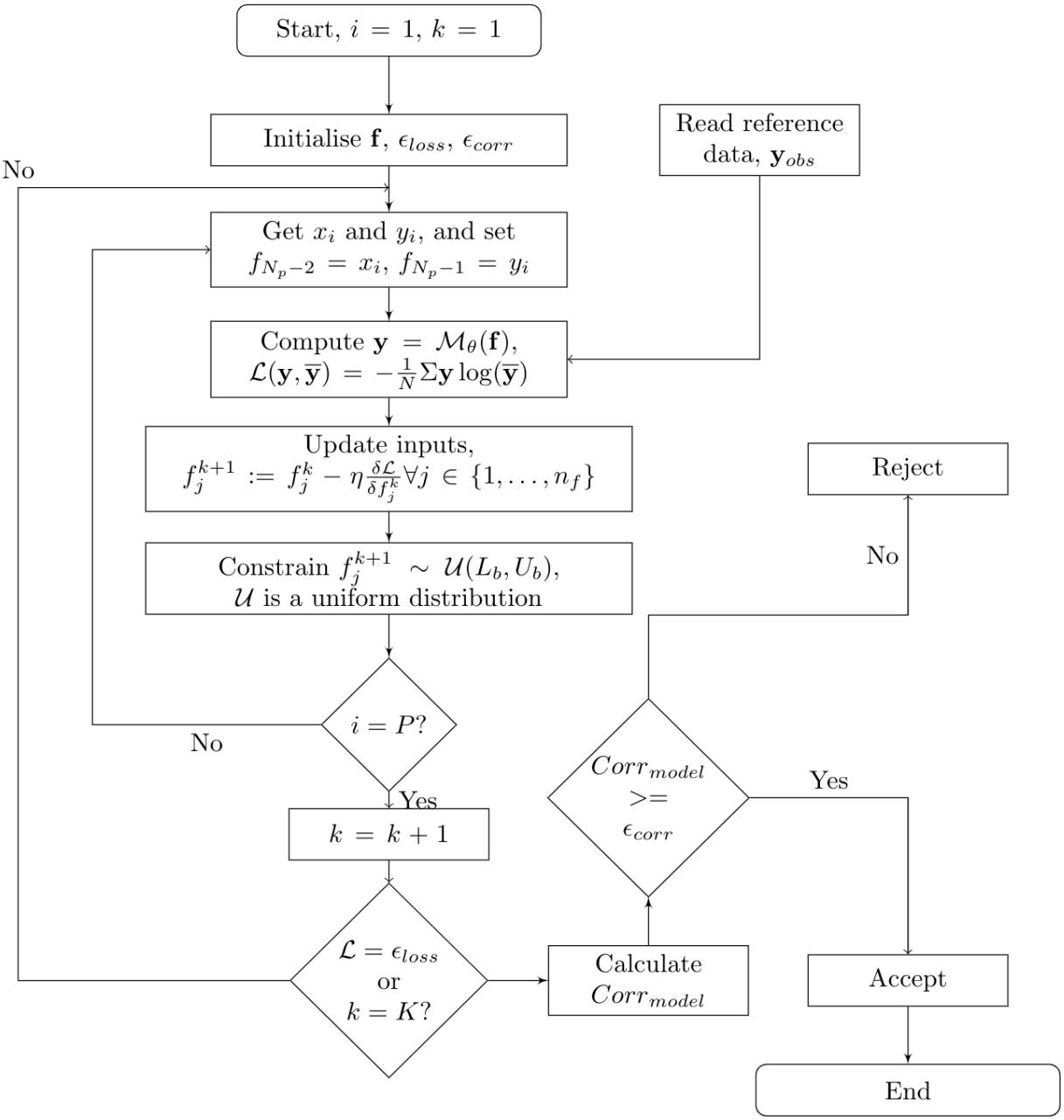


Figure 4.5: The workflow for conditioning SBGMs to data from multiple wells having an arbitrary number of data points.

4.4.1 Predicting Well Data Using Neural Networks

For the forward modelling step, we train the model with one point from 20×10^6 SBGMs with a train-test split of 0.8 in each of the 7 training cycle of 100epochs. The inputs to the neural network comprise the 240 parameterised inputs to the SBGM and 2 additional variables representing the location of each point in each model. Despite the large number of SBGMs to train the model, the number of training examples is arguably similar to or even less than other grid-based methods, where, for example, the 20×10^6 training examples will correspond to constructing a dataset of 20×10^2 images having a grid size of 100×100. We use a multilayer perceptron with 7 hidden layers in addition to the input and output layers (see Figure 4.4 and Table 4.1 for details of the network and hyperparameters used for optimising the network weights). Figure 4.6 shows the performance of the model on the training and test data over the 700 epochs. The loss is calculated from Equation (4.7) and the accuracy is the average number of correctly predicted facies types for all examples. As seen from Figure 4.6, the loss and accuracy for both the training and test data are comparable, implying that the model is not underfitting or overfitting. Table 4.2 gives a summary of the network performance.

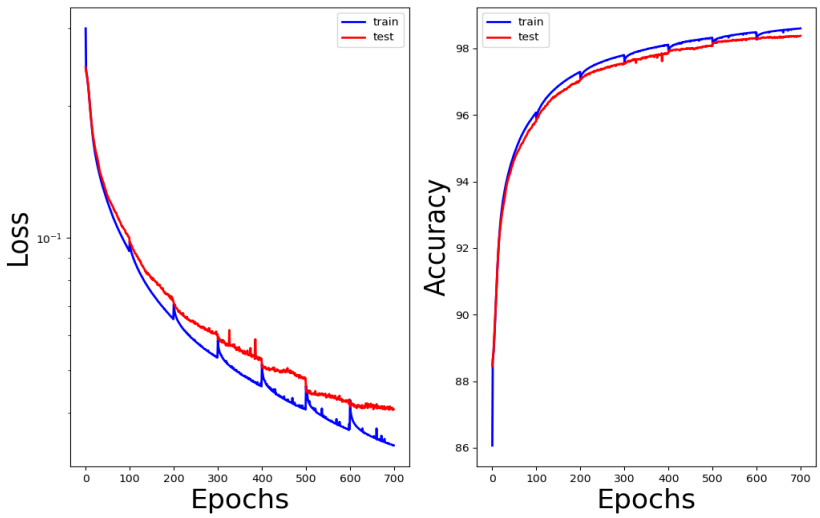


Figure 4.6: The loss and accuracy plots for the network presented in Figure 4.4 and hyperparameters given in Table 4.2 over 7 training runs each having 100 epochs, resulting in a total of 700 epochs. The curves provide information on the performance of the model on the training and test datasets. The results show the model is well fit as both datasets show similar performance.

Table 4.2: Performance metrics of the model evaluated for the training and test dataset

Performance Metric	**Value**
Final Training Loss	0.03367
Final Training Accuracy (%)	98.600
Final Test Loss	0.04071
Final Test Accuracy (%)	98.375
Training time on GPU (hrs)	25.32

The final test accuracy of ≈ 98.4% and training accuracy of ≈ 98.6% demonstrate that the network is able to predict the facies type at points within the model with high accuracy. The comparable loss between the training and test data (which consist of data from SBGM examples not included in the training dataset) indicate that the network is capable of generalising and therefore predicting the facies type for unseen data with reasonable accuracy. Figures 4.7a to 4.7f compare the ground truth and predicted facies type for three randomly generated SBGMs not used in training or testing the network to illustrate the capability of the model for generating very accurate SBGMs given the input parameters. Figures 4.8a and 4.9a show wells with different trajectories sampled from the same reference SBGM. The neural network predictions of the facies types at the well locations is compared with the ground truth in Figures 4.8b and 4.9b. The high network accuracy obtained in the forward modelling step allows for the application of the trained model in the conditioning step as previous work (Titus et al., 2022) reported that this is required for the inverse modelling in the conditioning step.

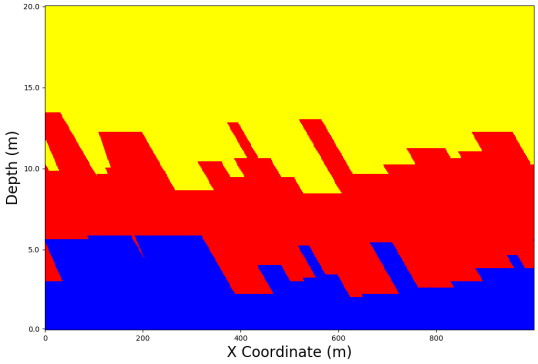
4.4.2 Conditioning SBGMs with Neural Networks

The results are presented here for conditioning SBGMs to facies data from wells with varying trajectories. The SBGM inputs used to generate the reference model given in Figure 4.10 were not used in training, testing, or validating the network. A randomly selected set of stochastic inputs is used as the initial guess for the algorithm shown in Figure 4.5. To condition the SBGM to data from all points along the reference well using the trained neural network, Equation (4.11) is used to calculate the mismatch (loss) between the facies type from the reference data and neural network prediction at points along the well. In the optimisation step, the input parameters are updated based on the mismatch calculated using SGD (see Equation (4.12)). The SBGM inputs are optimised for each point along the well and the results are improved as the loss is iteratively minimised. We use a tolerance <= 0.04, corresponding to the test loss or a maximum iteration of 500 (for the vertical wells case) and 600 (for the deviated wells case) as stopping criteria for the optimisation. The value of $\overline{Corr}_{model}$ is 98% ± 1%, corresponding to the test accuracy of the network. Figures 4.11a and 4.11b show the correlation for 100 conditioned realisations. The hyperparameters for the optimisation are given in Table 4.3.

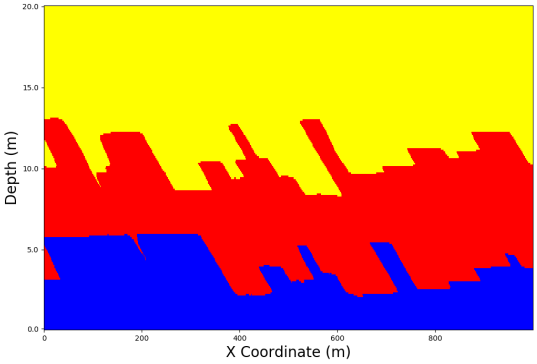
Table 4.3: Parameters for optimizing the SBGM inputs using the trained neural networks

Hyperparameter	**Value**
Learning rates	0.001
Weight decay	0.00
Maximum Iterations	1000
Tolerance	0.04±0.01
Well correlation cut-off (%)	95.0

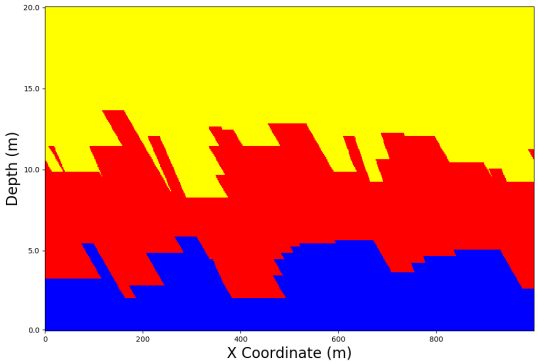
We evaluate uncertainties of the conditioned realisations for multiple wells with different trajectories (paths) using the frequency of predictions defined in Section 4.3.2 to demonstrate the applicability of the method for understanding subsurface uncertainty and risks. Using this approach, we generate the frequency of predictions



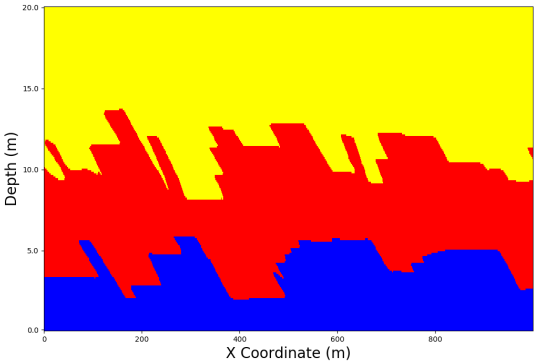
(a) Groundtruth - Example 1



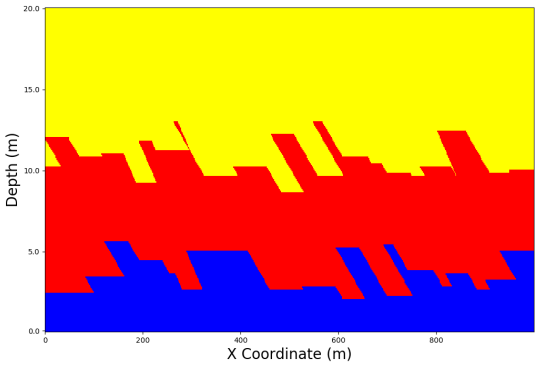
(b) Prediction - Example 1



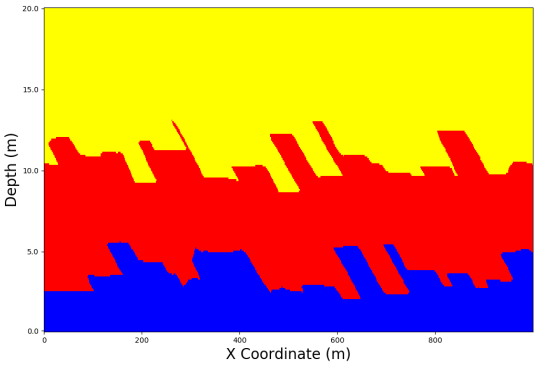
(c) Groundtruth - Example 2



(d) Prediction - Example 2

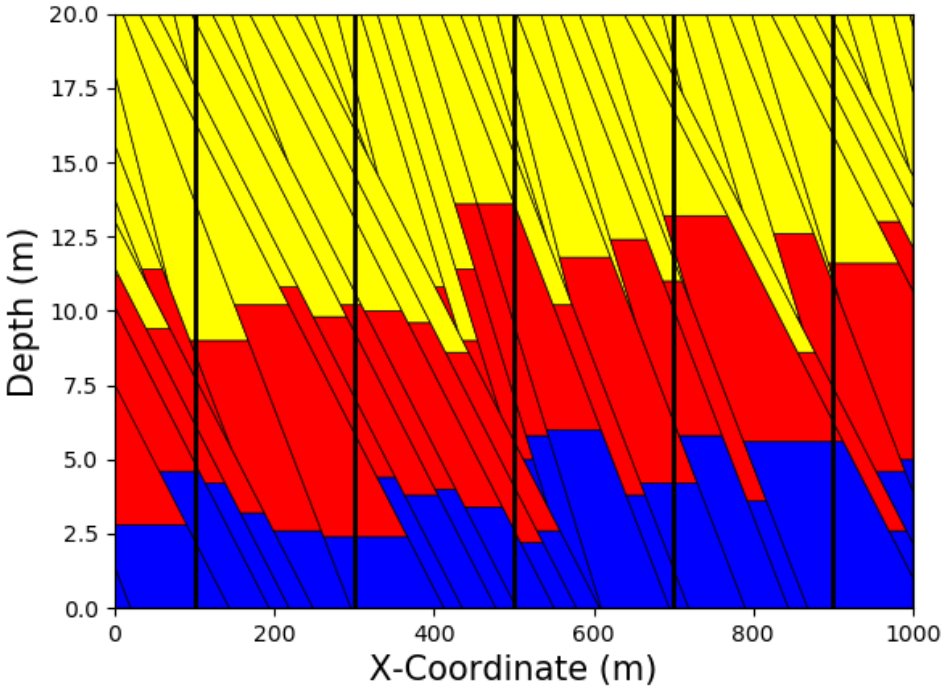


(e) Groundtruth - Example 3

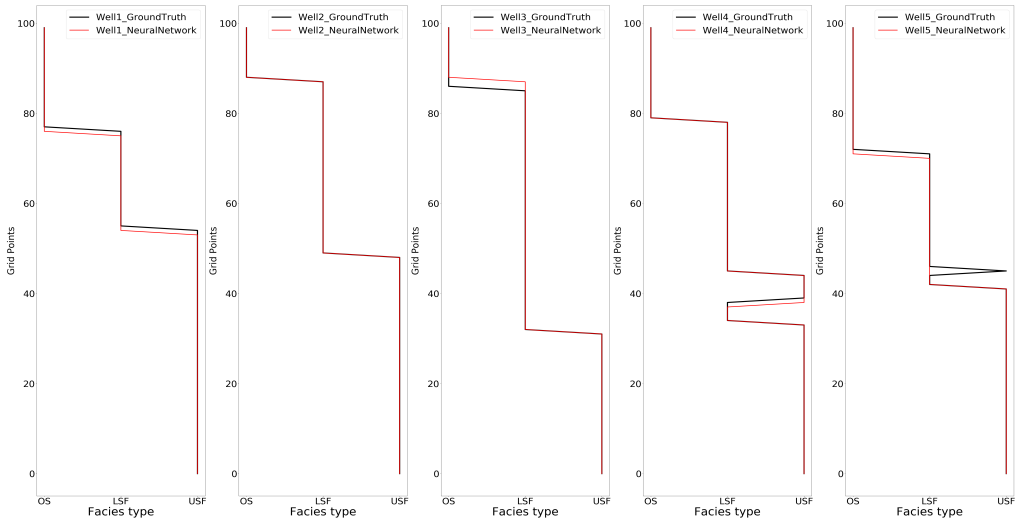


(f) Prediction - Example 3

Figure 4.7: Comparisons of ground truth SBGMs (left) with the neural network prediction (right) displayed on a 1000 by 200 grid. The models are gridded solely for visualisation and comparison, as the model is not restricted to a pre-defined grid and can accommodate any resolution.

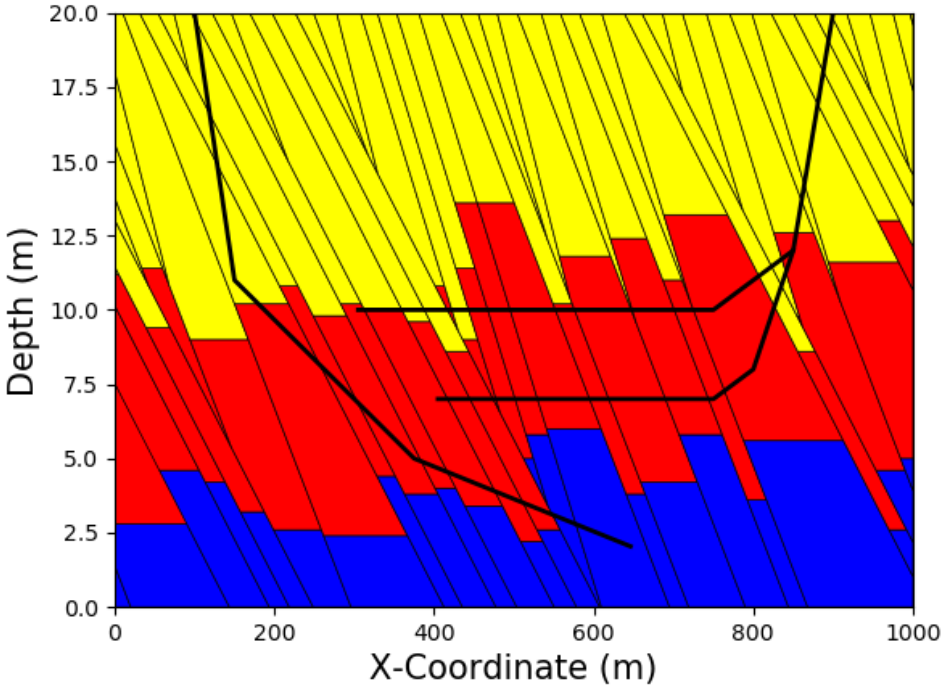


(a) Sampled vertical wells

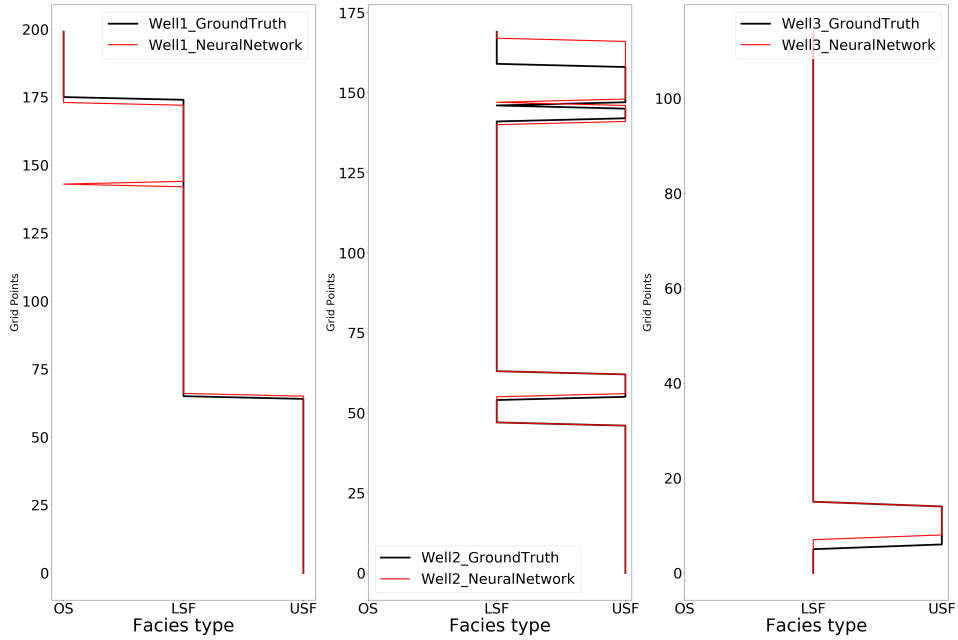


(b) A comparison of the predicted neural network facies type (SurfNet) and the actual facies type (GroundTruth) for five vertical wells shown in Figure 4.8a

Figure 4.8: The reference SBGM showing sampled vertical wells (top) and a comparison of the predicted neural network facies type (SurfNet) and the actual facies type (GroundTruth).



(a) Sampled deviated and multilateral wells



(b) A comparison of the predicted neural network facies type (SurfNet) and the actual facies type (GroundTruth) for deviated well and multilateral shown in Figure 4.9a

Figure 4.9: The reference SBGM showing sampled deviated and multilateral wells (top) and a comparison of the predicted neural network facies type (SurfNet) and the actual facies type (GroundTruth).

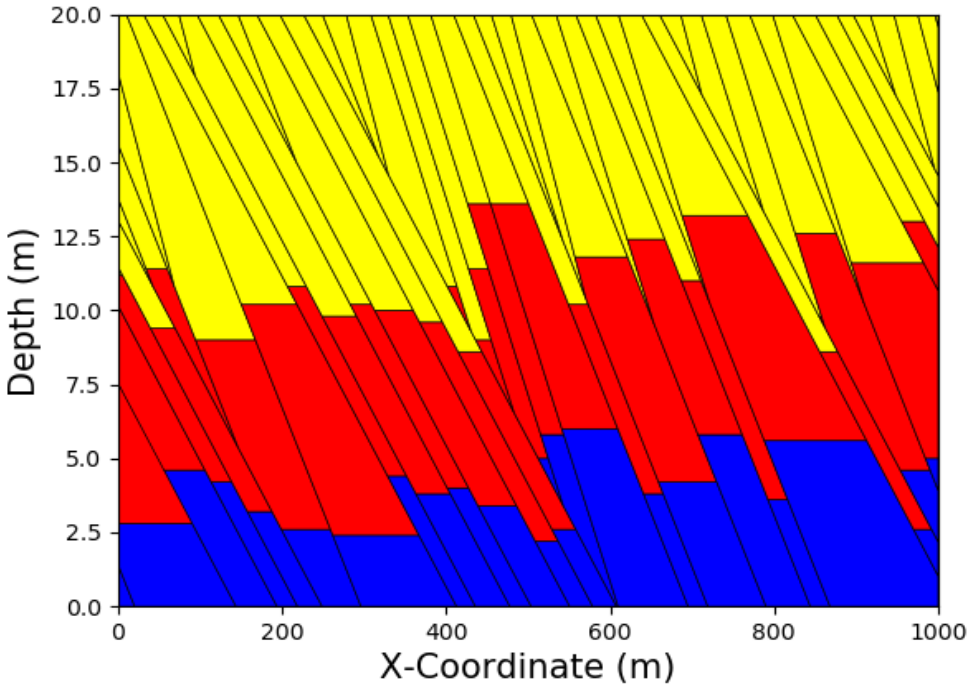
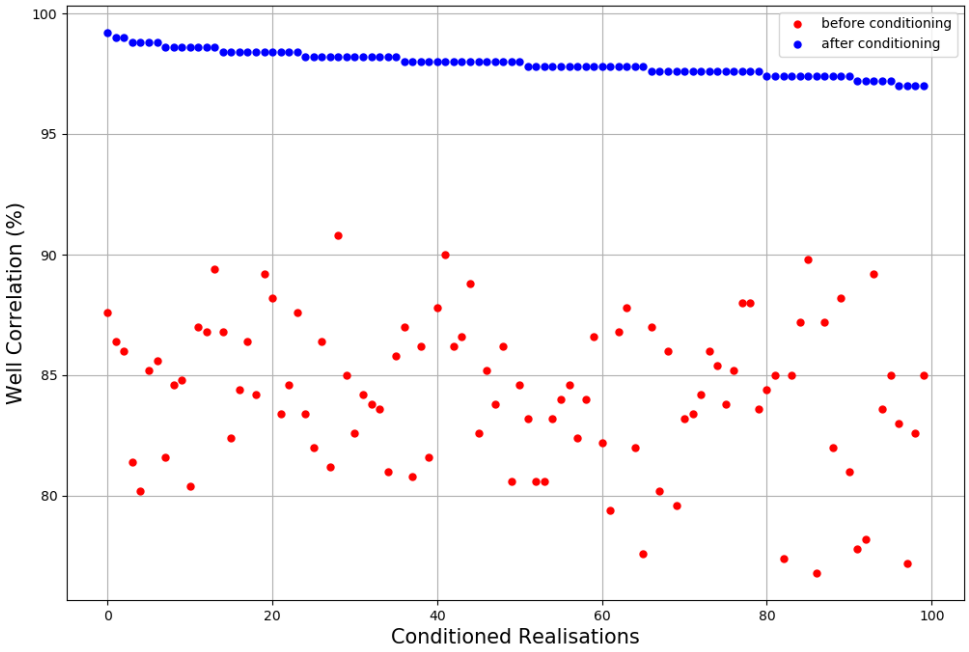


Figure 4.10: The reference surface-based geological model used for validating the predicted and optimised SBGMs.

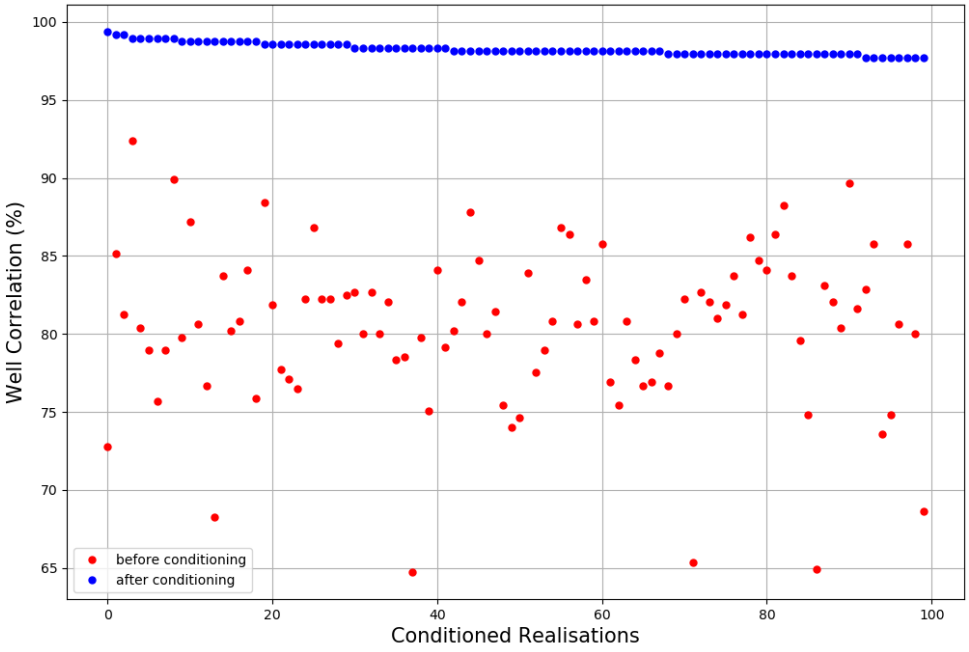
with 100 conditioned realisations.

To obtain the ensemble of models, we generate realistic inputs into the model for a range of stochastic initial guesses using the optimisation algorithm (see Figure 4.5) to calculate the mismatch between the network prediction of the facies type along the well and the reference facies types along the well. This provides input parameters that generate predictions that more accurately match the reference data compared to the initial SBGM parameters generated. We perform the optimisation in parallel to improve the speed of generating the conditioned realisations by conditioning the models simultaneously. We present the results for two cases — 1) the reference model with five vertical wells, and 2) the reference model with one deviated and one multilateral with two branches.

Figures 4.12a and 4.13a give f_{pred} at grid points within the model domain for 100 models generated using the randomly selected stochastic inputs. Each model is discretised using a grid size of $1000x60$ to visualise the results. The discretisation does not impact the accuracy of the predictions of conditioning as this is only performed to generate images for understanding locations of higher and lower uncertainties using equation (4.14). While the initialised models show that the uncertainty in the model is constant laterally, Figures 4.12b and 4.13b depict reduction in uncertainties within the model along the well trajectory where the sampled points provide the reference data for conditioning the SBGMs. This validates the applicability of the model for understanding uncertainties in subsurface geological models as locations within close proximity to well data



(a) Vertical wells



(b) Deviated and multilateral well

Figure 4.11: The correlation of each of 100 initialised models and the optimised models to the reference data at well locations, sorted in descending order of the optimised model correlations.

can be more accurately characterised with available data. The developed method provides a practical method for applying data to constrain geological models and decrease uncertainties using sparse data, particularly in regions with high uncertainty such as transition zones as the method supports the addition of more wells which can further decrease uncertainties (Titus et al., 2022). However, it is important to note that the addition of more wells translates to longer time for conditioning the models due to the increase in data points for each epoch in the SGD algorithm. For the cases considered here, each SGD step using 700 points takes $\approx$ 30seconds.

Figures 4.14b and 4.15b show the facies with the highest predictions at the grid locations for the 100 conditioned realizations shown in the frequency of predictions plots. It can be seen that geological features and trends in the reference model (Figure 4.10) emerge in the ensemble of conditioned realisations compared to the ensemble of initialised models. In contrast, Figures 4.14a and 4.15a which show the maximum predicted facies for the initialised SBGM models do not preserve patterns in the reference model.

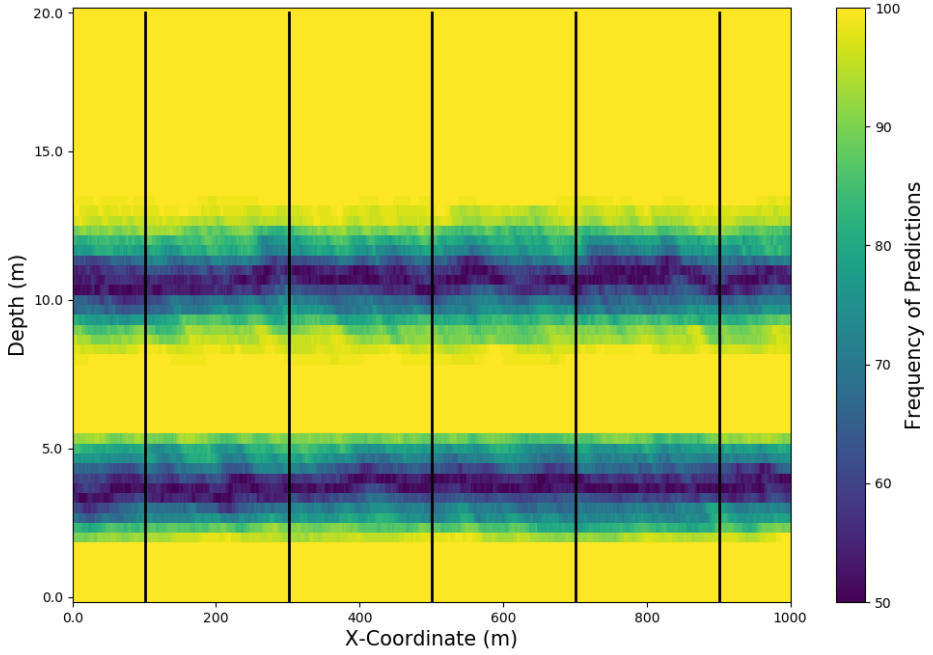
4.4.3 Considerations for Application to Field Cases

We have illustrated the approach here for 2D synthetic models. Nonetheless, the method explained in section 4.3.2 can be extended to 3D SBGMs and real field cases. The following subsections discuss the benefits of the presented method and we conclude this section by presenting a practical approach to conditioning SBGMs to well data for a reservoir deposited in a typical shallow marine depositional environment such as the Tarbet formation (Rønning and Steel (1987)), and highlight the challenges anticipated in applying the approach.

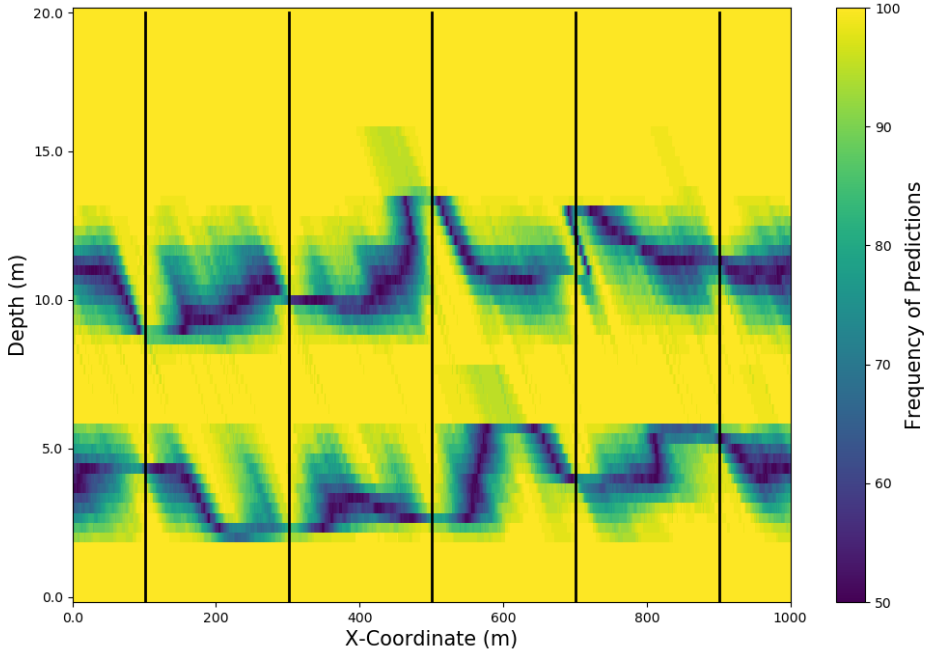
Benefits of a grid-free approach for modelling realistic subsurface reservoirs

Pixel-based approaches such as the multi-point statistics (MPS), rule-based (object-, process- or surface-based) models that adopt the use of pre-defined grids at any length scale present limitations to model resolution. In addition to the application of the grid-free approach described in Jacquemyn et al. (2019) to enable the rapid generation of multiple examples with which to train artificial neural networks, the conditioning methodology presented in Section 4.3.2 allows for greater flexibility and accuracy in capturing geological features that are representative of real field cases. The flexibility afforded by this approach can be extended to realistic and complex well trajectories such as multilaterals and highly inclined or deviated wells.

Section 4.4.2 demonstrates the method for assessing the uncertainties associated with the generated models. Analysing uncertainties is valuable for evaluating spatial variations away from control data in subsurface reservoirs and to inform decision making in subsurface field development projects.

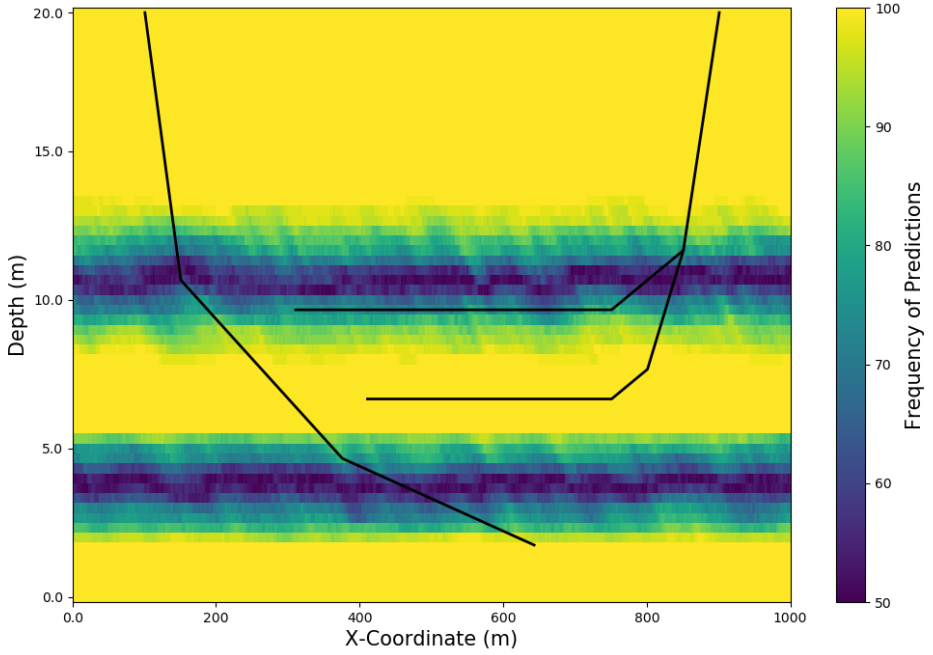


(a) Initialised models

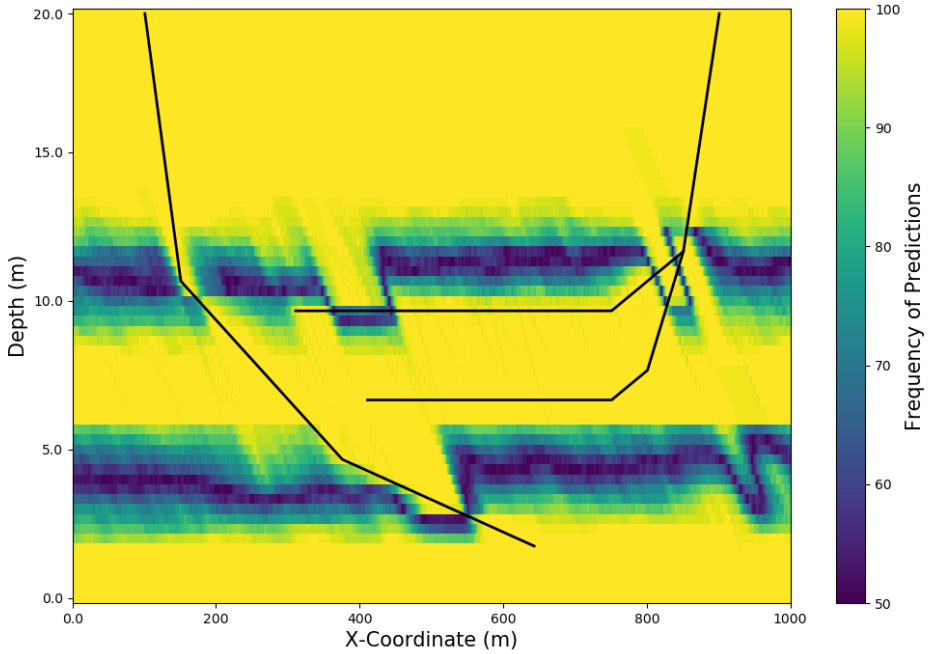


(b) Optimised models

Figure 4.12: Frequency of predictions for the initialised (top) and optimised (bottom) models with five vertical wells showing higher number of predictions of the same facies type among the 100 conditioned realisations, compared to regions away from the control data.

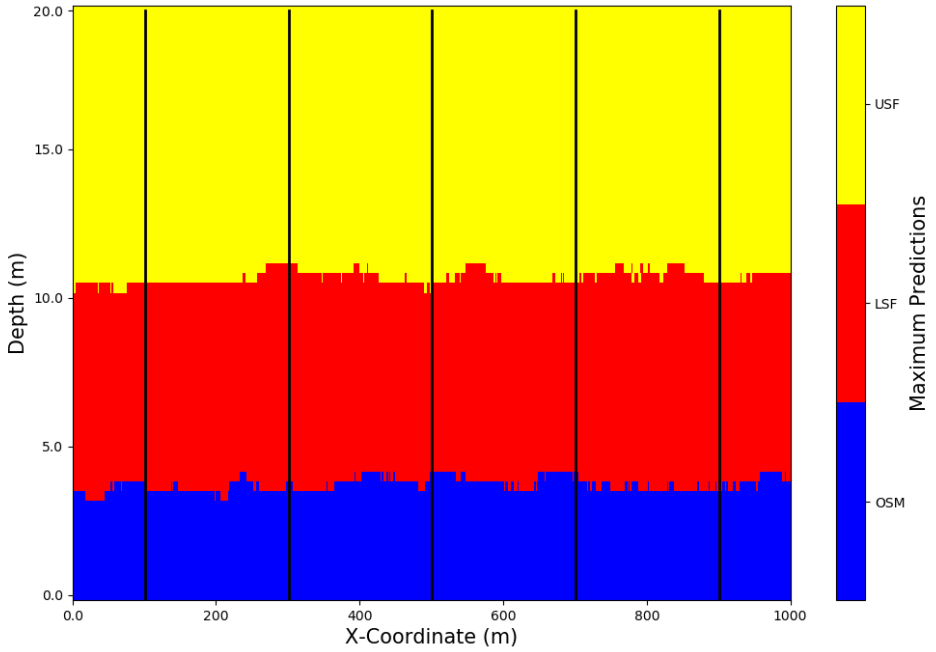


(a) Initialised models

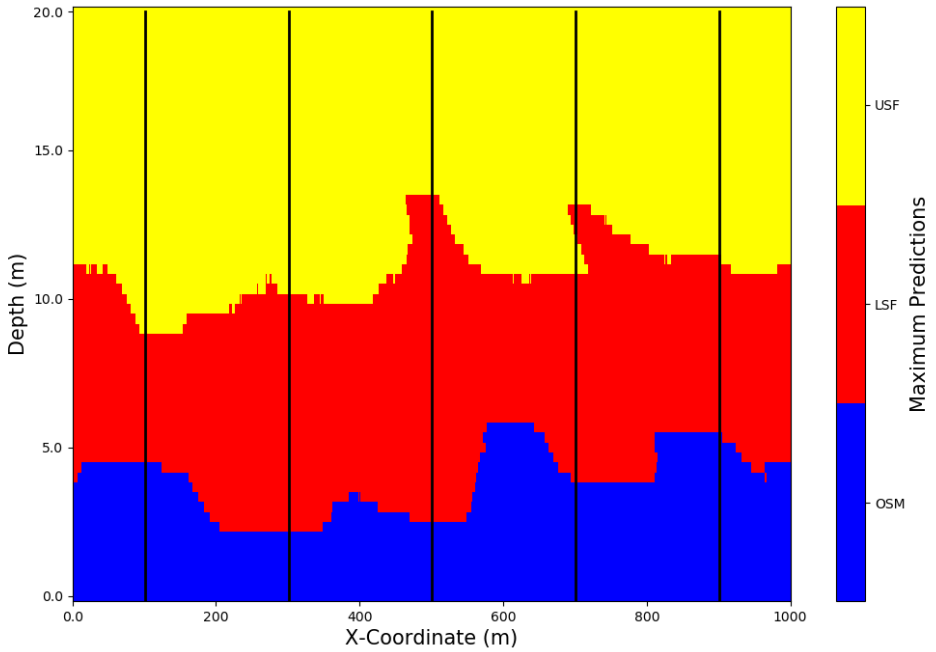


(b) Optimised models

Figure 4.13: Frequency of predictions for the initialised (top) and optimised (bottom) models with one deviated well and one multilateral well with two branches showing higher number of predictions of the same facies type among the 100 conditioned realizations, compared to regions away from the control data.

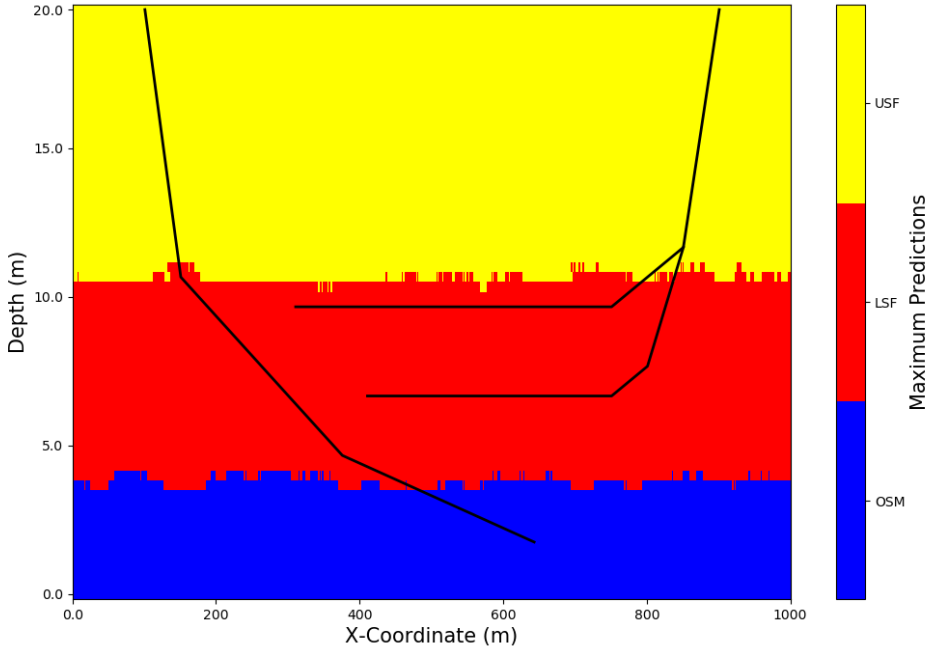


(a) Initialised models

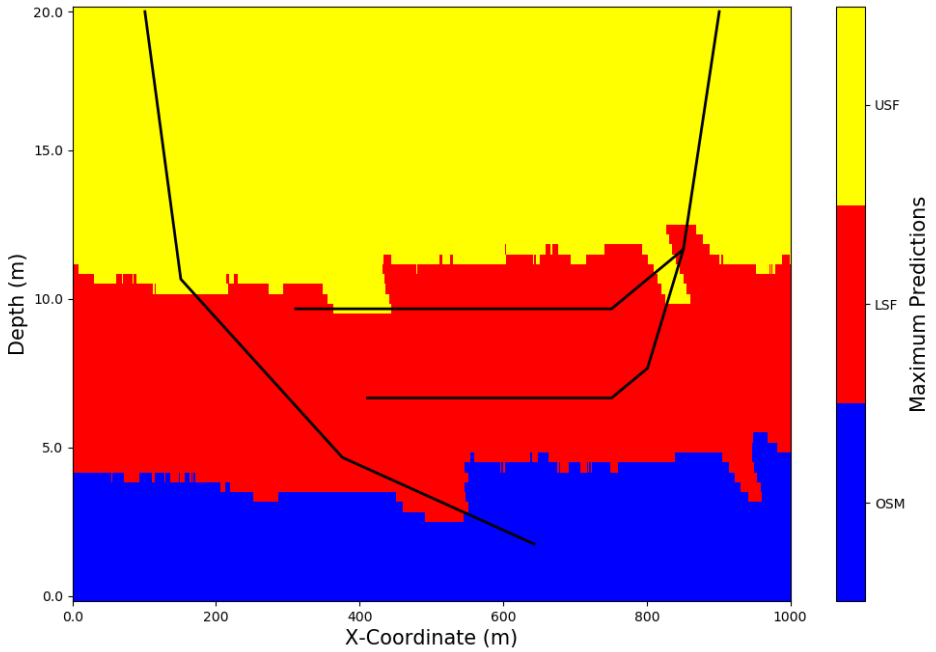


(b) Optimised models

Figure 4.14: Maximum predictions for the initialised (top) and optimised (bottom) models with five vertical wells showing good match with reference model at well locations.



