Hierarchical Geological Realism in History Matching for Reliable Reservoir Uncertainty Predictions

Alexandra Anatolievna Kuznetsova

Submitted for the
Degree of Doctor of Philosophy
Institute of Petroleum Engineering
Heriot-Watt University
September 2017

The copyright in this thesis is owned by the author. Any quotation from the thesis or use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information.
ABSTRACT

The oil and gas industry has been always associated with huge risks. To minimise these risks, one is looking for the reliable reservoir performance predictions to make better field development decisions. The great challenge associated with reliable predictions is to account for the essential geological uncertainties and propagate them through the engineering model validation process.

In this thesis, we propose a new methodology to improve the reliability of reservoir predictions under the Bayesian framework. The first step of the methodology applies the new hierarchical approach to account for essential geological uncertainties from different levels of geological data in facies modelling. As the result of the hierarchical approach, we evaluate the prior range of different geological uncertainties.

Facies models greatly affect simulation results but it’s a great challenge to history match them whilst maintaining geological realism. Therefore, next step of the methodology is aiming to improve geological realism during history matching. We propose to combine metric space approach and machine learning classification to evaluate geological relations between multiple geological scenarios and parameters combination and propagate them into history matching. Multidimensional scaling was used to analyse the similarity of the facies models in the metric space. Results of different machine learning classification methods – k-means clustering, Support Vector Machines, Random Forest – were compared to include the ones that performed better into history matching.

The reservoir predictions under uncertainty were performed by evaluating the Posterior Probability Distribution under the Bayesian framework and estimating the Credible Intervals (P10, P50, P90).

The methodology was applied to a synthetic case study based on a real reservoir of the West Coast of Africa (offshore turbidite reservoir). The main results show that the proposed methodology was able to improve the geological realistic facies model representation during history matching and uncertainty quantification.

Some additional controls of facies architecture and facies connectivity modelling could be introduced to improve the quality of the facies realisations.
DEDICATION

To my family, my fiancé, and my supervisors
ACKNOWLEDGMENTS

I would like to thank Dr Vasily Demyanov for inviting me to the Uncertainty group and opening a whole new world for me, for his guidance and great research ideas during all these years.

I would like to express my deepest gratitude to the best security guard ever, Prof Mike Christie, for all his wise and useful advice, for his help in shaping not only my skills (research and communication) but also my personality, for all the support during the “dark times”. Thank you very much!

I would like to thank Dr Dan Arnold for his enthusiasm and passion that inspired me a lot.

I would like to thank Jack Talbot for his creative advice, technical support, and exclusive everyday magic.

I would like to thank Prof Mikhail Kanevski from University of Lausanne for his recommendations on the use of machine learning, Dr. Alexandre Boucher from Stanford University for providing the latest version of SGeMS software, and a big thanks to Marco Thiele from Streamsim Technologies for VIP access to 3DSL simulator.

I would like to thank the examiners of this thesis: Dr Pauline Collon from Université de Lorraine and Dr Karl Stephen from Heriot-Watt University, for their recommendations that substantially improved the thesis.

I would like to thank Iain Davidson for being inspiring, supportive and guiding me through efficient communication.

I would like to thank by the sponsors of the Uncertainty Project (Phase IV) at Heriot-Watt: BG Group, E.ON, JOGMEC and RFD for the financial support provided for this research.

I would like to thank my colleagues Doaa, Junko, Behzad, Ilya, Shing Cheng and Nagoor.
I would like to thank the IPE IT team: Chris, Alan, Chis and Andy for always being fantastic.

Last but not least I would like to thank my fiancé, his family, my family, and our cat for being there for me all the time. I could not have done it without you all.
PUBLICATIONS


### ACADEMIC REGISTRY
### Research Thesis Submission

<table>
<thead>
<tr>
<th>Name:</th>
<th>ALEXANDRA ANATOLIEVNA KUZNETSOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>School/PGI:</td>
<td>EGIS</td>
</tr>
<tr>
<td>Version: (i.e. First, Resubmission, Final)</td>
<td>Final</td>
</tr>
<tr>
<td>Degree Sought (Award and Subject area)</td>
<td>PhD Petroleum Engineering</td>
</tr>
</tbody>
</table>

### Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself
2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.
4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

* Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.