(a) Initialised models



(b) Initialised models

Figure 4.15: Maximum predictions for the initialised (top) and optimised (bottom) models with one deviated well and one multilateral well with two branches deviated wells showing good match with reference model at well locations.

Limitations and challenges for real field cases

In contrast to the applications to synthetic models demonstrated in this work, the computational requirements for real field cases will be greater. This is due to the longer run time for generating SBGMs that represent complex geometries such as channelised fluvial sandbodies, structural configurations of folds and faults, and highly inclined well trajectories. This also results in a wide range of facies codes that are representative of subsurface geology by modelling petrophysical heterogeneity using discrete volumes of homogeneous properties that depict geologically consistent domains. Furthermore, increased memory requirements for storing training data due to the increased number of model inputs and more facies associations characteristic of field cases. Importantly, deeper or more complex neural networks will be required to train the model for representing SBGMs of real cases. These networks will be required to achieve training and testing accuracy particularly at stratigraphic boundaries where uncertainties are highest, as seen in the case for deviated wells, where there is an equal likelihood of locations at geological boundaries to be classified into any of the contact geological categories.

However, it is important to note that the challenges outlined here are not peculiar to SBGMs or the method applied to conditioning geological models. We have demonstrated that the use of SBGMs and neural networks for conditioning these models to well data presents opportunities to solve some of the most difficult problems in reservoir modelling. Furthermore, in comparison to other methods for generating subsurface reservoir models, the use of parametric, grid-free surface-based models for representing geological heterogeneity of interest such as features which influence the flow conditions has many advantages which outweigh the anticipated computational limitations. For example, the number of surfaces required to model realistic reservoirs using surfaces has been shown by Osman et al. (2021) to be low (a few tens in some cases) to generate representative models. This is because although petrophysical properties could vary on a point-to-point basis, such variability is not necessary to capture fluid flow in geological reservoir models. The computationally convenient approach to conditioning SBGMs to well data presented here can contribute to the widespread application of surface-based modelling.

4.5 Conclusions

An approach for conditioning grid-free surface-based geological models (SBGMs) to data from wells with varying trajectory has been proposed, demonstrated, and evaluated. The method is based on the use of neural networks, specifically multilayer perceptrons to represent the underlying process for generating the SBGMs in the forward modelling step. The trained network replaces the SBGM in the inverse modelling step. Using stochastic gradient descent, the network inputs are iteratively updated to generate the inputs to an SBGM that honour observations from wells. The proposed methodology provides a practical approach for conditioning geological models to data from wells with realistic trajectory, such as deviated, horizontal, and multilateral well paths.

Uncertainties among the conditioned models are evaluated, where a number of plausible geological models are generated and used to evaluate spatial variations and uncertainties by computing the frequency of predictions for the conditioned realisations.

The results show that neural networks can be applied to predict facies type at any location within the SBGM with remarkable success. The network is also capable of inferring a set of input parameters that generate a model of a shallow marine reservoir that honours observations at locations with well data. However, further work is required to demonstrate the approach for more realistic 3D geological models, taking into account the computational challenges to be anticipated.

Chapter 5

Data Assimilation with Machine Learning for History Matching

5.1 Abstract

A key step in developing representative models of subsurface reservoirs is calibrating fluid flow models with observed data to produce physically consistent numerical models of the reservoir. A major limitation to the widespread practical application of iterative data assimilation methods in history matching is the large dimensionality of subsurface models, particularly in grid-based models where thousands of cells are used to capture geological heterogeneity, thereby making the history matching process impractical in some cases. Our proposed workflow builds on a previous neural network-based methodology for conditioning parametric, grid-free surface-based geological models (SBGMs) to static data.

This study focuses on integrating dynamic data in a machine learning ensemble-based approach for data assimilation in the history matching process to improve the accuracy of reservoir models for forecasting and reservoir management strategies. By first ensuring that the geological model is conditioned to well log data, the solution space for history matching is constrained. We use adversarial learning to impose a data distribution and enable the network to gain knowledge of the statistical properties of the training dataset. The history matching process involves two main steps: 1)a neural network is trained to learn the underlying relationship between geological inputs and dynamic outputs in a forward modelling process, 2)the trained neural network replaces the forward model and the inputs are optimised by minimising the mismatch between the observed and predicted dynamic data from the network. The resulting model therefore honours geological information and production data. Using this approach, we can generate geological models that are consistent with geological processes, geophysical information and fluid flow data from wells.

5.2 Introduction

The safe, efficient, and sustainable production of energy from subsurface resources, such as geothermal systems, can contribute significantly to addressing global energy poverty and achieving the sustainable development goals (SDGs) (Vargas et al., 2022). Generating representative models of subsurface reservoirs will significantly improve fluid recovery and reservoir performance. Reservoir simulation models are built from an initial set of parameters, which are tuned to produce models that honour the observed fluid flow data. The conditioned models are then used for predicting, optimising, developing, and improving the recovery of subsurface fluids. In geoscience and geophysics, data assimilation has been widely used as a powerful tool for combining observations with model output to obtain more realistic representations of physical systems (Asch et al., 2016). Integrating observed data in building reservoir models is a key step in validating and correcting numerical simulations due to the uncertainty in model parameters.

In subsurface reservoir systems, data assimilation is commonly referred to as history matching. The process of history matching involves finding the set of input parameters to the reservoir simulation (such as the petrophysical properties) that minimise the difference between the observed performance and the simulation output over a historic production period (Chen and Oliver, 2011). Similar to data assimilation, history matching requires building models, of a subsurface reservoir in this case, that closely match the past behaviour described by historical data, commonly acquired from producing wells. However, the simulation of reservoir-well systems involves solving a computationally demanding problem given the spatial heterogeneity and anisotropy of reservoir rocks. This is a complex process that involves modelling large-scale, high-dimensional, nonlinear physical behaviour that characterises the flow of fluids in the subsurface. Consequently, reservoir history matching is considered an ill-posed inverse modelling exercise, where the aim is to derive the underlying spatial distribution of petrophysical properties which correspond to the observed fluid flow behaviour (Abdulrazzaq and Hasan, 2023). History-matched models can be used to optimise well placement and geometry, maximise fluid recovery, and study reservoir behaviour under multiple scenarios.

5.2.1 Data Assimilation for History Matching Subsurface Reservoir Models

Over the past two decades, data assimilation has attracted huge attention in geoscience and specifically in reservoir history matching for accumulating observed data into subsurface models for recursively updating the model state (Asch et al., 2016). The rationale for the application of data assimilation techniques in reservoir modelling is attributed to the characteristic limited data, large-scale nonlinearities, and high uncertainty common to other geophysical systems that apply data assimilation. In reservoir history matching, static data such as petrophysical properties from well logs (usually porosity and permeability) represent the model parameters, while dynamic data such as production data (well flow rate, temperature, pressure, and fluid saturation) are

the observed data integrated into reservoir models. The main objective of history matching is to calibrate subsurface reservoir simulation models with measured field production data to obtain a reliable model of the fluid flow process (Schiozer, 1999).

Traditional history matching involves manually adjusting static variables and running the reservoir simulator to ensure that the results match the measured dynamic data. This requires successively altering initial inputs and assumptions to geological models and generating multiple equivalent solutions to obtain a reasonable match at different stages of production and under different operating conditions (Brantson et al., 2018). The time requirements of traditional history matching, being a trial and error process, led to the development of automatic history matching for greater accuracy and less time-consuming calculations (Chen and Oliver, 2011). In automatic history matching, an objective function, designed as the mismatch between the measurements and model data is minimised in an optimisation framework using techniques such as gradient-based optimisation, response surface, genetic algorithms, particle swarm optimisation, simulated annealing, and variogram-based sensitivity to solve the inverse problem (Brantson et al., 2018). Data assimilation methods, particularly ensemble techniques (ensemble Kalman filter and ensemble smoother), have been widely employed in reservoir engineering to generate multiple reservoir model realisations that approximately characterise the reservoir (Emerick, 2016, 2018; Evensen, 2003). However, the computational expense, time requirements, and quality of results still pose a limitation to the application of these methods, for example, when considering forward-looking real-time reservoir management.

In this study, we adopt a data-driven approach to solving the history matching problem, where data assimilation is performed using machine learning to iteratively match the prediction from reservoir models to observed fluid flow data. This is similar to the iterative ensemble smoothing methods for data assimilation. The use of data-driven methods, and particularly machine learning, has been thoroughly explored in reservoir history matching over the past twenty years. Integrating machine learning in data assimilation improves the efficiency and accuracy of numerical simulations even for real field cases by interpreting machine learning as a dynamical system and applying data assimilation methods to train machine learning models (Buizza et al., 2022). This provides capabilities for more robust data assimilation and, by extension, history matching in reservoir simulation.

5.2.2 Machine Learning in Subsurface Reservoirs for History Matching

Advancements in big data and machine learning, such as automatic differentiation for backpropagation and high-dimensional parameter optimisation, and graphics processing units (GPUs) for massive parallelisation have presented opportunities for developing techniques that enable the application of machine learning and artificial intelligence for solving the history matching problem. Neural networks have been applied with significant success to solve ill-posed inverse problems in geoscience characterised by sparse data, spatio-temporal hetero-

geneity, multiresolution data, high dimensionality, noise, and uncertainty (Karpatne et al., 2019). In reservoir engineering, artificial neural networks have been used to predict the performance of CO_2 sequestration (Gorucu et al., 2005), pressure transient analysis in fracture networks (Alajmi and Ertekin, 2007), shale gas reservoir modelling (Klie and Florez, 2018), surrogate reservoir modelling (Shahkarami et al., 2015), and generation of geological models (Chan and Elsheikh, 2017; Laloy et al., 2018; Pan et al., 2021; Zhang et al., 2019).

Using a generalised regression neural network (GRNN), a one-pass learning algorithm based on non-parametric regression (Specht, 1991), Brantson et al. (2018) performed history matching of tight gas carbonate reservoir production profiles, although the GRNN failed to match the test data. Srinivasan et al. (2021) applied machine learning for history matching of unconventional reservoirs using physics-informed machine learning to capture and preserve underlying physical phenomena and mechanisms of fluid flow in complex reservoirs such as fractured shale gas reservoirs. The authors also demonstrated the application of transfer learning, where fast low-fidelity models which provide the advantage of speed are combined with slower high-fidelity models to improve the accuracy of forecasting. Li and Misra (2021) applied reinforcement learning (deep Q network (DQN) and deep deterministic policy gradients (DDPG)) to replace the process of exploring the parameter space and running the inversion to predict reservoir permeability for history matching. The method involves iteratively adjusting the reservoir parameters and simulating the production history, which can be time-consuming for more complex and realistic reservoir models than the simple grid-based model presented in the study. Using long short-term memory (LSTM) networks, Pal (2021) trained a model on historical data from injection and production wells for forecasting production of oil and water from the wells. A key limitation of the method is the large amount of continuous production and injection data required to obtain high accuracy. Canchumuni et al. (2020) used generative models to construct parameterisations of the reservoir model that are updated using ensemble smooother with multiple data assimilation (ES-MDA). The authors noted the training time of the networks as a limiting factor considering the model used was relatively small. Jo et al. (2021) proposed the use of GAN-EnKF for assimilating dynamic fluid data into reservoir models where the generative adversarial network (GAN) enables the parameterisation of high dimensional features for updating the ensembles for iterative assimilation in the ensemble Kalman filter (EnKF).

We have demonstrated in a previous work (Titus et al., 2022), the capability of neural networks for conditioning surface-based geological models to well logs. The resulting geological model can be discretised for fluid flow simulation using unstructured meshes to preserve the architecture of the surfaces, capture key flow processes, and allow for the representation of heterogeneity over multiple length-scales (Jackson et al., 2015). In this work, we extend the conditioning framework to history matching fluid flow models simulated using conditioned geological models. We explore the use of neural networks, and in particular the concept of adversarial learning for history matching geologically-constrained subsurface models to production data by assimilating dynamic information using machine learning for data assimilation. The benefit afforded by the adversarial learning

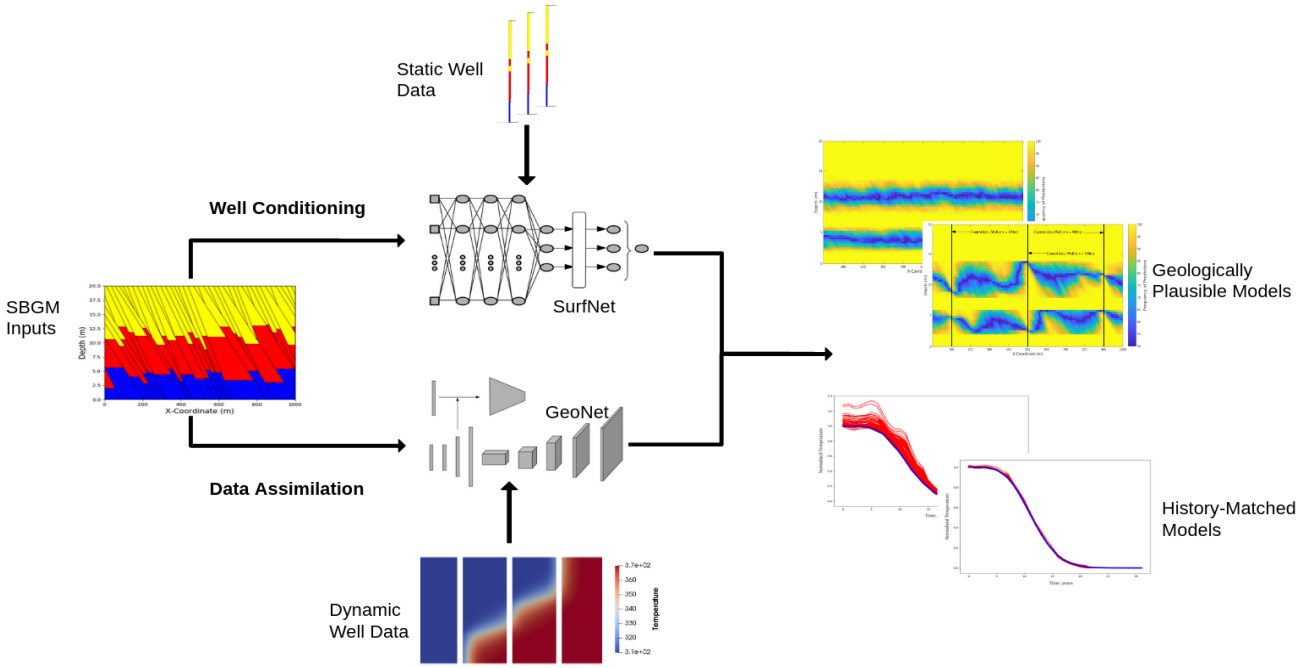


Figure 5.1: The data assimilation framework for history matching geologically constrained models proposed in this work.

allows for the use of small datasets for training the network and learning the relationship between the SBGM input parameters and production data by preserving the geometric and statistical characteristics of the training dataset through the pre-conditioning of the models to static well data. Similar to Li et al. (2021), a limit to the domain and data distribution has now been imposed in the well conditioning step and can allow for the use of small datasets for training the model. Figure 5.1 illustrates the proposed framework. We summarise our workflow as follows:

1. Generate an ensemble of geological models that have been conditioned to static data (facies type from well logs)

2. Train a neural network to map the inputs to the conditioned geological models to observed production data at the well

3. Using the trained network, perform history matching by generating the inputs to the geological model that match a given reference production data in an inverse modelling approach. This is the data assimilation step.

4. Assess the uncertainties associated with the conditioned geological and history-matched dynamic models.

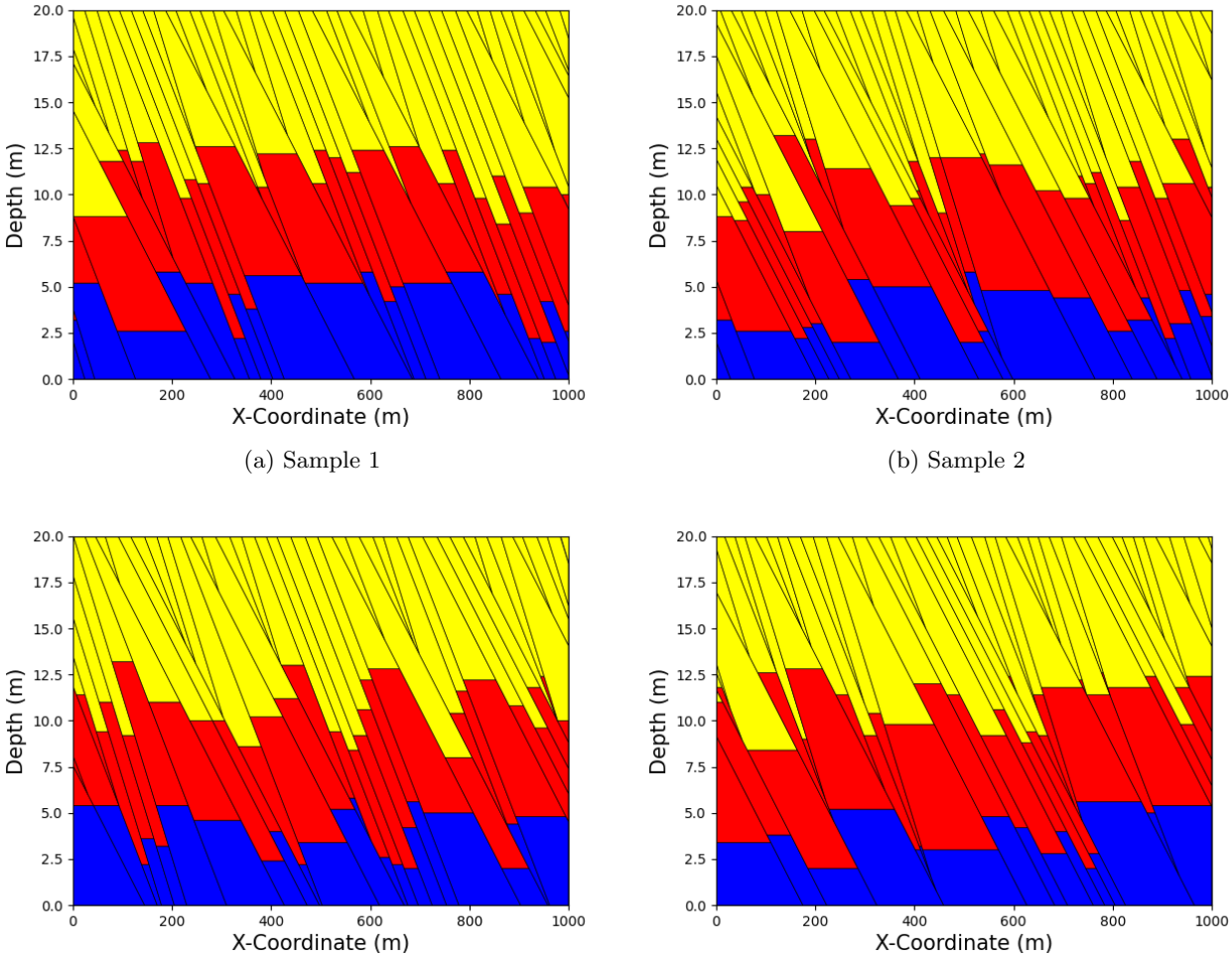
5.3 Methodology

The following sections describe the proposed method by first discussing the importance of SBGMs for generating realistic geological features for a 2D shallow marine depositional environment, along with a brief summary of the method for conditioning these geological models to static data in Section 5.3.1. Section 5.3.2 describes the fluid flow models used in reservoir simulation. Section 5.3.3 presents the theoretical framework for training the adversarial network and explains the technique applied for running the inversion to generate optimised inputs to the geological modelling that produce reservoir models that match a given reference production data. The application of the methodology is demonstrated in Sections 5.3.3 to 5.3.4 and the results are presented in Section 5.4. A discussion of the outcomes from this study and the considerations for future work is presented in Section 5.4.4 preceding the conclusions.

5.3.1 Conditioning SBGMs to Well Data

The first step in history matching involves building numerical models that simulate the dynamic behaviour of the reservoir. These models are developed from static geological models that represent the spatial distributions of petrophysical properties. The numerical models generally do not represent the observed performance or characteristics of the reservoir and must be calibrated to production data such as fluid flow rate. In this study, we perform history matching of reservoir models that are pre-conditioned to data from well logs. Ensuring that the geological models used honor static data serves as a constraint that can improve/accelerate the history matching process. The following paragraphs describes the geological models and the process of conditioning these to well data.

Surface-based geological models (SBGMs) predict depositional sequences and model geological features using surfaces that define volumes within which petrophysical properties are homogeneous (Jackson et al., 2014). If heterogeneity exists within these volumes, they are further subdivided by additional surfaces (Jackson et al., 2015). These models can be deterministic and/or stochastic, empirical and/or theoretical (Karssenberg et al., 2001). We use SBGMs that model geometries with parametric, grid-free surfaces using NURBS interpolation to describe even complex geological features and preserve geological realism such as the spatial continuity of channelised bodies (Jacquemyn et al., 2019). These SBGMs require fewer parameters compared to existing conventional geostatistical or other rule-based methods, and are not constrained to a pre-defined grid over any length-scale. Thus, the resolution of the model is not impacted. In this study, our approach is demonstrated with a 2D SBGM of a subsurface reservoir in a shallow marine depositional environment. This is a simplified geological model with a 20m thick and 1000m long single parasequence including dipping clinoforms and facies boundaries as shown in Figures 5.2a to 5.2d. The parameterised inputs to the SBGMs define geological geometries at multiple levels of surface hierarchy. These inputs are derived from stochastically sampled variables with uniform



(a) Sample 1

(b) Sample 2

(c) Sample 3

(d) Sample 4

Figure 5.2: Examples of 2D shallow marine surface-based geological models with one parasequence and 60 clinoforms (∼50x vertical exaggeration) generated using the method developed by Jacquemyn et al. (2019). Three facies are shown: Upper Shoreface (USF; yellow), Lower Shoreface (LSF) and Offshore Mudstone (OM).

distributions that describe the shape, orientation, and relative location of the surfaces (depositional or erosional).

In previous work, we demonstrated the capability of neural networks for conditioning these SBGMs to static well data (Titus et al., 2022). Information from well logs (such as facies associations) can be used to calibrate geological models. We condition SBGMs to observed data at the well location using neural networks to map the SBGM inputs, $\mathbf{p}$, to the facies type at well locations, $\tilde{\mathbf{y}}$, in the forward modelling step, and to generate the optimised SBGM inputs, $\mathbf{p}^*$, using a backpropagation algorithm to update SBGM inputs. The resulting SBGM generated from the updated inputs has been shown to honour the observed facies type along the well. Generating the inputs to the SBGM modelling process that will produce models that honour the observed geological information at the well location is an inverse modelling process. Given the trained network, $\mathcal{S}$, we find a set of input feature vectors $\mathbf{p}^*$ whose outputs from the network, $\tilde{\mathbf{y}}$, match the reference data or observations, $\mathbf{y}_{obs}$ by minimising the mismatch, $\mathcal{L}$, between the network prediction and the given observations.

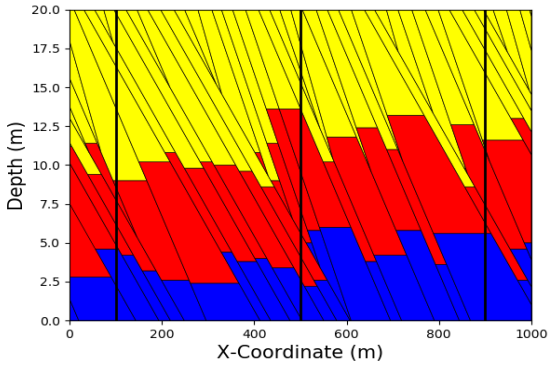


Figure 5.3: The reference surface-based geological model used for history matching with three wells located at x=100m, 500m and 900m.

This requires minimising the objective function $\mathcal{J}_c$ defined as:

$$\mathcal{J}_c(\mathbf{p}) = \mathcal{L}(\mathbf{z}_{i,obs}, \tilde{\mathbf{z}}_\mathbf{i}) := -\frac{1}{n_p}\sum_{i=1}^{n_p}\left(z^l_{i,obs}\log(\tilde{z}^l_i)\right), \tag{5.1}$$

where $\mathbf{z}_{i,obs}$ is the one-hot encoding of the observed facies type at each discretized location, n_p is the number of surface-based model input parameters, l is such that $z^l_{i,obs} \neq 0$, the mismatch or loss $\mathcal{L}$ remains the multi-class cross-entropy used in training the model and $\tilde{\mathbf{z}}_\mathbf{i}$ is the predicted facies type by the trained network at that location. In this work, we apply the trained network, SurfNet3, from Titus et al. (2022), hereafter referred to as SurfNet, denoted as $\mathcal{S}$, to perform well conditioning in both the data generation and inversion steps. We refer the reader to Titus et al. (2022) for more details on the method of conditioning SBGMs to well data using neural networks.

Using this approach allows for the generation of geological models that are consistent with the geological modelling process as features are well connected, relationships and the hierarchical arrangement of surfaces remain unchanged, and spatial continuity is preserved. Furthermore, the use of parametric, grid-free SBGMs that require only few parameters to model makes this class of geological models suitable candidates for training machine learning models for data assimilation due to the reduced computational expense of training the network even for complex geological geometries and architecture.

5.3.2 Fluid Flow Modelling

Reservoir simulation models are derived from modelling fluid flow in geological models by solving numerical problems that describe the physical behaviour of fluids in the subsurface. Since geological models which use parameterised architectural geometries to describe geological features provide superior geological realism compared to grid-based methods (Jackson et al., 2015), we use SBGMs in this study to represent geometries and heterogeneity encountered in shallow marine depositional environments where the surfaces define discrete rock

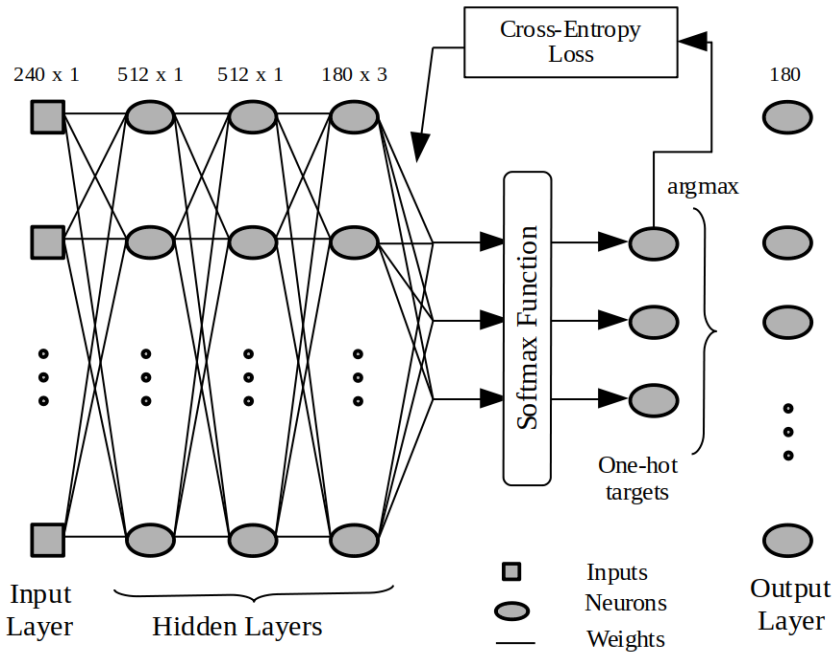


Figure 5.4: The neural network, SurfNet, $\mathcal{S}$ used in conditioning SBGMs to well data. The multi-layer perceptron maps a vector representation of the 240 SBGM input variables onto a discrete vector of the facies type with the maximum probability at all discretized location, $\hat{\mathbf{y}}$, from Titus et al. (2022).

volumes within which petrophysical properties are assumed to be uniform.

Here, we demonstrate our method for history matching in geothermal reservoirs. The flow of fluids in the subsurface is simulated by solving a set of non-linear partial differential equations that describe the transient behaviour of the fluid phases and the governing heat flow equations based on a finite element method (FEM). In this approach, velocity is discretised using finite elements (FE) while pressure and temperature are discretised using control volumes (CVs) with a dual mesh created based on the double control volume finite element method (DCVFEM) proposed in Salinas et al. (2018) to improve convergence rate. Petrophysical properties such as permeability and porosity are represented element-wise. Thermal conductivity, heat capacity and rock density are represented CV-wise. Fluid properties such as viscosity and density are represented CV-wise. Time is discretised using an implicit θ-method.

We represent the production and injection wells as the boundaries of the domain, where the injector, I, is at 0 m and the producer, P, is at 1000 m of the domain. Within the domain are three exploratory wells which serve as control points for data collection. For the reservoir simulation, chemical reactions, phase changes, geomechanical effects, multiphase flow, and other complex physical behaviour present in typical geothermal settings are neglected. Assuming conservation of mass and momentum, the mass balance equation is given as:

$$\frac{\partial(\phi\rho_f)}{\partial t} + \nabla \cdot \boldsymbol{u}\rho_f = s_c, \tag{5.2}$$

where ϕ is the porosity, ρ_f is the pressure and temperature-dependent fluid density, $\boldsymbol{u}$ is the Darcy velocity, and s_c is a source term. The momentum balance equation is:

$$\mu(\boldsymbol{K})^{-1}\boldsymbol{u} = \nabla p + \rho_f \boldsymbol{g} + s_m, \tag{5.3}$$

where $\boldsymbol{g}$ is the vector of the gravitational acceleration, $\boldsymbol{K}$ is the permeability tensor, p is the pressure, μ the fluid viscosity, s_m is a source term in the momentum equation. The temperature is described by the energy balance equation:

$$\xi \frac{\partial T}{\partial t} = -\nabla \cdot (\kappa \nabla T) - \phi \rho_f C_{pf} \nabla \cdot (\boldsymbol{u} T) + s_t, \tag{5.4}$$

$$\xi = (1-\phi)\rho_p C_{Pp} + \phi \rho_f C_{P_f}, \tag{5.5}$$

$$\kappa = (1-\phi)\kappa_p + \phi \kappa_f, \tag{5.6}$$

where C_{P_p} is the porous rock heat capacity, C_{P_f} is the fluid porosity, ρ_f is the fluid density, κ_p is the rock thermal conductivity, κ_f is the fluid thermal conductivity, ξ is the effective heat capacity, κ is the effective thermal conductivity and $\boldsymbol{u}$ is the Darcy velocity from the fluid flow model. Figure 5.1 illustrates the geothermal reservoir-well system simulation step. The underlying concepts that describe the application of SBGMs for simulating fluid flow in porous media, subsurface reservoirs and geothermal processes are presented in more detail in Salinas et al. (2021). An ensemble of models is created by mapping the generated facies distribution of the conditioned models to the permeability field applied to compute the predicted production data in the reservoir simulation models. The historical production data derived comprises: velocity (u_x and u_y), pressure (P) and temperature (T) in the reservoir over 32 years. Figure 5.5 illustrates the process of generating the training and test dataset and Table 5.1 describes the reservoir, fluid and well properties for the fluid flow models.

5.3.3 Data Assimilation With Machine Learning

Artificial neural networks (ANNs) are powerful data processing tools which can be used for extracting meaningful features that describe the relationship between a set of inputs and its outputs (Lecun et al., 1998). In reservoir simulation, deep learning approaches have been applied for approximating the complex non-linear relationship between input-output parameters of fluid flow in the subsurface to develop proxy models used for predicting reservoir behaviour (Amini and Mohaghegh, 2019). We perform history matching by using neural networks for data assimilation. This is formalised as an ill-posed inverse problem where the goal is to derive the inputs to the geological model (represented by a neural network) that account for the observed fluid flow behaviour.

Neural networks which use adversarial learning, particularly generative adversarial networks (GANs), have been

Table 5.1: Reservoir properties and description of the fluid flow models

Parameter	Value
Reservoir Properties	
Reservoir Extent (m)	1000×200
Initial Reservoir Pressure (kg/ms^2)	50×10^6
Left Boundary Pressure (kg/ms^2)	60×10^6
Right Boundary Pressure (kg/ms^2)	40×10^6
Initial Reservoir Temperature (K)	370
Porosity (%)	20
Rock Density (kg/m^3)	2650
Rock Heat Capacity (W/mK)	1000
Rock Thermal Conductivity (J/K)	2.9
Permeability** (m^2)	2×10^{-10} to 6×10^{-10}
Fluid Properties	
Fluid Density (kg/m^3)	1000
Viscosity (kg/ms)	0.001
Thermal Conductivity (W/mK)	0.15
Heat Capacity (J/K)	4185.5
Phase Volume Fraction (%)	20
Injector Fluid Temperature (K)	310
Well Properties	
Producer BHP (kg/ms^2)	40×10^6
Injector BHP (kg/ms^2)	60×10^6
Simulation Properties	
Timestep (days)	365
Simulation Period (yrs)	32

**The permeability field is populated by mapping the facies type at each point of the conditioned model to the corresponding value of permeability of the rock type, i.e. USF $= 2.0 \times 10^{-10}$, LSF $= 4.0 \times 10^{-10}$ and OS $= 6\,0\mathrm{E} - 10$

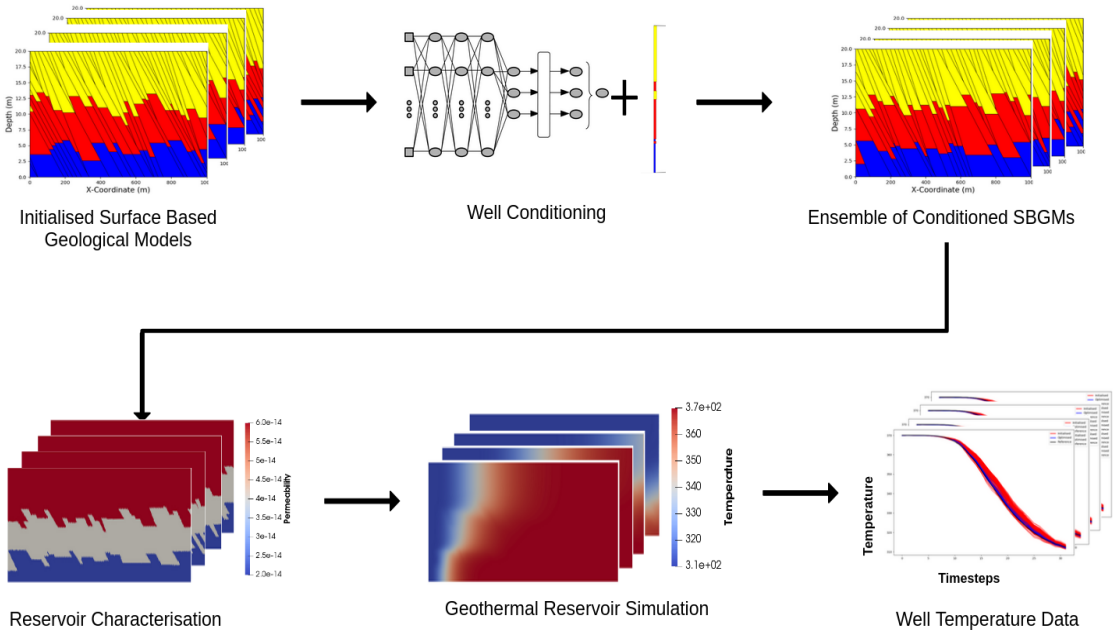


Figure 5.5: The workflow for generating the dataset used for training and testing GeoNet, $\mathcal{G}$. The process involves conditioning a randomly generated set of SBGMs using the trained conditioning network, $\mathcal{S}$, and generating the reservoir simulation models of the conditioned models to obtain the production data at the well locations.

applied for generating and calibrating static reservoir models (Chan and Elsheikh, 2019; Jo et al., 2020; Laloy et al., 2018; Mosser et al., 2018; Zhang et al., 2019), and for history matching by exploring the low-dimensional Gaussian latent space to improve the speed of the history matching process (Canchumuni et al., 2020; Jo et al., 2020; Mosser et al., 2019). In this work, we combine adversarial learning with data assimilation to generate subsurface reservoir models that are calibrated to production data from geothermal reservoir-well systems.

Adversarial Learning

In supervised learning, both inputs and outputs are presented to the neural network, and the model learns the mapping from the inputs to the output based on error feedback using the backpropagation technique (Rumelhart et al., 1986*b*). The concept of adversarial learning has been explored in generative models for matching an aggregated posterior to the prior to ensure that samples generated from any part of the prior space are meaningful. However, our approach adopts adversarial learning as a regularisation technique and not necessarily for sampling new, synthetic data from the distribution of the training data as in GANs. Typically, GANs find a map from a given distribution to the distribution underlying the training data (Goodfellow et al., 2014). In our adversarial model, the network relies on the training to capture the data distribution. We impose a Gaussian distribution to regularise the latent representation of the input data.

GeoNet: Adversarial Network for Reservoir History Matching

The adversarial network is made up of two convolutional networks: GeoNet, $\mathcal{G}$ and a discriminator, $\mathcal{D}$ as shown in Figure 5.6. GeoNet consists of two blocks; the first is a fully connected multilayer perceptron (MLP), $\mathcal{M}$ (consisting of four layers) that learns to convert the data distribution to the prior distribution. The second block, $\mathcal{C}$ (comprising a fully connected layer and five convolutional layers) matches the imposed prior to the production data. The input of $\mathcal{M}$ is a real data point $\mathbf{p} \in \mathbb{R}^{n_p}$. The adversarial variable $\mathbf{a}$ is the output of $\mathcal{M}$ i.e., $\mathbf{a} = \mathcal{M}(\mathbf{p})$. The input of $\mathcal{C}$ is $\mathbf{a}$ and the output of $\mathcal{C}$ is a three-dimensional array, $\hat{\mathbf{y}} \in \mathbb{R}^{N_y \times N_t \times N_c}$, where N_y is the number of points from all the observation wells, N_c is the number of of physical solution fields (components in the fluid production data, i.e. $\mathbf{T}, \mathbf{u}_x, \mathbf{u}_y$ and $\mathbf{P}$), and N_t is the number of timesteps.

Both $\mathcal{M}$ and $\mathcal{C}$ share the adversarial layer, denoted by $\mathbf{a}$. The blocks, $\mathcal{M}$ and $\mathcal{C}$ model the conditional distributions $p(\mathbf{a}|\mathbf{p})$ and $p(\mathbf{p}|\mathbf{a})$ respectively. The distribution of the adversarial variables in $\mathcal{G}$ denoted by $q(\mathbf{a})$ is the posterior distribution of adversarial variables. There is also a prior distribution on the adversarial variables, $p(\mathbf{d})$ which is determined by the user. We use a normal distribution in this work. Block $\mathcal{M}$ generates the adversarial variables from the posterior distribution:

$$\mathcal{M}(\mathbf{p}) = \mathbf{a} \sim q(\mathbf{a}). \tag{5.7}$$

The output layer of the discriminator $\mathcal{D}$ has a single neuron with a sigmoid activation function that classifies the adversarial variable as real from the prior distribution or fake and generated from $\mathcal{M}$ as follows:

$$\mathcal{D}(\mathbf{a}) := \begin{cases} 1 & \text{if } \mathbf{a} \text{ is real, i.e. } \mathbf{a} \sim p(\mathbf{a}) \\ 0 & \text{if } \mathbf{a} \text{ is generated, i.e. } \mathbf{a} \sim q(\mathbf{a}) \end{cases}. \tag{5.8}$$

Our model is regularised by matching the aggregated posterior, $q(\mathbf{a})$, to an arbitrary prior, $p(\mathbf{a})$. With this strategy, $\mathcal{M}$ ensures the aggregated posterior distribution can trick $\mathcal{D}$ into thinking that $q(\mathbf{a})$ is from $p(\mathbf{a})$. $\mathcal{M}$ and $\mathcal{D}$ are therefore constantly competing until the outputs of the adversarial layer conform to the distribution of the original data. It is the adversarial block that guides $q(\mathbf{a})$ to match $p(\mathbf{a})$ while $\mathcal{G}$ minimises the mismatch between the true and predicted production data. The networks play a min-max game, where $\mathcal{D}$ computes the probability that a point $\mathbf{p}$ in the data space is a sample from the data distribution (real samples) that we are trying to model, rather than a sample from $\mathcal{M}$ (fake samples). Table 5.2 describes the architecture for $\mathcal{D}$ and $\mathcal{G}$. Unlike the adversarial autoencoder (Makhzani et al., 2016), Geonet, $\mathcal{G}$ does not impose a distribution on a much smaller space in the adversarial layer to learn the identity map, or have two related subnetworks on either side of the adversarial layer. In our application of adversarial training, the underlying distribution of the training data is imposed on the output using an adversarial layer having a larger dimension than the input layer to allow for learning more information about the training data since data compression is not the goal of

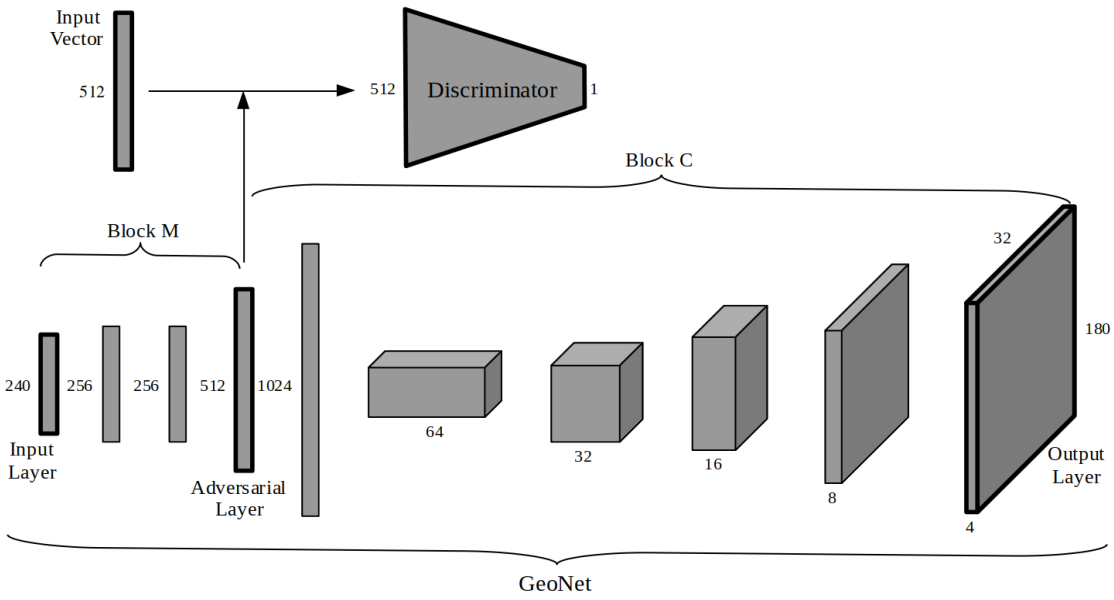


Figure 5.6: The architecture of the network showing GeoNet, the Discriminator and the adversarial layer.

the method.

Training the network

Adversarial learning involves alternating optimisation where the approximation and regularisation phases are repeated iteratively using gradient descent with backpropagation (Ghojogh et al., 2021). The adversarial network is trained in two phases- regularisation and approximation, with different objective functions: an adversarial training criterion (Goodfellow et al., 2014) that matches the aggregated posterior distribution of the adversarial variables to an arbitrary prior distribution, and a least squares functional that provides information on the mismatch between predicted and desired outputs.

The network is trained by leveraging the gradient of $\mathcal{D}(\mathbf{a})$ with respect to $\mathbf{p}$, and using that to modify its parameters in the regularisation phase. Thus, the adversarial network first updates the discriminator to tell apart the true samples (generated using the prior $p(\mathbf{p})$) from the generated samples, and then updates $\mathcal{G}$ to "fool" the discriminator. In the approximation step, $\mathcal{G}$ with the constrained adversarial variables is trained to minimise the mismatch between network predictions $\hat{\mathbf{y}}$ and the production data, $\mathbf{y}$. The loss functions used in the training are:

$$\mathcal{L}_{\mathcal{D}} = -\mathbf{E}_{\mathbf{a}\sim p(\mathbf{a})}[log(\mathcal{D}(\mathbf{a}))] - \mathbf{E}_{\mathbf{p}\sim p(\mathbf{p})}[log(1-\mathcal{D}(\mathcal{M}(\mathbf{p})))], \tag{5.9}$$

$$\mathcal{L}_{\mathcal{M}} = \mathbf{E}_{\mathbf{p}\sim p(\mathbf{p})}[log(1-\mathcal{D}(\mathcal{M}(\mathbf{p})))], \tag{5.10}$$

$$\mathcal{L}_{\mathcal{G}} = \mathbf{E}_{\mathbf{p}\sim p(\mathbf{p})}[\| \mathcal{G}(\mathbf{p}) - \mathbf{y} \|_2]. \tag{5.11}$$

Table 5.2: Architecture of the adversarial network applied for mapping the geological inputs, $\mathbf{p}$ to production data, $\mathbf{y}$.

Geonet($\mathcal{G}_\theta$)							
Layer	Type	Input Size	Filters	Kernel	Stride	Padding	Activation
1	Linear	240	-	-	-	-	LeakyReLU
2	Linear	256	-	-	-	-	LeakyReLU
3	Linear	256	-	-	-	-	LeakyReLU
4	Linear†	512	-	-	-	-	LeakyReLU
5	Linear	1024	-	-	-	-	LeakyReLU
6	ConvTransp2D	4×4	64	16×4	1×1	1×1	LeakyReLU
7	ConvTransp2D	17×5	32	8×4	2×2	1×1	LeakyReLU
8	ConvTransp2D	38×10	16	8×4	1×1	1×2	LeakyReLU
9	ConvTransp2D	43×9	8	8×4	2×2	1×2	LeakyReLU
10	ConvTransp2D	90×16	8	8×4	2×2	3×1	Tanh
Discriminator($\mathcal{D}_\psi$)							
Layer	Type	Input Size	Filters	Kernel	Stride	Padding	Activation
1	Linear	512	-	-	-	-	LeakyReLU
2	Linear	64	-	-	-	-	LeakyReLU
3	Linear	16	-	-	-	-	LeakyReLU
4	Linear	4	-	-	-	-	Sigmoid

Slope of LeakyReLU = 0.01
ConvTransp2D = Transpose Convolution
†Adversarial Layer

Equations (5.9) and (5.10) are for the regularisation phase while Equation (5.11) is for the approximation phase.

Framework for data assimilation

Data assimilation integrates model predictions with observations to reduce the uncertainty in the model. This involves solving an inverse problem. For example, variational data assimilation requires minimising a cost function using numerical optimisation techniques (Asch et al., 2016). Similar to the technique for training neural networks, variational data assimilation involves applying gradient descent to minimise a functional. While in machine learning, the objective function (known as the loss function) minimised during training is the difference between network predictions and the desired output, in variational data assimilation, the objective function describes the mismatch between observations and the model predictions. Thus, data assimilation can be efficiently performed using machine learning as the back-propagation technique for optimising neural networks is a suitable alternative for the adjoint models used in variational data assimilation (Buizza et al., 2022).

Given the trained neural network, $\mathcal{G}^*$, that approximates the underlying relationship between the geological inputs, $\mathbf{p}$ and the evolution of the system over a period of time $\mathbf{y}$, i.e., the production data, we perform history matching by using neural networks for data assimilation. Using this approach, we can learn the set of input

feature vectors, $\mathbf{p}$, to the trained neural network, which match the given or observed production history such that $\mathcal{G}(\mathbf{p}) = \mathbf{y}_{obg}$. The problem is formulated as finding a $\mathbf{p}^*$ that minimises the mismatch, $\mathcal{L}$, between the network prediction and the given observations,

$$\mathbf{p}^* = \underset{\mathbf{p}}{\operatorname{argmin}}\, \mathcal{J}_g(\mathbf{p}), \tag{5.12}$$

where $\mathcal{J}_g$ is the objective function to be minimized defined as:

$$\mathcal{J}_g(\mathbf{p}) = \mathcal{L}(\mathbf{y}_{obg}, \mathcal{G}^*(\mathbf{p}^*)) := \mathbf{E}_{\mathbf{p}\sim p(\mathbf{p})}[\| \mathbf{y}_{obg} - \mathcal{G}^*(\mathbf{p}) \|_2], \tag{5.13}$$

where $\mathbf{y}_{obg}$ is the reference fluid production data, $\mathbf{y}_{obg} \in \mathbb{R}^{N_y \times N_t \times N_c}$. The set of optimised inputs is constrained to the limits of the uniform distribution, $\mathcal{U}(L_b, U_b)$, of the stochastic inputs after every optimisation step:

$$\tilde{p}_j^{k+1} = \begin{cases} \tilde{p}_j^{k+1} & \text{if } L_b \leqslant \tilde{p}_j^{k+1} \leqslant U_b \\ L_b & \text{if } \tilde{p}_j^{k+1} < L_b \\ U_b & \text{if } \tilde{p}_j^{k+1} > U_b \end{cases} \tag{5.14}$$

Similar to the ES-MDA method, we perform the optimisation iteratively, where the proposed history matching method using the neural networks for data assimilation can be summarised as follows:

1. Randomly generate $\mathcal{K}$ SBGM models with stochastic inputs, $\mathbf{p}$.

2. Given the trained network $\mathcal{S}$, apply Equation(5.1) as described in Section 5.3.1, to generate $\mathbf{p}^*$, which is the set of SBGM inputs that match a given reference for each SBGM with initial input $\mathbf{p}$.

3. Given the trained network $\mathcal{G}$, apply Equation (5.13) to obtain the set of inputs $\hat{\mathbf{p}}$ that generate the static models that produce the historical production data.

4. Repeat steps two and three until the produced inputs generate SBGMs that match the well log and dynamic models that match the production data at the well locations.

This presents an efficient and robust approach to condition SBGMs through the use of an iterative method in an optimisation framework with neural networks, where the predictions $\tilde{\mathbf{y}}$ and $\hat{\mathbf{y}}$ from $\mathcal{S}$ and $\mathcal{G}$ respectively, are compared with the observed data, $\mathbf{y}_{obs}$ and $\mathbf{y}_{obg}$, and the inputs, $\mathbf{p}^*$ are updated at each iteration according to the data mismatch. Algorithm 5.1 and Figure 5.7 illustrate the details of the data assimilation framework using $\mathcal{S}$ and $\mathcal{G}$.

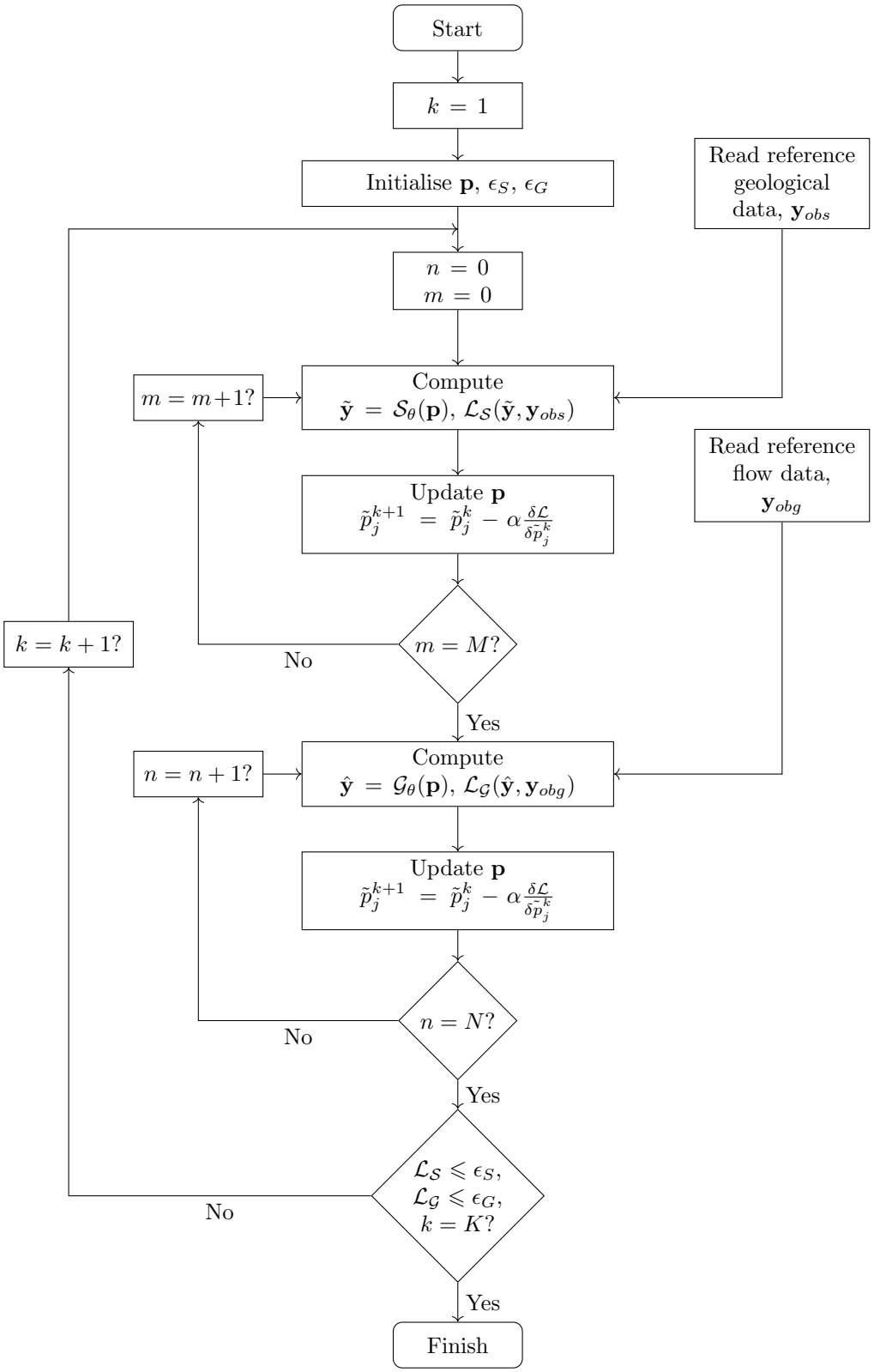


Figure 5.7: The framework for generating SBGM inputs that generate geologically plausible models that match static well data and are history-matched to production data.

Algorithm 5.1 Framework for history-matching subsurface reservoirs using machine learning

initialization: get random stochastic input parameters $\mathbf{p}^0 \in \mathbb{R}^{n_p}$
initialisation: $k = 1$, $m = 0$, $n = 0$
while $k \leqslant K$ **do** ▷ K = number of optimisation steps
▷ start well conditioning cycle
    $\mathbf{p}^k = \mathbf{p}^{k-1}$
    read $\mathbf{y}_{obs}$ and encode output data at each discretized location as $\mathbf{z}_{obs,i}$
    scale $\mathbf{p}$ to a normal distribution $\tilde{\mathbf{p}} \sim \mathcal{N}(0, 1)$
    while $m \leq M$ **do** ▷ M = number of well conditioning cycles
        compute $\tilde{\mathbf{y}} = \mathcal{S}_{\boldsymbol{\theta}}(\tilde{\mathbf{p}}^k)$ ▷ $\tilde{\mathbf{y}} \in \mathbb{R}^{n_p}$
        calculate $\mathcal{L}_{\mathcal{S}}(\tilde{\mathbf{y}}, \mathbf{y}_{obs})$
        $\frac{\delta \mathcal{L}}{\delta \tilde{p}_j^k} \forall j \in \{1, \ldots, n_p\}$
        $\tilde{p}_j^{k+1} := \tilde{p}_j^k - \frac{\eta \hat{m}}{\sqrt{\hat{v}} + \epsilon}$ ▷ update inputs with the Adam optimizer
        constrain $\tilde{\mathbf{p}}^k$ as in Equation (5.14)
    denormalise $\tilde{\mathbf{p}}^k$
    $m = 0$
▷ start data assimilation cycle
    read $\mathbf{y}_{obg}$ and apply min-max scaling
    scale $\mathbf{p}^k$ to a normal distribution, $\hat{\mathbf{p}}^k \sim \mathcal{N}(0, 1)$
    while $n \leq N$ **do** ▷ N = number of data assimilation cycles
        compute $\hat{\mathbf{y}} = \mathcal{G}_{\boldsymbol{\theta}}(\hat{\mathbf{p}}^k)$ ▷ $\hat{\mathbf{y}} \in \mathbb{R}^{N_y \times N_t \times N_c}$
        calculate $\mathcal{L}_{\mathcal{G}}(\hat{\mathbf{y}}, \mathbf{y}_{obg})$
        $\frac{\delta \mathcal{L}}{\delta \hat{p}_j^k} \forall j \in \{1, \ldots, n_p\}$
        $\hat{p}_j^{k+1} := \hat{p}_j^k - \frac{\eta \hat{m}}{\sqrt{\hat{v}} + \epsilon}$ ▷ update inputs with NAdam optimizer
        constrain $\hat{\mathbf{p}}$ as in equation 5.14
    denormalise $\hat{\mathbf{y}}$
    $n = 0$, $k = k + 1$
return $\hat{\mathbf{p}}$

5.3.4 Uncertainty Analysis

A good understanding of the uncertainties associated with the history-matched models informs reservoir performance prediction (forecasting), economic and risk assessment of different development scenarios. A group of models known as an ensemble is matched to the available historical production data. Ensemble based methods are employed to generate multiple reservoir model realizations that approximately characterise the reservoir (Jung et al., 2018). The information from ensembles (suite of model possibilities) provides an estimate of sensitivities and uncertainties.

To generate realisations conditioned to the observed well data, we apply the steps outlined in Section 5.3.3 for each initialised model. Similar to the method used for assessing the variation among the static models in Titus et al. (2022), we calculate the uncertainties associated with the conditioned realisations using the frequency of predictions (FOP) to provide a qualitative and quantitative measure of the agreement among the conditioned realisations and the reference data. The frequency of predictions, f^{pred} is given as the difference between the maximum frequency and the minimum frequency at each discretised location within the domain defined as:

$$f_{i,j}^{pred} = f_{i,j}^{max} - f_{i,j}^{min}, \tag{5.15}$$

where $f_{i,j}^{max}$ is the frequency of the facies with the highest occurrence (also defined as the occurrences of the facies type with the most predictions), and $f_{i,j}^{min}$ is the frequency of the facies with the least occurrence at each discretized location within the domain.

5.4 Results and Discussions

5.4.1 Forward Modelling: Predicting the Production Data

We present here the results of using neural networks to generate the production data using the method described in Section 5.2.1. Using the trained network $\mathcal{G}$, we are able to map the 240 SBGM inputs to the production data, $\mathbf{y}$ consisting of the pressure, velocity (in x- and y- directions) and the temperature for 60 discretised locations along each well. Thus, we map $\mathbf{p} \in \mathbb{R}^{n_p} \rightarrow \hat{\mathbf{y}} \in \mathbb{R}^{N_y \times N_t \times N_c}$, where N_c is 4 (temperature, pressure and velocity in the x and y-directions), N_t is 32 years and N_y is 180, corresponding to 60 grid points from each of the three exploratory wells. Similar to our previous work (Titus et al., 2022), we have chosen the number of grids to correspond to the vertical resolution of leading industry gamma well logging tools (0.3048 m $\pm$5%) (Schlumberger, 2015) and the maximum number of geological volumes that can intersect a well when generating the geological model. This yields a vertical resolution of 0.33m 0.33 m.

The dataset used for training $\mathcal{G}$ consists of 1000 geothermal reservoir simulations (as described in section 5.3.2),

with 900 examples used for training and 100 for testing. Despite the small number of training samples, the network is able to reasonably represent the production data given the SBGM inputs as the training loss shown in Figure 5.8 shows a reduction of more than 3 orders of magnitude over 400 epochs. The use of adversarial training for $\mathcal{G}$ constrains the network predictions to more closely match the desired distribution. This distribution is embedded in the inputs to the network, which have been preconditioned to a reference geological data. It was essential to ensure that $\mathcal{G}$ can be trained to a reasonable accuracy with far fewer samples compared to the trained network $\mathcal{S}$ used for conditioning wells to geological data, which uses three times more data. This is particularly important because the process of generating reservoir simulations can be time consuming in comparison to the generation of the SBGMs which take a fraction of the time. The use of fewer examples for training $\mathcal{G}$ improves the applicability and practicality of the method for more complex cases. Table 5.3 provides the training hyperparameters and metrics for $\mathcal{G}$ and Figures 5.9a to 5.9f show examples of the network predictions of the temperature over the simulation period. Although we have trained the model to predict temperature, pressure, and velocity, we focus on the validation of the method for the temperature data, which is of greater significance for evaluating thermal recovery from geothermal reservoirs.

Table 5.3: Hyperparameters used for training GeoNet and the performance metrics of the model

	Value
Hyperparameters*	
Learning rate	0.0005
Train-test split	0.90
Number of epochs	2000
Optimizer	NAdam
Weight decay	0.0
Betas	(0.8,0.999)
Annealing	0.95
Annealing step size	500
Discriminator iterations per generator	10
Weight initialisation	Kaiming
Number of epochs	3000
Metrics	
GeoNet Training Loss	0.00037
GeoNet Test Loss	0.00049
Generator Loss	6.1781
Discriminator Loss	0.0.721
Training time on GPUs (hrs)	5.05

*The hyperparameters of $\mathcal{D}$ are similar to $\mathcal{G}$

5.4.2 Inverse modelling: Optimising the Geological Model Inputs

We apply the method presented in Section 5.3.3 and described in Figure 5.7 to invert for the set of SBGM inputs, $\mathbf{p}^*$ that will generate the static and dynamic models that match the given reference geological and production data, $\mathbf{y}_{obs}$ and $\mathbf{y}_{obg}$ respectively. The reference production data was obtained from the simulation

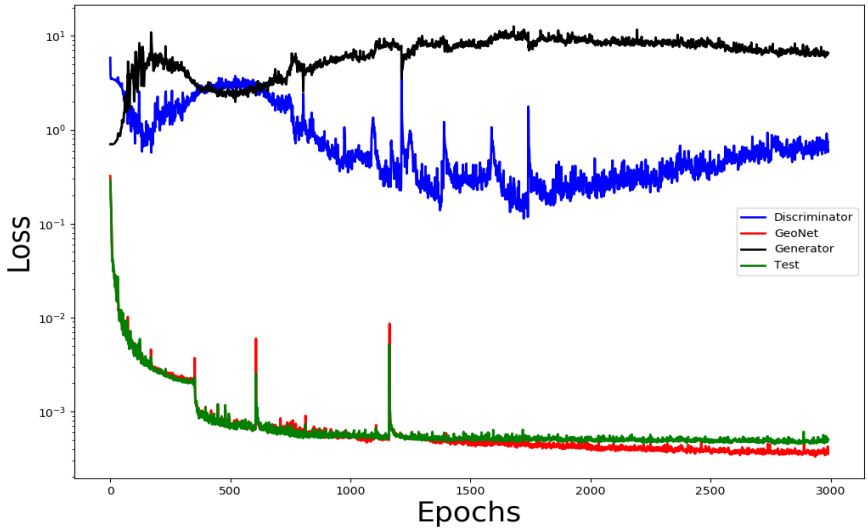
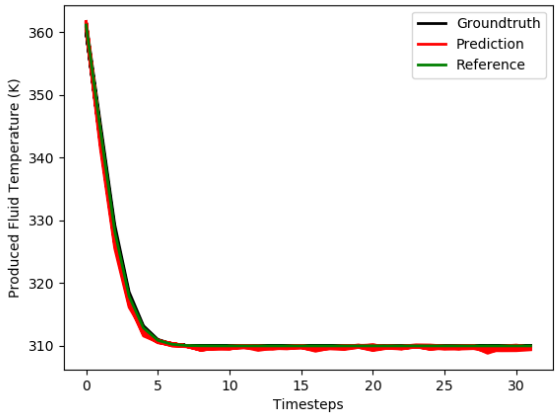


Figure 5.8: The loss plot of the adversarial network, GeoNet, as well as the discriminator, $\mathcal{D}$.

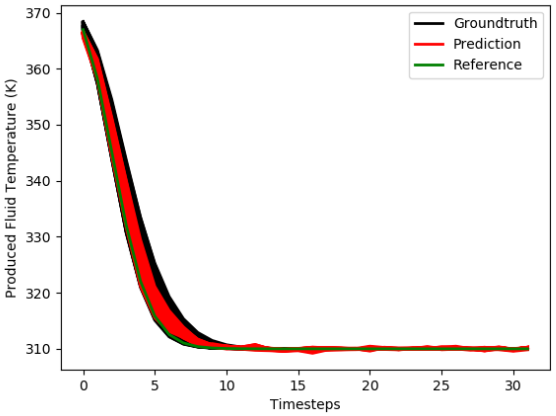
of a geothermal reservoir, whose geological model is shown in Figure 5.3, generated using a randomly selected set of stochastic SBGM inputs that was not used in training or testing $\mathcal{S}$ and $\mathcal{G}$.

Using Algorithm 5.1, we optimised the initial SBGMs to match the given reference well data. Figures 5.10 and 5.11 present the loss plots for two examples, depicting the decreasing well conditioning and data assimilation losses as the mismatch between the initialised and optimised inputs reduces. After 10000 iterations, with one cycle each for well conditioning and data assimilation, the initialised SBGM inputs have been optimised based on a comparison with the achieved training/test loss in the forward modelling. Due to the sparsity of data, the non-uniqueness of the solution, and the stochastic nature of the process of generating the SBGMs, there are several possible realisations of the SBGMs. Nonetheless, since the method lends itself to local optimisation techniques, we are able to generate 1000 conditioned realisations using 100 processors, as the optimisation takes less than thirty-five minutes (for the 10000 iterations) and the method lends itself to massive parallelism. This shows that through a process of continuously sampling and optimising SBGM inputs, multiple conditioned realisations can be rapidly generated. Table 5.4 shows the hyperparameters for $\mathcal{S}$ and $\mathcal{G}$ used to optimise the initialised SBGM inputs.

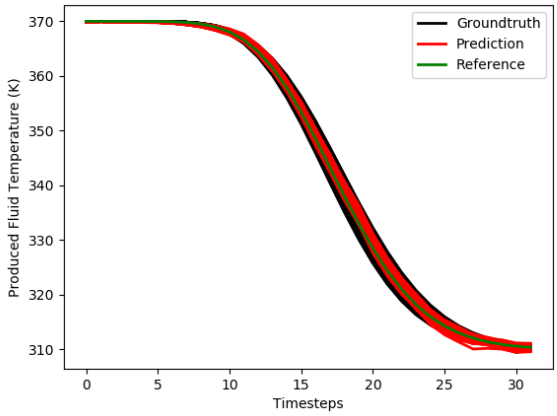
Figures 5.12a to 5.12f show examples of the predicted production data using the trained $\mathcal{G}$ for the optimised, initialised, and reference models. The results show reduced variation, and, by extension, decreased uncertainty of the optimised models in comparison to the initialised models. However, to validate the accuracy of the method, we generate the geological models with the optimised SBGMs, and run the geothermal reservoir simulation given these static models to evaluate the match of the dynamic models to the reference models at the well locations. We impose a well correlation criterion to determine models that are calibrated to the reference



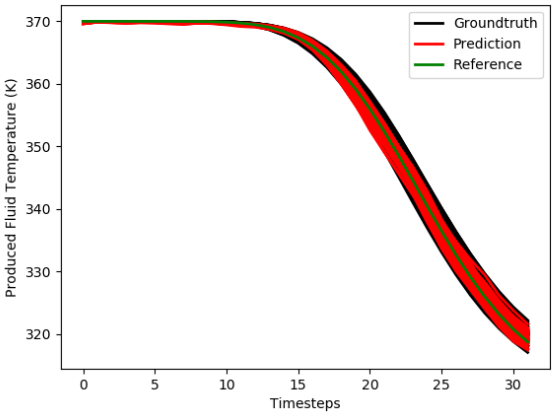
(a) Test Samples - Point in Well 1



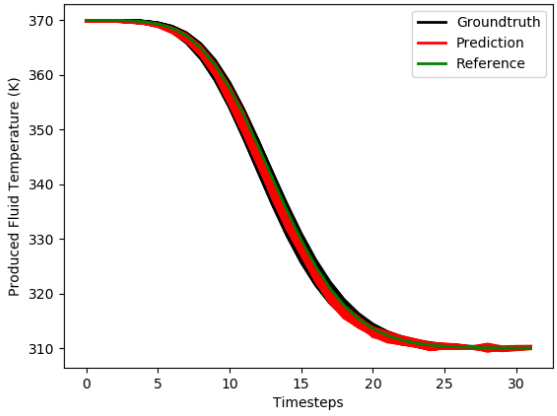
(b) Test Samples - Point in Well 1



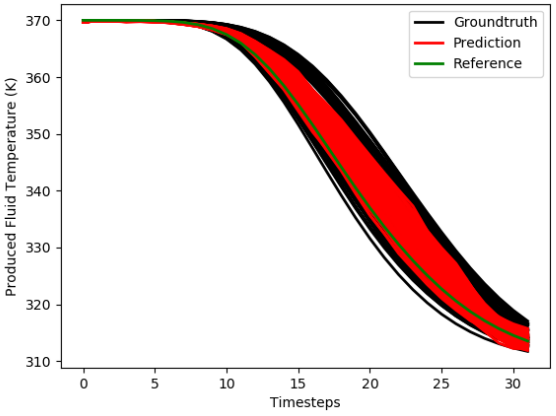
(c) Test Samples - Point in Well 2



(d) Test Samples - Point in Well 2



(e) Test Samples - Point in Well 3



(f) Test Samples - Point in Well 3

Figure 5.9: Sample predictions of the geothermal well simulations against the ground truth for unseen samples by the network.

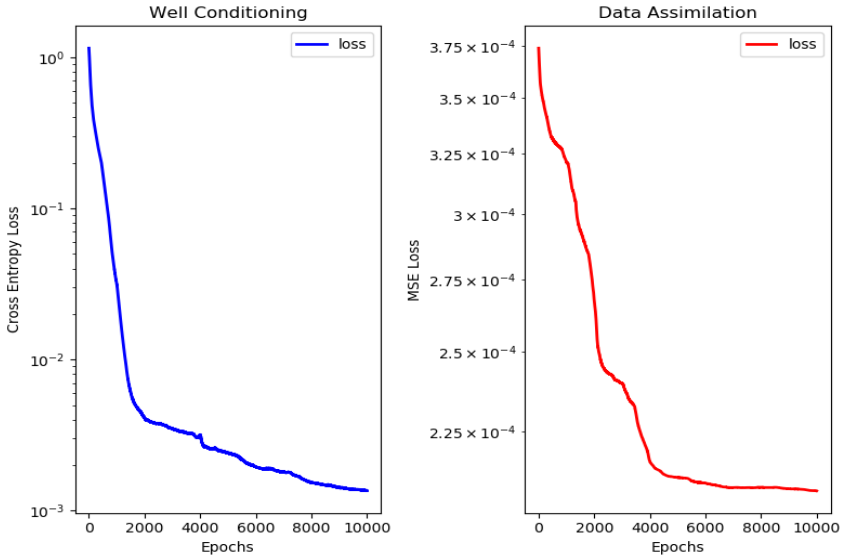


Figure 5.10: Plots showing the mismatch between the predicted and observed data using the trained networks $\mathcal{S}$ and $\mathcal{G}$ for the well conditioning and data assimilation respectively - Example 1

reservoir simulation by computing an error tolerance E_{well}, a mean absolute error given as:

$$E_{well} = \frac{1}{N_y \times N_t} \sum_{t=1}^{N_t} \sum_{i=1}^{N_y} (\mathbf{y}_{obg,i}^t - \hat{\mathbf{y}}_i^t). \tag{5.16}$$

We select models that are within ± 0.001 of the training loss of $\mathcal{G}$. From Figures 5.12a to 5.12f, it can be seen that a better fit to the reference model is obtained with the optimised models compared with the initialised models. Figures 5.13a to 5.13f shows the comparison of the temperature at the well locations in the forward models generated using the initialised, optimised, and reference geological models for two examples and four timesteps. Figure 5.14 shows the mean absolute difference of the initialised and optimised models from the reference model for all timesteps. The average correlation with the facies type at the well locations is presented in Figure 5.15 for the optimised and initialised models. The results further validate the improved match of the optimised models with the reference models for the static and dynamic data.

Table 5.4: Parameters for optimizing the SBGM inputs using the trained neural networks

	SurfNet	GeoNet
Learning rate	0.001	0.0005
Number of Cycles	1	1
Number of Iterations	10^4	10^4
Optimizer	Adam	NAdam
Well correlation cut-off	97.5	-
Well error tolerance	-	0.0007

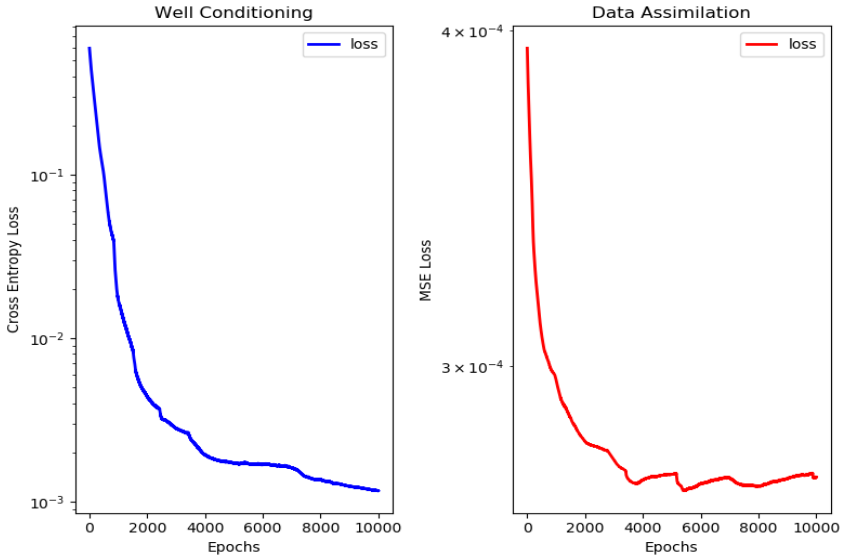
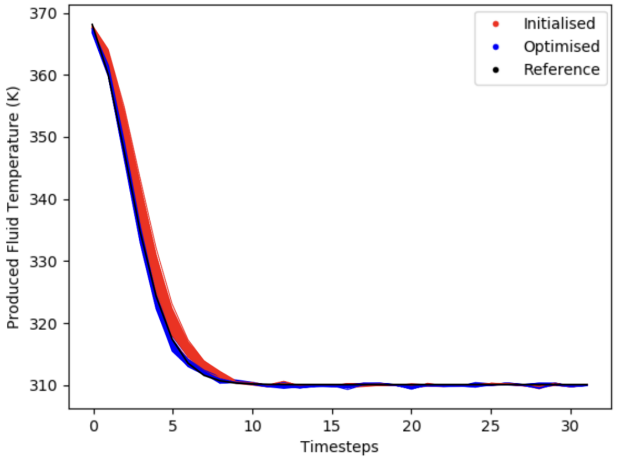


Figure 5.11: Plots showing the mismatch between the predicted and observed data using the trained networks $\mathcal{S}$ and $\mathcal{G}$ for the well conditioning and data assimilation respectively - Example 2

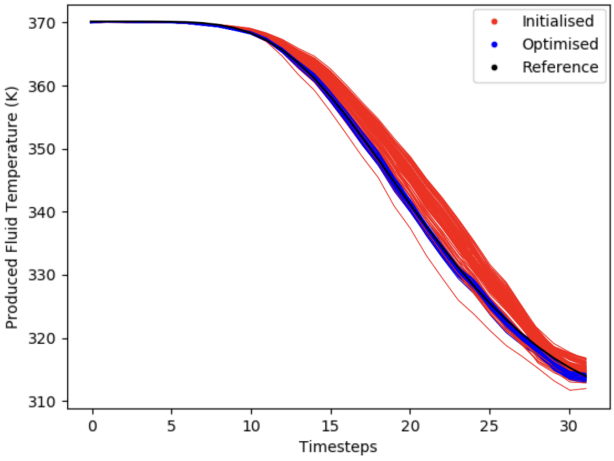
5.4.3 Uncertainty Assessments: Evaluating Spatial Variations and Well Correlations

In history matching, a set of models calibrated to the dynamic behaviour of the reservoir is required to evaluate uncertainties. Uncertainty quantification is commonly performed in a Bayesian framework by resampling the search space using Markov chain Monte-Carlo sampling algorithms. Various algorithms are used to search and sample the parameter space in history matching and uncertainty quantification, the choice of which impacts not only the quality of the result but also the speed of the process (Abdulrazzaq and Hasan, 2023). We use the frequency of predictions (FOP) described in Section 5.3.4 to assess and understand spatial variations and uncertainties associated with the generated realisations. The FOP plots are shown in Figures 5.16a to 5.16f.

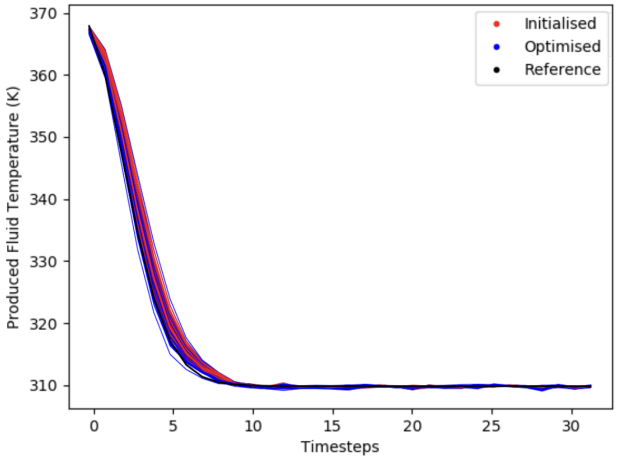
The FOP for the initialised models shows high uncertainty, as lateral variations across the domain even at well locations do not decrease. For the history-matched models, the FOP shows lower uncertainties and spatial variations at well locations, and other locations where the conditioned realisations show a high frequency of predictions. Figures 5.17a to 5.17f present the maximum prediction (MP) plots, which show the facies types with the highest predictions from all conditioned realisations within the geological model. Comparison of MP plots with the reference model (Figure 5.3 shows large-scale similarities, such as varying lateral thicknesses. Thus, the suite of optimised models has been able to capture and preserve general trends and patterns, particularly at the well location and its vicinity. Since the spatial distribution/variation of the rock facies and, in particular, the impact of large-scale heterogeneity is an important determinant of fluid flow behaviour (Osman et al., 2021), generating history-match models using the proposed approach can be a powerful tool for reservoir modelling. Figures 5.18a and 5.18b further validate this by providing a qualitative measure of the similarity of the initialised



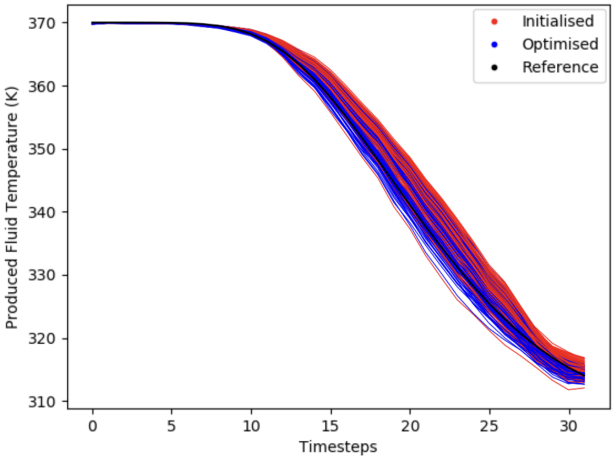
(a) Uncertainties with DA - Point in Well1



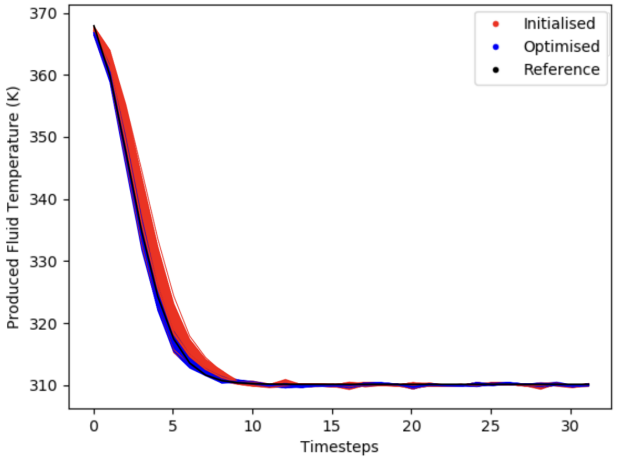
(b) Uncertainties with DA - Point in Well3



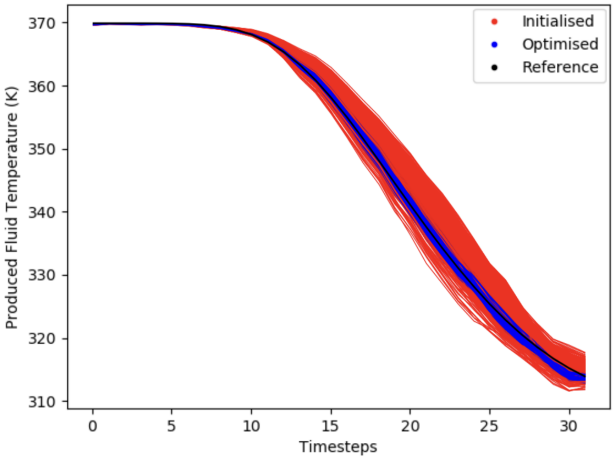
(c) Uncertainties with WC - Point in Well1



(d) Uncertainties with WC - Point in Well3

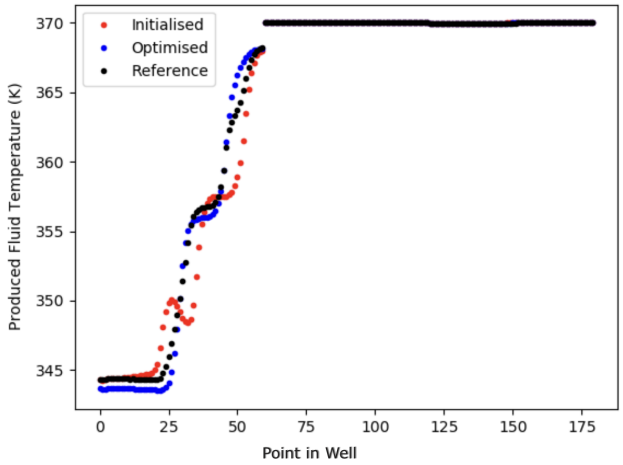


(e) Uncertainties with DA + WC - Point in Well1

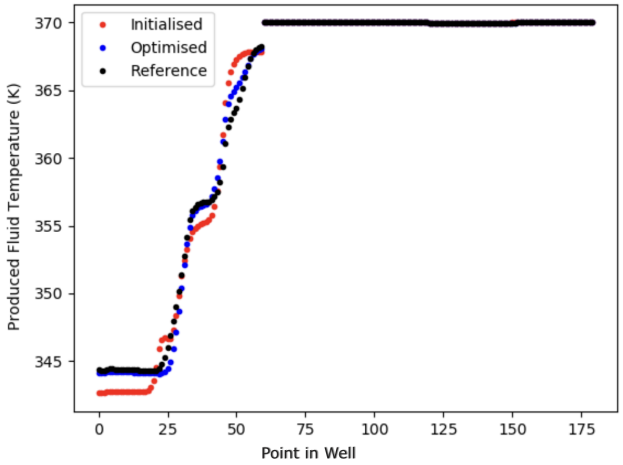


(f) Uncertainties with DA + WC - Point in Well3

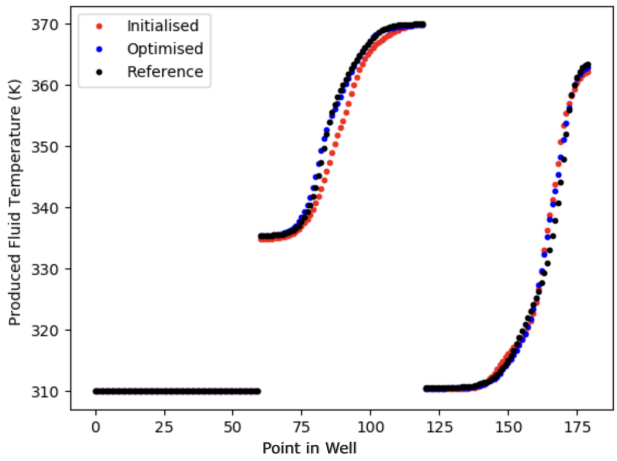
Figure 5.12: Uncertainties from the predictions with the optimised inputs. Top — optimising inputs with only data assimilation. Middle — optimising with only well conditioning. Bottom — optimising with data assimilation and well conditioning.



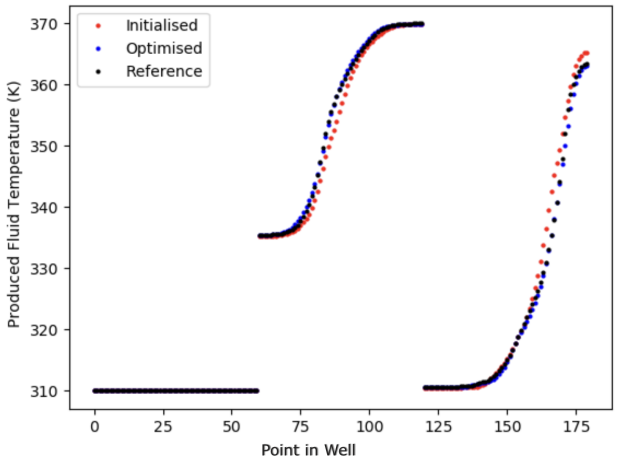
(a) Example 1 - Timestep 1



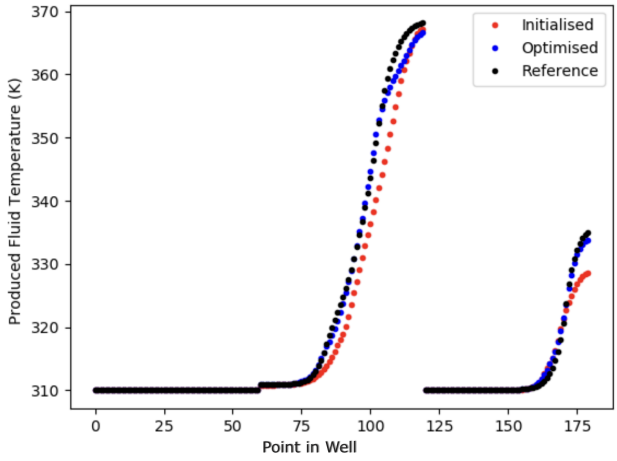
(b) Example 2 - Timestep 1



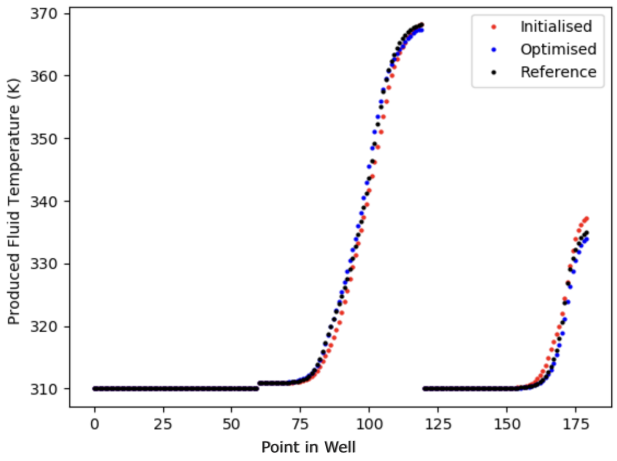
(c) Example 1 - Timestep 17



(d) Example 2 - Timestep 17



(e) Example 1 - Timestep 26



(f) Example 2 - Timestep 26

Figure 5.13: Temperature data at well1 (points 1 to 60), well2 (points 61 to 120) and well3 (points 121 to 180) from the reservoir simulation with the initialised, reference and optimised models.

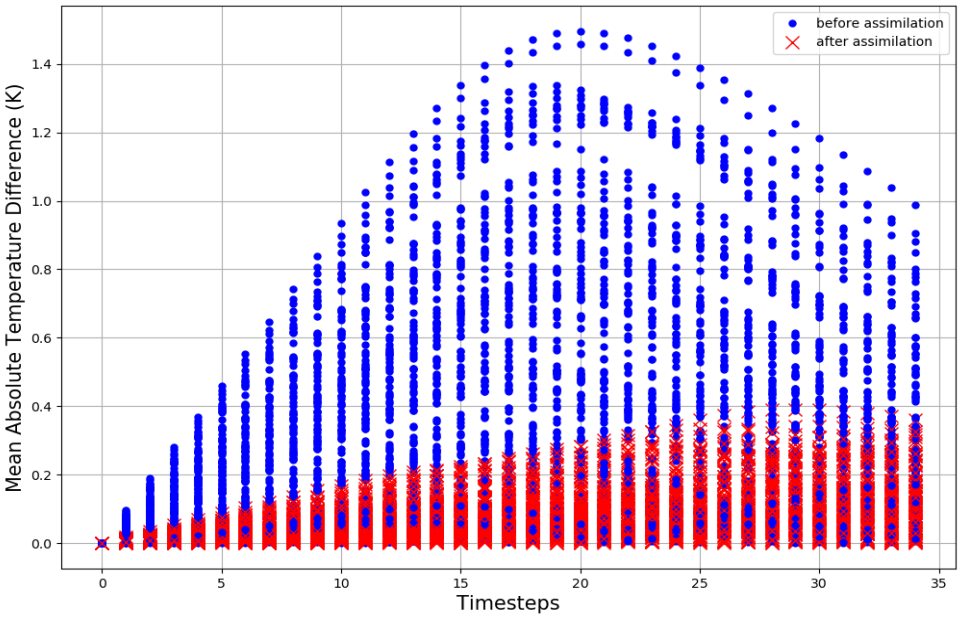


Figure 5.14: The mean absolute temperature difference of the optimised and initialised forward models with the reference forward models showing that the optimised models have a better match to the reference model in comparison to the initialised models.

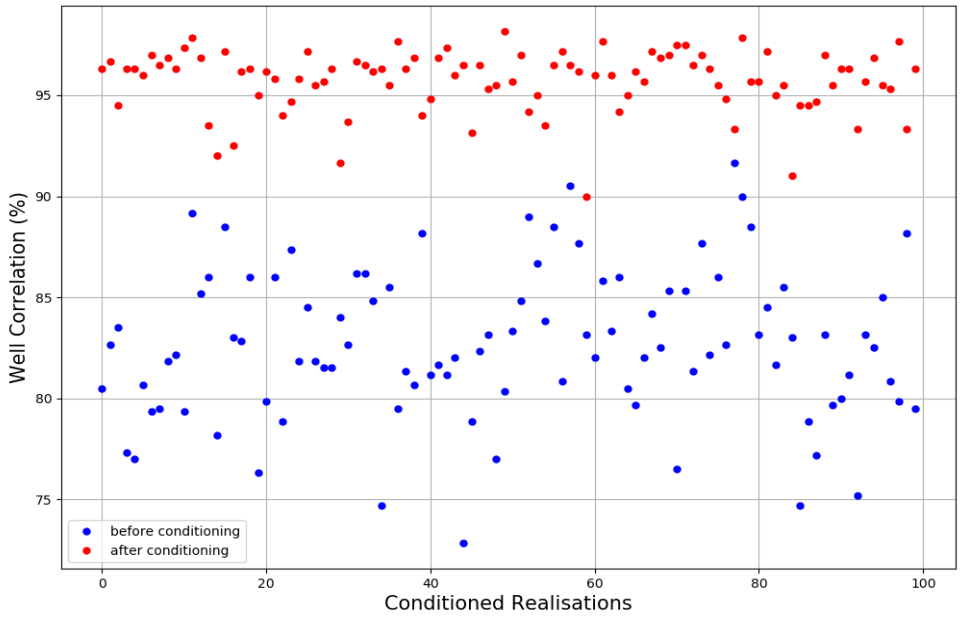


Figure 5.15: Correlation of the facies types at the wells with the reference well facies type for the initialised and optimised.

and optimised models to the reference model.

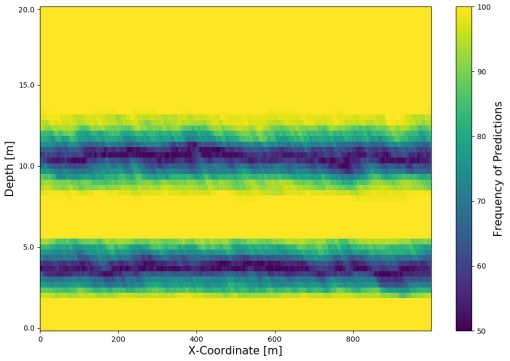
5.4.4 Discussions

We have demonstrated the applicability of our method for history matching reservoir simulation models to production data whilst still honouring geological data. Using the trained networks, history matching only takes approximately 200s for each example. The method also lends itself to massive parallelisation for the data assimilation with the neural network and low dimensionality of the SBGM inputs aids rapid convergence during the optimisation. Although SBGMs incorporate qualitative information such as stacking patterns, spatial and temporal depositional sequences, the addition of quantitative information such as production data to generate more representative dynamic models and further constrain the geological models achieves the primary aims of history matching. Calibrating the geology of the wellbore in a static model to field data provides a reliable model for developing, monitoring, optimising and managing the production from reservoirs. The resulting geological model is produced utilising geological rules constrained by well data, reservoir architecture, and associated petrophysical property distributions.

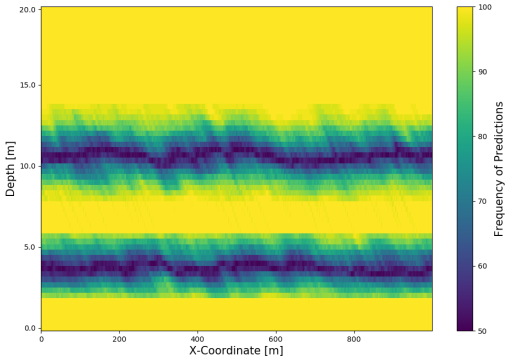
Unlike other geostatistical and rule-based modelling, we are able to generate geological models which present the expected geological features displayed in the initialised models, thus preserving geological plausibility of the ensemble. In contrast, most applications of neural networks are for grid-based geological models which impact the resolution and do not always preserve geological realism. In addition, the use of an adversarial layer to regularise the model reduces the number of training examples required to obtain reasonable accuracy. By essentially using the discriminator to constrain the adversarial vector to a Gaussian distribution, we ensure that samples in the adversarial space are consistent with the conditional data and, by extension, the reference model to which the training data is conditioned to.

In geological reservoir modelling and simulation, it is desirable to have multiple plausible geological and dynamic models to assess and quantify risks and uncertainties associated with sparse subsurface data. The method described here allows for the rapid generation of an ensemble of models. Looking into future applications such as real-time reservoir monitoring and management, data assimilation can be performed in a reduced space where the smaller number of variables will enable computations to run orders of magnitude faster. Since the evolution of the latent space defined in time is also a dynamic process, it is possible to represent the system as a time series data and model this using neural networks.