<table>
<thead>
<tr>
<th>Signature of Candidate:</th>
<th>Date:</th>
</tr>
</thead>
</table>

### Submission

<table>
<thead>
<tr>
<th>Submitted By (name in capitals):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature of Individual Submitting:</td>
</tr>
<tr>
<td>Date Submitted:</td>
</tr>
</tbody>
</table>

### For Completion in the Student Service Centre (SSC)

<table>
<thead>
<tr>
<th>Received in the SSC by (name in capitals):</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Method of Submission (Handed in to SSC; posted through internal/external mail):</td>
</tr>
<tr>
<td>1.2 E-thesis Submitted (mandatory for final theses)</td>
</tr>
<tr>
<td>Signature:</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

Contents
LISTS OF FIGURES ........................................................................................................... iv
LISTS OF TABLES ............................................................................................................... xiv
1 CHAPTER ONE. Introduction .......................................................................................... 18
  1.1 Thesis Motivation ........................................................................................................ 18
  1.2 Thesis objectives ......................................................................................................... 21
  1.3 Thesis Overview ......................................................................................................... 25
2 CHAPTER TWO. Geological Reservoir Modelling and Uncertainty Quantification .... 27
  2.1 Introduction ................................................................................................................ 27
  2.2 Static reservoir modelling ......................................................................................... 28
  2.3 Dynamic reservoir modelling .................................................................................... 35
  2.4 Bayesian Approach to Model Uncertainty ............................................................... 40
  2.5 History matching approaches ................................................................................... 42
  2.6 Similarity distance and metric space ........................................................................ 51
3 CHAPTER THREE. The Impact of Geological Parametrisation and Interpretation on History Matching and Uncertainty Quantification ........................................... 55
  3.1 Introduction ................................................................................................................. 55
  3.2 Ainsa II Case Study ................................................................................................... 57
    3.2.1 Outcrop interpretation .......................................................................................... 57
    3.2.2 Simulation models for history matching ............................................................. 60
  3.3 History Matching ....................................................................................................... 62
    3.3.1 Case study parametrisation ............................................................................... 62
    3.3.2 History matching algorithm .............................................................................. 62
    3.3.3 History matching results .................................................................................... 66
    3.3.4 Geological realism during history matching ...................................................... 69
  3.4 Uncertainty Quantification in Reservoir Predictions ................................................ 71
    3.4.1 Posterior probability distribution ....................................................................... 71
    3.4.2 NA-Bayes algorithm .......................................................................................... 73
    3.4.3 NA-Bayes prediction results ............................................................................. 76
  3.5 Summary .................................................................................................................... 78
4 CHAPTER FOUR Hierarchical Geological Realism in Multiple Training Images
  History Matching and Uncertainty Quantification ........................................................ 80
CHAPTER FIVE Metric Space Approach and Machine Learning Classification

5.1 Introduction ................................................................. 132
      5.1.1 Metric space approach ........................................... 132
      5.1.2 Machine learning metric space classification .............. 134
5.2 Multidimensional scaling ............................................. 135
      5.2.1 Principle and method used ....................................... 135
      5.2.2 Validation of the number of dimensions: the scree plot .... 137
      5.2.3 Multidimensional scaling results .............................. 138
      5.2.3 Identifying key number of dimensions ...................... 144
5.3 Machine Learning – Learning from Data .......................... 145
      5.3.1 Motivation and challenges ..................................... 145
      5.3.2 Types of learning process ..................................... 146
# Contents

5.4 Solving Classification Problem with Machine Learning ........................................... 152  
    5.4.1 k-means clustering ......................................................................................... 153  
    5.4.2 Support Vector Machines ............................................................................... 155  
    5.4.3 Random Forest .............................................................................................. 167  
5.5 Assessing Multi-class Classification Performance Quality ....................................... 173  
5.6 Geological Metric Space Classification Results ..................................................... 176  
    5.6.1 k-means clustering ......................................................................................... 176  
    5.6.2 Support Vector Machines ............................................................................... 179  
    5.6.3 Random Forest .............................................................................................. 184  
    5.6.4 Classification results comparison between different algorithms ....................... 199  
5.7 Conclusions ......................................................................................................... 200  
6 CHAPTER SIX Hierarchical Geological Realism and Classification in History  
    Matching and Uncertainty Quantification ............................................................... 202  
    6.1 Introduction ...................................................................................................... 202  
    6.2 Assisted History Matching (AHM) ................................................................. 202  
        6.2.1 AHM set up ............................................................................................... 202  
        6.2.2 Results of history matching ..................................................................... 203  
    6.3 Uncertainty Quantification in Reservoir Prediction ........................................... 216  
    6.4 Geological Realism in Predictions ..................................................................... 224  
    6.5 Metric Space and Multidimensional Scaling Comparison ................................... 227  
    6.6 Discussion ....................................................................................................... 231  
    6.7 Conclusions ...................................................................................................... 237  
7 CHAPTER SEVEN Conclusions and Future Work Suggestions .................................... 239  
    7.1 Overview .......................................................................................................... 239  
    7.2 Thesis Conclusions ............................................................................................ 239  
    7.3 Future Work Suggestions .................................................................................. 242
LISTS OF FIGURES