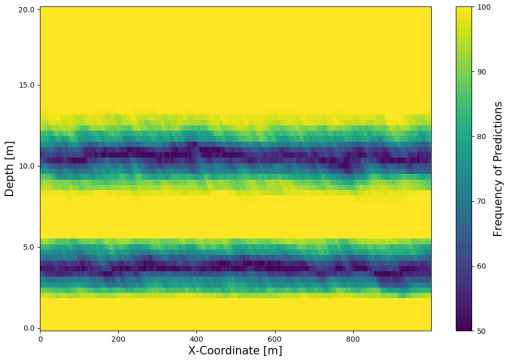
Our method uses simplified models that demonstrate fluid flow in geothermal reservoirs as a proof of concept that can be extended to more complex reservoir models given larger computational resources and more complex/advanced neural networks. Our approach for conditioning surface-based models has shown promising results. Integrating other machine learning models, such as transformers that have been shown to encode posi-



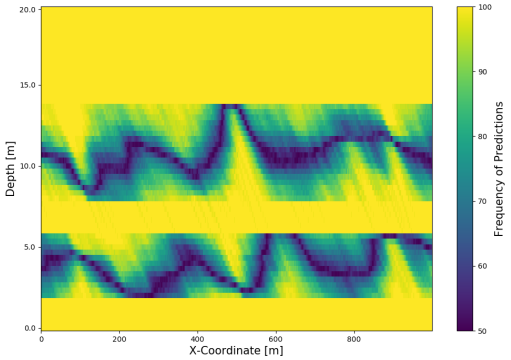
(a) Initialised Models - DA only



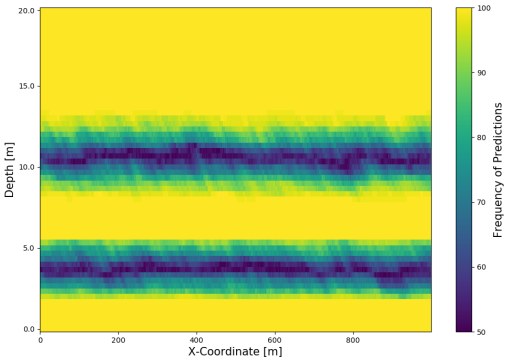
(b) Optimised Models - DA only



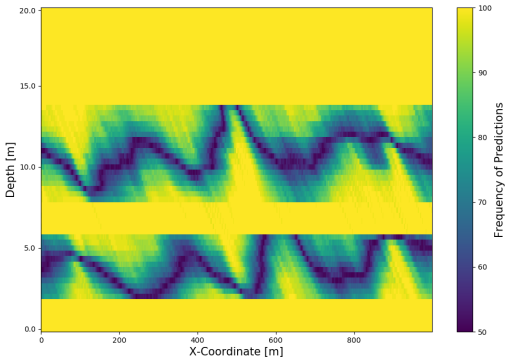
(c) Initialised Models - WC only



(d) Optimised Models - WC only

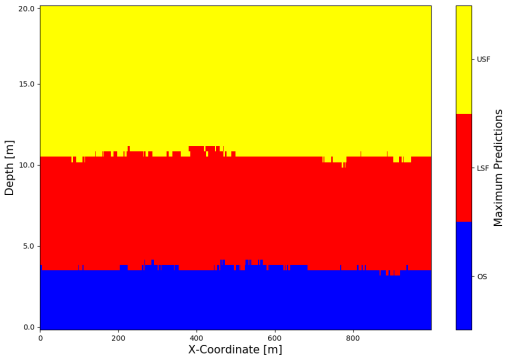


(e) Initialised Models - DA + WC

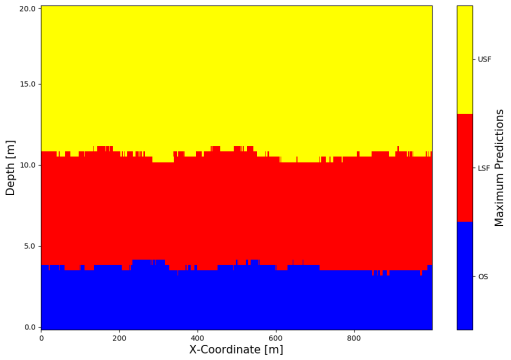


(f) Optimised Models - DA + WC

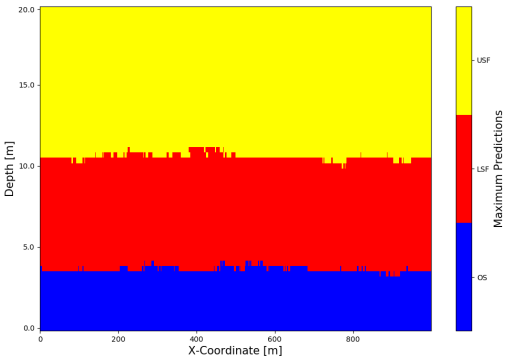
Figure 5.16: Frequency of Predictions (FOPs) from the predictions with the optimised inputs. Top- optimising inputs with only data assimilation. Middle - optimising with only well conditioning. Bottom- optimising with data assimilation and well conditioning.



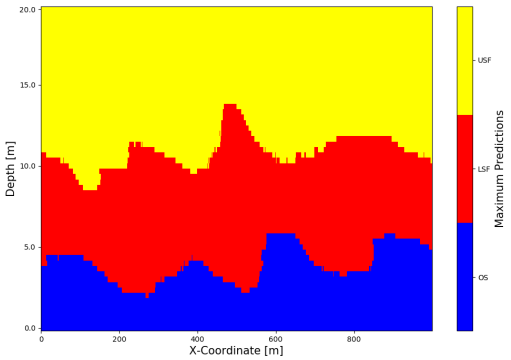
(a) Initialised Models -DA only



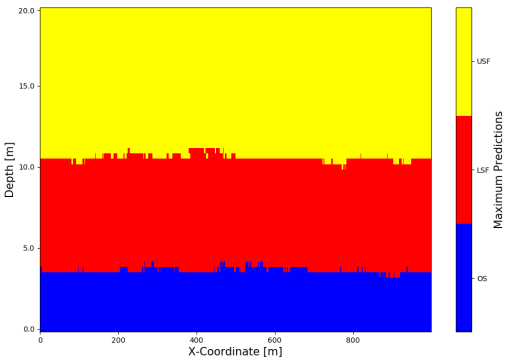
(b) Optimised Models - DA only



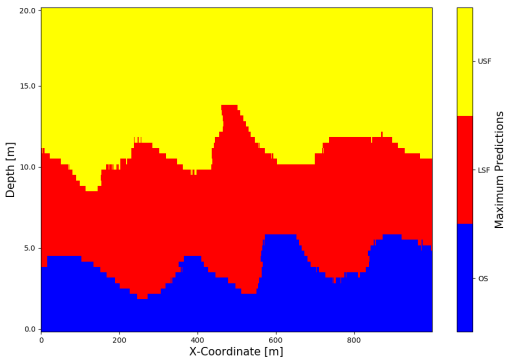
(c) Initialised Models - WC only



(d) Optimised Models - WC only

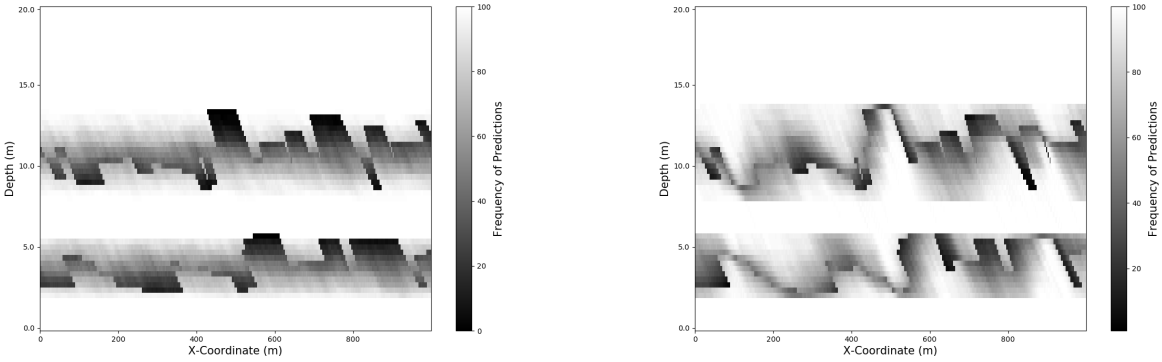


(e) Initialised Models - DA + WC only



(f) Optimised Models - DA + WC

Figure 5.17: Maximum Predictions (MaxPreds) of the geological models generated with the optimised inputs. Top— optimising inputs with only data assimilation. Middle— optimising with only well conditioning. Bottom— optimising with data assimilation and well conditioning.



(a) Initialised Models - DA + WC

(b) Optimised Models - DA + WC

Figure 5.18: Geology distribution plot showing the number of models with same facies type as the reference model at each grid point before and after conditioning and data assimilation. The plots show that the models have lower correlation to the reference data at well locations (x=100m, 500m and 900m) before well conditioning and data assimilation, compared to after conditioning.

tional data and inform spatial correlations, might prove beneficial. Furthermore, incorporating recurrent neural networks (RNNs) to condition individual surfaces based on temporal sequences can significantly improve the results by accounting for variations along well locations. Our use of neural networks for generating conditioned, grid-free, surface-based models could be extended to grid-based, object- and process-based modelling, and 3D models.

5.5 Conclusions

Representative models of the subsurface support decision-making in reservoir management for developing production strategies, predicting future performance, and performing economic analysis. In geothermal reservoir-well models, the number of parameters is orders of magnitude higher than the observed temperature measurements. Generating models that honour observed data requires more efficient data assimilation techniques that can solve such ill-posed problems rapidly in place of traditional data assimilation techniques that can be computationally expensive. This study demonstrates the application of neural networks for history-matching reservoir models by leveraging the computational convenience and geological realism of surface-based geological models, and adversarial learning for regularising a neural network to map geological models to historical production for use in a machine learning-based data assimilation framework for history-matching geothermal reservoir models.

Chapter 6

Transformer-Based Reduced Order Models of Geothermal Reservoirs with Unstructured Meshes

6.1 Abstract

Reduced order modelling allows for the representation of high-dimensional problems in reduced space where the number of parameters or controls is decreased, and the physical behaviour of dynamic systems is preserved with reasonable fidelity. We apply neural networks to develop reduced order models (ROMs) of geothermal systems that represent the high-order model and can run at a fraction of the computational expense of the full-physics model. This study presents a preliminary application of data compression using convolutional autoencoders with space-filling curves on unstructured meshes for generating reduced order models of subinterpolation,voirs. The technical contributions of this work are:

1. The use of space-filling curves to represent high-dimensional geothermal reservoir simulation data on unstructured meshes in one-dimension to aid the application of convolutional autoencoders for compressing data over two orders of magnitude lower.
2. Temporal evolution of the compressed variables in the reduced space using a transformer-based model as a predictive neural network to replace the geothermal reservoir simulator.

The massive speed-up from ROMs reduces the computational expense of reservoir forecasting and uncertainty quantification, where multiple reservoir simulations are needed for exploring controls and the parameter space.

6.2 Introduction

The simulation of geothermal reservoirs is a complex process that involves modelling large-scale, nonlinear physical behaviour that characterises the flow of fluids in the subsurface. This involves solving a high-dimensional, computationally demanding problem. Direct numerical simulations (DNS) yield high-fidelity models that approximate the real-world processes we observe. However, high-fidelity models, also known as full-physics models or high-order models, can be computationally expensive to develop for high-dimensional systems or even intractable where hundreds or thousands of models are required for use cases such as sensitivity analysis, uncertainty quantification, and design optimisation (Quarteroni and Rozza, 2014). Reduced order models (ROMs) have emerged as a class of models which can preserve the large-scale behaviour of a system with reasonable fidelity, at a fraction of the computational expense of the high-fidelity models from numerical simulations (Lucia et al., 2004). Also known as surrogate or proxy models, ROMs approximate the evolution of a system by using low-order representations of the high-fidelity model.

Reduced order modelling techniques have been explored in computational fluid dynamics with significant success to speed-up simulations (Eivazi et al., 2020; Gonzalez and Balajewicz, 2018; Hesthaven and Ubbiali, 2018). To develop reduced-order models (ROMs), snapshots or solutions from the full-physics model are projected onto a lower dimensional space to approximate the evolution of the full-order model.

In geological reservoir simulation, the commonly applied technique involves using the Galerkin method to project the governing equations onto the reduced basis space generated by applying a proper orthogonal decomposition (POD) approach on snapshots of the full order simulation. In POD, a singular value decomposition (SVD) (Henry and Hofrichter, 1992) is performed on state vectors derived from snapshots of simulation runs at various times to obtain the singular vectors with the largest singular values to represent the projection of the model to the reduced space. In combination with piecewise linearisation, POD has also been applied with trajectory piecewise linearisation (TPWL) for developing reduced order models for thermal simulation of in situ heating of unconventional reservoirs (Cardoso and Durlofsky, 2010). However, a key limitation of projection-based methods using POD for dimensionality reduction is that the subspace (low-dimensional space) suffers from not being able to represent physical systems that are highly nonlinear, such as advection-dominated flows and sharp-gradient problems like those encountered in subsurface reservoir simulations. Although techniques such as the discrete empirical interpolation method (DEIM) (Xiao et al., 2014), Petrov-Galerkin (Carlberg et al., 2017; Choi et al., 2020), adaptive algorithms (Carlberg, 2015) and shifted POD (Reiss et al., 2018) have been developed and applied to handle these limitations, the applications have usually been successful on a case-by-case basis, or in problems where some prior knowledge of the system is applied to derive a special treatment that enhances the subspace representation of the full-order simulations.

Data-driven models are particularly useful in fields such as Global Economic Markets, Human Behaviour, and

Social Dynamics, which lack governing equations that can be used to describe the real-world processes observed in systems. Data-based/data-directed/data-controlled/data-driven methods, particularly non-intrusive ROMs (NIROMs) have also been used for generating low-order models of physical systems. NIROMs perform this without the need to depend on the governing equations or source code of the high-fidelity model (Xiao et al., 2017) as the reduced order models are generated in a data-driven approach using samples of model inputs and outputs as training data. The evolution of the model is approximated by interpolating the reduced snapshots using projection-based methods such as radial basis functions, cubic interpolation and data-driven methods such as machine learning techniques. Methods such as POD have been applied for dimensionality reduction, and radial basis functions (RBF) for the interpolation of the reduced snapshots of the training data (Kostorz et al., 2020). Samuel and Muggeridge (2022) have also demonstrated the application of NIROMs to generate ROMs of gas reservoirs that run at least an order of magnitude faster than the high-fidelity models. However, Kostorz et al. (2021) reported that the application of NIROMs in two phase porous media is difficult to model accurately due to non-smooth behaviours characterised by a propagating shock such as in immiscible displacements. Yu et al. (2019) presents a review of NIROM.

6.2.1 Deep Learning and Low-Order Models

Since the simulation of geothermal reservoirs remains an area of active research, the application of ROMs to geothermal reservoir simulation is limited to traditional methods such as POD (Li et al., 2020), TPWL (Cardoso and Durlofsky, 2010) and high-order regression-based model order reduction methods (Mudunuru et al., 2017). Although neural networks have been applied over the past two decades in exploration, drilling, well engineering, reservoir characterisation, and engineering of geothermal systems (Okoroafor et al., 2022), applications of neural networks for reduced order modelling in geothermal reservoirs have been limited to simulating wellbore and reservoir pressures and temperatures (Ariturk, 2018; Aydin et al., 2020; Duplyakin et al., 2022; Porkhial et al., 2015; Shi, Song and Song, 2021).

In subsurface reservoirs, Mo et al. (2018) developed ROMs of carbon storage process-based multiphase flow using image-to-image regression with a convolutional encoder-decoder network. Jin et al. (2020) applied autoencoders for data compression and a linear transition model for state evolution in a heterogeneous hydrocarbon reservoir. Recurrent Neural Networks (RNNs), commonly used for processing sequential data such as time series data have also been applied for generating proxy models for CO_2 sequestration (Amini and Mohaghegh, 2019). By combining deep residual recurrent neural networks (DR-RNN) for time-series prediction with PODGalerkin and DEIM for model order reduction, Kani and Elsheikh (2019) derived reduced order models for uncertainty quantification (UQ) tasks. Long short-term memory (LSTM) networks (Gers et al., 1999; Hochreiter and Schmidhuber, 1997) which have been proposed for mitigating issues with RNNs such as the vanishing gradient problem have also been applied in developing ROMs of subsurface reservoirs. Temirchev et al. (2020) proposed

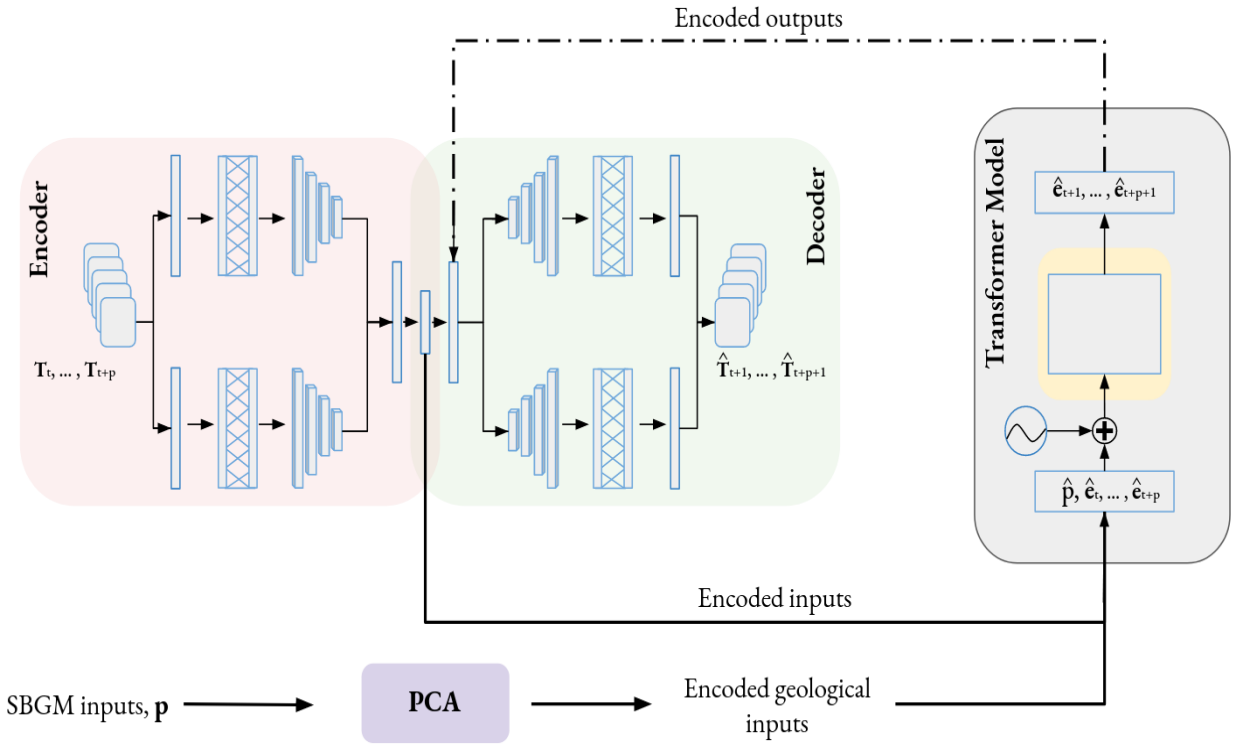


Figure 6.1: The proposed method for generating Reduced order models using space-filling curves with convolutional autoencoder (SFC-CAE) and a transformer-based predictive network. The SFC-CAE performs the data transformation and compression. The predictive network evolves the model state in the latent space. The predicted model states are fed into the decoder to be reconstructed to the high-dimensional space.

the use of variational autoencoders (VAEs) to represent the low-dimensional state of a hydrocarbon reservoir, and gated recurrent units (GRUs) for the dynamic simulation to construct ROMs of 3D reservoirs. Tang et al. (2020) applied convolutional and recurrent neural networks for predicting dynamic pressure and saturation maps and well rates of new geological realisation in a history matching framework. Recently, transformer-based models (Vaswani et al., 2017), used in Natural Language processing, have been shown to address issues with recurrent networks as transformers use self-attention that do not require inputs to be processed sequentially or need recurrent connections that can result in vanishing gradient. Furthermore, the parallelisation afforded by the self-attention in transformer based model allows for processing huge amount of data more rapidly. In one of the first approaches, Zhang et al. (2023) have applied transformers to generate proxy models of hydrocarbon reservoirs on structured grids.

In this work, we explore the use of neural networks for generating ROMs of 2D geothermal reservoir where geological heterogeneity corresponds to the rock permeability, and the temperature of the fluid flowing within the reservoir is modelled over time. We adopt the use of space-filling curves on unstructured meshes for obtaining the low dimensional space. This provides the advantage of preserving geological realism, capturing key fluid flow behaviours, and greater flexibility for resolving complex geometries. We use surface-based geological models (SBGMs) that more accurately capture geological architectures, reduce computational costs, and are not limited

to a predefined grid or require upscaling (Jacquemyn et al., 2019) to generate reservoir models of geothermal reservoirs. The SBGMs are discretised using unstructured meshes that more accurately represent complex geometries. This motivates the use of space-filling curves (SFCs) that transform multi-dimensional solutions on unstructured meshes into one-dimensional (1D) representation, to which neural networks such as convolutional networks that rely on exploring spatial correlation in structured data can be applied. The application of space-filling curves with convolutional autoencoders (SFC-CAEs) for dimensionality reduction has been demonstrated with remarkable success for fluid flow problems (Heaney et al., 2024). The method of generating SFCs here involves ordering sequences of partitions such that a path traced through the partitions is continuous and the locality of the points traversed is preserved. As with convolutional neural networks (CNNs), this allows for the detection of structural coherence and spatial correlation in the data.

We apply convolutional autoencoders with space filling curves and transformer-based models for unstructured meshes to develop reduced order models that represent the spatio-temporal behaviour of geothermal reservoirs with realistic geological architectures in the following steps:

1. First, we apply SFCs for data transformation. This is necessary to handle the unstructured grids required for generating realistic geological reservoir models. The SFCs allow for a mapping from a high-dimensional space to a one-dimensional space, where deep learning networks that rely on exploring structural/spatial correlations such as convolutional neural networks can subsequently be applied for dimensionality reduction.

2. Using 1D convolutional autoencoders, we compress the data orders of magnitude lower and obtain the latent variables, which represent the low-order model of the full physics simulation.

3. Finally, to generate the reduced order models, we use the latent representation with a transformer-based network for simulating the reservoir in reduced space. The reconstructed simulation is then obtained by decompressing the predicted latent variables using the decoder block of the convolutional autoencoder.

Figure 6.1 highlights the proposed method. The following section describes the methodology: the process of the geothermal reservoir simulation for generating training data for the neural networks, the method of applying the SFCs for transforming the 2D data into 1D, the convolutional autoencoder applied for data compression, and the transformer-based model for the predictive network. Section 6.4 presents the results of the proposed model and discusses the benefits, limitations, and future applications of the methodology. The final section presents the summary and conclusions of this study.

6.3 Methodology

We present the method in detail for generating reduced-order models using convolutional autoencoders with space-filling curves. First, we describe the geothermal reservoir simulation process for generating the training and test dataset in Section 6.3.1. We explain the space-filling curve approach in Section 6.3.2 as a data transformation step in the compression of the full-physics models using the convolutional autoencoders given in Section 6.3.3. The final part, Section 6.3.4, provides the details of the predictive model applied for modelling the temporal evolution of the compressed variables.

6.3.1 Data Generation - Geothermal Reservoir Simulation

Simulating reservoir models begins with generating a representative geological model of the subsurface. We use geological models known as SBGMs that represent geometries with parametric, grid-free surfaces using non-uniform rational B-splines (NURBS) interpolation based on the work of Jacquemyn et al. (2019). Using this approach, the resulting surfaces represent realistic geometries and determine the distribution of heterogeneity as the surfaces define discrete rock volumes within which petrophysical properties such as porosity and permeability are homogeneous. The SBGMs are discretised using an unstructured, triangular mesh. The mesh elements are populated with values of petrophysical properties such as porosity and permeability. We solve the governing porous-media fluid flow equations in the discretised SBGM simulation mesh by using the control volume finite element method (CV-FEM) as implemented in IC-FERST (Jackson et al., 2015; Salinas et al., 2018, 2021).

Considering a 2D single phase fluid flow and heat problem in a geothermal system with two vertical wells (one producer and one injector, as shown in Figure 6.2), and assuming thermal equilibrium and conservation of mass and momentum, the mass balance equation is given as:

$$\frac{\partial(\phi\rho_f)}{\partial t} + \nabla \cdot \boldsymbol{u}\rho_f = s_c, \tag{6.1}$$

where ϕ is the porosity, ρ_f is the pressure and temperature-dependent fluid density, $\boldsymbol{u}$ is the Darcy velocity, and s_c is a source term. The momentum balance equation is:

$$\mu(\boldsymbol{K})^{-1}\boldsymbol{u} = \nabla p + \rho_f \boldsymbol{g} + s_m, \tag{6.2}$$

where $\boldsymbol{g}$ is the vector of the gravitational acceleration, $\boldsymbol{K}$ is the permeability tensor, p is the pressure, μ the fluid viscosity, s_m is a source term in the momentum equation. The temperature field is described by the energy balance equation:

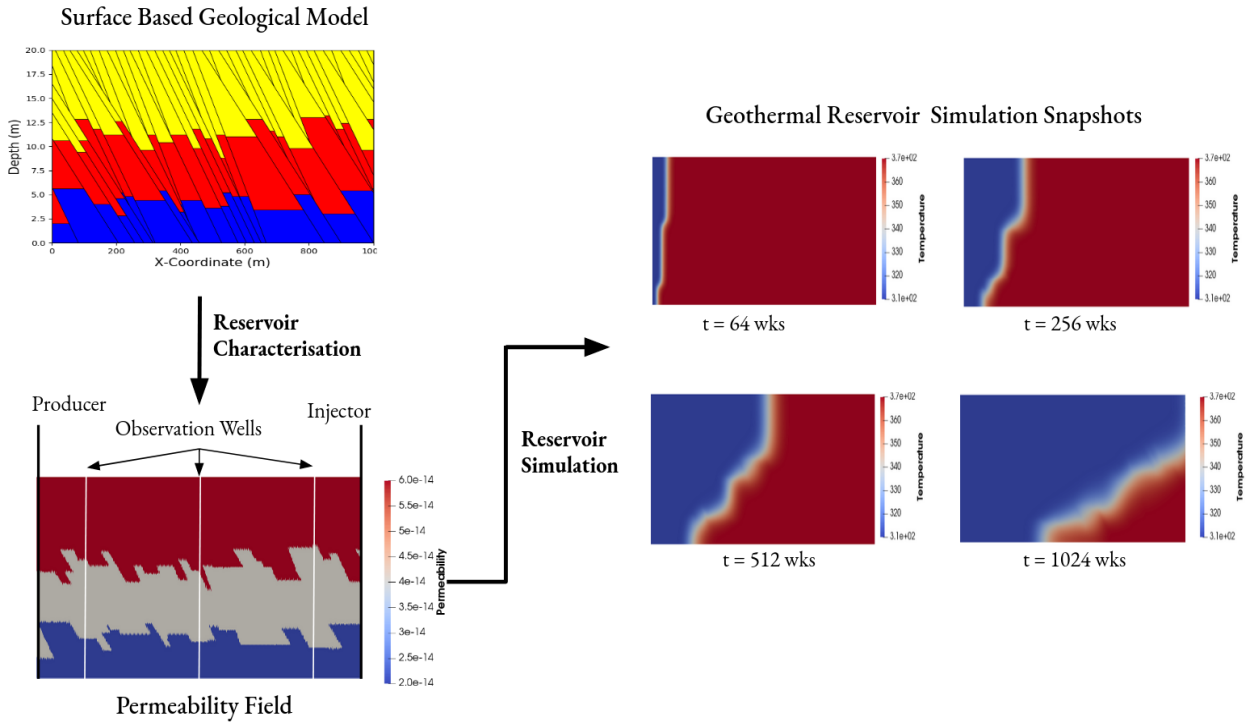


Figure 6.2: The process of creating the dataset for training and testing the compression and prediction neural networks involves generating SBGMs and mapping the permeability field to the facies distribution in the geological model. The resulting geological model is used to develop reservoir simulation models.

$$\xi \frac{\partial T}{\partial t} = -\nabla \cdot (\kappa \nabla T) - \phi \rho_f C_{P_f} \nabla \cdot (\boldsymbol{u} T) + s_t, \tag{6.3}$$

$$\xi = (1-\phi)\rho_p C_{P_p} + \phi \rho_f C_{P_f}, \tag{6.4}$$

$$\kappa = (1-\phi)\kappa_p + \phi \kappa_f, \tag{6.5}$$

where C_{P_p} is the porous rock heat capacity, C_{P_f} is the fluid porosity, ρ_f is the fluid density, κ_p is the rock thermal conductivity, κ_f is the fluid thermal conductivity, ξ is the effective heat capacity, T is the temperature field, $\boldsymbol{u}$ is the velocity of the fluid obtained using Darcy's law and κ is the effective thermal conductivity. We represent the production and injection wells as the boundaries of the domain, where the injector, I, is at 0 m and the producer, P, is at 1000 m of the domain. Within the domain are three exploratory wells that serve as control points for data collection. For the idealised model considered here, chemical reactions, phase changes, geomechanical effects, and other complex physical behaviours present in typical geothermal settings are neglected.

We discretise Equation (6.3) in space by testing the weak-form using control volume (CV) basis functions and discretizing in time using the Θ-method (see Salinas et al. (2021) for more details). For large number of degrees of freedom, this is computationally intensive and impractical for time-critical problems such as reservoir monitoring where the solutions are required rapidly. Furthermore, applications that require sampling a large parameter

space, such as in uncertainty quantification and sensitivity analysis, will greatly benefit from the speed-up in running multiple models. Reduced-order models, low-order models, or surrogate models are developed to run at a fraction of the computational cost of the full-order model. To develop the reduced order models, we only require the temperature data, $\mathbf{X}$ generated from solving Equation (6.3). As in NIROMs, we assume we do not have access to Equation (6.3), only the data set. The data, $\mathbf{X}$ referred to as the snapshots is derived by sampling the temperature, $\mathbf{T}_t$ within the reservoir at every time step, t of the simulation. The matrix of the data is written as

$$\mathbf{X} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_{ts-1}, \mathbf{T}_{ts}\}, \tag{6.6}$$

where $\mathbf{T}_t \in \mathbb{R}^{N_n}$ is the spatially discretised state variable of the 2D problem (where N_n is the number of nodes of the unstructured mesh) for each time step, given that the total number of time steps is ts.

The training and test dataset, $\mathcal{X}_T = \{\mathbf{X}_1, \mathbf{X}_2 \ldots, \mathbf{X}_{M-1}, \mathbf{X}_M\}$ is composed of M examples of geothermal simulations, where the examples are generated from stochastic SBGMs with varying geometries and architectures whose rock types at the nodes of the computational mesh map to the permeability in the reservoir, given that this petrophysical property is the dominant factor influencing fluid flow in geothermal reservoirs. Figure 6.2 illustrates the data generation process, and Table 6.2 describes the reservoir, fluid, and well properties for the geothermal reservoir fluid flow models.

6.3.2 Data Transformation - Space-Filling Curves

Using recursive construction, SFCs are able to transform multidimensional data into one-dimensional data, while preserving the locality and coherency of data points (Zhou et al., 2021). However, classical space-filling curves such as the Hilbert, Moore, and Sierpinski curves (Hilbert, 1891; Moore, 1990; Platzman and Bartholdi, 1989) cannot be applied directly to unstructured grids without modifications. To generate space-filling curves on unstructured meshes, we apply the method described in Heaney et al. (2024) in the following steps:

1. Construct a graph where the vertices represent the nodes or cells, and edges represent connections between nodes or cells as determined by the computational mesh or discretisation stencil.

2. Create intermediate level of partitions at different levels of refinement using a domain decomposition approach based on nested bisection until almost all of the partitions contain one vertex. The number of levels of decomposition, L, is obtained by finding the smallest integer L that satisfies

$$2^L \geqslant N, \tag{6.7}$$

where N is the total number of nodes or vertices in the original graph of the discretisation stencil (computational mesh).

3. Order the partitions beginning at the coarsest level and ending at the final level to obtain the desired graph partitioning by applying MFT-RNN, a local search method based on recurrent neural networks (RNNs) for graph partitioning using the mean field theorem (MFT) (Pain et al., 1999) which seeks to minimize the functional:

$$F = \frac{1}{2}\left(\boldsymbol{x}^T\mathbf{K}\boldsymbol{x} + \alpha\boldsymbol{x}^T\mathbf{C}\boldsymbol{x}\right), \tag{6.8}$$

where $\boldsymbol{x}$ represents the distance between nodes or cells, $\mathbf{K}$, the communication cost, is a high-order matrix associated with inter-partition communications of the graph and $\mathbf{C}$, the balancing, is a high-order matrix that ensures the number of vertices in each partition at every level are equal. The scalar α controls the balance between the importance of the distribution of vertices throughout the partitions and the importance of reducing the communication costs. The number of partitions created at each level is 2, determined by the dimension of the problem in our case.

The space-filling curve ordering or numbering can be obtained from the vertex orderings. Based on the findings from Heaney et al. (2024), we use two space-filling curves to ensure connectivity between partitions by weighting $\mathbf{K}$ to discourage partitions from being made across the same edges in both SFCs. The combination of two SFCs ensures greater connectivity because although points close together on the one-dimensional space-filling curve are close together in the high-dimensional physical space, the reverse is not true. Applying the SFC transformation, $\mathcal{F}_s$ to the input features yields

$$\mathcal{X}^s = \mathcal{F}_s(\mathcal{X}), \tag{6.9}$$

where the training and test dataset $\mathcal{X}^s = \{\mathbf{X}_1^s, \mathbf{X}_2^s \ldots, \mathbf{X}_{M-1}^s, \mathbf{X}_M^s\}$ and the data for each example $\mathbf{X}_m^s = \{\mathbf{T}_1^s, \mathbf{T}_2^s, \ldots, \mathbf{T}_{ts-1}^s, \mathbf{T}_{ts}^s\}$, given the one-dimensional representation of the temperature for each snapshot as

$$\mathbf{T}_t^s = \{T_{1,t}, \ldots, T_{2,t}, T_{n-1,t}, T_{N,t}\} \in \mathbb{R}^N. \tag{6.10}$$

Feature-scaling

This is performed to remove any bias to values of large magnitudes. For each data point, a feature scaled variable is given as

$$\tilde{T}_{n,t} = \frac{T_{n,t}^s - T_{n,mean}^s}{T_{n,std}^s}, \tag{6.11}$$

where $T_{n,mean}^s$ is the mean temperature at node n and $T_{n,std}^s$ is the standard deviation for node n. Therefore the resulting dataset is of the form

$$\tilde{\mathcal{X}}_T^s = \{\tilde{\mathbf{X}}_1^s, \tilde{\mathbf{X}}_2^s, \ldots, \tilde{\mathbf{X}}_{M-1}^s, \tilde{\mathbf{X}}_M^s\}, \tag{6.12}$$

where $\tilde{\mathbf{X}}_m^s = \{\tilde{\mathbf{T}}_1^s, \tilde{\mathbf{T}}_2^s \ldots, \tilde{\mathbf{T}}_{ts-1}^s, \tilde{\mathbf{T}}_{ts}^s\}$ consists of the scaled input features.

6.3.3 Data Compression - Convolutional Autoencoders

Autoencoder networks were described by Hinton and Salakhutdinov (2006) as a nonlinear generalisation of principal component analysis (PCA) for dimensionality reduction. These neural networks consist of an encoder and a decoder. The encoder compresses the data to a lower dimensional latent space and the decoder reconstructs the input data from the reduced space to its initial high-order dimensional space. Variants of autoencoders including fully connected autoencoders (Zhang et al., 2021), variational autoencoders (Canchumuni et al., 2019*b*) and convolutional autoencoders (Mo et al., 2018) have been applied in reservoir modelling and simulation to learn the low-dimensional representation of state variables.

We choose convolutional autoencoders (CAEs) for model order reduction due to the ability of CNNs to preserve spatial coherence by representing local features through its convolutional filters (Xu and Duraisamy, 2020). However, this limits the application of CAEs for data compression and generation of ROMs to data on structured grids. In a conventional convolutional autoencoder, the encoder network gives the low-dimensional representation

$$\hat{\mathbf{e}} = \mathcal{F}_{\mathcal{E}}(\tilde{\mathbf{T}}^s; \boldsymbol{\theta}_{\mathcal{E}}), \tag{6.13}$$

given that $\hat{\mathbf{e}} \in \mathbb{R}^{N_e}$ is the low-dimensional representation of the input state, also known as the latent state, with $N_e \ll N$, and $\boldsymbol{\theta}_{\mathcal{E}}$ refers to the trainable parameters of the encoder network consisting of the weights, $\mathbf{W}$, and biases $\mathbf{b}$. The decoder $\mathcal{F}_{\mathcal{D}}$, outputs the reconstructed input, $\hat{\mathbf{T}}^s$ from the latent state as follows:

$$\hat{\mathbf{T}}^s = \mathcal{F}_{\mathcal{D}}(\hat{\mathbf{e}}; \boldsymbol{\theta}_{\mathcal{D}}). \tag{6.14}$$

Training the network in supervised learning mode involves finding the parameters, $\boldsymbol{\theta}_{\mathcal{E}}$ and $\boldsymbol{\theta}_{\mathcal{D}}$ of the encoder and decoder respectively, that minimize the data mismatch, also referred to as reconstruction error, over all training examples. This is given as

$$\boldsymbol{\theta}_{\mathcal{E}}, \boldsymbol{\theta}_{\mathcal{D}} = \underset{\boldsymbol{\theta}^*_{\mathcal{E}}, \boldsymbol{\theta}^*_{\mathcal{D}}}{\operatorname{argmin}} [\mathcal{L}_{\mathcal{D}}(\hat{\mathbf{T}}^s, \tilde{\mathbf{T}}^s)], \tag{6.15}$$

where $\mathcal{L}_{\mathcal{D}}$ is the loss function which measures the discrepancy between $\hat{\mathbf{T}}^s$ and $\tilde{\mathbf{T}}^s$ as a mean squared error (L^2-norm)

$$\mathcal{L}_{\mathcal{D}}(\hat{\mathbf{T}}^s, \tilde{\mathbf{T}}^s) = \|\hat{\mathbf{T}}^s - \tilde{\mathbf{T}}^s\|_2^2. \tag{6.16}$$

To use CAEs with data on unstructured meshes, we apply modifications to this classic convolutional autoencoder using the space-filling curve approach described in Section 6.3.2 which involve the addition of the SFC transformation and sparse layers. The method of developing the SFC-CAE used in this study is described in

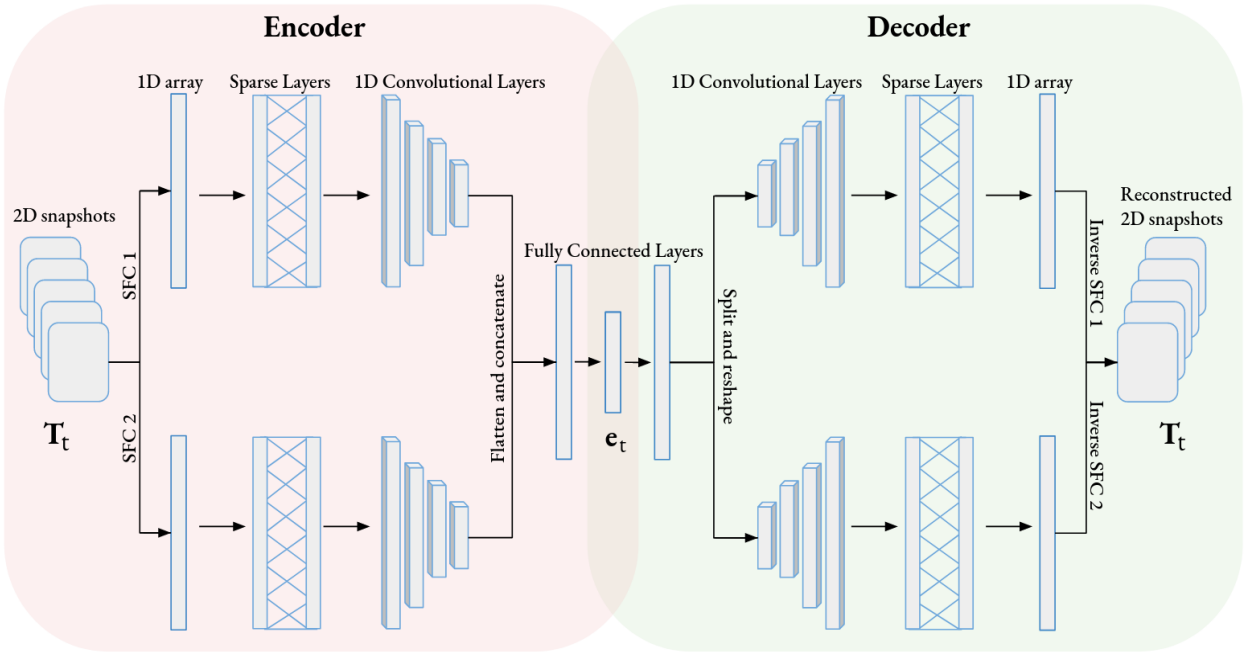


Figure 6.3: The space-filling curve convolutional autoencoder (SFC-CAE) applied for compressing the high-dimensional geothermal reservoir simulation snapshots, $\mathbf{T}$ into the low-dimensional latent space, $\mathbf{e}$. The SFC-CAE uses two SFCs to transform the 2D snapshots to 1D arrays that feed into convolutional and fully connected layers in the encoder network, $\mathcal{E}$. The decoder, $\mathcal{D}$ is a mirror image of $\mathcal{E}$, with the inverse SFC transformation performed to transform the 1D array into the reconstructed 2D snapshots

detail in Heaney et al. (2024) and the modifications to the implementation of Yu (2021) are used in this work. Figure 6.3 illustrates the SFC-CAE applied in this work for compressing geothermal reservoir simulation data and Table 6.1 provides the hyperparameters of the network.

6.3.4 Temporal Evolution of the Latent State - Transformer-Based Model

The state evolution of the low-dimensional dynamics of the geothermal reservoir is represented using a neural network, $\mathcal{T}$

$$\hat{\mathbf{e}}_{t+1} = \mathcal{T}(\hat{\mathbf{e}}_t), \forall t \in \{1, 2, 3, \ldots, N_t\}, \tag{6.17}$$

where the vector $\hat{\mathbf{e}}_t$ is the low-dimensional encoding of $\tilde{\mathbf{T}}$ from the SFC-CAE, given as the output of $\mathcal{E}$.

Inspired by the work of Vaswani et al. (2017), we use a transformer-based model as the predictive network to simulate the low-order geothermal reservoir. Transformers have been widely applied in natural language processing, and in computer vision for learning long-term dependencies in text and videos respectively. More recently, transformers have also been applied to model physical dynamics (Geneva and Zabaras, 2022). We use the generative pre-trained transformer (GPT) (Radford and Narasimhan, 2018; Radford et al., 2019) models based on the implementation in Wolf et al. (2020). The GPT model is a transformer decoder model that consists

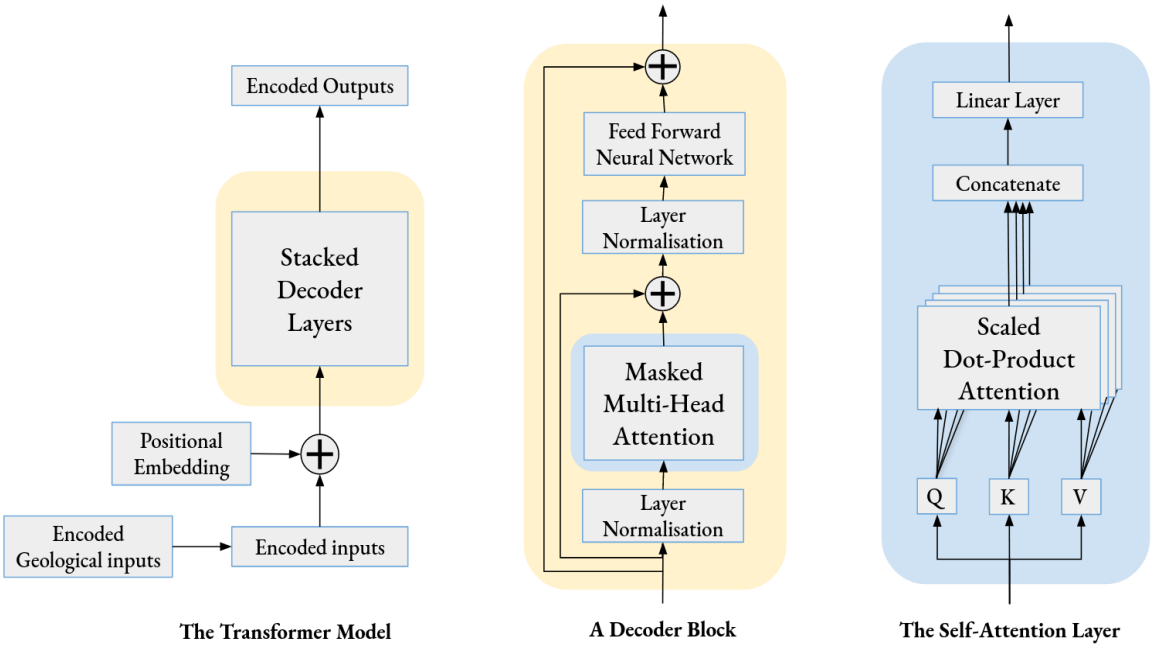


Figure 6.4: The Generative pre-trained transformer (GPT) model applied for temporal evolution (left). The architecture of the network comprises of stacked self-attention and fully connected feed forward layers in the decoder (middle). The absence of recurrence in the models necessitates the use of positional information. The positional encodings are added to the input embeddings (encodings). The positional encodings have the same dimension as the input embeddings which are the low-dimensional representations of the state variables from the encoder. The self-attention layers (right) in the decoder allow each position in the decoder to attend to all positions in the decoder up to and including that position. The masked multi-head attention prevents leftward information flow in the decoder to preserve the auto-regressive property (adapted from Vaswani et al. (2017)).

of a stack of transformer decoder layers that use the masked attention mechanism (see Figure 6.4). The input to our transformer-based predictive model, $\mathcal{T}$ is a set of compressed SBGM parameterised inputs and the set of low-dimensional vectors, $\mathbf{E} = \{\tilde{\mathbf{e}}_0, \tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_{N_t-1}, \tilde{\mathbf{e}}_{N_t}\}$ from $\mathcal{E}$ in the SFC-CAE which represent the encoded geothermal reservoir simulation state variables at each time step. This encoding replaces the input embedding in NLP applications. We use the sinusoidal positional encoding proposed in Vaswani et al. (2017) to allow the model to preserve and learn the relative positions of the input variables and order of the input sequences. This is given as:

$$P_{p,2j} = \sin(p/10000^{2j/N_e}), \tag{6.18}$$

$$P_{p,2j+1} = \cos(p/10000^{2j/N_e}), \tag{6.19}$$

where p is the position of the encoded input $\tilde{\mathbf{e}}$ in the sequence, j is the index of the element in $\tilde{\mathbf{e}}$ and N_e is the encoding dimension.

To capture contextual relationships between input variables, transformers use self-attention and recurrence to allow information to persist and to learn temporal dependencies. Using the attention mechanism, transformers learn longer-term dependency compared to RNNs and LSTMs. Self-attention, also referred to as intra-attention captures the representation of a sequence by relating the different positions of the sequence. Attention models are capable of amplifying the signals in the decoder from the relevant part of the input sequence to produce better results. In a similar way to which RNNs allow information to persist by capturing representations of the previous time step in the hidden states and CNNs extract different features using convolution filters, attention heads perform different tasks related to the structure of the input sequences and incorporate an understanding of relevant inputs to that of the input being currently processed. We describe the self-attention mechanism used in $\mathcal{T}$ in more details.

Self-attention

The self-attention model (see Figure 6.4) used in this work is the scaled dot product, where for each input data point, $\mathbf{e}_n$ we calculate:

$$\mathbf{q}_i = \mathcal{F}_q(\mathbf{e}_n), \quad \mathbf{k}_i = \mathcal{F}_k(\mathbf{e}_n), \quad \mathbf{v}_i = \mathcal{F}_v(\mathbf{e}_n), \tag{6.20}$$

where $\mathcal{F}_q$, $\mathcal{F}_k$ and $\mathcal{F}_v$ are neural networks which map $\mathbf{e}_n$ to vectors $\mathbf{q}_i$, $\mathbf{k}_i$ and $\mathbf{v}_n$, which represent the query, key and values respectively for calculating the attention score, $\alpha_{n,i}$, which provides a measure of the relevance of each input variable, to other inputs in the sequence given as:

$$\alpha_{n,i} = \frac{\exp(\mathbf{q}_n^T \mathbf{k}_i / \sqrt{d})}{\sum_{j=1}^{k} \exp(\mathbf{q}_n^T \mathbf{k}_i / \sqrt{d})}, \tag{6.21}$$

where d is the dimension of the input encoding. The context output $\mathbf{c}_{n,i}$ given as

$$\mathbf{c}_{n,i} = \sum_{i=1}^{k} \alpha_{n,i} \mathbf{v}_i, \tag{6.22}$$

is the output of the attention head. The attention for an entire sequence length can be given as

$$\mathbf{C} = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V}. \tag{6.23}$$

One benefit of the self-attention mechanism is that Equation (6.22) does not need to be processed in the order of the input sequence. This allows massive parallelisation and improved model performance over a shorter period of time. Furthermore, the masked attention applies attention for time steps up to and including the current time step, further reducing computational expense.

Geological context

We include geological information of the reservoir in the neural network by concatenating a compressed set of the parameterised SBGM inputs, $\mathbf{p}$, having a similar dimension to the latent variables to the beginning of each input sequence. The compressed SBGM inputs, $\hat{\mathbf{p}}$ are derived by performing principal component analysis (PCA), a common machine learning approach for data compression based on singular value decomposition (SVD):

$$\hat{\mathbf{p}} = \mathcal{F}_p \mathbf{p}, \tag{6.24}$$

where $\mathcal{F}_p$ defines the PCA transform from the high dimensional SBGM inputs to the reduced variables.

The insertion of $\hat{\mathbf{p}}$ at the beginning of the sequence acts similarly to a start token in natural language processing which provides useful information to decoder networks such as in the transformer-based model in this work. The use of geological constraints as side information in transformers has been suggested for improving model performance in transformer-based models for subsurface reservoir application (Zhang et al., 2023).

Training - Online

We train the transformer-based prediction network, $\mathcal{T}$ using the auto-regressive approach of Radford and Narasimhan (2018), where the set of current and previous time steps, $\mathbf{E} = \{\tilde{\mathbf{e}}_0, \tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_{N_t-1}, \tilde{\mathbf{e}}_{N_t}\}$ is used to present the next time step,$\mathbf{e}_{t+1}$. The loss, $\mathcal{L}_\mathcal{T}$ is the standard mean squared error, given as:

$$\mathcal{L}_{\mathcal{D}} := \| \mathcal{T}(\hat{\mathbf{e}}_t) - \hat{\mathbf{e}}_{t+1} \|_2 \, . \tag{6.25}$$

The hyperparameters and modifications to the GPT2 implementation of Wolf et al. (2020) used in this work are given in Table 6.1.

Table 6.1: Hyperparameters and performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$).

	SFC-CAE	**GPT**
Hyperparameters		
Learning rate	5e-5	5e-5
Train-Val-test split	0.90-0.03-0.07	0.75-0.20-0.05
Number of epochs	400	26
Optimizer	NAdam	AdamW
Weight decay	0.0	-
Betas	(0.8,0.999)	-
Warm-up steps	-	100
LayerNorm epsilon	-	10^{-8}
Weight initialisation	Kaiming (He et al., 2015)	Pre-TrainedGPT2

Prediction - Offline

The SFC-CAE and Transformer-based network represent the spatio-temporal high-dimensional dynamics of the temperature field by integrating the temporal evolution of the low-dimensional latent dynamics. Using the trained network, $\mathcal{T}$ with parameters $\boldsymbol{\theta}_{\mathcal{T}}$, and given input temperature within a context length, c, $\mathbf{T}^0 \ldots \mathbf{T}^c \in [-1, 1]^{N_n}$, we obtain a low-dimensional representation of the initial condition, $\mathbf{v}^0 \ldots \mathbf{v}^c$ using the encoder network. The predictive network is then applied to evolve the latent variables or low-dimensional representation for N_t steps. The decoder is applied to reconstruct the full-dimensional representation of the temperature, $\hat{\mathbf{T}}$ at every time step. The procedure is illustrated in Algorithm 6.1.

6.4 Results and Discussions

6.4.1 Data Compression

We generate the training, validation and test dataset as described in Section 6.3.1. Table 6.2 describes the reservoir, fluid, and well properties for the fluid flow models. The training, validation and test data sets consist of 90, 3 and 7 fluid flow simulations, respectively, each having a different geology (obtained from the permeability field) and having 512 randomly selected timesteps between $t = 0$ and $t = 1024$ with time step size, $\Delta t = 604\,800\,\mathrm{s}$ or 1 week. Therefore, the final dataset consisted of 90×1024 training examples, 3×1024 validation examples, and 7×1024 test examples. The SFC-CAE presented in Section 6.3.3 is trained to reconstruct the original

Algorithm 6.1 Generating ROM with SFC-CAE and Transformer-based model

set number of prediction steps, N_t
read initial conditions $\mathbf{T}_0$
load and freeze the trained $\mathcal{E}, \mathcal{D}$ and $\mathcal{T}$
apply SFC transformation: $\mathbf{T}_0^s \leftarrow \mathcal{F}_S(\mathbf{T}_0)$ ▷ $\mathbf{T}_0^s \in \mathbb{R}^N$
apply feature scaling: $\tilde{\mathbf{T}}_0^s \in [-1, 1]$
compress SBGM inputs: $\hat{\mathbf{p}} = \mathcal{F}_p \mathbf{p}$,

encoder forward pass: $\hat{\mathbf{e}}^0 \leftarrow \mathcal{F}_{\mathcal{E}}(\tilde{\mathbf{T}}_0^s)$
▷ $\hat{\mathbf{e}}^0 \in \mathbb{R}^{N_e}$

for $t \in 1, 2, \ldots, N_{t-1}$ **do**
  $\hat{\mathbf{e}}^t \leftarrow \mathcal{F}_{\mathcal{T}}(\hat{\mathbf{e}}^{t-1})$
get $\hat{\mathbf{E}} = \{\hat{\mathbf{e}}^1, \hat{\mathbf{e}}^2, \ldots, \hat{\mathbf{e}}^{N_t-1}, \hat{\mathbf{e}}_t^N\}$

decoder forward pass: $\hat{\mathbf{X}}^s \leftarrow \mathcal{F}_{\mathcal{D}}(\hat{\mathbf{E}})$ ▷ $\hat{\mathbf{X}}^s = \{\hat{\mathbf{T}}_1^s, \hat{\mathbf{T}}_2^s, \ldots, \hat{\mathbf{T}}_{N_t}^s \hat{\mathbf{T}}_{N_t}^s\} \in \mathbb{R}^N$

apply inverse SFC transformation: $\hat{\mathbf{X}} \leftarrow \mathcal{F'}_s(\hat{\mathbf{X}}^s)$

return $\hat{\mathbf{X}} = \{\hat{\mathbf{T}}^1, \hat{\mathbf{T}}_2, \ldots, \hat{\mathbf{T}}_{N_t-1}, \hat{\mathbf{T}}_{N_t}\}$

high-fidelity geothermal reservoir simulation snapshots with reasonable accuracy. Figure 6.5 shows the training loss over 400 epochs. The trained model encodes the temperature field to an embedding dimension of 128, $\mathcal{F}_{\mathcal{S}} : \mathbb{R}^{26990} \rightarrow \mathbb{R}^{26990}; \mathcal{F}_{\mathcal{E}} : \mathbb{R}^{26990} \rightarrow \mathbb{R}^{128}; \mathcal{F}_{\mathcal{D}} : \mathbb{R}^{128} \rightarrow \mathbb{R}^{26990}$. We choose the encoding dimension, $N_e = 128$ which is large enough to preserve key flow behaviour as the SFC-CAE discards information while extracting dominant features from the high-dimensional flow field into the dimension of latent space. The final mean squared error from the training is 0.001 which is a reduction of over two orders in magnitude from the starting loss.

Figures 6.6a and 6.6b show the reconstructed snapshots at four different time steps for two randomly chosen examples not used in training the convolutional autoencoder. These results validate the efficiency and accuracy of the method for compressing data. The pointwise difference in the temperature fields between the original and reconstructed snapshots is also given as a quantitative measure of the data compression. The SFC-CAE is capable of compressing and reconstructing the high-dimensional data in less than 0.01 s.

6.4.2 State Prediction

Similar to the SFC-CAE, the transformer-based model, $\mathcal{T}$ described in Section 6.3.4 and illustrated in Figure 6.4 is trained with 75, 20 and 5 fluid flow simulations each having 1024 time steps for the training, validation and test datasets respectively. A sequence length (context length) of 128 is chosen for $\mathcal{T}$ with 8 decoder layers, each having 8 masked attention heads. The transformer is trained for 12 epochs. Figure 6.7 shows the loss curve for the training. In Figures 6.8a and 6.8b, we present the predictions for two randomly chosen unseen examples (i.e. samples with permeability field and temperature evolution not used in training the SFC-CAE

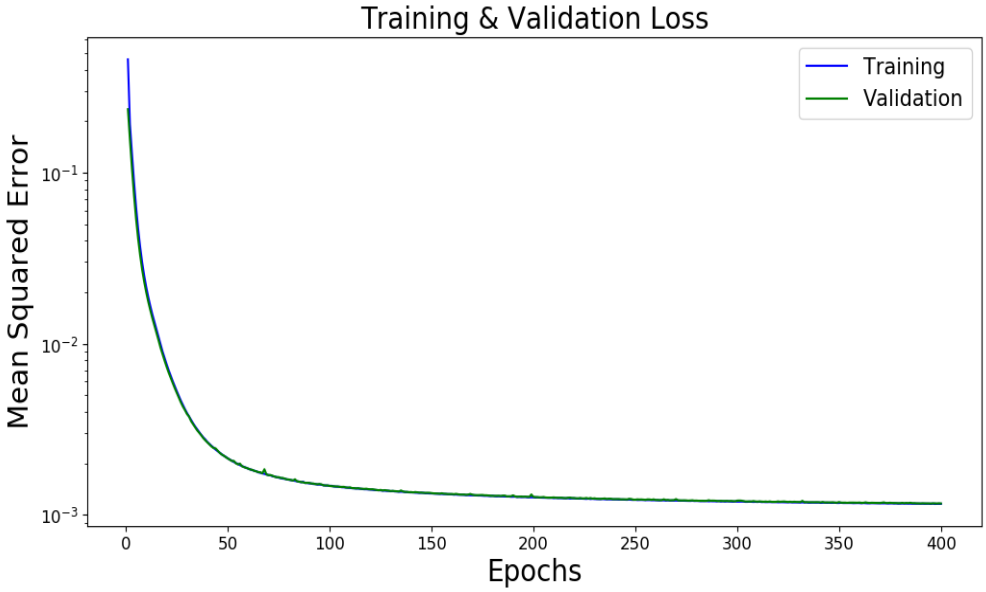
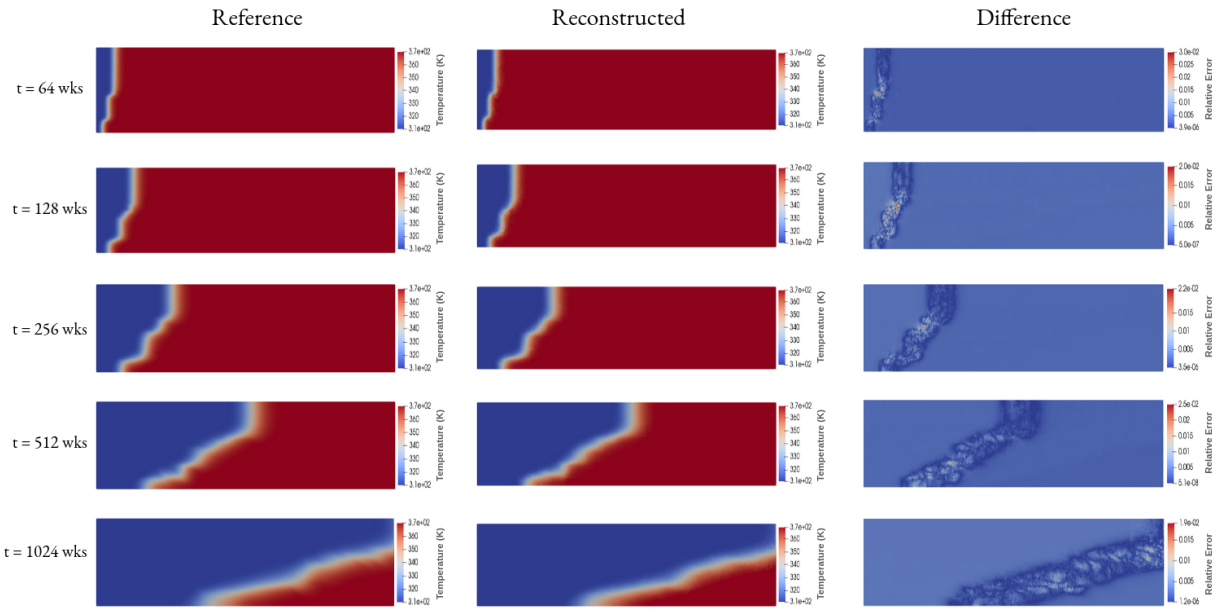


Figure 6.5: The loss curve showing the mean squared error over 400 training epochs of the SFC-CAE. The loss reduces by about three orders of magnitude. A very close match is seen between the training and validation loss showing that the the model has learnt to generalise well.

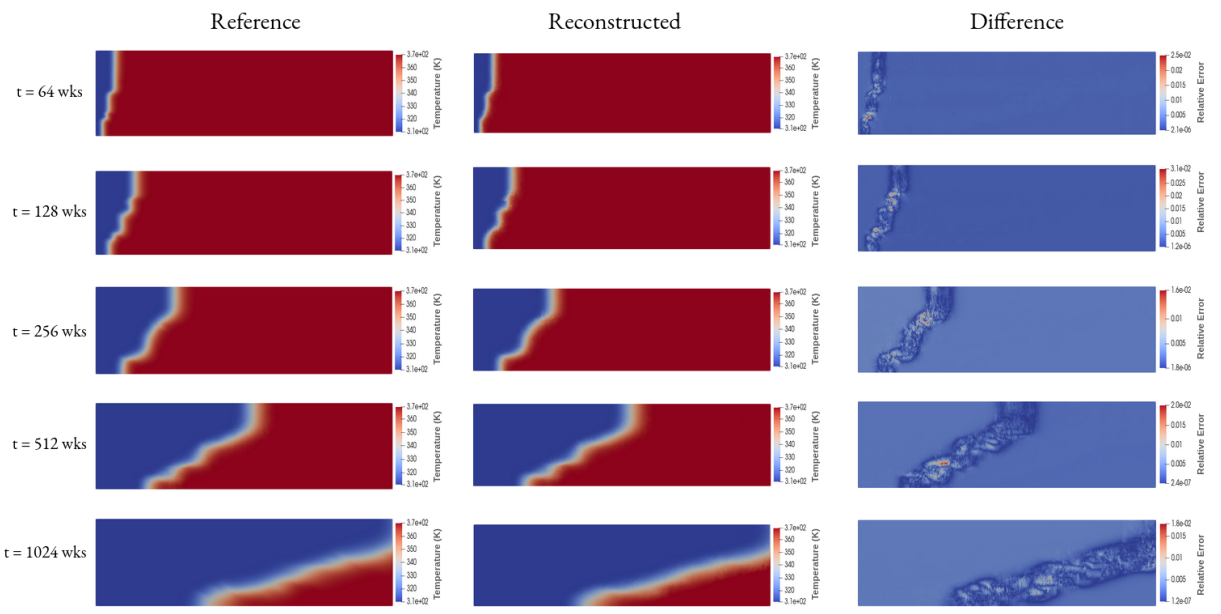
or the GPT2 model). To generate the predictions, the trained $\mathcal{T}$ is provided with some context. This includes the start token, which is the compressed geological inputs, $\hat{\mathbf{p}}$ and the compressed temperature for 255 timesteps $\hat{\mathbf{e}}_0, ..., \hat{\mathbf{e}}_{255}$ (which are concatenated to give the model sequence length of 256 series data points) and allowed to predict the time steps to $t = 1024$.

With the context length provided, the neural network is provided information of the unseen permeability field and temperature. The difference between the predicted model and the high-fidelity simulation for each time step is given as the relative error between the reference and predicted and reconstructed model for two examples not used in training the model (see Figures 6.8a and 6.8b). From these results, it can be seen that the transformer model yields reasonably accurate predictions beyond its context length of 256 data points consisting of the compressed SBGM inputs and 255 temperature data. Figures 6.9a and 6.9b show the root mean squared error between $\mathbf{T}_t$ and $\hat{\mathbf{T}}_t$, $\forall t \in 1, 2, 3, \ldots, N_t$.

The capability of the transformer for predicting reasonably well beyond the provided context length of 256 indicates that it has learned the physical dynamics of the geothermal reservoir models considering the use of only 75 reservoir simulations for training the model. Transformers have been shown to outperform other commonly applied neural networks for generating ROMs such as LSTMs (Geneva and Zabaras, 2022; Zhang et al., 2023). The GPT2 model is capable of predicting the temperature distribution for each time step in 0.01 s given the preceding timesteps of length equal to the model context or sequence length. Although the proposed method requires the use of a sequence of initial time steps, this approach is valuable in subsurface reservoir forecasting, where an ensemble of history-matched models generated from assimilating observed data,



(a) Example 1



(b) Example 2

Figure 6.6: The results of the SFC-CAE for two examples not used in training the model at five different time steps. The original high fidelity model is shown on the left, the reconstructed geothermal simulation is given in the middle and the difference shown as the relative error between the reference and reconstructed model is presented on the right.

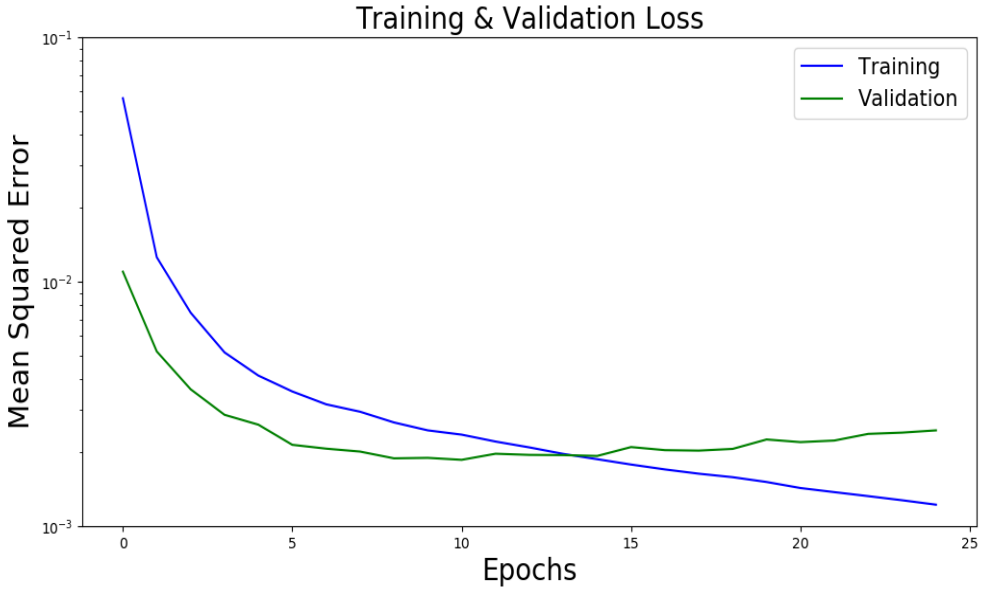


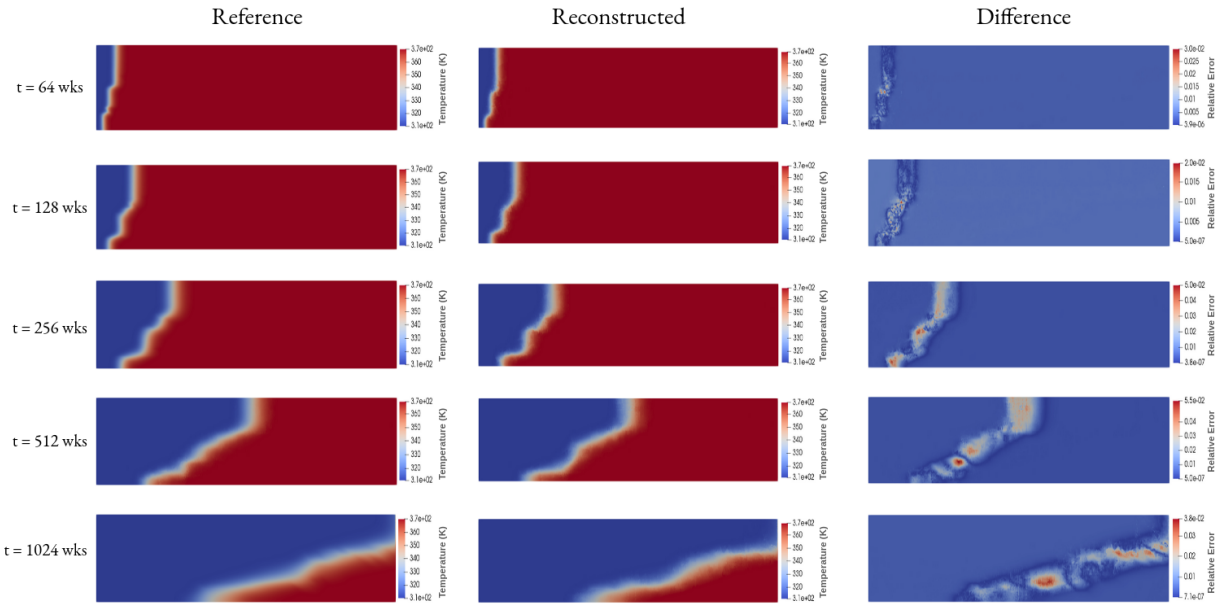
Figure 6.7: The loss curve showing the mean squared error over 12 training epochs of the transformer-based model. Beyond 12 epochs, the model begins to overfit and is therefore unable to generalise to unseen data.

can provide the sequence length for generating predictions far into the future using this method. The advantages of using the transformer for forecasting are immense in terms of the time saved. The GPT model applied in this work is capable of generating reservoir simulation models more than three hundred times faster than the high fidelity model.

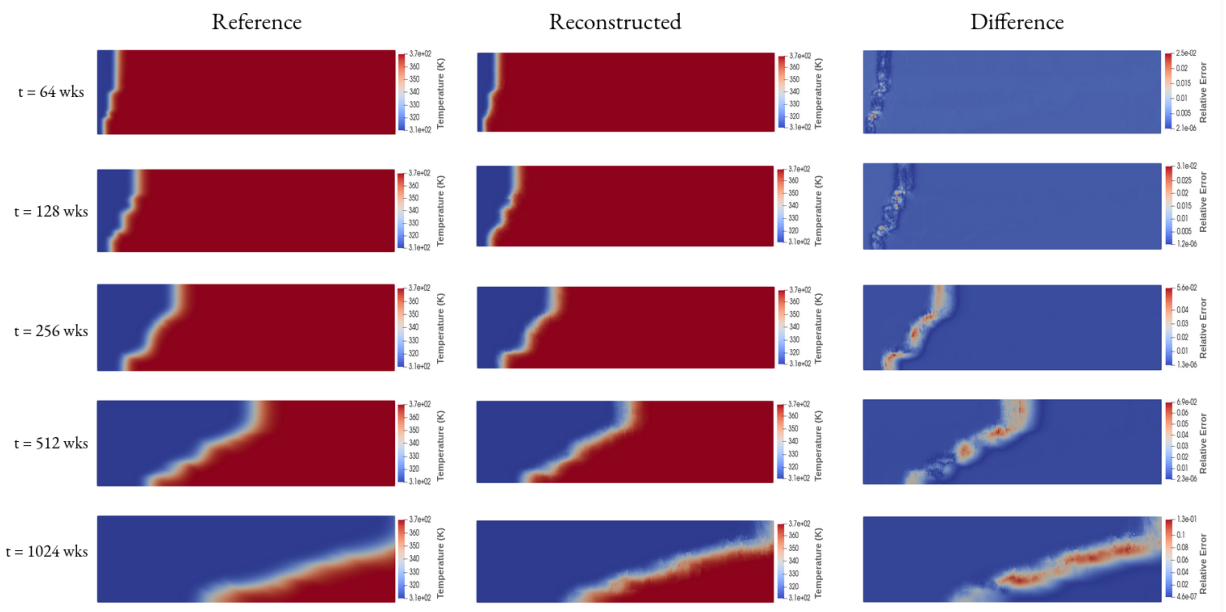
Another benefit of using a transformer-based model as our predictive network, $\mathcal{T}$ is that the use of transformer layers allows for learning complex temporal dependencies since the time integration scheme in numerical methods are also a linear weighting of features from past time steps. Geneva and Zabaras (2022) have shown the mathematical similarity of this commonality between self-attention and numerical time-integration methods and suggests that this may make transformer-based models more suitable for learning dynamical systems than alternative deep learning approaches. Furthermore, the inclusion of time as an input through the use of the positional encoding as required in modelling physical systems with explicit time-dependent terms is useful for transformer-based models.

6.4.3 Discussions and Future Work

Subsurface models have historically been developed using physics-based methods which model the relationship between inputs and outputs of systems that can be described with ordinary or partial differential equations using numerical methods. The use of data-driven methods can be beneficial when physics-based models are computationally expensive or difficult to formulate. For example, in reservoir history matching, it is desirable to have multiple plausible geological and dynamic reservoir models to better evaluate risks. A key benefit of

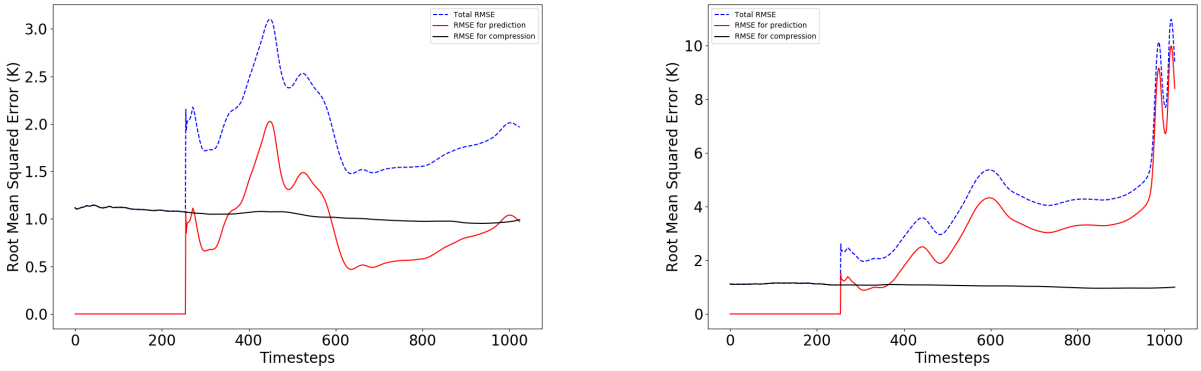


(a) Example 1



(b) Example 2

Figure 6.8: The predictions from the GPT model for two examples not used in training the model at five different time steps. The original high-fidelity model is shown on the left. The predicted geothermal simulation is given in the middle and the difference shown as the relative error between the reference and reconstructed model is presented on the right.



(a) Example 1

(b) Example 2

Figure 6.9: The root mean squared error from the reconstruction and compression step for each time step for the two examples presented in the compression and state prediction.

Table 6.2: Reservoir properties and description of the fluid flow models

Parameter	Value
Reservoir Properties	
Reservoir Extent (m)	1000×200
Initial Reservoir Pressure ($kg/m\,s^2$)	50×10^6
Left Boundary Pressure ($kg/m\,s^2$)	60×10^6
Right Boundary Pressure ($kg/m\,s^2$)	40×10^6
Initial Reservoir Temperature (K)	370
Porosity (%)	20
Rock Density (kg/m^3)	2650
Rock Heat Capacity (W/m K)	1000
Rock Thermal Conductivity (J/K)	2.9
Permeability** (m^2)	2×10^{-10} to 6×10^{-10}
Fluid Properties	
Fluid Density (kg/m^3)	1000
Viscosity (kg/m s)	0.001
Thermal Conductivity (W/m K)	0.15
Heat Capacity (J/K)	4185.5
Phase Volume Fraction (%)	20
Injector Fluid Temperature (K)	310
Well Properties	
Producer Bottom Hole Pressure (BHP) ($kg/m\,s^2$)	40×10^6
Injector Bottom Hole Pressure (BHP) ($kg/m\,s^2$)	60×10^6
Simulation Properties	
Time step (days)	365
Simulation Period (yrs)	32

**The permeability field is populated by mapping the facies type at each point of the conditioned model to the corresponding value of permeability of the rock type, i.e USF = 2.0E-10, LSF = 4.0E-10 and OS = 6.0E-10

Table 6.3: Performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$).

	SFC-CAE	**GPT**
Metrics		
Training Loss	0.001	0.001
Test Loss	0.001	0.002
Training time on GPU (hrs)	5.81	0.63

this method is its potential application for real-time data analysis of field observations and forecasting for rapid reservoir management, such as prediction under uncertainties, to gain insights on optimal decisions to improve efficiency and productivity. The developed reduced order models can also be applied to speed up optimisation of well placements where multiple simulations are required and sensitivities to controls analysed or to carry-out initial explorations in the hyperparameter space in the optimisation step where the full-physics or high-fidelity model can also subsequently be used to refine the solution in the vicinity of the optimal solution determined from the initial explorations of the controls.

The proof of concept proposed in this work provides a foundation for developing ROMs of geothermal reservoirs that allow for the rapid generation of an ensemble of models. However, extending the applications of this work to more complex, field cases will require some improvements and modifications. Similar work by Zhang et al. (2023) for a 3D hydrocarbon reservoir over a twenty-year production period with 80 days for each time step and $24 \times 25 \times 15$ blocks (having a dimension of 9000 and over three times less than the illustrated models in this study) required 500 realizations to represent spatial heterogeneity. Based on the combinations of permeability, porosity, well locations and type (injector or producer), the authors used 24 000 K training examples. Since transformers, primarily used in natural language processing for training models require large datasets, model performance for more complex reservoirs could be improved with more training data or the use of strategies such as weight initialisation and learning rate update schemes (Xu et al., 2021), and the addition of other reservoir and well properties as additional constraints to the model. Furthermore, a physics-informed learning (Karniadakis et al., 2021) approach to training the transformer-based network could improve the results, especially to extrapolate the solution far into the future beyond the simulation period of the training dataset.

6.5 Conclusions

Convolutional autoencoders (CAEs) with space-filling curves (SFCs) that transform unstructured data into 1D representations can be applied to perform data compression effectively. The use of the SFC-CAE as an encoding model provides a unique and powerful method for projecting high-dimensional dynamical systems into a low-dimensional vector representation that can be used in predictive models to generate reduced order models. Since the evolution of the latent space defined in time is also a dynamic process, it is possible to represent the

system as a time series data and model this using transformer-based models such as the Generative Pre-trained Transformer(GPT). The commonality between the self-attention mechanisms in transformers and numerical time-integration methods makes transformer-based models more suited for learning dynamical systems than alternative deep learning approaches.

This study presents the first application of 1) SFC-CAE to subsurface reservoirs for the compression of data on unstructured meshes which better capture irregular geometries and architectures typical of subsurface reservoirs, and 2) GPT transformer-based models for predicting and extrapolating spatio-temporal data in geothermal reservoir models. Future work will aim to enable the application of the developed approach with physics-informed learning for rapid reservoir management and optimisation for field cases.

Chapter 7

A Proposed Workflow for the Simulation of Geothermal Reservoirs with Data Assimilation and Reduced Order Modelling

7.1 Abstract

Geothermal resources in geological settings similar to hydrocarbon reservoirs present new possibilities for the development of geothermal energy. By leveraging the knowledge, technology, and data from the oil and gas industry, reservoir modelling and simulation of geothermal systems can be significantly improved.

A key step in developing representative models of subsurface reservoirs is the history matching process. This is required to produce physically consistent and reliable geological and fluid flow models of the reservoir. However, the high dimensionality and nonlinearity of large-scale models such as geothermal reservoirs can make the history matching process impractical in some cases. Using neural networks, we perform history matching in a data assimilation approach, and reduced-order modelling to generate representative models of geothermal reservoirs that are computationally convenient. We combine the benefits of surface-based modelling, adversarial learning, space-filling curves, and transformer-based models in an integrated framework that can aid rapid reservoir forecasting and management strategies. The technical contributions of this work are as follows:

1. The use of neural networks for generating reservoir models that are consistent with geological processes,

geophysical information, and fluid flow data from wells using parametric, grid-free surface-based geological models (SBGMs) that better preserve geological features and capture fluid flow behaviour in subsurface reservoirs.

2. The application of space-filling curves to represent high-dimensional geothermal reservoir simulation data on unstructured meshes in one dimension to facilitate the application of convolutional autoencoders for data compression.

3. Temporal evolution of the compressed variables in reduced space using a transformer-based model as a predictive neural network to replace the geothermal reservoir simulator for reservoir forecasting.

7.2 Introduction

The production of energy from geothermal resources has the potential to contribute to sustainable energy development as part of a global energy mix (Goldstein et al., 2011; Vargas et al., 2022). Sedimentary basins in geological settings similar to those encountered in oil and gas production are becoming attractive for the production of geothermal energy (Sanyal and Butler, 2010). Reservoir rocks in sedimentary basins, having sufficient temperature, thickness, porosity, and permeability, can extend the development of geothermal energy to areas that were not previously explored (Anderson, 2013; Vivas et al., 2023). Many reservoirs in oil and gas sedimentary basins have a wealth of exploration data such as core samples, seismic data, and logs, which can be used for reservoir characterisation. An additional benefit of exploring sedimentary basins is the opportunity to use existing or abandoned oil and gas wells. This presents new opportunities for the co-production and production of geothermal energy from hydrocarbon reservoirs in various regions across the globe. Furthermore, the extensive experience gained by oil and gas companies and other industry participants through the exploration of hydrocarbon reserves in past and recent years can contribute significantly to the advancement of geothermal projects (Clegg and Krase, 2022). Geoscientists and reservoir engineers can take advantage of their understanding of seismic surveys, deep wells, and knowledge of drilling sedimentary basins to evaluate and exploit geothermal resources while minimising subsurface risks and uncertainties (Hardman et al., 2020). Drawing on extensive research and applications of data assimilation for history matching and reduced-order modelling for reservoir simulation in the oil and gas sector, we apply these techniques to geothermal reservoirs using an artificial intelligence-driven approach.

Data assimilation is applied to integrate observed data into numerical models to produce more realistic and reliable representations of physical systems that can be used for optimisation or forecasting future behaviour (Carrassi et al., 2018). Data assimilation is commonly applied in geophysical problems characterised by limited data, large-scale nonlinearities, and high uncertainty (Asch et al., 2016). Various formulations of data assimilation techniques have been applied in reservoir history matching to identify input parameters to reservoir

simulation, such as petrophysical properties, that minimise the discrepancy between observed and simulated output over a historical production period (Chen and Oliver, 2011). Reservoir history matching aims to calibrate subsurface reservoir simulation models to measured field production data to obtain a reliable model of the fluid flow process (Schiozer, 1999). However, applying traditional history matching methods or conventional data assimilation is computationally expensive due to the sparse data, non-uniqueness or even the intractability of solving the ill-posed inverse modelling task of deriving the underlying spatial distribution of petrophysical properties that correspond to the observed fluid flow behaviour (Abdulrazzaq and Hasan, 2023). Furthermore, the limitations posed by the computational cost, time constraints, and quality of the results of conventional history matching frameworks restrict the application of these methods to real-time reservoir management scenarios.

In this study, a data-driven methodology is applied for history matching, in which well conditioning and data assimilation are performed using machine learning to iteratively condition geological models to observed well log data, and match predictions from reservoir models to fluid flow data, similar to iterative ensemble smoothing methods of data assimilation (ES-MDA) (Emerick and Reynolds, 2013). Data-driven approaches have previously been applied for history matching. For example, generalised regression neural network (GRNN) for history-matching of tight gas carbonate reservoir production profiles (Brantson et al., 2018), physics-informed machine learning for history matching of unconventional reservoirs (Srinivasan et al., 2021), reinforcement learning (deep Q network (DQN) and deep deterministic policy gradients (DDPG)) to predict reservoir permeability for history matching (Li and Misra, 2021), long short-term memory (LSTM) networks for forecasting production of oil and water from the wells (Pal, 2021), generative models to construct parameterisations of the reservoir model in an ES-MDA framework (Canchumuni et al., 2020), and also generative adversarial network (GAN) and ensemble Kalman filter for assimilating dynamic fluid data into reservoir models (Jo et al., 2021). We propose the use of neural networks for history matching with grid-free, computationally convenient surface-based geological models (SBGMs) with compact parameterisation. The use of adversarial learning for history matching reservoir models to production data using machine learning for data assimilation is also demonstrated. The benefit of adversarial learning allows for the use of small datasets to train the network when constraints and/or limits to the domain and data distribution have been been imposed (Li et al., 2021), such as in the well conditioning step to aid the model in learning the relationship between the input parameters of SBGM and the production data. The developed history matching framework also establishes a strong basis for reservoir forecasting using reduced order modelling.

Neural networks have been applied to generate reduced order models (ROMS) that preserve large-scale behaviours of the direct numerical simulation (DNS) and can run orders of magnitude faster than the high-fidelity models from numerical simulations. Applications of neural networks for generating ROMs in hydrocarbon reservoirs include: variational autoencoders (VAEs) and gate recurrent units (GRUs) (Temirchev et al., 2020) for 3D reservoirs, convolutional and recurrent neural networks (Tang et al., 2020), convolutional encoder-decoder

network (Mo et al., 2018), physics-based autoencoders (Jin et al., 2020), deep residual recurrent neural networks (DR-RNN) (Kani and Elsheikh, 2019). Although neural networks have been widely applied over the past two decades in exploration, drilling, well engineering, reservoir characterisation, and engineering of geothermal systems (Okoroafor et al., 2022), applications of neural networks for reduced-order modelling in geothermal reservoirs have been limited to simulating wellbore and reservoir pressures and temperatures (Ariturk, 2018; Aydin et al., 2020; Porkhial et al., 2015; Shi, Song and Song, 2021). By combining history matching with reduced-order modelling in an integrated data-driven workflow, our proposed method provides a tool with the potential for real-time reservoir management of geothermal reservoirs.

The proposed method uses surface-based geological models that offer the advantage of geological realism, reduced computational costs, and are not limited to a predefined grid (Jacquemyn et al., 2019). These allow for the use of unstructured meshes that capture key fluid flow behaviours in reservoir simulation models (Jackson et al., 2015) and motivate the use of space-filling curves (SFCs) that transform multidimensional solutions on unstructured meshes into one-dimensional (1D) representation, to which neural networks such as convolutional networks that rely on exploring spatial correlation in structured data are then applied. We apply convolutional autoencoders with space-filling curves (SFC-CAEs) and a transformer-based network for state evolution to develop reduced-order models that represent the spatio-temporal behaviour of geothermal reservoirs. Figure 7.1 illustrates the proposed framework. We summarise our workflow for history matching geologically constrained reservoir models and developing reduced order models for forecasting the spatio-temporal behaviour of geothermal reservoirs with realistic geological architectures as follows:

1. Generate an ensemble of geological models that have been conditioned to static data (facies type from well logs).

2. Using the trained networks for modelling SBGMs and well production data, perform history matching by generating the inputs to the SBGM that match a given reference geological and production data in an inverse modelling approach.

3. Assess the uncertainties associated with the conditioned geological and history-matched models.

4. Generate the high-fidelity models for the history-matched timesteps and validate this with the reference data.

5. Apply SFCs for data transformation to map the reservoir simulation from a high-dimensional space to a one-dimensional space.

6. Using 1D convolutional autoencoders, compress the data orders of magnitude lower and obtain the latent variables, which represent the low-order model of the full-physics simulation.

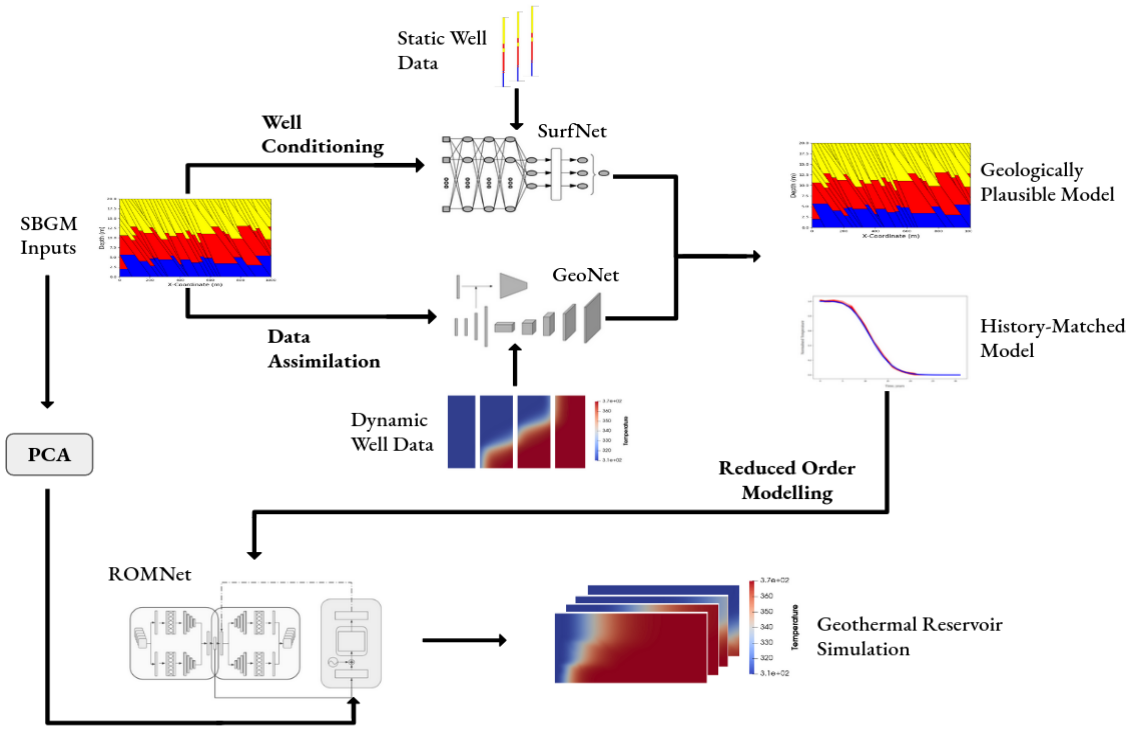


Figure 7.1: The proposed method for rapid generation of history-matched geothermal reservoir models using neural networks for well conditioning, data assimilation and reduced-order modelling.

7. Finally, use the latent representation of the history-matched timesteps with the predictive network to simulate the reservoir in reduced space. The reconstructed simulation is obtained by decompressing the predicted latent variables using the decoder block of the convolutional autoencoder.

We describe the proposed methodology in more detail in Section 7.3 by first discussing the method for generating SBGMs with realistic geological features for the shallow marine depositional system in Section 7.3.1 followed by a description of the geothermal reservoir simulation process in Section 7.3.2. Section 7.3.3 explains the history matching process using well conditioning and data assimilation. Section 7.3.4 explains the method for generating reduced order modelling using neural networks for data compression and state evolution. The results are presented in section 7.4. A discussion of the results of this study along with the benefits, limitations and considerations for future work is presented in Section 7.5.

7.3 Methods

History matching involves building numerical models that simulate the behaviour of the reservoir observed from well data. Dynamic models are developed from static geological models that represent the spatial distributions of the petrophysical properties that govern fluid flow in the subsurface. The following sections describe the method of generating geologically constrained geological models that honour static data and can improve/accelerate the history matching of numerical models to production data.

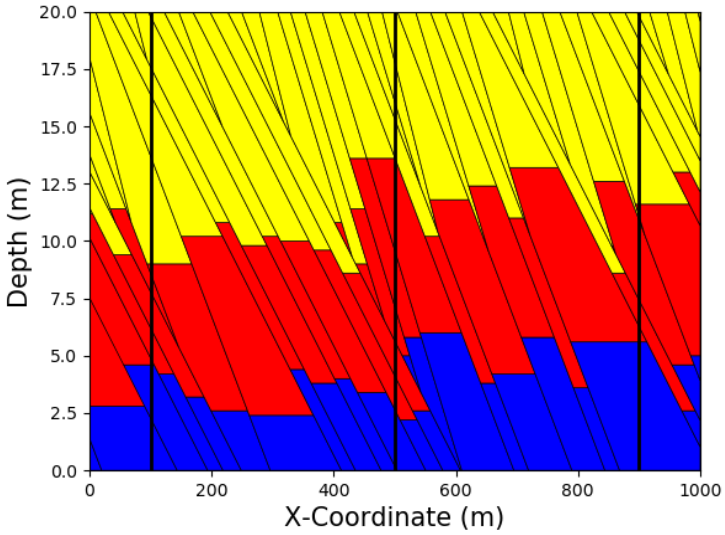


Figure 7.2: An example of a 2D surface-based geological model. This is the reference surface-based geological model used for history matching with three wells located at x=100m, 500m and 900m.

7.3.1 Geological Reservoir Modelling

We apply surface-based geological models (SBGMs) that predict depositional sequences and model geological features using surfaces that define volumes within which petrophysical properties are uniform (Jackson et al., 2014, 2015). These SBGMs model geometries with parametric, grid-free surfaces using NURBS interpolation to describe even complex geological features and preserve geological realism such as the spatial continuity of channelised bodies (Jacquemyn et al., 2019). Two benefits of these SBGMs exist which make them suitable for our applications with neural networks: 1) fewer parameters are required to generate and manipulate even complex geometries compared to existing conventional geostatistical or other rule-based methods, 2) the models are not constrained to a pre-defined grid over any length-scale and thus, the resolution of the model is not impacted.

In this study, 2D SBGMs of a simplified subsurface reservoir in a shallow marine depositional environment with a 20m thick and 1000m long single parasequence including dipping clinoforms and facies boundaries such as shown in Figure 7.2 are used. Three facies are considered: upper shoreface (USF; yellow), lower shoreface (LSF; red) and offshore mudstone (OM; blue). The inputs to the SBGMs are derived from stochastically sampled variables from uniform distributions that parameterize the shape, orientation, and relative location of the surfaces (depositional or erosional), which are sequentially placed to represent the clinoform surfaces and facies boundaries. In total, 60 clinoforms are constructed to fill the parasequence; each clinothem is subdivided by two facies boundaries and a spacing so that the model requires 240 input values (i.e. 60 $\times$ clinoform dip angle, clinoform spacing, lower facies boundary position, upper facies boundary position).

7.3.2 Reservoir Simulation

We generate high-fidelity numerical models with the SBGMs to produce historical production data for the history matching step. The SBGMs are discretised using an unstructured, triangular mesh and more accurately and realistically capture fluid flow behaviour in the subsurface (Jackson et al., 2015). The mesh elements are populated with values of petrophysical properties such as porosity and permeability. A geothermal system is modelled by solving partial differential equations that describe the physical behaviour of fluids in the subsurface. The mathematical model applied in this research is for a confined and saturated reservoir, predominantly composed of consolidated sedimentary material. Flow of a single-phase, Newtonian fluid in an isotropic, non-deformable reservoir is modelled. Geochemical and geomechanical effects are neglected and thermal equilibrium between rock and fluid is assumed.

The governing heat and momentum transport equations are solved using the double control volume finite element method (DCVFEM) proposed by Salinas et al. (2018) to improve the convergence rate using a dual mesh. With this approach, velocity is discretised using Finite Elements (FE) while pressure and temperature are discretised using Control Volumes (CVs). Petrophysical properties such as permeability and porosity are represented element-wise. Thermal conductivity, heat capacity, and rock density are represented CV-wise. Fluid properties such as viscosity and density are represented CV-wise. Time is discretised using an implicit θ-method. Assuming conservation of mass and momentum, the mass balance equation is given as:

$$\frac{\partial(\phi\rho_f)}{\partial t} + \nabla \cdot \boldsymbol{u}\rho_f = s_c, \tag{7.1}$$

where ϕ is the porosity, ρ_f is the pressure and temperature-dependent fluid density, $\boldsymbol{u}$ is the Darcy velocity, and s_c is a source term. The momentum balance equation is:

$$\mu(\boldsymbol{K})^{-1}\boldsymbol{u} = \nabla p + \rho_f \boldsymbol{g} + s_m, \tag{7.2}$$

where $\boldsymbol{g}$ is the vector of the gravitational acceleration, $\boldsymbol{K}$ is the permeability tensor, p is the pressure, μ the fluid viscosity, s_m is a source term in the momentum equation. The temperature field, T is described by the energy balance equation:

$$\xi\frac{\partial T}{\partial t} = -\nabla \cdot (\kappa \nabla T) - \phi\rho_f C_{P_f} \nabla \cdot (\boldsymbol{u}T) + s_t, \tag{7.3}$$

$$\xi = (1-\phi)\rho_p C_{P_p} + \phi\rho_f C_{P_f}, \tag{7.4}$$

$$\kappa = (1-\phi)\kappa_p + \phi\kappa_f, \tag{7.5}$$

where C_{P_p} is the porous rock heat capacity, C_{P_f} is the fluid porosity, ρ_f is the fluid density, κ_p is the rock

thermal conductivity, κ_f is the fluid thermal conductivity, ξ is the effective heat capacity, $\boldsymbol{u}$ is the velocity of the fluid obtained using Darcy's Law and κ is the effective thermal conductivity. The simulation of geothermal reservoirs using surface-based modelling is presented in more detail in Salinas et al. (2021). Since permeability is the main petrophysical control on fluid flow and consequently on heat and mass transport in geothermal reservoirs (Blank et al., 2021), the ensemble of reservoir models is generated by mapping the permeability of the facies in the SBGMs to the pemeability at the nodes of the computational mesh. The production and injection wells are represented as the left and right boundaries of the reservoir model, where the injector, I, is at 0m and the producer, P, is at 1000m. Within the domain are three exploratory wells at $x = 100m$, $x = 500m$, and $x = 900m$ that serve as control points for data collection. The historical production data derived comprises: velocity ($\boldsymbol{u}_x$ and $\boldsymbol{u}_y$), pressure (p) and temperature (T) over 30 years. Table 7.1 describes the reservoir, fluid, and well properties for the geothermal reservoir fluid flow models.