Figure 1-1 A toy example of underestimated forecast uncertainty (From Vink et al., 2015) .................................................................................................................. 20

Figure 1-2 Geologically unrealistic rectangular shapes around the wells resulted by applying porosity multipliers during history matching (From Valjak, 2008) .............. 20

Figure 1-3 Schematic view of the high-level objectives of the thesis......................... 22

Figure 1-4 Schematic representation of the hierarchical approach to account for the geological uncertainties.............................................................................................. 23

Figure 1-5 Schematic view on the main difference between two proposed techniques to include geological uncertainties into history matching................................. 23

Figure 1-6 Schematic representation of the first workflow proposed to include geological uncertainties from hierarchical approach into history matching process............. 24

Figure 1-7 Schematic view of the proposed metric space classification step to differentiate between different geological scenarios in metric space........................................ 25

Figure 1-8 Schematic view of the first workflow proposed to include geological uncertainties from hierarchical approach into history matching process................. 25

Figure 1-9 Schematic view of the final step of the proposed methodology to perform reservoir forecast under uncertainty................................................................. 25

Figure 2-1 Process-based facies modelling input parameters in a cross-section of a point bar. Lateral accretion surfaces and coarse channel lag deposits are indicated in the pointbar. (2) Map view of facies in the depositional environment (From Willems et al., 2017) ........................................................................................................ 29

Figure 2-2 An example illustrating the Sequential Indicator Simulation using a variogram (Deutsch and Journel, 1998) and Multiple-point simulation, performed with the method IMPALA (Straubhaar et al., 2011) using a training image. The simulated values come from a prior model based on the known cell configuration in the neighbourhood of the cell to simulate (From Rongier et al., 2016). ........................................................................ 32

Figure 2-3 Example of MPS SNESIM algorithm scanning a training image to obtain a conditional (modified from Caers, 2005)........................................................................................................ 34

Figure 2-4 3D 3 phase flow simulation of a volatile oil system with gas and water Injectors (50x30x7 Reservoir Grid System) (From the (“Marathon Center of Excellence for Reservoir Studies, Showcase,” 2014) .................................................................................. 36

Figure 2-5 Reservoir connectivity described by streamline-trajectories between injector and producers before history matching (coloured by injector-producer pair). (From Singh et al., 2014)................................................................................................................... 39

Figure 2-6 Gradient-based optimisation combined with tunnelling method. Local minimization at each valley followed by a tunnelling to ignore the local minima (from Gomez et al., 1999 ) ........................................................................................................ 46
List of figures

Figure 2-7 Conceptual illustration of similarity distance in metric space between model realisations .......................................................... 52

Figure 2-8 Classification results of metric space: each class represents a different geological scenario (training image). Red square represents the region with low misfit models identified during history matching (from Rojas, 2013) ... 53

Figure 3-1 Comparison of Ainsa II outcrop photo, high-resolution model, and facies interpretation models produced by Team A, Team B, and Team C. Each 2D model has an injector (I) and a producer (P) well at each side of the model (Courtesy of A. Gardiner, Heriot-Watt University) .................................................. 58

Figure 3-2 Visual comparison between fluvial meandering and deep-marine channels. A: Meandering channel, Chubut River, Argentina (from Foix et al., 2012); B: Deep-marine channel, De Soto Canyon, Gulf of Mexico (from Posamentier and Kolla, 2003) ........ 59

Figure 3-3 Visual comparison between the fluvial channel and lacustrine deposits. Left: Ancient Fluvial channel deposit in the Stellarton Formation (Pennsylvanian), Coalburn Pit, near Thorburn, Nova Scotia (from Rygel, 2001); Right: Lacustrine deposits of the Eocene Green River Formation, Wyoming (from Suthren, 2004) ......................... 60

Figure 3-4 Velocity component construction (From Mohamed et al., 2010b) ........ 64

Figure 3-5 Example of porosity variation for each facies during history matching (Team A). Other parameters varied are not shown .................................................................................. 67

Figure 3-6 Example of porosity variation for each facies during history matching (Team B). Other parameters varied are not shown .................................................................................. 68

Figure 3-7 Example of porosity variation for each facies during history matching (Team C). Other parameters varied are not shown .................................................................................. 68

Figure 3-8 Results of automated history matching (five best models, models are very similar, therefore only one model is visible) for oil and water production rates of three different model interpretations from Team A, Team B, and Team C. These models have no water production and have pressure within the limits, therefore were further discriminated because of a slight pressure effect on the surface rates ....................... 70