Table 7.1: Reservoir properties and description of the fluid flow models

Parameter	Value
Reservoir Properties	
Reservoir Extent (m)	1000×200
Initial Reservoir Pressure (kg/ms^2)	50×10^6
Left Boundary Pressure (kg/ms^2)	60×10^6
Right Boundary Pressure (kg/ms^2)	40×10^6
Initial Reservoir Temperature (K)	370
Porosity (%)	20
Rock Density (kg/m^3)	2650
Rock Heat Capacity (W/mK)	1000
Rock Thermal Conductivity (J/K)	2.9
Permeability** (m^2)	2×10^{-10} to 6×10^{-10}
Fluid Properties	
Fluid Density (kg/m^3)	1000
Viscosity (kg/ms)	0.001
Thermal Conductivity (W/mK)	0.15
Heat Capacity (J/K)	4185.5
Phase Volume Fraction (%)	20
Injector Fluid Temperature (K)	310
Well Properties	
Producer BHP (kg/ms^2)	40×10^6
Injector BHP (kg/ms^2)	60×10^6
Simulation Properties	
Timestep (seconds)	604 800
Simulation Period (yrs)	30

**The permeability field is populated by mapping the facies type at each point of the conditioned model to the corresponding value of permeability of the rock type, i.e. USF = 2.0×10^{-10}, LSF = 4.0×10^{-10} and OS = 6.0×10^{-10}

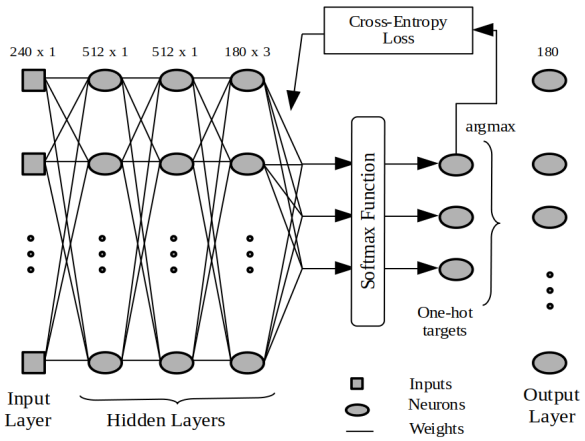


Figure 7.3: The neural network, SurfNet, $\mathcal{S}$ used in conditioning SBGMs to well data. The multi-layer perceptron maps a vector representation of the 240 SBGM input variables onto a discrete vector of the facies type with the maximum probability at all well locations, $\hat{\mathbf{y}}$, from Titus et al. (2022).

7.3.3 History Matching

We perform history matching by iteratively optimising the SBGM inputs in the well conditioning and data assimilation steps. The method is described in the following sections.

Well Conditioning

Geological models that are conditioned to static well data (information from well logs such as facies associations) using the technique presented in Titus et al. (2022) can be applied to aid history matching. Using the trained neural network, SurfNet3 from Titus et al. (2022), hereafter referred to as SurfNet, $\mathcal{S}$, we find input feature vectors $\mathbf{p}^*$ whose outputs from the network, $\tilde{\mathbf{y}}$, match the reference data or observations, $\mathbf{y}_{obs}$ by minimising the mismatch, $\mathcal{L}$, between the network prediction and the given observations given as

$$\mathcal{L}(\mathbf{z}_{i,obs}, \tilde{\mathbf{z}}_\mathbf{i}) := -\frac{1}{n_p} \sum_{i=1}^{n_p} \left(z^l_{i,obs} \log(\tilde{z}^l_i) \right), \tag{7.6}$$

where $\mathbf{z}_{i,obs}$ is the one-hot encoding of the observed facies type at each discretised location, n_p is the number of SBGM input parameters, l is such that $z^l_{i,obs} \neq 0$, the mismatch or loss $\mathcal{L}$ is the multiclass cross-entropy used in training $\mathcal{S}$, and $\tilde{\mathbf{z}}_\mathbf{i}$ is the predicted facies type by the trained network at that location. Details of the method are provided in Titus et al. (2022) and the architecture of $\mathcal{S}$ is shown in Figure 7.3.

The approach allows the generation of geological models that are consistent with the geological modelling process. Furthermore, the use of parametric, grid-free SBGMs that require few parameters to model makes the method applicable for a neural-network-based data assimilation due to the reduced computational expense of training the network.

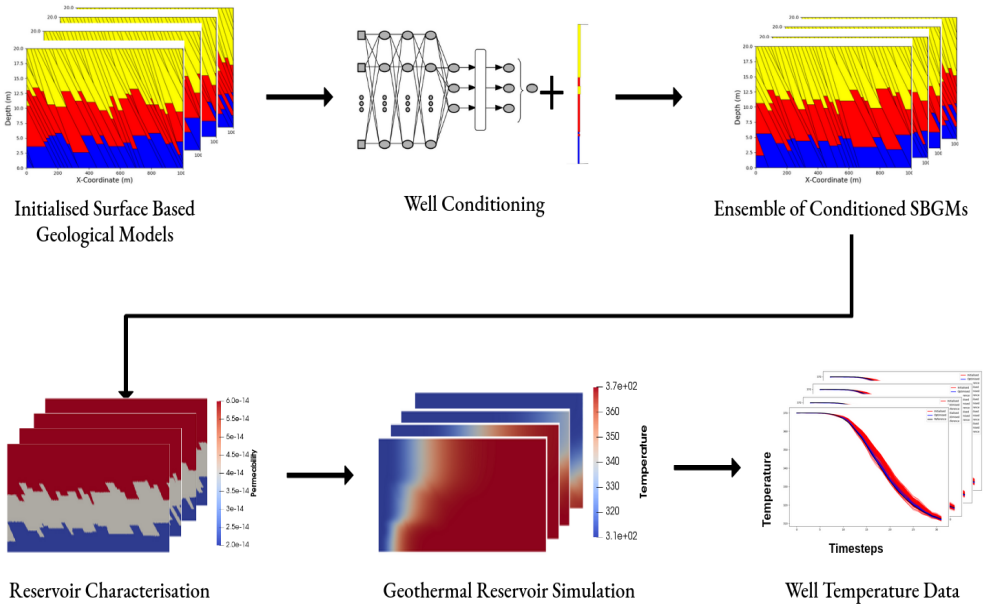


Figure 7.4: The process of generating the dataset for training and testing the neural network for the data assimilation. This involves conditioning a randomly generated set of SBGMs using the trained conditioning network and generating the reservoir simulation models of the conditioned models to obtain the production data at the well locations

Data Assimilation

Neural networks have been applied to approximate complex nonlinear relationships between input-output parameters of fluid flow in the subsurface (Amini and Mohaghegh, 2019). Data assimilation can be performed using the backpropagation technique to optimise neural networks as an alternative to adjoint models used in variational data assimilation (Buizza et al., 2022). We perform history matching with neural networks where the goal is to derive the inputs to the geological model (represented by a neural network) that account for the observed fluid flow behaviour. We apply adversarial learning with data assimilation to generate subsurface reservoir models that are calibrated to production data from a geothermal reservoir. The neural network relies on adversarial learning as a regularisation technique to capture the data distribution. Figure 7.4 illustrates the data generation process for the training examples.

Given the trained neural network, $\mathcal{G}$, (see Figure 7.5) that approximates the underlying relationship between geological inputs, $\mathbf{p}$ and the evolution of the system over a period of time $\mathbf{y}$, i.e, the production data, we perform history matching using neural networks for data assimilation. In this approach, the set of input feature vectors, $\mathbf{p}$, to the trained neural network, $\mathcal{G}$, which match the historical data, $\mathbf{y}_{obg}$ such that $\mathcal{G}(\mathbf{p}) = \mathbf{y}_{obg}$ is obtained. The problem is formulated as finding a $\mathbf{p}^*$ that minimises the mismatch, $\mathcal{L}$, between the network prediction and the given observations,

$$\mathbf{p}^* = \underset{\mathbf{p}}{\operatorname{argmin}}\, \mathcal{J}_g(\mathbf{p}), \tag{7.7}$$

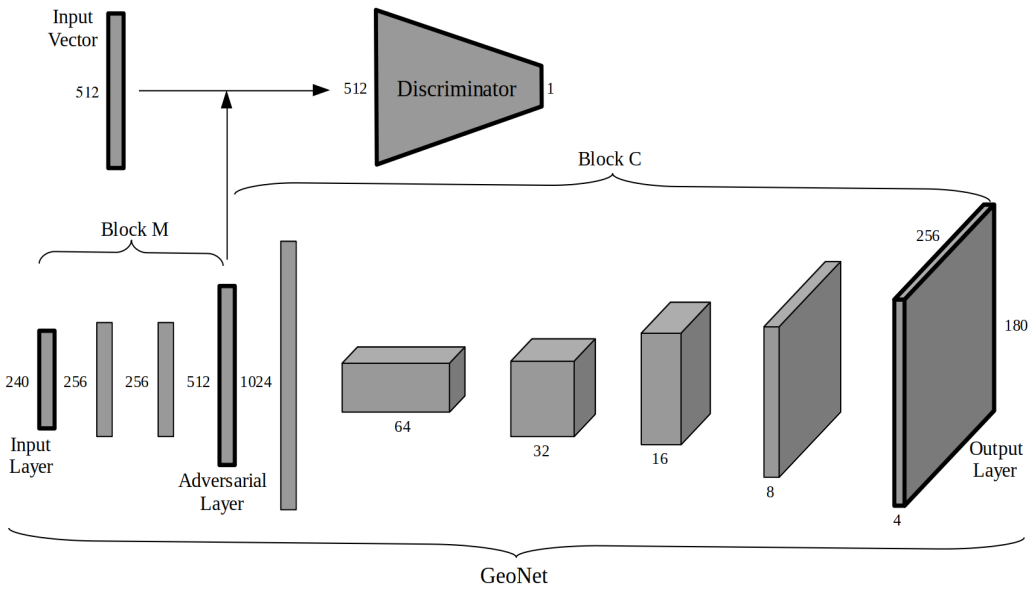


Figure 7.5: The architecture of the network showing GeoNet, the Discriminator and the adversarial layer.

where $\mathcal{J}_g$ is the objective function to be minimized defined as:

$$\mathcal{J}_g(\mathbf{p}) = \mathcal{L}(\mathbf{y}_{obg}, \mathcal{G}^*(\mathbf{p}^*) := \mathbf{E}_{\mathbf{p}\sim p(\mathbf{p})}[\| \mathbf{y}_{obg} - \mathcal{G}^*(\mathbf{p}) \|_2]. \tag{7.8}$$

Similar to the ensemble smoother with multiple data assimilation (ES-MDA) (Emerick and Reynolds, 2013), we perform the optimisation iteratively, where the proposed history matching method using neural networks for well conditioning and data assimilation is performed as follows:

1. Randomly generate $\mathcal{K}$ SBGM models with stochastic inputs, $\mathbf{p}$.

2. Given the trained network $\mathcal{S}$, apply Equation (7.6), to generate a set of SBGM inputs that match a given reference for each SBGM with initial input $\mathbf{p}$.

3. Given the trained network $\mathcal{G}$, apply Equation (7.8) to obtain the set of inputs $\hat{\mathbf{p}}$ that generate the static models which produce the historical production data.

4. Repeat steps two and three until the produced inputs generate SBGMs that match the well log and dynamic models that match the production data at the well locations.

The proposed method for conditioning SBGMs to geological data, and fluid flow models to production data using neural networks provides a robust approach to history matching where the predictions $\tilde{\mathbf{y}}$ and $\hat{\mathbf{y}}$ from $\mathcal{S}$ and $\mathcal{G}$, respectively, are compared with the observed data, $\mathbf{y}_{obs}$ and $\mathbf{y}_{obg}$, and the inputs, $\mathbf{p}^*$ are simultaneously updated at each iteration. The hyperpameters for training $\mathcal{S}$ are provided in Titus et al. (2022) and for $\mathcal{G}$ in Table 7.2.

The parameters for optimisation in the inverse modelling step are provided in Table 7.3. Algorithm 7.1 provides the history matching framework using $\mathcal{S}$ and $\mathcal{G}$.

Table 7.2: Hyperparameters used for training GeoNet and the performance metrics of the model

	Value
Hyperparameters*	
Learning rate	0.0005
Train-test split	0.90
Number of epochs	2000
Optimizer	NAdam
Weight decay	0.0
Betas	(0.8,0.999)
Annealing	0.95
Annealing step size	500
Discriminator iterations per generator	10
Weight initialisation	Kaiming
Number of epochs	3000
Metrics	
GeoNet Training Loss	0.00037
GeoNet Test Loss	0.00049
Generator Loss	6.1781
Discriminator Loss	0.0.721
Training time on GPUs (hrs)	5.05

*The hyperparameters of $\mathcal{D}$ are similar to $\mathcal{G}$

Table 7.3: Parameters for optimizing the SBGM inputs using the trained neural networks

	SurfNet	GeoNet
Learning rate	0.001	0.0005
Number of Cycles	1	1
Number of Iterations	10^4	10^4
Optimizer	Adam	NAdam
Well correlation cut-off	97.5	-
Well error tolerance	-	0.0007

7.3.4 Reduced Order Modelling

Reduced order models are useful for modelling systems with a large number of degrees of freedom. For example, in time-critical problems such as reservoir forecasting and monitoring, where solutions are required rapidly, generating the full-physics model can be computationally intensive. Furthermore, applications that require sampling a large parameter space, such as uncertainty quantification and sensitivity analysis, will greatly benefit from the speed-up in running multiple models. Figure 7.6 illustrates the process for generating snapshots of the temperature field to train and test reduced-order modelling neural networks.

Algorithm 7.1 Framework for history-matching subsurface reservoirs using machine learning

initialization: get random stochastic input parameters $\mathbf{p}^0 \in \mathbb{R}^{n_p}$
initialisation: $k = 1$, $m = 0$, $n = 0$
while $k \leq K$ **do** ▷ K = number of optimisation steps
▷ start well conditioning cycle
    $\mathbf{p}^k = \mathbf{p}^{k-1}$
    set the trainable parameters of the pretrained network $\mathcal{S}$ to $\mathbf{p}^k$
    read $\mathbf{y}_{obs}$ and encode output data at each discretized location as $\mathbf{z}_{obs,i}$
    scale $\mathbf{p}$ to a normal distribution $\tilde{\mathbf{p}} \sim \mathcal{N}(0, 1)$
    while $m \leq M$ **do** ▷ M = number of well conditioning cycles
        compute $\tilde{\mathbf{y}} = \mathcal{S}_{\boldsymbol{\theta}}(\tilde{\mathbf{p}}^k)$ ▷ $\tilde{\mathbf{y}} \in \mathbb{R}^{n_p}$
        calculate $\mathcal{L}_{\mathcal{S}}(\tilde{\mathbf{y}}, \mathbf{y}_{obs})$
        $\frac{\delta \mathcal{L}}{\delta \tilde{p}_j^k} \forall j \in \{1, \ldots, n_p\}$
        $\tilde{p}_j^{k+1} := \tilde{p}_j^k - \eta \frac{\hat{m}}{\sqrt{\hat{v}} + \epsilon}$ ▷ update inputs with the Adam optimizer
        constrain $\tilde{\mathbf{p}}^k$ $\tilde{\mathbf{p}}^k \sim \mathcal{U}([\boldsymbol{a}, \boldsymbol{b}])$, where a and b are the vectors of the upper and lower limits of the SBGM inputs.
    denormalise $\tilde{\mathbf{p}}^k$
    $m = 0$
▷ start data assimilation cycle
    set the trainable parameters of the pretrained network $\mathcal{G}$ to $\mathbf{p}^k$
    read $\mathbf{y}_{obg}$ and apply min-max scaling
    scale $\mathbf{p}^k$ to a normal distribution, $\hat{\mathbf{p}}^k \sim \mathcal{N}(0, 1)$
    while $n \leq N$ **do** ▷ N = number of data assimilation cycles
        compute $\hat{\mathbf{y}} = \mathcal{G}_{\boldsymbol{\theta}}(\hat{\mathbf{p}}^k)$ ▷ $\hat{\mathbf{y}} \in \mathbb{R}^{N_y \times N_t \times N_c}$
        calculate $\mathcal{L}_{\mathcal{G}}(\hat{\mathbf{y}}, \mathbf{y}_{obg})$
        $\frac{\delta \mathcal{L}}{\delta \hat{p}_j^k} \forall j \in \{1, \ldots, n_p\}$
        $\hat{p}_j^{k+1} := \hat{p}_j^k - \eta \frac{\hat{m}}{\sqrt{\hat{v}} + \epsilon}$ ▷ update inputs with NAdam optimizer
        constrain $\tilde{\mathbf{p}}^k$ $\tilde{\mathbf{p}}^k \sim \mathcal{U}([\boldsymbol{a}, \boldsymbol{b}])$, where a and b are the vectors of the upper and lower limits of the SBGM inputs.
    denormalise $\hat{\mathbf{y}}$
    $n = 0$, $k = k + 1$
return $\hat{\mathbf{p}}$

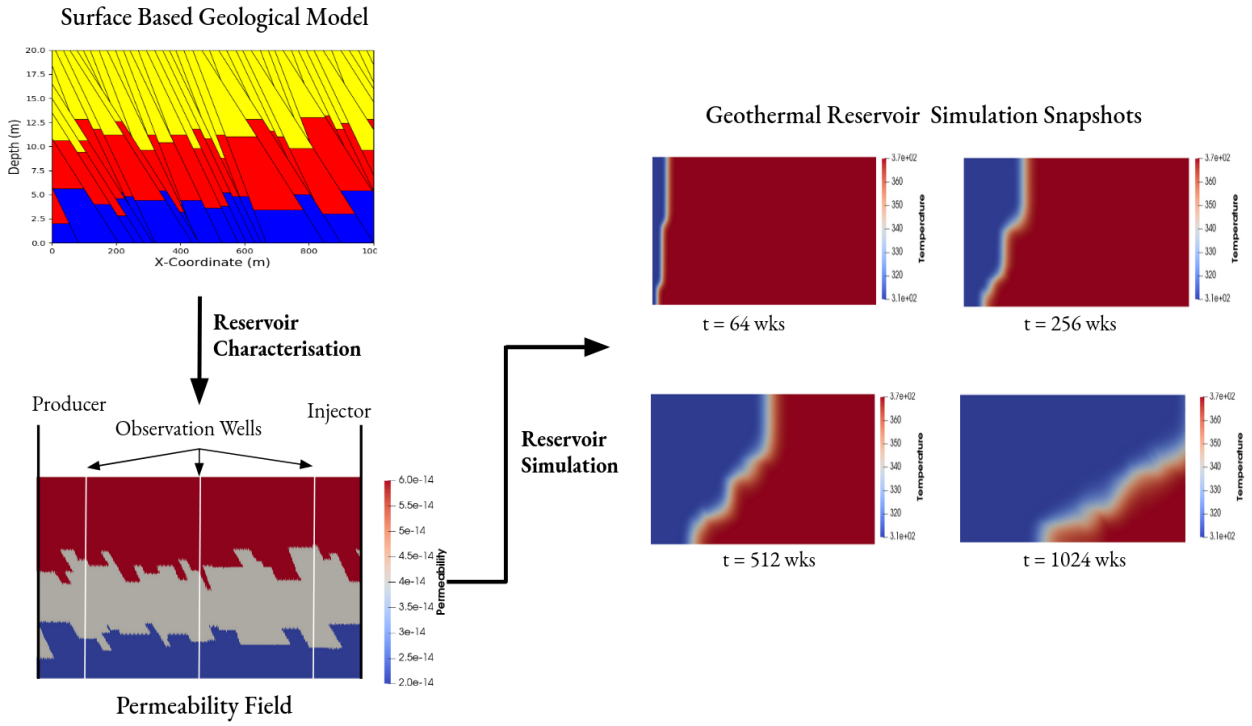


Figure 7.6: The process of generating the data for training and testing the compression and prediction neural networks involves generating SBGMs and mapping the permeability field to the facies distribution in the geological model. The resulting geological model is used to produce the reservoir simulation models from which the 2D snapshots are obtained.

Data Compression

Convolutional neural networks (CNNs) learn feature spaces by preserving the spatial coherence of data (Xu and Duraisamy, 2020). However, this limits the application of convolutional autoencoders (CAEs) for data compression and generation of ROMs to data on structured grids. We use CAEs with space-filling curves (SFC-CAEs) for unstructured meshes to develop reduced-order models that represent the spatio-temporal behaviour of geothermal reservoirs with realistic geological architectures in the following steps:

1. First, we apply SFCs for data transformation of the feature-scaled temperature data: $\tilde{\mathbf{T}}^s \leftarrow \mathcal{F}_S(\mathbf{T}^s)$ where $\mathcal{F}_S : \mathbb{R}^{26990} \rightarrow \mathbb{R}^{26990}$. This is necessary to handle the unstructured grids required to generate realistic geological reservoir models. The SFCs allow a mapping from a high-dimensional space to a one-dimensional space, where deep learning networks that rely on exploring structural/spatial correlations such as CNNs can be applied for dimensionality reduction.

2. Using the encoder part, $\mathcal{E}$ of the 1D convolutional autoencoder, we compress the feature-scaled data orders of magnitude lower and obtain the latent variables: $\hat{\mathbf{e}} \leftarrow \mathcal{F}_{\mathcal{E}}(\tilde{\mathbf{T}}^s)$ which represent the low-order model of the full physics simulation, where $\mathcal{F}_{\mathcal{E}} : \mathbb{R}^{26990} \rightarrow \mathbb{R}^{128}$.

3. To generate reduced order models, we use the latent representation with a transformer-based network, $\mathcal{T}$ to simulate the reservoir in reduced space: $\hat{\mathbf{e}}^t \leftarrow \mathcal{F}_{\mathcal{T}}(\hat{\mathbf{e}}^{t-1})$.

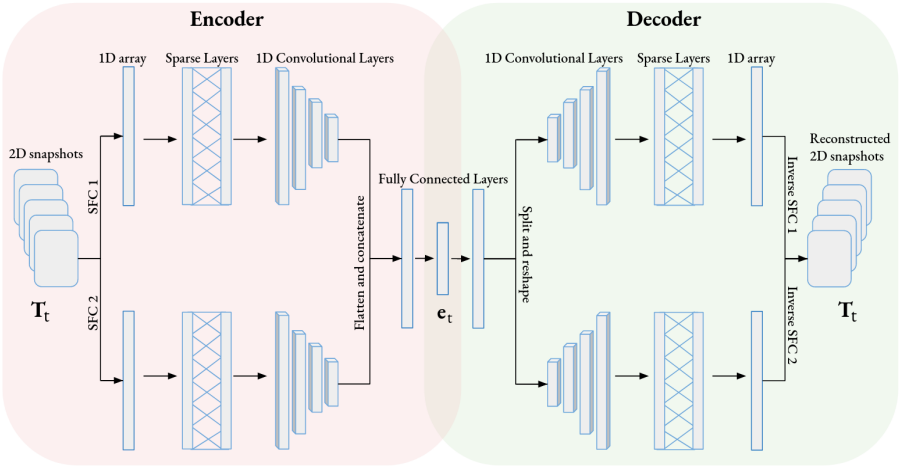


Figure 7.7: The space-filling curve convolutional autoencoder (SFC-CAE) applied for compressing the high-dimensional geothermal reservoir simulation snapshots, $\mathbf{T}$ into the low-dimensional latent space, $\mathbf{e}$. The SFC-CAE uses two SFCs to transform 2D snapshots into 1D arrays that feed into convolutional and fully connected layers in the encoder network, $\mathcal{E}$. The decoder, $\mathcal{D}$ is a mirror image of $\mathcal{E}$, with the inverse SFC transformation performed to transform the 1D array from $\mathcal{D}$ into the final reconstructed 2D snapshots

4. The reconstructed simulation is then obtained by decompressing the predicted latent variables using the decoder block, $\mathcal{D}$ of the convolutional autoencoder: $\hat{\mathbf{T}}^s \leftarrow \mathcal{F}_{\mathcal{D}}(\hat{\mathbf{e}})$ where $\mathcal{F}_{\mathcal{D}} : \mathbb{R}^{128} \rightarrow \mathbb{R}^{26990}$ and the inverse SFC can be use to obtain the ordered reconstructed solution: $\hat{\mathbf{T}} \leftarrow \mathcal{F'}_s(\hat{\mathbf{T}}^s)$.

The method of developing the SFC-CAE used in this study is described in detail in Heaney et al. (2024) and based on modifications for subsurface modelling applications to the implementation of Yu (2021). Figure 7.7 illustrates the SFC-CAE applied in this work to compress geothermal reservoir simulation data and Table 7.4 provides the hyperparameters of the network.

The SFC-CAE is trained in supervised learning mode to find the parameters, $\boldsymbol{\theta}_{\mathcal{E}}$ and $\boldsymbol{\theta}_{\mathcal{D}}$ of the encoder and decoder, respectively, that minimise the data mismatch, also referred to as reconstruction error, over all training examples. This is given as

$$\boldsymbol{\theta}_{\mathcal{E}}, \boldsymbol{\theta}_{\mathcal{D}} = \underset{\boldsymbol{\theta}^*_{\mathcal{E}}, \boldsymbol{\theta}^*_{\mathcal{D}}}{\operatorname{argmin}} [\mathcal{L}_{\mathcal{D}}(\hat{\mathbf{T}}^s, \tilde{\mathbf{T}}^s)], \tag{7.9}$$

where $\hat{\mathbf{T}}^s = \mathcal{F}_{\mathcal{D}}(\hat{\mathbf{e}}; \boldsymbol{\theta}_{\mathcal{D}})$ is the reconstructed output from the decoder given that $\hat{\mathbf{e}} = \mathcal{F}_{\mathcal{E}}(\tilde{\mathbf{T}}^s; \boldsymbol{\theta}_{\mathcal{E}})$ is the latent representation of the feature scaled input $\tilde{\mathbf{T}}^s$ with SFC transformation applied and $\mathcal{L}_{\mathcal{D}}$ is the loss function which measures the discrepancy between $\hat{\mathbf{T}}^s$ and $\tilde{\mathbf{T}}^s$:

$$\mathcal{L}_{\mathcal{D}}(\hat{\mathbf{T}}^s, \tilde{\mathbf{T}}^s) = \|\hat{\mathbf{T}}^s - \tilde{\mathbf{T}}^s\|_2^2. \tag{7.10}$$

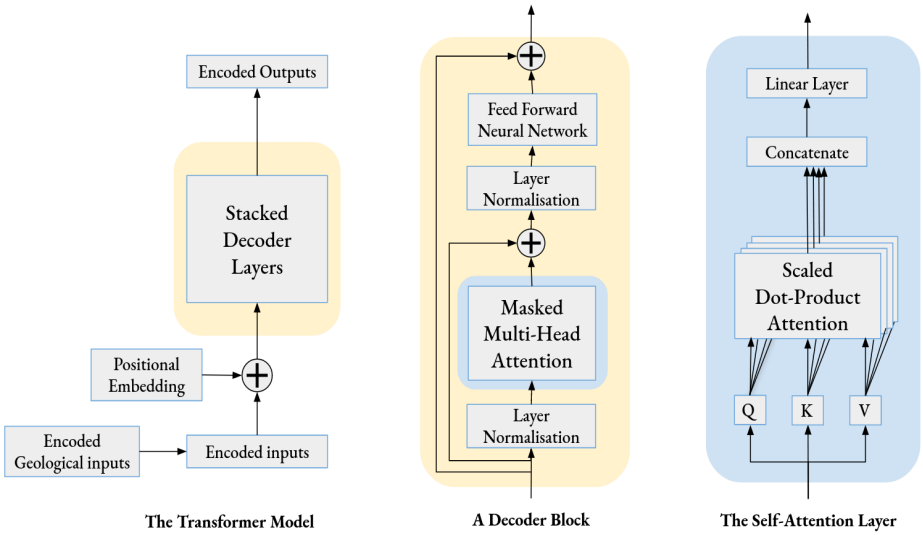


Figure 7.8: The neural network applied in this study based on the Generative Pre-trained Transformer (GPT) architecture (depicted on the left). Its network structure is composed of stacked self-attention and fully connected feed-forward layers within the decoder (shown in the middle). Positional encodings are introduced and added to the input embeddings, which are essentially low-dimensional representations of the state variables obtained from the encoder. The self-attention layers (illustrated on the right) enable each position within the decoder to attend to all other positions, including itself. To maintain the auto-regressive property, a masked multihead attention mechanism is used, which restricts the flow of information in a leftward direction within the decoder (adapted from Vaswani et al. (2017)).

Temporal Evolution of the Latent State

The state evolution of the low-dimensional dynamics of the geothermal reservoir, $\hat{\mathbf{e}}$ is represented by using a transformer-based neural network, $\mathcal{T}$ (Vaswani et al., 2017) as a predictive network to simulate the low-order geothermal reservoir. We use the Generative Pre-trained Transformer (GPT) (Radford and Narasimhan, 2018; Radford et al., 2019) based on the implementation of Wolf et al. (2020). The GPT model is a transformer decoder model that consists of a stack of transformer decoder layers that use the masked attention mechanism. The input to our transformer-based predictive model, $\mathcal{T}$ is the set of low-dimensional vectors, $\mathbf{E} = \{\tilde{\mathbf{e}}_0, \tilde{\mathbf{e}}_1, ..., \tilde{\mathbf{e}}_{N_t-1}, \tilde{\mathbf{e}}_{N_t}\}$ from $\mathcal{E}$ in the SFC-CAE that represent the encoded geothermal reservoir simulation state variables at each time step. The transformer-based predicted network, $\mathcal{T}$ is trained using the auto-regressive approach of Radford and Narasimhan (2018), where the set of current and previous time steps, $\mathbf{E} = \{\tilde{\mathbf{e}}_0, \tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_{N_t-1}, \tilde{\mathbf{e}}_{N_t}\}$ is used to predict the next time step,$\mathbf{e}_{t+1}$. The loss, $\mathcal{L}_{\mathcal{T}}$ is the standard mean squared error:

$$\mathcal{L}_{\mathcal{T}} := \| \mathcal{T}(\hat{\mathbf{e}}_t) - \hat{\mathbf{e}}_{t+1} \|_2 \,. \tag{7.11}$$

The hyperparameters and modifications to the GPT2 implementation of Wolf et al. (2020) used in this work are given in Table 7.4. The procedure for generating ROMS using SFC-CAE for compression and GPT for temporal evolution is provided in algorithm 7.2.

Table 7.4: Hyperparameters and performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$).

	SFC-CAE	**GPT**
Hyperparameters		
Learning rate	5e-5	5e-5
Train-Val-Test split	0.90-0.03-0.07	0.70-0.29-0.01
Number of examples	105	512
Number of epochs	400	380
Optimizer	NAdam	AdamW
Weight decay	0.0	-
Betas	(0.8,0.999)	-
Warm-up steps	-	100
LayerNorm epsilon	-	10^{-8}
Weight initialisation	Kaiming (He et al., 2015)	Kaiming (He et al., 2015)

Algorithm 7.2 Generating ROM with SFC-CAE and Transformer-based model

set number of prediction steps, N_t
read initial conditions $\mathbf{T}_0$
load and freeze pretrained $\mathcal{E}, \mathcal{D}$ and $\mathcal{T}$
apply SFC transformation: $\mathbf{T}_0^s \leftarrow \mathcal{F}_S(\mathbf{T}_0)$ ▷ $\mathbf{T}_0^s \in \mathbb{R}^N$
apply feature scaling: $\tilde{\mathbf{T}}_0^s \in [-1, 1]$

encoder forward pass: $\hat{\mathbf{e}}^0 \leftarrow \mathcal{F}_{\mathcal{E}}(\tilde{\mathbf{T}}_0^s)$ ▷ $\hat{\mathbf{e}}^0 \in \mathbb{R}^{N_e}$

for $t \in 1, 2, ..., N_{t-1}$ **do**
  $\hat{\mathbf{e}}^t \leftarrow \mathcal{F}_{\mathcal{T}}(\hat{\mathbf{e}}^{t-1})$
get $\hat{\mathbf{E}} = \{\hat{\mathbf{e}}^1, \hat{\mathbf{e}}^2, ..., \hat{\mathbf{e}}^{N_t-1}, \hat{\mathbf{e}}_t^N\}$

decoder forward pass: $\hat{\mathbf{X}}^s \leftarrow \mathcal{F}_{\mathcal{D}}(\hat{\mathbf{E}})$ ▷ $\hat{\mathbf{X}}^s = \{\hat{\mathbf{T}}_1^s, \hat{\mathbf{T}}_2^s, ..., \hat{\mathbf{T}}_{N_t}^s \hat{\mathbf{T}}_{N_t}^s\} \in \mathbb{R}^N$

apply inverse SFC transformation: $\hat{\mathbf{X}} \leftarrow \mathcal{F'}_s(\hat{\mathbf{X}}^s)$

return $\hat{\mathbf{X}} = \{\hat{\mathbf{T}}^1, \hat{\mathbf{T}}_2, ..., \hat{\mathbf{T}}_{N_t-1}, \hat{\mathbf{T}}_{N_t}\}$

7.4 Application to Geothermal Reservoirs

The results of the history matching and reduced order modelling using neural networks for optimisation and modelling state evolution are presented in the following sections.

7.4.1 History Matching with Well Conditioning and Data Assimilation

We train the adversarial network, $\mathcal{G}$ with geothermal reservoir simulations generated as described in Section 7.3.2. The network maps the SBGM inputs, $\mathbf{p}$, to the production data, $\mathbf{y}$, consisting of the pressure, velocity (in the x and y directions), and the temperature for 60 discretised locations along three wells: $\mathbf{p} \in \mathbb{R}^{n_p} \rightarrow \hat{\mathbf{y}} \in \mathbb{R}^{N_y \times N_t \times N_c}$, where N_c is 4 (temperature, pressure and velocity in the x and y directions), N_t is 256 weeks, and N_y is 180, corresponding to 60 grid points from each of the three exploratory wells. The vertical resolution along the well is 0.33m, which corresponds to the resolution of leading industry gamma well logging

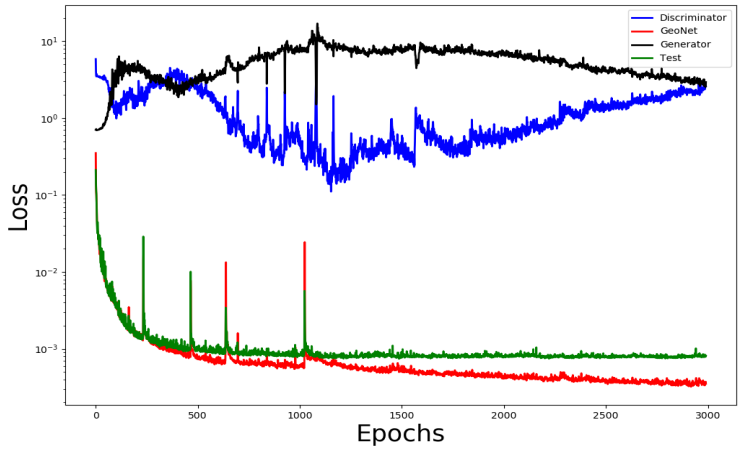


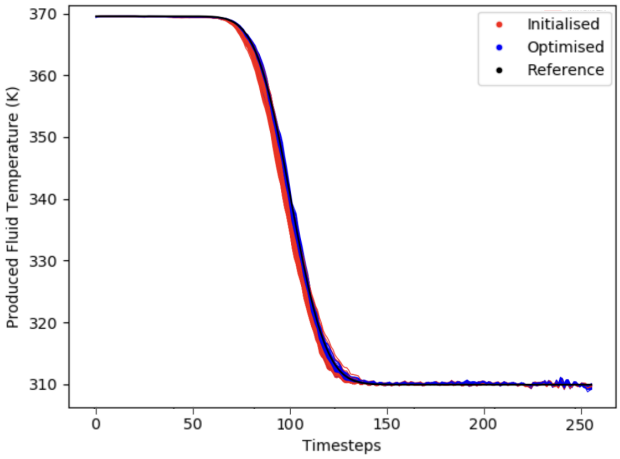
Figure 7.9: The loss plot of the adversarial network, GeoNet, $\mathcal{G}$ as well as the discriminator, $\mathcal{D}$.

tools (0.3048m ±5%) (Schlumberger, 2015). The results show that $\mathcal{G}$ is able to learn the relationship between the production data and the SBGM inputs to a reasonable degree of accuracy as the training loss in Figure 7.9 shows a reduction of three orders of magnitude over 3000 epochs.

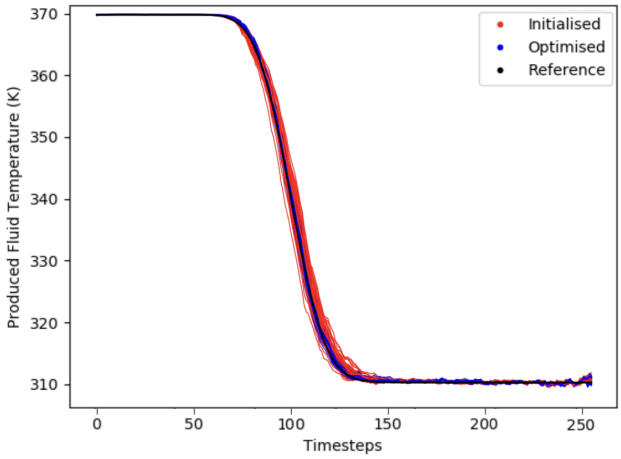
Using the trained models and the method presented in Section 7.1, we invert for the set of SBGM inputs, $\mathbf{p}^*$ that generate static and dynamic models that match the given reference geological, $\mathbf{y}_{obs}$ and production data, $\mathbf{y}_{obg}$ that $\mathcal{S}$ and $\mathcal{G}$ have not seen during training or testing. The reference model is shown in Figure 7.2. Applying algorithm 7.1, we optimised 100 SBGMs to match the given reference well data with 10000 iterations and one cycle for the well conditioning and data assimilation phases. Table 7.3 shows the hyperparameters for $\mathcal{S}$ and $\mathcal{G}$ used to optimise the initialised SBGM inputs. Figures 7.10a to 7.10f show examples of the predicted production data using the trained $\mathcal{G}$ for the optimised, initialised, and reference models. The results show a closer match to the reference model for the optimised model compared to the initialised models, which have a higher variation from the reference model and, by extension, greater uncertainties.

To validate the inversion results, we generate the geological models using the optimised SBGMs from the history matching process and apply these to run the geothermal reservoir simulation (as described in Section 7.3.2 to compare with the reference model at the well locations. Figure 7.11 shows the results for randomly selected timesteps of a model with a mean absolute error of 0.004 for the optimised model and 0.177 for the initialised model at the well locations.

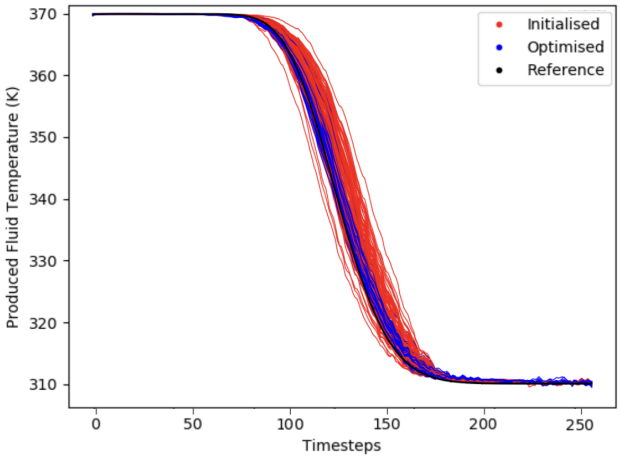
We apply the frequency of predictions (FOP) and maximum predictions (described in Titus et al. (2022)) to evaluate spatial variations and uncertainties associated with the optimised geological models as shown in Figure 7.12. The FOP for the initialised models shows high uncertainty, with significant lateral variations observed across the domain, even at well locations. However, for the history-matched models, the FOP has lower uncertainties and spatial variations, particularly at well locations. Figure 7.13 illustrates the maximum



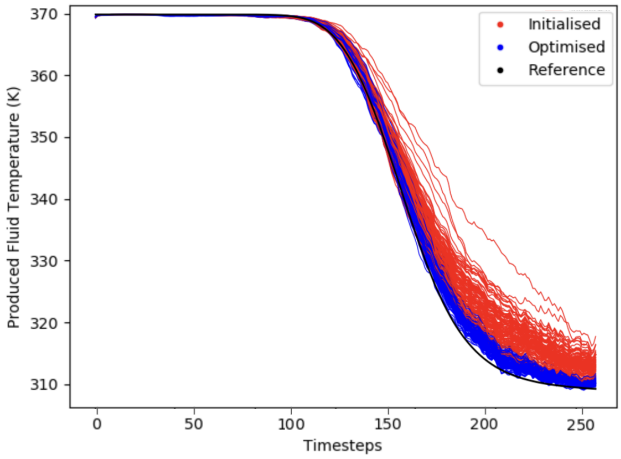
(a) Point 9 in Well1



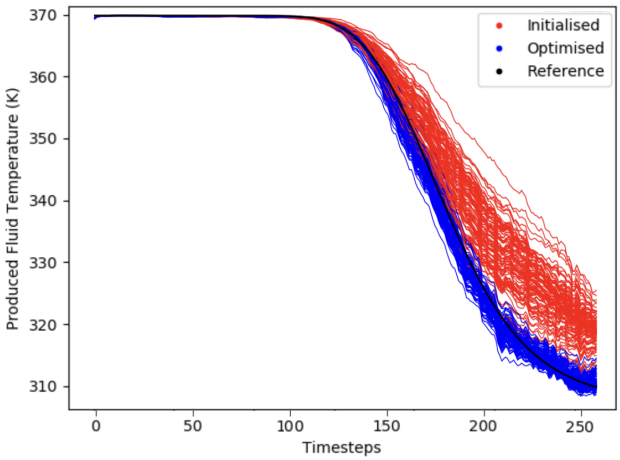
(b) Point 21 in Well1



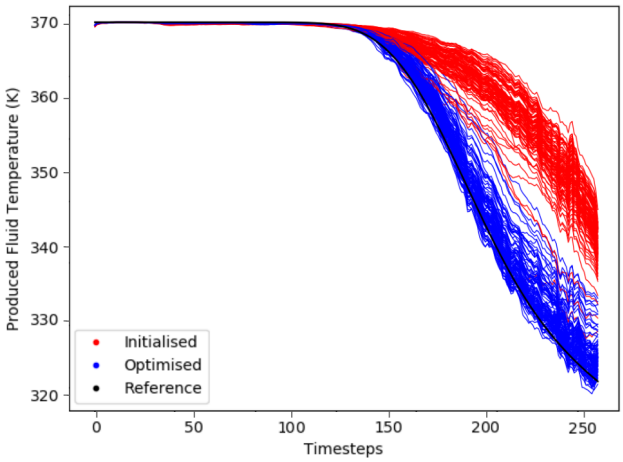
(c) Point 31 in Well1



(d) Point 45 in Well1

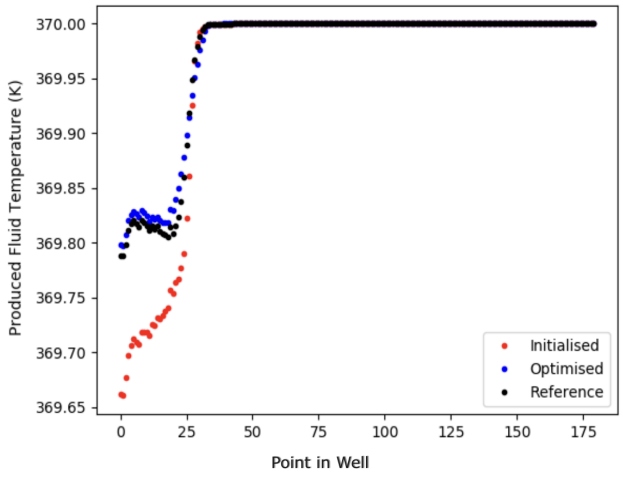


(e) Point 47 in Well1

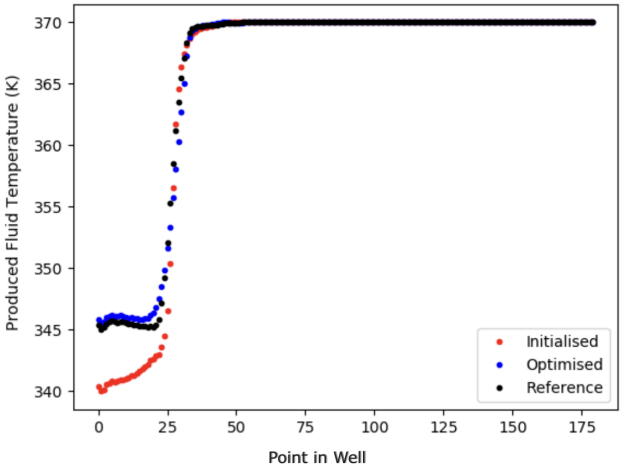


(f) Point 51 in Well1

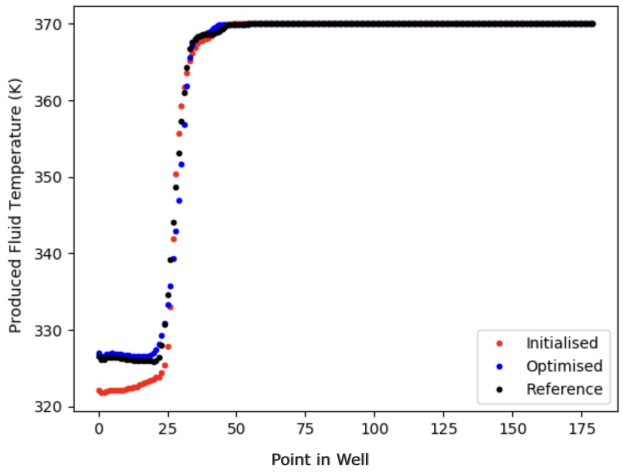
Figure 7.10: Neural network predictions with the initialised, optimised and reference SBGM inputs for points in Well1 as a measure of uncertainties among models.



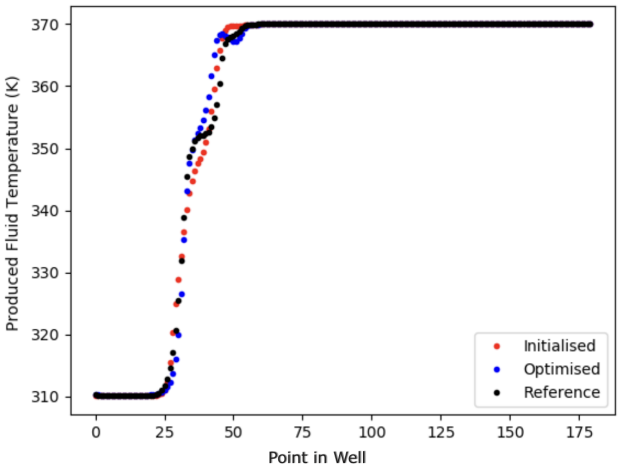
(a) Timestep 63



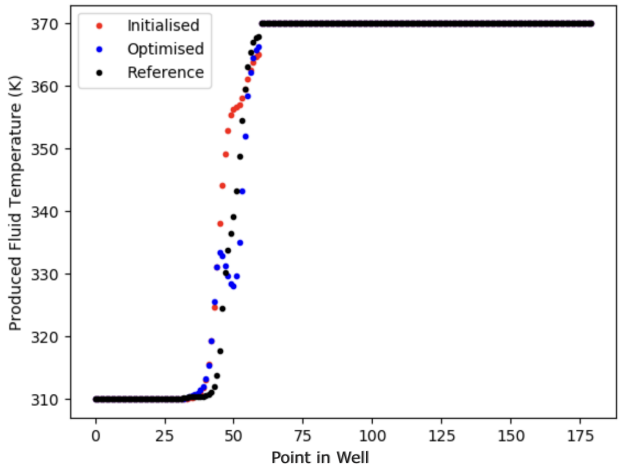
(b) Timestep 97



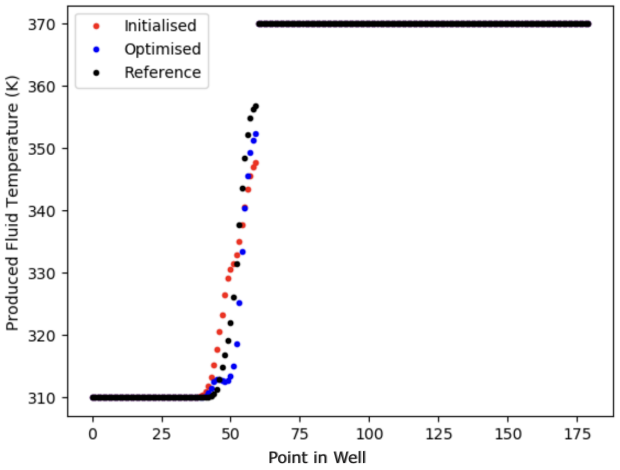
(c) Timestep 109



(d) Timestep 139



(e) Timestep 205



(f) Timestep 255

Figure 7.11: Temperature data at well1 (points 1 to 60), well2 (points 61 to 120) and well3 (points 121 to 180) from the reservoir simulation with the initialised, reference and optimised models. The results show a good match for the temperature at the well locations in the forward models of the history-matched (optimised) models and the reference model, compared to the initialised models.

predicted facies at each grid point from the 100 conditioned models. When comparing Figure 7.13b with the reference model (Figure 7.2), similar trends can be observed particularly at well locations, in contrast to Figure 7.13a. This shows that the set of optimised models is capable of preserving geological features, especially in the vicinity of the wells. The average correlation with the facies type of the reference at the well locations is also presented in Figure 7.14 for the optimised and initialised models.

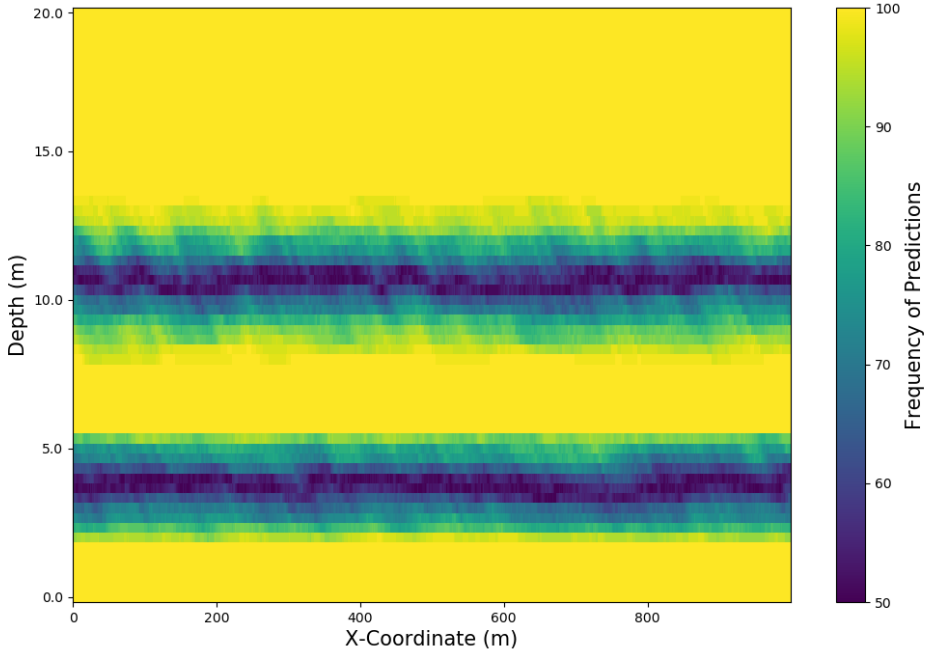
7.4.2 Reservoir Forecasting with the History-Matched Models

The SFC-CAE presented in Section 7.3.4 was trained and applied to simulate and reconstruct the original high-fidelity geothermal reservoir simulation snapshots. Figure 7.15a displays the training loss over 400 epochs. The trained model encodes the temperature field into a low-dimensional embedding, using the encoder $\mathcal{E}$, and reconstructs the high-dimensional temperature field using the decoder and an encoding dimension, $N_e = 128$, which is large enough to retain critical flow characteristics while still computationally efficiency for rapidly generating reduced order models.

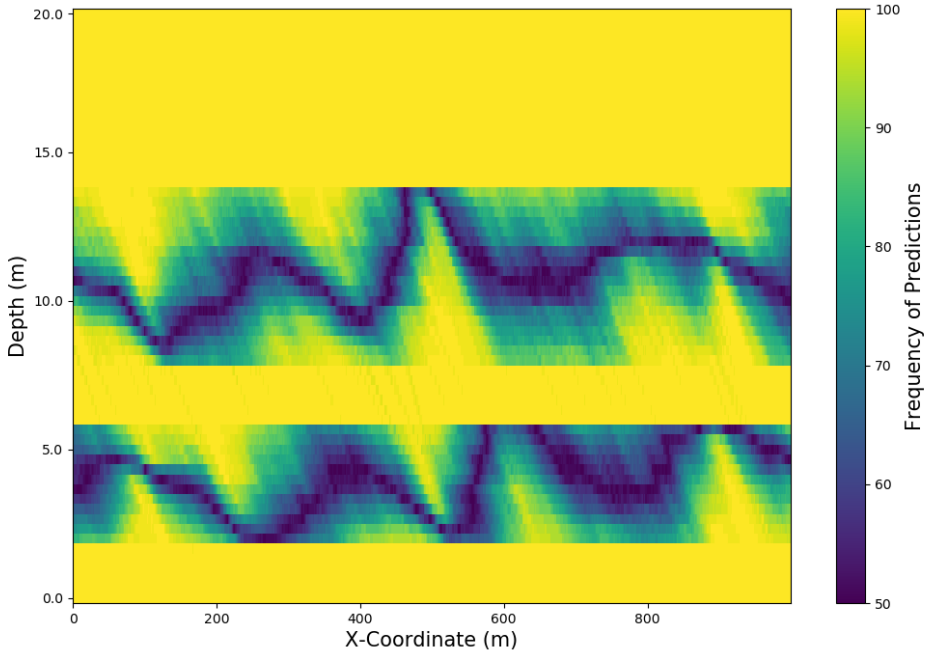
The transformer-based predictive model, $\mathcal{T}$ illustrated in Figure 7.8 was trained using the hyperparameters given in Table 7.4 with each fluid flow simulations having 1024 time steps for the training, validation, and test datasets. The network architecture comprises 16 decoder layers, each containing one masked attention head. A sequence length of 256 was used, corresponding to the number of timesteps of the production data used in the history matching step. Figure 7.16 shows the forecast for the history-matched model from Section 7.4.1, along with the predictions for the corresponding initialised models. Here, the trained $\mathcal{T}$ is given the temperature information from the history-matched models for a context length of 256 time steps and is allowed to predict to $t = 1024$. The root mean squared error (RMSE) of the predictions in Figure 7.17 shows the difference between the reference and temperature forecast for the optimised and initialised models. The GPT2 model is capable of predicting the temperature distribution for each time step in 0.01 s given the preceding timesteps of length equal to the model context or the sequence length.

Table 7.5: Performance metrics for the SFC-CAE ($\mathcal{E}$ and $\mathcal{D}$), and GPT model ($\mathcal{T}$).

	SFC-CAE	**GPT**
Metrics		
Training Loss	0.0011	0.0014
Test Loss	0.0011	0.0012
Training time on GPU (hrs)	5.81	22.29

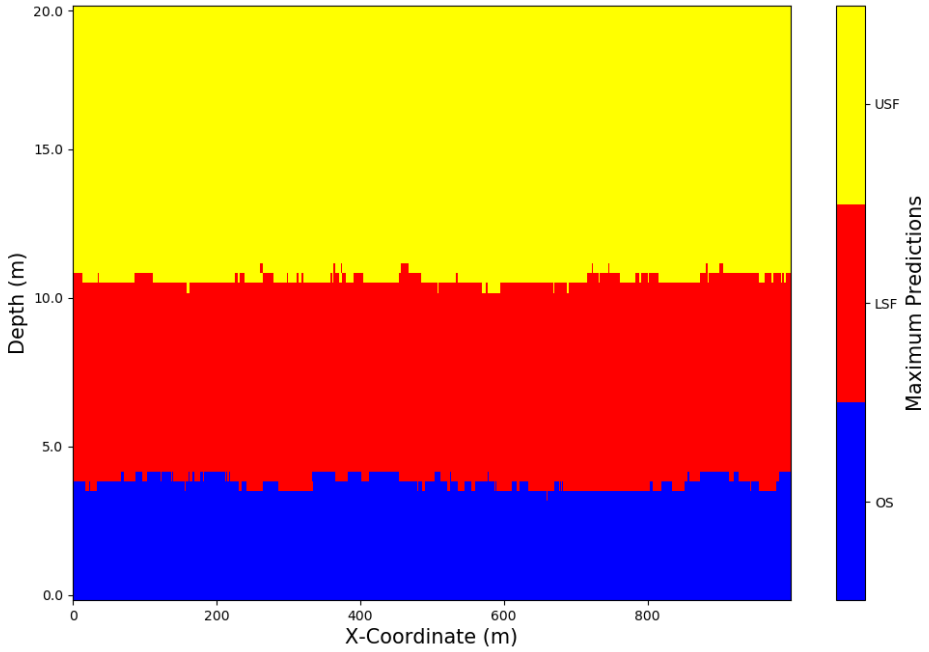


(a) Initialised models

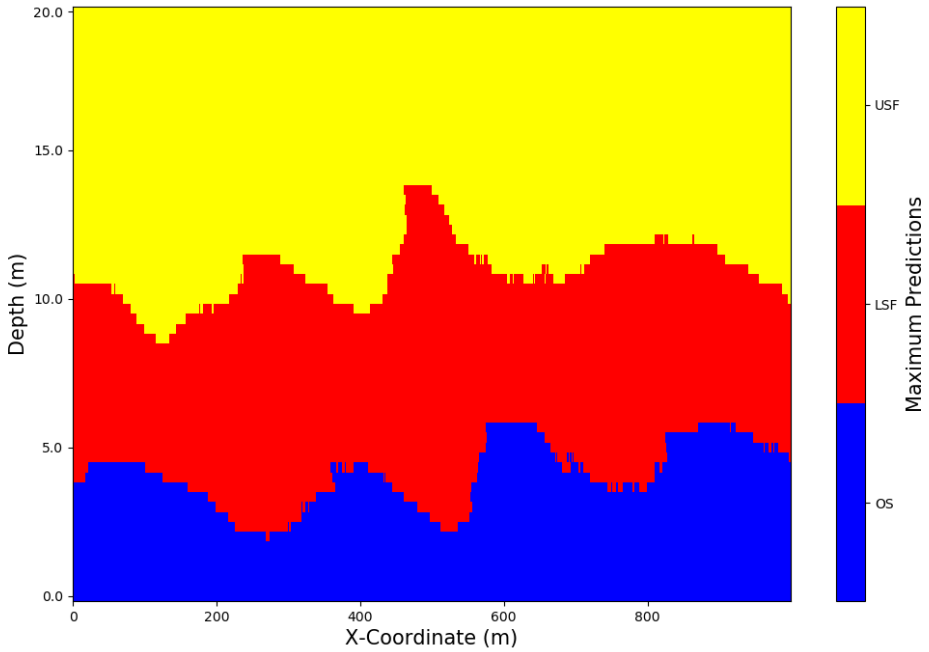


(b) Optimised models

Figure 7.12: Frequency of predictions for the initialised (top) and optimised (bottom) models for 100 conditioned models.



(a) Initialised models



(b) Optimised models

Figure 7.13: Maximum predictions for the 100 initialised (top) and optimised (bottom) models.

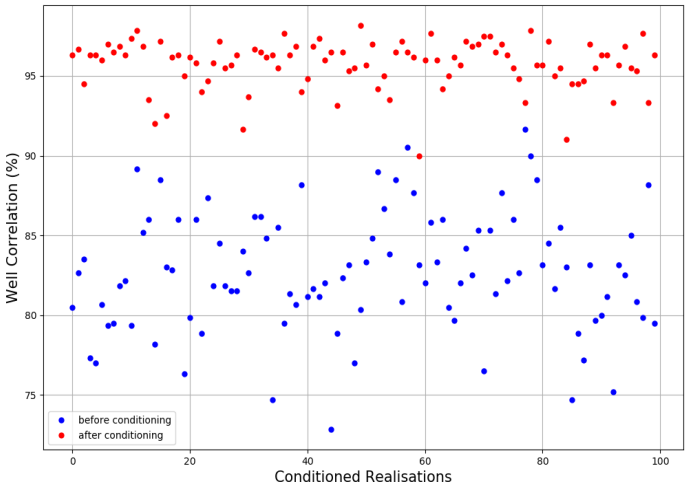
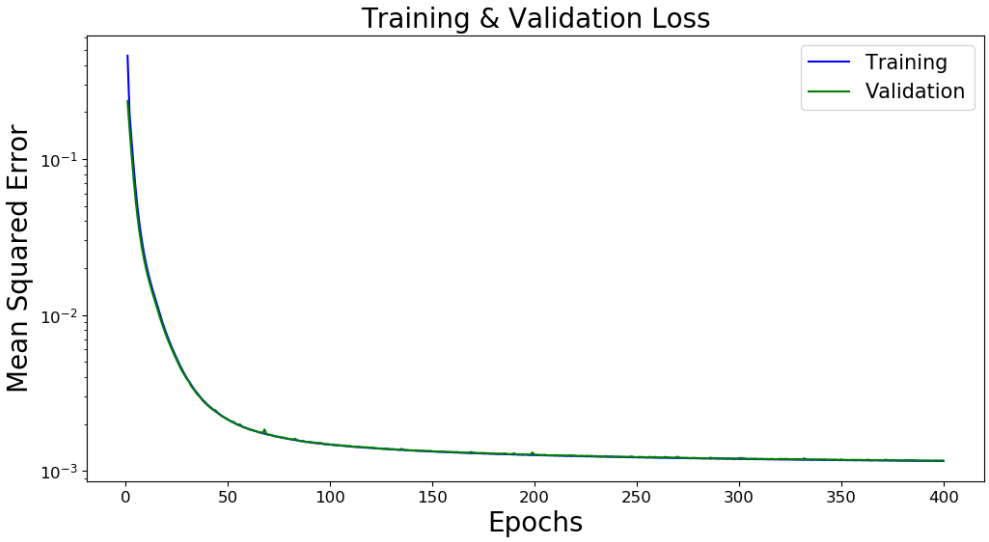


Figure 7.14: Correlation of the facies types at the wells with the reference well facies type for the initialised and optimised models.

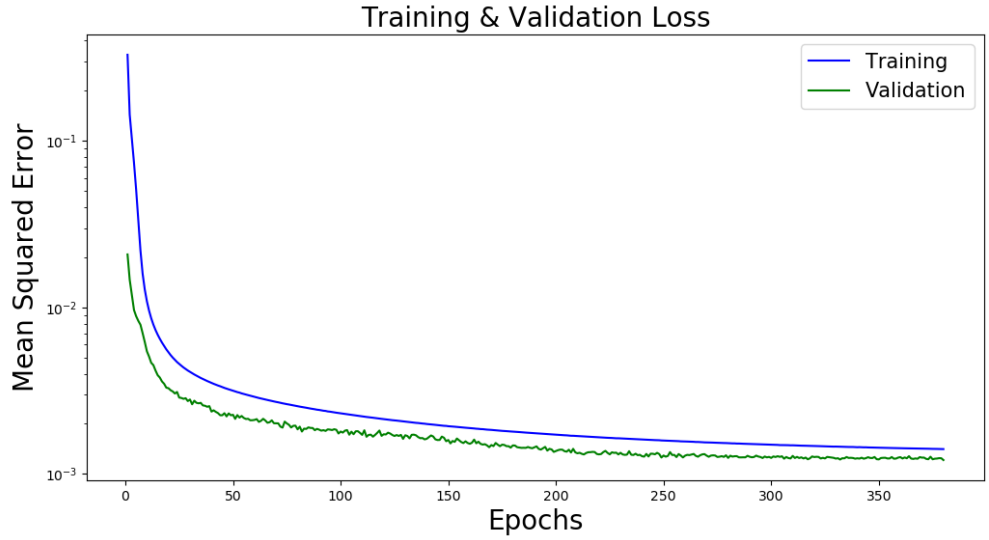
7.5 Discussions

The results presented demonstrate a new approach for rapidly generating representative reservoir models that honour geological and production data from wells. The key benefits of the proposed method is the computational efficiency and applicability, and the potential for real-time reservoir management. The developed method offers significant advantage in terms of computational speed, with history matching taking only about 200 seconds and forecasting about 0.01 seconds for each time step for the test cases considered here. This efficiency is further enhanced by the potential for massive parallelisation using the trained networks for well conditioning and data assimilation to generate multiple models in the history matching workflow. The developed reduced-order modelling approach could significantly accelerate tasks such as well optimisation and sensitivity analyses in field development planning, where multiple simulations are required to assess and quantify risks and uncertainties. The presented method does not compromise resolution or geological realism by eliminating the need for a pre-defined grid in the geological modelling step. The study also introduces the use of space-filling curves with unstructured meshes for data compression in subsurface reservoirs. The space-filling curves preserve structural coherence and spatial correlation of the compressed data, allowing for the use of unstructured meshes in reservoir simulation to better capture geological features and architectures.

However, more research is required to improve the accuracy of the reduced order modelling, where the error in the workflow is seen to be the highest. The use of small time windows where sequential data assimilation and reservoir forecasting are performed over short periods of time, similar to the filtering approach in ensemble-based filtering in data assimilation (Jung et al., 2018) might improve forecasting accuracy. Developing the transformer-based model presented here using a physics-informed neural network approach could also significantly improve



(a) The loss curve showing the mean squared error over 400 training epochs of the SFC-CAE.



(b) The loss curve showing the mean squared error over 380 training epochs of the transformer-based model.

Figure 7.15: The loss curves of the SFC-CAE (top) and GPT model (bottom) showing the mean squared error over 400 and 380 training epochs respectively. The loss reduces by about three orders of magnitude in the SFC-CAE. A very close match is seen between the training and validation loss in both models.

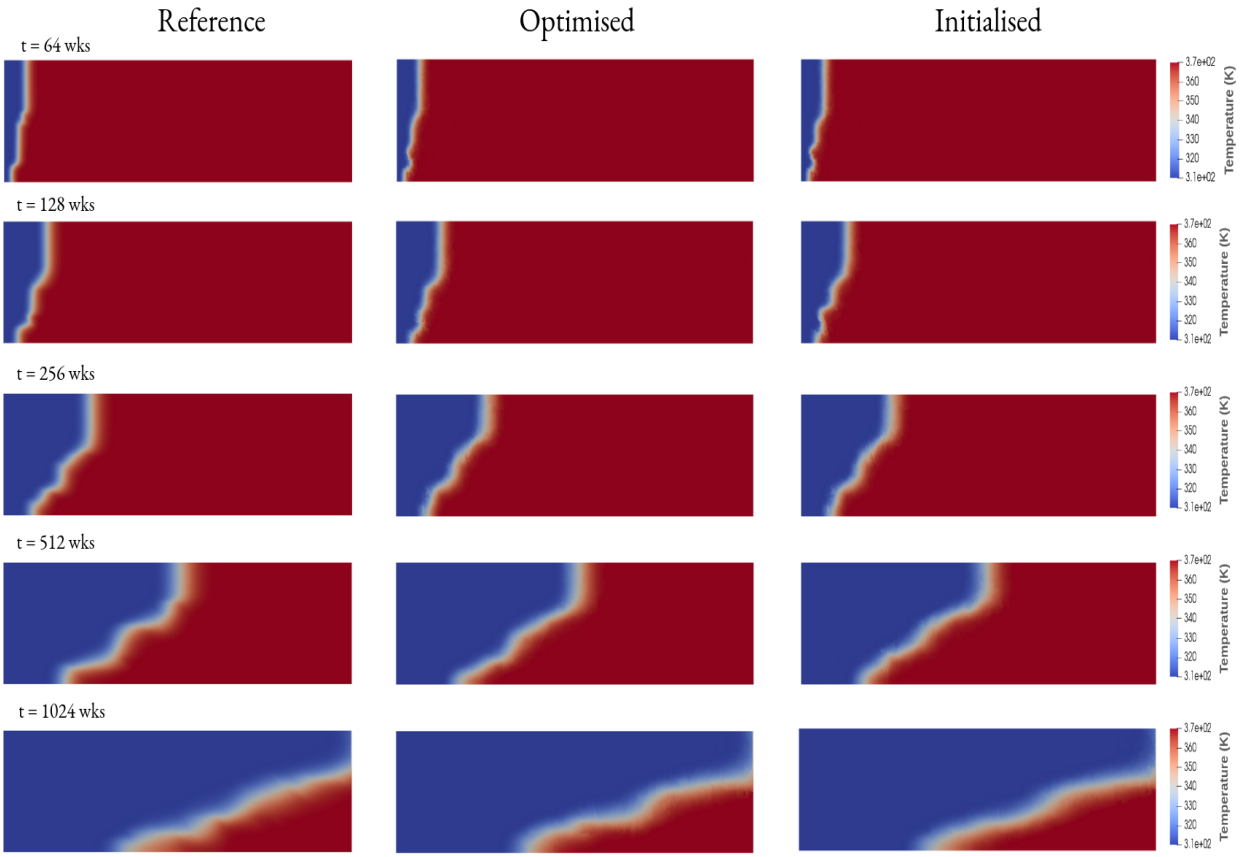


Figure 7.16: The results of the SFC-CAE for one example not used in training the model at five different time steps. The original high-fidelity model is shown on the left, the reconstructed geothermal simulation from forecasting the history-matched (optimised) models is given in the middle and the reconstructed forecast for the initialised models are presented on the right.

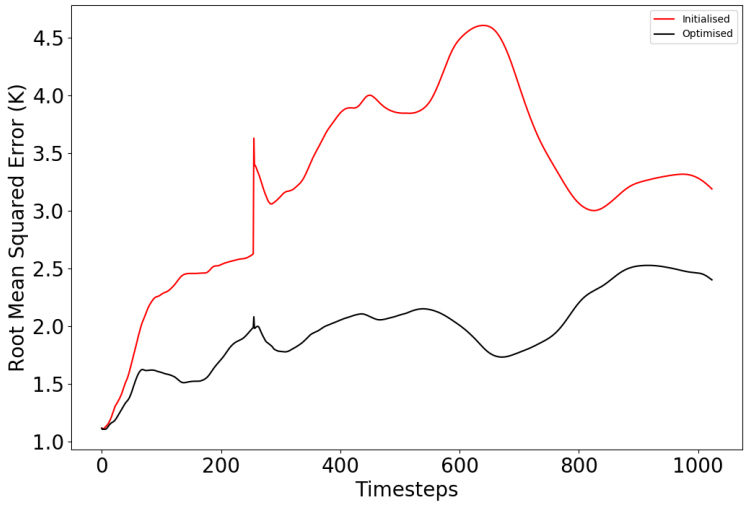


Figure 7.17: The mean squared error for each time step in the reconstructed simulation for the history-matched (optimised) and initialised model.

the results. Furthermore, the method can benefit from a more rigourous and robust uncertainty quantification method in a probabilistic context to better understand risks and support decision-making.

Although our proof of concept focused on simplified models for geothermal reservoirs, the use of surface-based geological models to preserve geological realism, space-filling curves for unstructured meshes, neural networks for history matching, and transformer-based models to model reservoir dynamics pave the way for the application of the proposed methodology for field cases. Data-driven methods can be beneficial for more accurate and practical reservoir simulations in cases where physics-based models are computationally expensive or difficult to formulate.

7.6 Conclusions

Although geothermal resources in settings similar to hydrocarbon reservoirs have the potential to contribute to clean energy resources, generating representative geothermal reservoir models that honour observed data and can aid optimisation, forecasting, and the understanding of risks and uncertainties remains an area of active research. This study proposes a neural network-based approach for history matching that integrates well conditioning and data assimilation in an iterative workflow. Using this approach, reservoir models that are consistent with geological processes, geophysical information, and fluid flow data from wells can be rapidly generated. Convolutional autoencoders with space-filling curves and transformer-based networks allow for the development of reduced-order models (ROMs) of geothermal systems that represent the high-order model with sufficient accuracy, at a fraction of the computational expense of running the full-physics model. The speed-up from history matching and reduced-order modelling with neural networks holds potential for application to real-time reservoir management and decision-making processes.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

Geothermal resource development presents new opportunities for sustainable energy production, economic growth, and a more environmentally friendly energy landscape. However, subsurface modelling is crucial to ensuring efficient heat extraction for decades to come.

Recent progress in deep learning has motivated the application of optimisation techniques with backpropagation for inverse modelling to condition static and dynamic reservoir models to well data. The emergence of advanced machine learning methods commonly used in computer vision and natural language processing has also inspired research into modelling physical behaviours in computational geoscience. This thesis integrates deep learning models such as multilayer perceptrons, convolutional neural networks, adversarial networks, and transformer-based networks into a framework that allows for the application of artificial intelligence in geological modelling, history matching, and reservoir forecasting.

In Chapter 3, a neural network is trained to understand the process of generating surface-based geological models (SBGMs) by learning the relationship between the SBGM inputs and the facies at well locations. SBGMs provide an efficient way to represent geological heterogeneity and require fewer parameters for producing realistic geological architectures, but conditioning them to well data is challenging due to the ill-posed nature of the problem. Conditioning the SBGMs to observed data is achieved through inverse modelling, where the trained neural network replaces the forward model. The inputs are optimised using backpropagation, allowing for the generation of multiple realisations that fit well data and help to understand uncertainties in reservoir models. The method enables the rapid generation of parametric, grid-free surfaces in SBGMs that are calibrated to observed well data. The uncertainty assessment shows low variations in models close to the control data and higher variations farther away, aiding the understanding of spatial uncertainties.

Chapter 4 extends the method to condition SBGMs to well data, particularly for wells with complex trajectories. The method leverages neural networks to simulate the process of generating and calibrating SBGMs to well data. A significant advantage of this method is its flexibility due to the grid-free approach that allows the representation of various geological architectures and well trajectories, including deviated, horizontal, and multilateral paths, compared with other methods that require a pre-defined grid. The method ensures geological consistency in the generated models and provides a practical tool for reservoir modelling and characterisation. The approach also includes a method for evaluating uncertainties, making it a valuable diagnostic and prognostic tool for geologists and reservoir engineers.

A workflow for developing representative models of subsurface reservoirs by assimilating dynamic data using a machine learning ensemble-based approach is presented in Chapter 5. The method is demonstrated with geothermal reservoir-well models, where the number of parameters greatly exceeds the observed temperature measurements. The proposed solution builds on the neural network-based methodology for conditioning parametric grid-free surface-based geological models (SBGMs) to static data. This study expands the approach to include dynamic data from reservoir simulation. The resulting reservoir models are consistent with geological processes, geophysical information from well logs, and fluid flow data from wells. The use of neural networks offers a more efficient solution to the ill-posed inverse problem of history matching, compared to traditional computationally expensive methods. The developed methodology is proposed for reservoir management as it supports rapid decision-making.

Chapter 6 applies advanced deep learning models to generate reduced-order models (ROMs) of geothermal reservoir-well systems, focusing on efficient data representation and the preservation of the system dynamics to a significant degree of accuracy. The chapter introduces the use of convolutional autoencoders (CAEs) with space-filling curves (SFCs) for data compression in the generation of ROMs of subsurface reservoirs having unstructured meshes that better capture the geometries and architectures of subsurface geological bodies. The study demonstrates the application of transformer-based models, with self-attention mechanisms, to learn and predict temperature in geothermal reservoirs. This is the first application of SFC-CAE for data compression in subsurface reservoirs with transformer-based models for predicting spatio-temporal data. The developed ROMs are suitable for significantly speeding up reservoir forecasting and uncertainty analysis, where multiple reservoir simulations are needed for rapid reservoir management and optimisation.

In Chapter 7, a framework is presented that integrates the methods discussed in the preceding chapters: conditioning SBGMs with neural networks, history matching with adversarial learning, and reduced order models with convolutional autoencoders, space-filling curves, and transformer-based networks for reservoir forecasting. Although future work is required to improve the accuracy of the results by decreasing errors that propagate to the final reservoir forecasting step, the proof of concept presented contributes to history matching and reservoir simulation in subsurface reservoirs, particularly for geothermal energy, providing a rapid, robust, and reliable

method for reservoir modelling and management using neural networks. This has great potential for real-time reservoir monitoring and management.

8.2 Future Work

The subsurface industry is embracing artificial intelligence, particularly advanced deep learning methods, to streamline operations, gain valuable insights, and address challenges in both the upstream and downstream sectors. The field of subsurface modelling and development has witnessed a growing interest in the research and implementation of artificial intelligence since the beginning of this Ph.D. Several reviews highlight this trend (Dramsch, 2020; Mishra, 2022; Okoroafor et al., 2022; Samnioti and Gaganis, 2023; Sircar et al., 2021). Factors such as enhanced computing capabilities, improved data accessibility, open source codes, more robust models, refined training approaches, and the industry's growing need for automation and efficiency are all contributing to the increasing adoption of machine learning in this sector. There are still opportunities to be explored to advance the techniques developed in this research and enhance the practicality of the proposed approach to generate subsurface reservoir models of geothermal systems that are representative and computationally efficient.

8.2.1 Conditioning SBGMs to Stratigraphic Surfaces

Representing geological information requires models that best capture the inherent relationships and interpretations of the depositional processes that produce the observed sequences and patterns. Models of geological depositional settings can be seen as products of a sequence modelling task. Conditioning SBGMs to stratigraphic surfaces has the potential to greatly enhance the accuracy of the generated SBGMs by incorporating more geological constraints to improve facies classification and reduce uncertainties. Furthermore, since perturbing previous events can lead to variations among the conditioned realisations at the boundaries of facies associations, incorporating information about variations in the number and thickness of layers using information from surfaces may improve the results of conditioning models to a reference data.

Traditional machine learning algorithms, such as decision trees, feedforward networks, and convolutional neural networks, are unable to effectively capture the sequential information present in depositional sequences, layers, or hierarchical surfaces. Sequence modelling networks are more effective in handling processing tasks that involve hierarchical relationships and have been applied to predict stratigraphic units and well logs by interpreting borehole data as oriented spatial sequences or series of data that are processed using recurrent neural networks (RNNs) (Zhang et al., 2018; Zhou et al., 2019). Although RNNs and long short-term memory (LSTM) (Hochreiter and Schmidhuber, 1997) have become the conventional methods for sequence modelling,

transformer-based networks improve limitations of RNNS and LSTMs, such as the requirement to input data at each time step, which leads to increased computational time.

In the case of the geological models considered in this research, the sequences, which can be given as the parameterized inputs for each parasequence, can reach hundreds. This has a significant impact on computational efficiency when using RNNs or LSTMs. Transformers employ attention mechanisms to construct sequential models that do not depend on the distance between the input and output. By applying self-attention, global dependencies between inputs and outputs are obtained or derived, enabling parallelisation and improving model performance. Transformers with self-attention mechanisms have been used to mitigate the challenges associated with learning long-term or global dependencies (Vaswani et al., 2017). Furthermore, the use of positional embeddings to encode geological meanings based on location within the sequence, and the bidirectional nature of certain transformer-based models, facilitates the modelling of geological features by examining their contextual associations.

8.2.2 Physics-Informed Neural Networks (PINNs)

Machine learning models are usually trained on historical observed data, and while these models can provide accurate predictions for scenarios similar to the training dataset, machine learning models struggle to generalise well to unseen or significantly differing conditions. Reservoir systems are complex and can exhibit unique characteristics, making it challenging even for state-of-the-art neural networks, such as transformers, to handle novel or rare situations. This can be attributed to the fact that machine learning models do not incorporate prior knowledge or governing laws specific to physical systems (Karniadakis et al., 2021; Raissi et al., 2019). Integrating domain knowledge into machine learning models is crucial to maintaining physical consistency and ensuring realistic simulation results.

In future work, neural network-based parameterizations of the forward-problem in the transformer-based model could accelerate and improve the results of the proposed reduced modelling approach presented in Chapters 6 and 7. Physics-informed neural networks (PINNs) show the potential to constrain neural networks to a solution space that honours the governing physical in computational science (Karniadakis et al., 2021; Kashinath et al., 2021; Wu et al., 2018). In recent years, PINNs have been successfully applied to model fluid flow in porous media (Bi et al., 2024; Han et al., 2023; Pachalieva et al., 2022; Yan et al., 2022; Zhang, 2022). Since transformers learn longer-term dependencies between inputs and outputs by relying on its attention mechanisms to prevent the loss of contextual information, including physical constraints might enhance the capability of transformers to allow for information to persist far beyond training scenarios and conditions.

Furthermore, the improved reduced-order models with PINNs can be used for the rapid generation of ensembles to obtain sensitivities using a goal-based approach as alternatives to adjoints in variational data assimila-

tion (Heaney et al., 2018). This could ensure that small localised errors in the initial conditions or controls do not affect the model solution far in time and space to improve the results of the history matching process.

8.2.3 Data Assimilation in the Latent Space

To further improve the speed of computations and allow applications such as real-time data analysis, data assimilation can be performed in a reduced space where the smaller number of variables reduces computational expense and enables computations to run orders of magnitude faster. Several approaches to data assimilation in the latent space have been developed: truncated singular vector decomposition (TSVD) to represent a low-dimensional approximation of the covariance matrix, control variable transform (CVT) and reduced order modelling where data compression is typically performed using autoencoders (Mack et al., 2020) and time evolution of system dynamics is performed using various types of neural networks, such as fully connected networks with skip connections (Peyron et al., 2021) or LSTMs (Amendola et al., 2021; Cheng et al., 2022). Mosser et al. (2019) performed history-matching by performing gradient descent on the latent variables that control the output of the generative model. Yoon et al. (2022) have demonstrated the use of convolutional neural networks (CNNs), LSTM, and deep neural networks (DNNs) for surrogate modelling and variational autoencoders (VAE) for data assimilation in the latent space within a histroy matching framework of multiphase CO_2 flow and pressure propagation in a heterogeneous reservoir.

Deep learning has been applied for ensemble data assimilation in low-dimensional latent spaces in subsurface systems (Razak et al., 2022). Wang et al. (2023) demonstrated the use of machine learning-based surrogate models with active learning in iterative ensemble smoothing (IES) for inversion. The surrogate model is continuously adjusted and improved with selected posterior samples in each iteration of IES, allowing it to adapt to the search path of the solution. This benefit of constraining the search space is similar to that presented in Chapter 5 where the history matching step consists of well-conditioning and data assimilation cycles in each iteration. Performing data assimilation in reduced space can improve computational efficiency, particularly for real cases.

8.2.4 Application to Real Field Cases

While the techniques described in this thesis are demonstrated using a 2D synthetic case study, the methods described in this work can also be applied to real cases, where 3D surface-based geological models (SBGMs) can be generated using the method presented in Jacquemyn et al. (2019). For field cases, the workflow outlined in Chapter 7 can be applied. However, in addition to the aforementioned future work, the following modifications will be necessary to make the implementation more practical and applicable:

1. Constructing 3D SBGMs and modelling complex geometries such as channelised fluvial sandbodies, structural configurations of folds and faults, and highly inclined well trajectories will require the use of surfaces with higher degrees to represent real-world geological architectures and features.

2. A broad range of facies codes that are representative of subsurface geology will be necessary for modelling petrophysical heterogeneity.

3. The modelling of coupled hydraulic-thermal-mechanical-chemical processes (HTMC). This implies that reservoir simulation models must consider heat exchange via conduction to surrounding rocks even in regions where there is no flow, and stresses induced by temperature and pressure changes may create or modify fracture networks and cause deposition or dissolution of minerals. Thus, geochemical effects due to fluid flow or thermal effects will need to be modelled.

4. The presented data assimilation applications make the assumption that model errors in time are uncorrelated. However, it has been demonstrated in typical geoscience models that errors are correlated in time (Evensen, 1994). Therefore, it will beneficial to incorporate error modelling to better quantify uncertainties and risks.

The proposed modifications will require the following improvements to reduce the limitations of the developed method for field cases:

Increased computational requirements and longer run time Generating SBGMs that represent complex geometries and reservoir models that model HTMC processes will require more computational resources. Nevertheless, the use of SBGMs with parameterized inputs allows for the rapid generation of geological models, and the adoption of techniques such as dynamic mesh optimisation could significantly reduce simulation time whilst preserving reasonable accuracy of the fluid flow models.

Greater model complexity and training time Longer time requirements for training neural networks due to increased input parameters which could require more complex and deeper networks. However, the type of SBGMs used, which provide the inputs for many of the models trained, have many advantages, as highlighted throughout this research. The ability of the models to accurately represent intricate subsurface heterogeneity with fewer parameters could also potentially significantly decrease the computational cost for more complex geologies or 3D models due to fewer optimisation parameters. Furthermore, Osman et al. (2021) has shown that the number of surfaces required to model realistic reservoirs using surfaces is low to generate representative models.

Increased data requirements Deep learning methods such as those presented in this research are powerful methods for generating parametric representations of physical behaviours in geoscience and other large-scale applications. With increased input parameters, model complexity, and real-world behaviours such as high nonlinearities and high dimensionality, more training data will be required to obtain reasonable and satisfactory accuracy for field cases. One potential approach is to fine-tune pre-trained models that have been trained on extensive geological datasets, as a foundation for field scenarios with limited data. The methods presented in this research could be of great benefit, as the geothermal resources considered are in settings similar to oil and gas reservoirs and will likely have extensive data for machine learning applications.

Bibliography

Abdulrazzaq, F. N. and Hasan, O. F. (2023), 'A review of automatic history matching', *Materials Today: Proceedings* **80**(3), 3817–3822.

Alajmi, M. and Ertekin, T. (2007), The development of an artificial neural network as a pressure transient analysis tool for applications in double-porosity reservoirs, *in* 'SPE Asia Pacific Oil and Gas Conference and Exhibition, Jakarta, Indonesia'. SPE-108604-MS.

Alguliyev, R., Aliguliyev, R., Imamverdiyev, Y. and Sukhostat, L. (2022), 'History matching of petroleum reservoirs using deep neural networks', *Intelligent Systems with Applications* **16**, 200128.

Alimonti, C., Soldo, E. and Scrocca, D. (2021), 'Looking forward to a decarbonized era: Geothermal potential assessment for oil and gas fields in Italy', *Geothermics* **93**, 102070.

Amendola, M., Arcucci, R., Mottet, L., Casas, C. Q., Fan, S., Pain, C., Linden, P. and Guo, Y.-K. (2021), Data assimilation in the latent space of a convolutional autoencoder, *in* 'Computational Science – ICCS 2021', Springer International Publishing, Cham, pp. 373–386.

Amini, S. and Mohaghegh, S. (2019), 'Application of machine learning and artificial intelligence in proxy modeling for fluid flow in porous media', *Fluids* **4**(3), 126.

Aminzadeh, F. and Dasgupta, S. N. (2013), *Geophysics for petroleum engineers*, Elsevier Science, pp. 151–189.

Anderson, J. L. (1996), 'A method for producing and evaluating probabilistic forecasts from ensemble model integrations', *Journal of Climate* **9**, 1518–1530.

Anderson, T. C. (2013), Geothermal potential of deep sedimentary basins in the United States, *in* 'Unconventional Resources Technology Conference, Denver, Colorado'. URTEC-1579729-MS.

Anzellini, S., Dewaele, A., Mezouar, M., Loubeyre, P. and Morard, G. (2013), 'Melting of iron at Earth's inner core boundary based on fast x-ray diffraction', *Science* **340**, 464–466.

Apte, A., Jones, C. K. R. T., Stuart, A. M. and Voss, J. (2008), 'Data assimilation: Mathematical and statistical perspectives', *International Journal for Numerical Methods in Fluids* **56**(8), 1033–1046.

Ariturk, M. S. (2018), Optimizing the production and injection wells flow rates in geothermal field using artificial intelligence, Master's thesis, West Virginia University.

Asch, M., Bocquet, M. and Nodet, M. (2016), *Data assimilation: methods, algorithms, and applications*, Society for Industrial and Applied Mathematics, Philadelphia.

Auroux, D., Blum, J. and Ruggiero, G. (2016), *Data Assimilation for Geophysical Fluids: The Diffusive Back and Forth Nudging*, Springer International Publishing, Cham, pp. 139–174.

Aydin, H., Akin, S. and Senturk, E. (2020), 'A proxy model for determining reservoir pressure and temperature for geothermal wells', *Geothermics* **88**, 101916.

Bahrami, P., Sahari Moghaddam, F. and James, L. A. (2022), 'A review of proxy modeling highlighting applications for reservoir engineering', *Energies* **15**(14). 5247.

Bai, Z. (2002), 'Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems', *Applied Numerical Mathematics* **43**(1), 9–44.

Barkaoui, A., Zarhloule, Y., Verdoya, M., Pasquale, V. and Lahrach, H. (2014), 'Progress in understanding the geothermal sedimentary basins in northeastern Morocco', *Journal of African Earth Sciences* **97**, 1–8.

Bertoncello, A., Sun, T., Li, H., Mariethoz, G. and Caers, J. (2013), 'Conditioning surface-based geological models to well and thickness data', *Mathematical Geosciences.* **45**, 873–893.

Bethke, C. M., Altaner, S. P., Harrison, W. J. and Upson, C. (1988), 'Supercomputer analysis of sedimentary basins', *Science* **239**(4837), 261–267.

Bi, J., Li, J., Wu, K., Chen, Z., Chen, S., Jiang, L., Feng, D. and Deng, P. (2024), 'A physics-informed spatial-temporal neural network for reservoir simulation and uncertainty quantification', *SPE Journal* **29**(04), 2026–2043. SPE-218386-PA.

Bielicki, J. M., Leveni, M., Johnson, J. X. and Ellis, B. R. (2023), 'The promise of coupling geologic CO_2 storage with sedimentary basin geothermal power generation', *iScience* **26**(2), 105618.

Bishop, C. M. (2006), *Pattern recognition and machine learning*, Information science and statistics, pp. 209, Springer-Verlag, Berlin, Heidelberg.

Blank, L., Meneses Rioseco, E., Caiazzo, A. and Wilbrandt, U. (2021), 'Modeling, simulation, and optimization of geothermal energy production from hot sedimentary aquifers', *Computational Geosciences* **25**(1), 67–104.

Bottou, L. (2010), Large-scale machine learning with stochastic gradient descent, *in* 'Proceedings of COMPSTAT'2010', Physica-Verlag HD, Heidelberg, pp. 177–186.

Bouttier, F. and Courtier, P. (2002), *Data assimilation concepts and methods*, Meteorological Training Course Lecture Series, European Centre for Medium Range Weather Forecasts (ECMWF), pp. 1–5, 30–35.

Brantson, E. T., Ju, B., Omisore, B. O., Wu, D., Selase, A. E. and Liu, N. (2018), 'Development of machine learning predictive models for history matching tight gas carbonate reservoir production profiles', *Journal of Geophysics and Engineering* **15**(5), 2235–2251.

Brcković, A., Kovačević, M., Cvetković, M., Kolenković Močilac, I., Rukavina, D. and Saftić, B. (2017), 'Application of artificial neural networks for lithofacies determination based on limited well data', *Central European Geology* **60**, 299–315.

Buizza, C., Casas, C. Q., Nadler, P., Mack, J., Marrone, S., Titus, Z., Cornec, C. L., Heylen, E., Dur, T., Ruiz, L. B., Heaney, C., Lopez, J. A. D., Kumar, K. S. and Arcucci, R. (2022), 'Data learning: Integrating data assimilation and machine learning', *Journal of Computational Science* **58**, 101525.

Busby, J. (2014), 'Geothermal energy in sedimentary basins in the UK', *Hydrogeology Journal* **22**(1), 129–141.

Bdard, K., Comeau, F.-A., Raymond, J., Malo, M. and Nasr, M. (2018), 'Geothermal characterization of the St. Lawrence lowlands sedimentary basin, Qubec, Canada', *Natural Resources Research* **27**(4), 479–502.

Campbell, S. A., Mielke, P. and Gtz, A. E. (2016), 'Geothermal energy from the Main Karoo Basin? New insights from borehole KWV-1 (Eastern Cape, South Africa', *Geothermal Energy* **4**(1).

Canchumuni, S. A., Emerick, A. A. and Pacheco, M. A. (2017), Integration of ensemble data assimilation and deep learning for history matching facies models, *in* 'Offshore Technology Conference Brazil, Rio de Janeiro, Brazil'. OTC-28015-MS.

Canchumuni, S. W., Castro, J. D., Potratz, J., Emerick, A. A. and Pacheco, M. A. C. (2020), 'Recent developments combining ensemble smoother and deep generative networks for facies history matching', *Computational Geosciences* **25**, 433–466.

Canchumuni, S. W., Emerick, A. A. and Pacheco, M. A. C. (2019*a*), 'History matching geological facies models based on ensemble smoother and deep generative models', *Journal of Petroleum Science and Engineering* **177**, 941–958.

Canchumuni, S. W., Emerick, A. A. and Pacheco, M. A. C. (2019*b*), 'Towards a robust parameterization for conditioning facies models using deep variational autoencoders and ensemble smoother', *Computers & Geosciences* **128**, 87 – 102.

Cano, N. A., Cspedes, S., Redondo, J., Foo, G., Jaramillo, D., Martinez, D., Gutirrez, M., ataquiba, J., Rojas, J., Corts, F. B. and Franco, C. A. (2022), 'Power from geothermal resources as a co-product of the oil and gas industry: A review', *ACS Omega* **7**(45), 40603–40624.

Cardoso, M. A. and Durlofsky, L. J. (2010), 'Linearized reduced-order models for subsurface flow simulation', *Journal of Computational Physics* **229**(3), 681–700.

Carlberg, K. (2015), 'Adaptive h-refinement for reduced-order models', *International Journal for Numerical Methods in Engineering* **102**, 1192–1210.

Carlberg, K., Barone, M. and Antil, H. (2017), 'Galerkin v. least-squares Petrov-Galerkin projection in nonlinear model reduction', *Journal of Computational Physics* **330**, 693–734.

Carrassi, A., Bocquet, M., Bertino, L. and Evensen, G. (2018), 'Data assimilation in the geosciences: An overview of methods, issues, and perspectives', *Wiley Interdisciplinary Reviews: Climate Change* **9**(5), e535.

Caulk, R. A. and Tomac, I. (2017), 'Reuse of abandoned oil and gas wells for geothermal energy production', *Renewable Energy* **112**(C), 388–397.

Chai, Z., Yan, B., Killough, J. and Wang, Y. (2018), 'An efficient method for fractured shale reservoir history matching: The embedded discrete fracture multi-continuum approach', *Journal of Petroleum Science and Engineering* **160**, 170–181.

Chan, S. and Elsheikh, A. H. (2017), 'Parametrization and generation of geological models with generative adversarial networks'. *ArXiv e-prints*. arXiv:1708.01810v2 [stat.ML].

Chan, S. and Elsheikh, A. H. (2019), 'Parametric generation of conditional geological realizations using generative neural networks', *Computational Geosciences.* **23**, 925–952.

Chatterjee, A. (2000), 'An introduction to the proper orthogonal decomposition', *Current Science* **78**(7), 808–817.

Chen, Y. and Oliver, D. (2014), 'History matching of the Norne full-field model using an iterative ensemble smoother', *SPE Reservoir Evaluation & Engineering* **17**(2), 244–256. SPE-164902-PA.

Chen, Y. and Oliver, D. S. (2011), 'Ensemble randomized maximum likelihood method as an iterative ensemble smoother', *Mathematical Geosciences* **44**(1), 1–26.

Chen, Y., Oliver, D. S. and Zhang, D. (2009), 'Data assimilation for nonlinear problems by ensemble kalman filter with reparameterization', *Journal of Petroleum Science and Engineering* **66**(1), 1–14.

Chen, Y., Smith, P., Carrassi, A., Pasmans, I., Bertino, L., Bocquet, M., Finn, T. S., Rampal, P. and Dansereau, V. (2024), 'Multivariate state and parameter estimation with data assimilation applied to sea-ice models using a Maxwell elasto-brittle rheology', *The Cryosphere* **18**(5), 2381–2406.

Cheng, S., Prentice, I. C., Huang, Y., Jin, Y., Guo, Y.-K. and Arcucci, R. (2022), 'Data-driven surrogate model with latent data assimilation: Application to wildfire forecasting', *Journal of Computational Physics* **464**, 111302.

Choi, Y., Coombs, D. and Anderson, R. (2020), 'SNS: A solution-based nonlinear subspace method for time-dependent model order reduction', *SIAM Journal on Scientific Computing* **42**(2), A1116–A1146.

Christie, M., Demyanov, V. and Erbas, D. (2006), 'Uncertainty quantification for porous media flows', *Journal of Computational Physics* **217**, 143–158.

Clegg, J. and Krase, S. (2022), New-generation geothermal wells - How the upstream oil and gas industry can bring its experience to bear, *in* 'SPE/IADC Drilling Conference and Exhibition, Galveston, Texas, USA'. SPE-208747-MS.

Cullick, A. S., Johnson, D. and Shi, G. (2006), Improved and more-rapid history matching with a nonlinear proxy and global optimization, *in* 'SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA'. SPE-101933-MS.

Daley, R. (1993), *Atmospheric data analysis*, Cambridge Atmospheric and Space Science Series, Cambridge University Press.

Darcy, H. (1856), *Les fontaines publiques de la ville de Dijon: Exposition et application des principes à suivre et des formules à employer dans les questions de distribution d'eau : Ouvrage terminé par un appendice relatif aux fournitures d'eau de plusieurs villes, au filtrage des eaux et à la fabrication des tuyaux de fonte, de plomb, de tôle et de bitume*, Victor Dalmont, éditeur.

Davis, A. P. and Michaelides, E. E. (2009), 'Geothermal power production from abandoned oil wells', *Energy* **34**(7), 866–872.

Denney, D. (2006), 'Treating ncertainties in reservoir-performance prediction with neural networks', *Journal of Petroleum Technology* **58**(06), 69–71. SPE-0606-0069-JPT.

Deutsch, C. V. and Wang, L. (1996), 'Hierarchical object-based stochastic modeling of fluvial reservoirs', *Mathematical Geology* **28**, 857–880.

Dogo, E. M., Afolabi, O. J., Nwulu, N. I., Twala, B. and Aigbavboa, C. O. (2018), A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks, *in* '2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)', pp. 92–99.

Dovera, L. and Della Rossa, E. (2011), 'Multimodal ensemble Kalman filtering using gaussian mixture models', *Computational Geosciences* **15**(2), 307–323.

Dramsch, J. S. (2020), '70 years of machine learning in geoscience in review', *Advances in Geophysics* **61**, 1–55.

Duplyakin, D., Beckers, K. F., Siler, D. L., Martin, M. J. and Johnston, H. E. (2022), 'Modeling subsurface performance of a geothermal reservoir using machine learning', *Energies* **15**(3). 967.

Dupont, E., Zhang, T., Tilke, P., Liang, L. and Bailey, W. (2018), 'Generating realistic geology conditioned on physical measurements with generative adversarial networks'. *ArXiv e-prints*. arXiv:1802.03065v3 [stat.ML].

Eivazi, H., Veisi, H., Naderi, M. H. and Esfahanian, V. (2020), 'Deep neural networks for nonlinear model order reduction of unsteady flows', *Physics of Fluids* **32**(10), 105104.

Eltahan, E., Ganjdanesh, R., Yu, W., Sepehrnoori, K., Drozd, H. and Ambrose, R. (2020), Assisted history matching using bayesian inference: Application to multi-well simulation of a huff-n-puff pilot test in the Permian Basin, *in* 'SPE/AAPG/SEG Unconventional Resources Technology Conference, Austin, Texas, USA'. URTEC-2020-2787-MS.

Emerick, A. A. (2016), 'Analysis of the performance of ensemble-based assimilation of production and seismic data', *Journal of Petroleum Science and Engineering* **139**, 219–239.

Emerick, A. A. (2018), 'Deterministic ensemble smoother with multiple data assimilation as an alternative for history-matching seismic data', *Computational Geosciences* **22**(5), 1175–1186.

Emerick, A. A. and Reynolds, A. C. (2013), 'Ensemble smoother with multiple data assimilation', *Computers and Geosciences* **55**, 3–15.

Esmaili, S. and Mohaghegh, S. D. (2016), 'Full field reservoir modeling of shale assets using advanced data-driven analytics', *Geoscience Frontiers* **7**, 1674–9871.

Evensen, G. (1994), 'Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics', *Journal of Geophysical Research* **99**(C5), 10143–10162.

Evensen, G. (2003), 'The ensemble Kalman filter: Theoretical formulation and practical implementation', *Ocean Dynamics* **53**, 343–367.

Evensen, G., Hove, J., Meisingset, H., Reiso, E., Seim, K. S. and Espelid, Ø. (2007), Using the EnKF for assisted history matching of a North Sea reservoir model, *in* 'SPE Reservoir Simulation Symposium, Houston, Texas, USA', Society of Petroleum Engineers (SPE). SPE-106184-MS.

Evensen, G., Vossepoel, F. C. and van Leeuwen, P. J. (2022), *Randomized-Maximum-Likelihood Sampling*, Springer International Publishing, Cham, pp. 73–77.

Ferguson, G. and Grasby, S. E. (2014), 'The geothermal potential of the basal clastics of Saskatchewan, Canada', *Hydrogeology Journal* **22**(1), 143–150.

Ferguson, G. and Ufondu, L. (2017), 'Geothermal energy potential of the western Canada sedimentary basin: Clues from coproduced and injected water', *Environmental Geosciences* **24**(3), 113–121.

Fitch, P. J., Lovell, M. A., Davies, S. J., Pritchard, T. and Harvey, P. K. (2015), 'An integrated and quantitative approach to petrophysical heterogeneity', *Marine and Petroleum Geology* **63**, 82–96.

Fu, H., Zhang, Y. and Ma, M. (2019), 'Seismic waveform inversion using a neural network-based forward', *Journal of Physics: Conference Series* **1324**(1), 012043.

Garcia-Gutierrez, A., Espinosa-Paredes, G. and Hernandez-Ramirez, I. (2002), 'Study on the flow production characteristics of deep geothermal wells', *Geothermics* **31**(2), 141–167.

Geneva, N. and Zabaras, N. (2022), 'Transformers for modeling physical systems', *Neural Networks* **146**, 272–289.

Gers, F., Schmidhuber, J. and Cummins, F. (1999), Learning to forget: continual prediction with LSTM, *in* '1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470)', Vol. 2, pp. 850–855.

Gharibi, S., Mortezazadeh, E., Hashemi Aghcheh Bodi, S. J. and Vatani, A. (2018), 'Feasibility study of geothermal heat extraction from abandoned oil wells using a u-tube heat exchanger', *Energy* **153**, 554–567.

Ghojogh, B., Ghodsi, A., Karray, F. and Crowley, M. (2021), 'Generative adversarial networks and adversarial autoencoders: Tutorial and survey'. *ArXiv e-prints*. arXiv:2111.13282 [cs.LG].

Gluyas, J., Adams, C., Busby, J., Craig, J., Hirst, C., Manning, D., McCay, A., Narayan, N., Robinson, H., Watson, S., Westaway, R. and Younger, P. (2018), 'Keeping warm: a review of deep geothermal potential of the UK', *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy* **232**(1), 115–126.

Goldstein, B., Hill, A., Long, A., Budd, A., Holgate, F. and Malavazos, M. (2009), Hot rock geothermal energy plays in Australia, *in* 'Proceedings of the Thirty-Fourth Workshop on Geothermal Reservoir Engineering', Stanford University, Stanford, California. SGP-TR-187.

Goldstein, B., Hiriart, G., Tester, J. B., Bertani, R., Bromley, G.-N. L. C., Huenges, E. H., Ragnarsson, A., Mongillo, M. M. and Zui, V. I. (2011), Great expectations for geothermal energy to 2100, *in* 'Proceedings of the Thirty-Sixth Workshop on Geothermal Reservoir Engineering'. SGP-TR-191.

Gonzalez, F. J. and Balajewicz, M. (2018), 'Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems'. *ArXiv e-prints*. arXiv:1808.01346v2 [math.DS].

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014), Generative adversarial nets, *in* 'Advances in Neural Information Processing Systems', Vol. 27, Curran Associates, Inc., pp. 2672–2680.

Gorucu, F., Gorucu, F. B., Ertekin, T., Bromhal, G. S., Smith, D. H., Sams, W. N. and Jikich, S. A. A. (2005), Neurosimulation tool for predicting performance in enhanced coalbed methane and CO_2 sequestration project, *in* 'SPE Annual Technical Conference, Dallas, Texas, USA', Vol. 28. SPE-97164-MS.

Grana, D., Azevedo, L. and Liu, M. (2020), 'A comparison of deep machine learning and Monte Carlo methods for facies classification from seismic data', *Geophysics* **85**(4), WA41–WA52.

Gratton, D. and Willcox, K. (2012), Reduced-order, trajectory piecewise-linear models for nonlinear computational fluid dynamics, *in* '34th AIAA Fluid Dynamics Conference and Exhibit, Portland, Oregon, USA'. AIAA 2004-2329.

Gu, Y. and Oliver, D. S. (2006), 'The ensemble Kalman filter for continuous updating of reservoir simulation models', *Journal of Energy Resources Technology* **128**(1), 79–87.

Gu, Y. and Oliver, D. S. (2007), 'An iterative ensemble Kalman filter for multiphase fluid flow data assimilation', *SPE Journal* **12**, 438–446. SPE-108438-PA.

Haji, S. H. and Abdulazeez, A. M. (2021), 'Comparison of optimization techniques based on gradient descent algorithm: A review', *PalArchs Journal Of Archaeology Of Egypt/Egyptology* **18**(4), 2715–2743.

Han, J.-X., Xue, L., Wei, Y.-S., Qi, Y.-D., lei Wang, J., Liu, Y.-T. and Zhang, Y.-Q. (2023), 'Physics-informed neural network-based petroleum reservoir simulation with sparse data using domain decomposition', *Petroleum Science* **20**(6), 3450–3460.

Hansen, J. A. and Penland, C. (2007), 'On stochastic parameter estimation using data assimilation', *Physica D: Nonlinear Phenomena* **230**(1), 88–98.

Hardman, J., Bolton, R., Drumm, E., Crossley, R., Olivares, C., MacInnes, E. and Jarvis, R. (2020), Transitioning from oil and gas to geothermal: A geoscience perspective, *in* '1st Geoscience & Engineering in Energy Transition Conference', Vol. 2020, European Association of Geoscientists & Engineers, pp. 1–5.

Hassanpour, M. M., Pyrcz, M. J. and Deutsch, C. V. (2013), 'Improved geostatistical models of inclined heterolithic strata for McMurray Formation, Alberta, Canada', *AAPG Bulletin* **97**(7), 1209–1224.

Haugen, V., Naevdal, G., Natvik, L.-J., Evensen, G., Margrethe Berg, A. and Flornes, K. (2008), 'History matching using the ensemble Kalman filter on a North Sea field case', *SPE Journal* **13**(04), 382–391. SPE-102430-PA.

He, K., Zhang, X., Ren, S. and Sun, J. (2015), 'Delving deep into rectifiers: Surpassing human-level performance on imagenet classification'. *ArXiv e-prints*. arXiv:1502.01852v1 [cs.CV].

Heaney, C. E., Li, Y., Matar, O. K. and Pain, C. C. (2024), 'Applying convolutional neural networks to data on unstructured meshes with space-filling curves', *Neural Networks* **175**, 106198.

Heaney, C. E., Salinas, P., Fang, F., Pain, C. C. and Navon, I. M. (2018), 'Goal-based sensitivity maps using time windows and ensemble perturbations'. *ArXiv e-prints*. arXiv:1804.04457v2 [cs.CE].

Henry, E. and Hofrichter, J. (1992), [8] Singular value decomposition: Application to analysis of experimental data, *in* 'Numerical Computer Methods', Vol. 210 of *Methods in Enzymology*, Academic Press, pp. 129–192.

Hesthaven, J. S. and Ubbiali, S. (2018), 'Non-intrusive reduced order modeling of nonlinear problems using neural networks', *Journal of Computational Physics* **363**, 55–78.

Hilbert, D. (1891), 'Ueber die stetige Abbildung einer Line auf ein Flächenstück', *Mathematische Annalen* **38**, 459–460.

Hinton, G. E. and Salakhutdinov, R. R. (2006), 'Reducing the dimensionality of data with neural networks', *Science* **313**(5786), 504–507.

Hirst, C. (2017), The geothermal potential of low enthalpy deep sedimentary basins in the UK, PhD thesis, Durham University.

Hochreiter, S. and Schmidhuber, J. (1997), 'Long short-term memory', *Neural Computation* **9**(8), 1735–1780.

Jackson, M. D., Hampson, G. J., Saunders, J. H., El-Sheikh, A., Graham, G. H. and Massart, B. Y. G. (2014), 'Surface-based reservoir modelling for flow simulation', *Geological Society, London, Special Publications* **387**(1), 271–292.

Jackson, M. D., Percival, J. R., Mostaghimi, P., Tollit, B. S., Pavlidis, D., Pain, C. C., Gomes, J. L. M. A., ElSheikh, A. H., Salinas, P., Muggeridge, A. H. and Blunt, M. J. (2015), 'Reservoir modeling for flow simulation by use of surfaces, adaptive unstructured meshes, and an overlapping-control-volume finite-element method', *SPE Reservoir Evaluation & Engineering* **18**(2), 115–132. SPE-163633-PA.

Jacquemyn, C., Jackson, M. D. and Hampson, G. J. (2019), 'Surface-based geological reservoir modelling using grid-free NURBS curves and surfaces', *Mathematical Geosciences* **51**, 1–28.

Jafarpour, B. and McLaughlin, D. B. (2008), 'History matching with an ensemble Kalman filter and discrete cosine parameterization', *Computational Geosciences* **12**(2), 227–244.

Jessop, A. M., Ghomshei, M. M. and Drury, M. J. (1991), 'Geothermal energy in Canada', *Geothermics* **20**(5), 369–385.

Jin, Z. L., Liu, Y. and Durlofsky, L. J. (2020), 'Deep-learning-based surrogate model for reservoir simulation with time-varying well controls', *Journal of Petroleum Science and Engineering* **192**, 107273.

Jo, H., Jung, H., Ahn, J., Lee, K. and Choe, J. (2017), 'History matching of channel reservoirs using ensemble Kalman filter with continuous update of channel information', *Energy Exploration & Exploitation* **35**(1), 3–23.

Jo, H., Pan, W., Santos, J. E., Jung, H. and Pyrcz, M. J. (2021), 'Machine learning assisted history matching for a deepwater lobe system', *Journal of Petroleum Science and Engineering* **207**, 109086.

Jo, H., Santos, J. E. and Pyrcz, M. J. (2020), 'Conditioning well data to rule-based lobe model by machine learning with a generative adversarial network', *Energy Exploration & Exploitation* **38**(6), 2558–2578.

Johnson, L. A., PE and Walker, E. D. (2010), Oil production waste stream, a source of electrical power, *in* 'Prooceedings of the Thirty-Fifth Workshop on Geothermal Reservoir Engineering', Stanford University, Stanford, California. SGP-TR-188.

Jung, S., Lee, K., Park, C. and Choe, J. (2018), 'Ensemble-based data assimilation in reservoir characterization: A review', *Energies* **11**(2). 445.

Kabanikhin, S. I. (2008), 'Definitions and examples of inverse and ill-posed problems', *Journal of Inverse and Ill-posed Problems* **16**(4), 317–357.

Kalman, R. E. (1960), 'A new approach to linear filtering and prediction problems', *Journal of Basic Engineering* **82**(1), 35–45.

Kalnay, E. (2012), *Atmospheric modeling, data assimilation and predictability*, Cambridge University Press, pp. 136–204.

Kalnay, E. and Yang, S.-C. (2010), 'Accelerating the spin-up of ensemble Kalman filtering', *Quarterly Journal of the Royal Meteorological Society* **136**(651), 1644–1651.

Kani, J. and Elsheikh, A. (2019), 'Reduced-order modeling of subsurface multi-phase flow models using deep residual recurrent neural networks', *Transport in Porous Media* **126**, 713–741.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S. and Yang, L. (2021), 'Physics-informed machine learning', *Nature Reviews Physics* **3**, 422 – 440.

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A. and Kumar, V. (2019), 'Machine learning for the geosciences: Challenges and opportunities', *IEEE Transactions on Knowledge and Data Engineering* **31**(8), 1544–1554.

Karssenberg, D., Törnqvist, T. E. and S., B. J. (2001), 'Conditioning a process-based model of sedimentary architecture to well data', *Journal of Sedimentary Research* **71**(6), 868–879.

Kashinath, K., Mustafa, M., Albert, A., Wu, J.-L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H. A., Marcus, P., Anandkumar, A., Hassanzadeh, P. and Prabhat, n. (2021), 'Physics-informed machine learning: Case studies for weather and climate modelling', *Philosophical Transactions of the Royal Society A* **379**(2194), 20200093.

Kingma, D. P. and Ba, J. (2014), 'Adam: A method for stochastic optimization'. *ArXiv e-prints*. arxiv:1412.6980v9 [cs.LG].

Klie, H. and Florez, H. (2018), Data-driven modeling of fractured shale reservoirs, *in* 'Proceedings of the ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery, Barcelona', Vol. 2018, European Association of Geoscientists & Engineers, pp. 1–18.

Kostorz, W. J., Muggeridge, A. H. and Jackson, M. D. (2020), 'An efficient and robust method for parameterized nonintrusive reduced-order modeling', *International Journal for Numerical Methods in Engineering* **121**(20), 4674–4688.

Kostorz, W. J., Muggeridge, A. H. and Jackson, M. D. (2021), 'Non-intrusive reduced order modeling: Geometrical framework, high-order models, and a priori analysis of applicability', *International Journal for Numerical Methods in Engineering* **122**(10), 2545–2565.

Kravaris, C. and Seinfeld, J. H. (1985), 'Identification of parameters in distributed parameter systems by regularization', *SIAM Journal on Control and Optimization* **23**(2), 217–241.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2017), 'Imagenet classification with deep convolutional neural networks', *Communications of the ACM* **60**(6), 84–90.

Laloy, E., Hrault, R., Jacques, D. and Linde, N. (2018), 'Training-image based geostatistical inversion using a spatial generative adversarial neural network', *Water Resources Research* **54**(1), 381–406.

LeCun, Y. A., Bottou, L., Orr, G. B. and Müller, K.-R. (2012), *Efficient BackProp*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 9–48.

Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998), 'Gradient-based learning applied to document recognition', *Proceedings of the IEEE* **86**(11), 2278–2324.

Lee, K. C. (1996), Classification of geothermal resources - an engineering approach, *in* 'Proceedings of the Twenty-First Workshop on Geothermal Reservoir Engineering', Stanford University, Stanford, California. SGP-TR-151.

Lee, K. C. (2001), 'Classification of geothermal resources by exergy', *Geothermics* **30**(4), 431–442.

Li, D.-C., Chen, S.-C., Lin, Y.-S. and Huang, K.-C. (2021), 'A generative adversarial network structure for learning with small numerical data sets', *Applied Sciences* **11**(22). 10823.

Li, H. and Misra, S. (2021), 'Reinforcement learning based automated history matching for improved hydrocarbon production forecast', *Applied Energy* **284**, 116311.

Li, H., Wan, B., Chu, D., Wang, R., Ma, G., Fu, J. and Xiao, Z. (2023), 'Progressive geological modeling and uncertainty analysis using machine learning', *ISPRS International Journal of Geo-Information* **12**(3). 97.

Li, T., Gao, Y., Han, D., Yang, F. and Yu, B. (2020), 'A novel POD reduced-order model based on EDFM for steady-state and transient heat transfer in fractured geothermal reservoir', *International Journal of Heat and Mass Transfer* **146**, 118783.

Liang, Y., Lee, H., Lim, S., Lin, W., Lee, K. and Wu, C. (2002), 'Proper orthogonal decomposition and its applications—Part I: Theory', *Journal of Sound and Vibration* **252**(3), 527–544.

Liu, Y., Sun, W. and Durlofsky, L. J. A. (2019), 'Deep-learning-based geological parameterization for history matching complex models', *Mathematical Geosciences* **51**, 725–766.

Liu, Y., Wang, L., Shuai, Y., Zhang, Y., Liu, Z. and Chen, Z. (2015), 'Quantitative seismic interpretations to detect biogenic gas accumulations: A case study from Qaidam basin, China', *Bulletin of Canadian Petroleum Geology* **63**(1), 108–121.

Lopez, S., Hamm, V., Le Brun, M., Schaper, L., Boissier, F., Cotiche, C. and Giuglaris, E. (2010), '40 years of Dogger aquifer management in Ile-de-France, Paris Basin, France', *Geothermics* **39**(4), 339–356.

Lorentzen, R., Flornes, K. and Nvdal, G. (2012), 'History matching channelized reservoirs using the ensemble Kalman filter', *SPE Journal* **17**(01), 137–151. SPE-143188-PA.

Lorenz, E. N. (2005), 'Designing chaotic models', *Journal of the Atmospheric Sciences* **62**(5), 1574–1587.

Lowrie, W. and Fichtner, A. (2020), *The Earth's Heat*, Cambridge University Press, pp. 253–287.

Lucia, D. J., Beran, P. S. and Silva, W. A. (2004), 'Reduced-order modeling: New approaches for computational physics', *Progress in Aerospace Sciences* **40**(1), 51–117.

Luo, X., Stordal, A. S., Lorentzen, R. J. and Naevdal, G. (2015), 'Iterative ensemble smoother as an approximate solution to a regularized minimum-average-cost problem: Theory and applications', *SPE Journal* **20**(5), 962–982. SPE-176023-PA.

Ma, X., Zhang, J. and Zhao, H. (2009), Application of artificial neural networks in lithofacies interpretation used for 3D geological modelling, *in* '2009 ISECS International Colloquium on Computing, Communication, Control, and Management', Vol. 4, pp. 451–454.

Ma, X., Zhang, K., Wang, J., Yao, C., Yang, Y., Sun, H. and Yao, J. (2022), 'An efficient spatial-temporal convolution recurrent neural network surrogate model for history matching', *SPE Journal* **27**, 1160–1175. SPE-208604-PA.

Mack, J., Arcucci, R., Molina-Solana, M. and Guo, Y.-K. (2020), 'Attention-based convolutional autoencoders for 3D-variational data assimilation', *Computer Methods in Applied Mechanics and Engineering* **372**, 113291.

Maiti, S., Tiwari, R. K. and Kümpel, H. (2007), 'Neural network modelling and classification of lithofacies using well log data: A case study from KTB borehole site', *Geophysical Journal International* **169**(2), 733–746.

Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I. and Frey, B. (2016), 'Adversarial autoencoders'. *ArXiv e-prints*. arXiv:1511.05644v2 [cs.LG].

Maschio, C. and Schiozer, D. J. (2014), 'Bayesian history matching using artificial neural network and Markov Chain Monte Carlo', *Journal of Petroleum Science and Engineering* **123**, 62–71.

Michael, H. A., Li, H., Boucher, A., Sun, T., Caers, J. and Gorelick, S. M. (2010), 'Combining geologic-process models and geostatistics for conditional simulation of 3D subsurface heterogeneity', *Water Resources Research* **46**(5), W05527.

Miller, J. K., Sun, T., Li, H., Stewart, J., Genty, C., Li, D. and Lyttle, C. (2008), Direct modeling of reservoirs through forward process-based models: Can we get there?, *in* 'IPTC International Petroleum Technology Conference'. IPTC-12729-MS.

Miller, R. B. and Castle, J. W. andTemples, T. J. (2000), 'Deterministic and stochastic modeling of aquifer stratigraphy, South Carolina', *Ground Water* **38**(2), 284 – 295.

Mishra, S., ed. (2022), *Machine Learning Applications in Subsurface Energy Resource Management: State of the Art and Future Prognosis*, CRC Press.

Mo, S., Zhu, Y., Zabaras, N., Shi, X. and Wu, J. (2018), 'Deep convolutional encoder-decoder networks for uncertainty quantification of dynamic multiphase flow in heterogeneous media', *Water Resources Research* **55**(1), 703–728.

Moeck, I. S. (2014), 'Catalog of geothermal play types based on geologic controls', *Renewable and Sustainable Energy Reviews* **37**, 867–882.

Mohmad, N. I., Mandal, D., Amat, H., Sabzabadi, A. and Masoudi, R. (2020), History matching of production performance for highly faulted, multi layered, clastic oil reservoirs using artificial intelligence and data analytics: A novel approach, *in* 'SPE Asia Pacific Oil and Gas Conference and Exhibition, Virtual'. SPE-202460-MS.

Moore, E. H. (1990), 'On Certain Crinkly Curves', *Transactions of the American Mathematical Society* **1**(1), 72–90.

Mosser, L., Dubrule, O. and Blunt, M. J. (2018), 'Conditioning of three-dimensional generative adversarial networks for pore and reservoir-scale models'. *ArXiv e-prints*. arXiv:1802.05622v1 [stat.ML].

Mosser, L., Dubrule, O. and Blunt, M. J. (2019), 'Deepflow: History matching in the space of deep generative models'. *ArXiv e-prints*. arXiv:1905.05749v2 [cs.LG].

Mosser, L., Dubrule, O. and Blunt, M. J. (2020), 'Stochastic seismic waveform inversion using generative adversarial networks as a geological prior', *Mathematical Geosciences* **52**, 53–79.

Mudunuru, M. K. K., Harp, S., R., D., Guthrie, J., D., G. and Viswanathan, H. S. (2017), 'Regression-based reduced-order models to predict transient thermal output for enhanced geothermal systems', *Geothermics* **70**, 192–205.

Muffler, P. and Cataldi, R. (1978), 'Methods for regional assessment of geothermal resources', *Geothermics* **7**(2), 53–89.

Naevdal, G., Johnsen, L. M., Aanonsen, S. I. and Vefring, E. H. (2005), 'Reservoir monitoring and continuous model updating using ensemble Kalman filter', *SPE Journal* **10**(01), 66–74. SPE-84372-PA.

Odumodu, C. F. R. and Mode, A. W. (2016), 'Geothermal gradients and heat flow variations in parts of the eastern Niger delta, Nigeria', *Journal of the Geological Society of India* **88**(1), 107–118.

Okoroafor, E. R., Smith, C. M., Ochie, K. I., Nwosu, C. J. and Gudmundsdottir, H. (2022), 'Machine learning in subsurface geothermal energy: Two decades in review', *Geothermics* **102**, 102401.

Oliver, D. S. and Chen, Y. (2011), 'Recent progress on reservoir history matching: A review', *Computational Geosciences* **15**(1), 185–221.

Osman, H., Graham, G. H., Moncorge, A., Jacquemyn, C. and Jackson, M. D. (2021), 'Is cell-to-cell scale variability necessary in reservoir models?', *Mathematical Geosciences* **53**(4), 571–596.

Pachalieva, A., OMalley, D., Harp, D. R. and Viswanathan, H. S. (2022), 'Physics-informed machine learning with differentiable programming for heterogeneous underground reservoir pressure management', *Scientific Reports* **12**, 18734.

Pain, C. C., de Oliveira, C. R. E. and Goddard, A. J. H. (1999), 'A neural network graph partitioning procedure for grid-based domain decomposition', *International Journal for Numerical Methods in Engineering* **44**(5), 593–613.

Pal, M. (2021), 'On application of machine learning method for history matching and forecasting of times series data from hydrocarbon recovery process using water flooding', *Petroleum Science and Technology* **39**(15-16), 519–549.

Pal, M., Patil, P. and Saxena, P. (2022), Supervised deep learning-based models for HM and forecasting of water flooding and enhanced oil recovery scenarios, *in* 'Fourth EAGE WIPIC Workshop', Vol. 2022, European Association of Geoscientists & Engineers, pp. 1–7.

Pan, W., Torres-Verdn, C. and Pyrcz, M. J. (2021), 'Stochastic pix2pix: A new machine learning method for geophysical and well conditioning of rule-based channel reservoir models', *Natural Resources Research* **30**(2), 1319–1345.

Parquer, M. N., Collon, P. and Caumon, G. (2017), 'Reconstruction of channelized systems through a conditioned reverse migration method', *Mathematical Geosciences* **49**, 965–994.

Peyron, M., Fillion, A., Grol, S., Marchais, V., Gratton, S., Boudier, P. and Goret, G. (2021), 'Latent space data assimilation by using deep learning', *Quarterly Journal of the Royal Meteorological Society* **147**(740), 3759–3777.

Piegl, L. and Tiller, W. (1997), *The NURBS book*, Monographs in visual communications, Springer-Verlag, Berlin, Heidelberg.

Pinto, O., Aguilera, P., Cideos, O. and Henriquez, J. L. (2021), Towards the use of geothermal resources available in oil and gas sedimentary basins in Colombia, *in* 'Proceedings World Geothermal Congress 2020+1', Reykjavik, Iceland.

Platzman, L. K. and Bartholdi, J. J. (1989), 'Spacefilling curves and the planar travelling salesman problem', *Journal of the Association for Computing Machinery* **36**(4), 719–737.

Porkhial, S., Salehpour, M., Ashraf, H. and Jamali, A. (2015), 'Modeling and prediction of geothermal reservoir temperature behavior using evolutionary design of neural networks', *Geothermics* **53**, 320–327.

Pyrcz, M. and Deutsch, C. (2014), *Geostatistical reservoir modeling*, Oxford University Press, pp. 209 – 315.

Pyrcz, M. J., Catuneanu, O. and Deutsch, C. V. (2005), 'Stochastic surface-based modeling of turbidite lobes', *AAPG Bulletin* **89**(2), 177–191.

Pyrcz, M. J., Sech, R. P., Covault, J. A., Willis, B. J., Sylvester, Z. and Sun, T. (2015), 'Stratigraphic rule-based reservoir modeling', *Bulletin of Canadian Petroleum Geology* **63**(4), 287–303.

Quarteroni, A. and Rozza, G., eds (2014), *Reduced Order Methods for Modeling and Computational Reduction*, Springer Cham, pp. 235–236.

Quintero-Núñez, M., Cuaya-Simbro, M. E., Canales-Rodriguez, M. A., Garcia-Cueto, R., Santillán-Soto, N., Ojeda-Bentez, S. and Velázquez-Limón, N. (2015), A comparative study on the environmental impacts of a binational geothermal system: Imperial Valley, CA., U.S.A. and Cerro Prieto, Mexicali Valley, BC, México, *in* 'World Geothermal Congress 2015', Melbourne, Australia.

Radford, A. and Narasimhan, K. (2018), 'Improving language understanding by generative pre-training'. Semantic Scholar Corpus ID: 49313245.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. and Sutskever, I. (2019), 'Language models are unsupervised multitask learners'. Semantic scholar Corpus ID: 160025533.

Raissi, M., Perdikaris, P. and Karniadakis, G. (2019), 'Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations', *Journal of Computational Physics* **378**, 686–707.

Razak, S. M., Jahandideh, A., Djuraev, U. and Jafarpour, B. (2022), 'Deep Learning for Latent Space Data Assimilation in Subsurface Flow Systems', *SPE Journal* **27**(05), 2820–2840. SPE-203997-PA.

Reichle, R. H. (2008), 'Data assimilation methods in the Earth sciences', *Advances in Water Resources* **31**(11), 1411–1418.

Reis, L. C. (2006), Risk analysis with history matching using experimental design or artificial neural networks, *in* 'SPE Europec/EAGE Annual Conference and Exhibition, Vienna, Austria'. SPE-100255-MS.

Reiss, J., Schulze, P., Sesterhenn, J. and Mehrmann, V. (2018), 'The Shifted Proper Orthogonal Decomposition: A Mode Decomposition for Multiple Transport Phenomena', *SIAM Journal on Scientific Computing* **40**(3), A1322–A1344.

Rex, R. W. (1973), 'Geothermal resources in the Imperial Valley of California', *Bulletin Volcanologique* **37**(3), 461–462.

Robbins, H. and Monro, S. (1951), 'A stochastic approximation method', *The Annals of Mathematical Statistics* **22**(3), 400 – 407.

Rønning, K. and Steel, R. J. (1987), Depositional sequences within a "transgressive" reservoir sandstone unit: The middle Jurassic Tarbert formation, Hild area, northern North Sea, *in* 'Proceedings of the North Sea oil and gas reservoirs seminar', pp. 169–176.

Ruder, S. (2017), 'An overview of gradient descent optimization algorithms'. *ArXiv e-prints*. arXiv:1609.04747v2 [cs.LG].

Ruiu, J., Caumon, G. and Viseur, S. (2016), 'Modeling channel forms and related sedimentary objects using a boundary representation based on non-uniform rational b-splines', *Mathematical Geosciences* **48**(3), 259–284.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986*a*), Learning internal representations by error propagation, *in* 'Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations', The MIT Press, pp. 318–362.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986*b*), 'Learning representations by back-propagating errors', *Nature* **323**, 533–536.

Rybach, L. (1981), Geothermal systems, conductive heat flow, geothermal anomalies, *in* 'Geothermal Systems: Principles and case histories', John Wiley & Sons, pp. 3–31.

Rybach, L. (2010), The future of geothermal energy and its challenges, *in* 'Proceedings World Geothermal Congress 2010', Bali, Indonesia.

Rybach, L. (2015), 'Classification of geothermal resources by potential', *Geothermal Energy Science* **3**(1), 13–17.

Salehi, S., Vivas, C., Nygaard, R. and Rehg, D. (2022), 'Scalable geothermal energy potential from sedimentary basins: Opportunities for energy transition and skilled workforce development', *Geothermal Resources Council Transactions* **46**(1), 964–975.

Salinas, P., Pavlidis, D., Xie, Z., Osman, H., Pain, C. C. and Jackson, M. D. (2018), 'A discontinuous control volume finite element method for multi-phase flow in heterogeneous porous media', *Journal of Computational Physics* **352**, 602–614.

Salinas, P., Regnier, G., Jacquemyn, C., Pain, C. C. and Jackson, M. D. (2021), 'Dynamic mesh optimisation for geothermal reservoir modelling', *Geothermics* **94**, 102089.

Samnioti, A. and Gaganis, V. (2023), 'Applications of machine learning in subsurface reservoir simulation—A review—Part I', *Energies* **16**(16). 6079.

Samuel, J.-S. and Muggeridge, A. H. (2022), 'Fast modelling of gas reservoir performance with proper orthogonal decomposition based autoencoder and radial basis function non-intrusive reduced order models', *Journal of Petroleum Science and Engineering* **211**, 110011.

Sanyal, S. K. and Butler, S. J. (2010), Geothermal power capacity of wells in non-convective sedimentary formations, *in* 'Proceedings World Geothermal Congress 2010', Bali, Indonesia.

Scherer, J., Amaya, A., Chandrasekar, H., Chin, H., Golla, G., Manuel, F., Klenner, R. and Angcoy, E. (2022), Progress for closed-loop geothermal projects in steam and two-phase dominated reservoirs, *in* '3rd Philippine International Geothermal Conference, Philippines'.

Schiozer, D. J. (1999), Use of reservoir simulation, parallel computing and optimization techniques to accelerate history matching and reservoir management decisions, *in* 'SPE Latin America and Caribbean Petroleum Engineering Conference, , Caracas, Venezuela'. SPE-53979-MS.

Schlumberger (2015), *Wireline Services Catalog*, Schlumberger, p. 26.
URL: *https://www.scribd.com/document/347849565/2015-Wireline-Services-Catalog-Full-Permission*

Schubert, G. (1997), *Mantle convection*, Springer Netherlands, Dordrecht, pp. 416–421.

Sech, R. P., Jackson, M. D. and Hampson, G. J. (2009), 'Three-dimensional modeling of a shoreface-shelf parasequence reservoir analog: Part 1. surface-based modeling to capture high-resolution fades architecture', *AAPG Bulletin* **93**, 1155–1181.

Shahkarami, A., Mohaghegh, S. D. and Hajizadeh, Y. (2015), Assisted history matching using pattern recognition technology, *in* 'SPE Digital Energy Conference and Exhibition, The Woodlands, Texas, USA'. SPE-173405-MS.

Sharmin, T., Khan, N. R., Akram, M. S. and Ehsan, M. M. (2023), 'A state-of-the-art review on geothermal energy extraction, utilization, and improvement strategies: Conventional, hybridized, and enhanced geothermal systems', *International Journal of Thermofluids* **18**, 100323.

Shi, Y., Rop, E., Wang, Z., Jiang, G., Wang, S. and Hu, S. (2021), 'Characteristics and formation mechanism of the Olkaria geothermal system, Kenya revealed by well temperature data', *Geothermics* **97**, 102243.

Shi, Y., Song, X. and Song, G. (2021), 'Productivity prediction of a multilateral-well geothermal system based on a long short-term memory and multi-layer perceptron combinational neural network', *Applied Energy* **282**, 116046.

Shuai, Y., White, C. D., Sun, T. and Feng, Y. (2016), 'A gathered EnKF for continuous reservoir model updating', *Journal of Petroleum Science and Engineering* **139**, 205–218.

Sircar, A., Yadav, K., Rayavarapu, K., Bist, N. and Oza, H. (2021), 'Application of machine learning and artificial intelligence in oil and gas industry', *Petroleum Research* **6**(4), 379–391.

Skjervheim, J.-A. A., Evensen, G., Hove, J. and Vabø, J. G. (2011), An ensemble smoother for assisted history matching, *in* 'SPE Reservoir Simulation Conference, The Woodlands, Texas, USA'. SPE-141929-MS.

Skjervheim, J.-A., Evensen, G., Aanonsen, S., Ruud, B. and Johansen, T. (2007), 'Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter', *SPE Journal* **12**(3), 282–292. SPE-95789-PA.

Smith, S., Elsen, E. and De, S. (2020), On the generalization benefit of noise in stochastic gradient descent, *in* 'Proceedings of the 37th International Conference on Machine Learning', Vol. 119 of *Proceedings of Machine Learning Research*, PMLR, pp. 9058–9067.

Specht, D. (1991), 'A general regression neural network', *IEEE Transactions on Neural Networks* **2**(6), 568–576.

Srinivasan, S., Malley, D. O., Mudunuru, M. K., Sweeney, M. R., Hyman, J. D., Karra, S., Frash, L., Carey, J. W., Gross, M. R., Guthrie, G. D., Carr, T., Li, L. and S., H. (2021), 'A machine learning framework for rapid forecasting and history matching in unconventional reservoirs', *Scientific Reports* **11**(1), 21730.

Tang, H., Volkov, O., Tchelepi, H. A. and Durlofsky, L. J. (2021), Reduced-order modeling in a general reservoir simulation setting, *in* 'SPE Western Regional Meeting, Virtual'. SPE-200794-MS.

Tang, M., Liu, Y. and Durlofsky, L. J. (2020), 'A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems', *Journal of Computational Physics* **413**, 109456.

Tarantola, A. (1986), 'A strategy for nonlinear elastic inversion of seismic reflection data', *Geophysics* **51**(10), 1893–1903.

Temirchev, P., Simonov, M., Kostoev, R., Burnaev, E., Oseledets, I., Akhmetov, A., Margarit, A., Sitnikov, A. and Koroteev, D. (2020), 'Deep neural networks predicting oil movement in a development unit', *Journal of Petroleum Science and Engineering* **184**, 106513.

Titus, Z., Heaney, C. E., Jacquemyn, C., Salinas, P., Jackson, M. D. and Pain, C. C. (2022), 'Conditioning surface-based geological models to well data using artificial neural networks', *Computational Geosciences* **26**(4), 779–802.

Vargas, C. A., Caracciolo, L. and Ball, P. J. (2022), 'Geothermal energy as a means to decarbonize the energy mix of megacities', *Communications Earth & Environment* **3**(1).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I. (2017), Attention is all you need, *in* 'Advances in Neural Information Processing Systems', Vol. 30, Curran Associates, Inc.

Vivas, C., Salehi, S., Nygaard, R. and Rehg, D. (2023), Scalable geothermal energy potential from sedimentary basins in Oklahoma, *in* 'SPE Oklahoma City Oil and Gas Symposium'. SPE-213094-MS.

Vocadlo, L. (2015), 2.06 - Earth's core: Iron and iron alloys, *in* 'Treatise on Geophysics (Second Edition)', Elsevier, Oxford, pp. 117–147.

Wang, N., Chang, H. and Zhang, D. (2023), 'Inverse modeling for subsurface flow based on deep learning surrogates and active learning strategies', *Water Resources Research* **59**(7), e2022WR033644.

Wang, S., Yan, J., Li, F., Hu, J. and Li, K. (2016), 'Exploitation and utilization of oilfield geothermal resources in China', *Energies* **9**(10). 798.

Wang, Y. C., Pyrcz, M. J., Catuneanu, O. and Boisvert, J. B. (2018), 'Conditioning 3D object-based models to dense well data', *Computers and Geosciences.* **115**, 1–11.

Wang, Z., Di, H., Shafiq, M. A., Alaudah, Y. and AlRegib, G. (2018), 'Successful leveraging of image processing and machine learning in seismic structural interpretation: A review', *The Leading Edge* **37**(6), 451–461.

Watson, S. M., Falcone, G. and Westaway, R. (2020), 'Repurposing hydrocarbon wells for geothermal use in the UK: The onshore fields with the greatest potential', *Energies* **13**(14). 3541.

Williams, A. F., Lawless, J. V., Ward, M. A., Holgate, F. L. and Larking, A. (2010), A code for geothermal resources and reserves reporting, *in* 'Proceedings World Geothermal Congress 2010', Bali, Indonesia.

Williams, C. F., Reed, M. J. and Anderson, A. F. (2011), Updating the classification of geothermal resources, *in* 'Proceedings of the Thirty-Sixth Workshop on Geothermal Reservoir Engineering', Stanford University, Stanford, California. SGP-TR-191.

Wingate, D., Kane, J., Wolinsky, M. and Sylvester, Z. A. (2016), 'A new approach for conditioning process-based geological models to well data', *Mathematical Geosciences* **48**, 371–397.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. L., Gugger, S., Drame, M., Lhoest, Q. and Rush, A. M. (2020), 'HuggingFace's Transformers: State-of-the-art Natural Language Processing'. *ArXiv e-prints*. arXiv:1910.03771v5 [cs.CL].

Wu, J.-L., Xiao, H. and Paterson, E. (2018), 'Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework', *Physical Review Fluids* **3**(7), 074602.

Xiao, D., Fang, F., Buchan, A. G., Pain, C. C., Navon, I. M., Du, J. and Hu, G. (2014), 'Non-linear model reduction for the navier-stokes equations using residual DEIM method', *Journal of Computational Physics* **263**, 1–18.

Xiao, D., Fang, F., Pain, C. C. and Navon, I. M. (2017), 'A parameterized non-intrusive reduced order model and error analysis for general time-dependent nonlinear partial differential equations and its applications', *Computer Methods in Applied Mechanics and Engineering* **317**, 868–889.

Xu, J. and Duraisamy, K. (2020), 'Multi-level convolutional autoencoder networks for parametric prediction of spatio-temporal dynamics', *Computer Methods in Applied Mechanics and Engineering* **372**, 113379.

Xu, P., Kumar, D., Yang, W., Zi, W., Tang, K., Huang, C., Cheung, J. C. K., Prince, S. J. and Cao, Y. (2021), Optimizing deeper transformers on small datasets, *in* 'Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)', Association for Computational Linguistics, Online, pp. 2089–2102.

Yan, B., Harp, D. R., Chen, B. and Pawar, R. (2022), 'A physics-constrained deep learning model for simulating multiphase flow in 3D heterogeneous porous media', *Fuel* **313**, 122693.

Yang, Y., Robinson, C., Heitz, D. and Mémin, E. (2015), 'Enhanced ensemble-based 4DVar scheme for data assimilation', *Computers and Fluids* **115**, 201–210.

Yin, Z., Feng, T. and Macbeth, C. (2019), 'Fast assimilation of frequently acquired 4D seismic data for reservoir history matching', *Computers and Geosciences* **128**, 30–40.

Yoon, H., Lee, J. and Kadeethum, T. (2022), Deep learning-based data assimilation in the latent space for real-time forecasting of geologic carbon storage, *in* 'Proceedings of the 16th Greenhouse Gas Control Technologies Conference (GHGT-16)', Lyon, France.

Yu, J. (2021), 'SFC-CAE-Ready-to-use', *GitHub Repository*. https://github.com/acse-jy220/SFC-CAE-Ready-to-use [Commit no. da29d4b].

Yu, J., Yan, C. and Guo, M. (2019), 'Non-intrusive reduced-order modeling for fluid problems: A brief review', *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* **233**(16), 5896–5912.

Yu, S. and Ma, J. (2021), 'Deep learning for geophysics: Current and future trends', *Reviews of Geophysics* **59**(3), e2021RG000742.

Zafari, M. and Reynolds, A. C. (2007), 'Assessing the uncertainty in reservoir description and performance predictions with the ensemble Kalman filter', *SPE Journal* **12**(03), 382–391. SPE-95750-PA.

Zangl, G., Giovannoli, M. and Stundner, M. (2006), Application of artificial intelligence in gas storage management, *in* 'SPE Europec/EAGE Annual Conference and Exhibition, Vienna, Austria', Vienna, Austria. SPE-100133-MS.

Zarrouk, S. J. and McLean, K. (2019), Chapter 2 - Geothermal systems, *in* 'Geothermal Well Test Analysis', Academic Press, pp. 13–38.

Zhang, D., Chen, Y. and Meng, J. (2018), 'Synthetic well logs generation via recurrent neural networks', *Petroleum Exploration and Development* **45**(4), 629–639.

Zhang, F., Nghiem, L. and Chen, Z. (2023), 'Evaluating reservoir performance using a transformer based proxy model', *Geoenergy Science and Engineering* **226**, 211644.

Zhang, K., Zhang, J., Ma, X., Yao, C., Zhang, L., Yang, Y., Wang, J., Yao, J. and Zhao, H. (2021), 'History matching of naturally fractured reservoirs using a deep sparse autoencoder', *SPE Journal* **26**(04), 1700–1721. SPE-205340-PA.

Zhang, T., Tilke, P., Dupont, E., Zhu, L., Liang, L. and Bailey, W. (2019), 'Generating geologically realistic 3D reservoir facies models using deep learning of sedimentary architecture with generative adversarial networks', *Petroleum Science* **16**, 541–549.

Zhang, X., Pyrcz, M. J. and Deutsch, C. V. (2009), 'Stochastic surface modeling of deepwater depositional systems for improved reservoir models', *Journal of Petroleum Science and Engineering* **68**(1-2), 118–134.

Zhang, Z. (2022), 'A physics-informed deep convolutional neural network for simulating and predicting transient darcy flows in heterogeneous reservoirs without labeled data', *Journal of Petroleum Science and Engineering* **211**, 110179.

Zhong, Z., Sun, A. Y., Wang, Y. and Ren, B. (2020), 'Predicting field production rates for waterflooding using a machine learning-based proxy model', *Journal of Petroleum Science and Engineering* **194**, 107574.

Zhou, C., Ouyang, J., Ming, W., Zhang, G., Du, Z. and Liu, Z. (2019), 'A stratigraphic prediction method based on machine learning', *Applied Sciences* **9**(17). 3553.

Zhou, L., Johnson, C. R. and Weiskopf, D. (2021), 'Data-driven space-filling curves', *IEEE Transactions on Visualization and Computer Graphics* **27**(2), 1591–1600.