Figure 3-9 (left) Linear correlation between porosity and permeability values before history matching; (right) No linear correlation between porosity and permeability values after history matching ................................................................. 71

Figure 3-10 Schematic view of the Bayesian framework used in this thesis to perform reservoir predictions under uncertainty (From Erbas, 2007) ......................................................... 74

Figure 3-11 Example of Gibbs sampler in NAB (Modified from Sambridge, 1999b) ... 76

Figure 3-12 NAB workflow for predicting uncertainty in reservoir forecast (Modified from Erbas, 2007) .................................................................................................................. 76

Figure 3-13 Forecast results (probability of exceedance) for the oil and water production rates for all three models (Team A, Team B, and Team C) ........................................................................ 78
List of figures

Figure 4-1 Examples of three turbidite categories based on the type of the feeder (Modified from (Stow and Mayall, 2000)) ................................................................. 84

Figure 4-2 Schematic view of a slope mini-basin and mud rich fine-grained submarine fans depicting main turbidite structural elements: channel, leveed channels, lobes and sheet lobes (modified from DeVay et al., 2000) ...................................................... 85

Figure 4-3 Conceptual model showing a turbidity current within a channel (from Shanmugam et al., 1993) ...................................................................................................................... 85

Figure 4-4 Schematic description of geological zones of slope-valley turbidite. This scheme also describes the location of the WCA (from Hoffman et al., 2006) .......................... 86

Figure 4-5 Interpreted seismic map of the top layer of the WCA reservoir (from Hoffman et al., 2006) ...................................................................................................................... 87

Figure 4-6 Four structural blocks of the WCA reservoir, separated by a number of faults (from Park et al., 2013) ......................................................................................................... 87

Figure 4-7 Geological scenarios of West Coast of Africa reservoir generated by geologists (courtesy of Stanford University and Chevron). Each colour represents different facies: dark red (brown) – channel sand (Facies 1), light blue and yellow - poorer quality sand (Facies 2 and Facies 3), dark blue – shale (Facies 4) ........................................................................ 90

Figure 4-8 Schematic representation of the hierarchical approach to account for the geological uncertainties in facies modelling for West Coast of Africa ...................... 92

Figure 4-9 An example of unconditional SNESIM facies realisation (right) based on the training image (left) (From Arpat, 2005) ................................................................. 94

Figure 4-10 WCA well data color-coded by facies and used for hard data conditioning with SNESIM (courtesy of Stanford University) ........................................................................ 96

Figure 4-11 WCA facies probability cubes for each facies used for soft data conditioning with SNESIM (courtesy of Stanford University) ......................................................... 96

Figure 4-12 Example of multiple different model realisations based on one training image: top left – training image, top right – new realisation with global rotation angle = 45°, bottom left - new realisation with global affinity x = 1 and y = 2, bottom right - new realisation with global affinity x = 1, y = 3 and global rotation angle = 30° (modified from Caers and Zhang, 2004) .................................................. 97

Figure 4-13 Examples of facies model realisation with SNESIM for 11 training images: the first realisation has no affinity or rotation parameters applied, and the second realisation has different values of affinity ................................................................. 102

Figure 4-14 Examples of geologically unrealistic visual facies representations created with MPS SNESIM .................................................................................................................. 103

Figure 4-15 Relative permeability for oil/water and oil/gas behaviour (from Hoffman et al., 2006) .................................................................................................................. 104

Figure 4-16 2D geological model of the Truth Case (top view), built from training image TI 9 with the following parameters: affx=1, affy=0.75, affz=1.25, rot=0°, facies
proportions: shale=58% (blue), poorer quality sand 1=11% (light blue), poorer quality sand 2=9% (yellow), channel sand=28% (dark red/brown), conditioned to wells and seismic data

Figure 4-17 Porosity and permeability of the Truth case model, 2D top view

Figure 4-18 Truth case fluid production profile

Figure 4-19 Saturation of water of the Truth case model at the end of the simulation

Figure 4-20 Schematic view of decision variable and objective function spaces for a two-objective optimisation problem with three decision variables

Figure 4-21 An example of two-objective optimisation with six solutions illustrating the Pareto dominance concept (modified from Hajizadeh et al., 2011)

Figure 4-22 Schematic view of a Pareto front for a two-objective optimisation

Figure 4-23 Pareto fronts for two misfit components of a multi-objective history matching with PSO for WCA case study

Figure 4-24 Represents the selected training images during history matching and the corresponding value of misfit (<100) for one of five history matching runs. The points are coloured based on iteration number of history matching process

Figure 4-25 Geological modelling parameters versus misfit. Models from different training images are represented by a different colour. Affx, affy, affz – affinity x, affinity y, affinity z, Rot – rotation, P1, P2, P3, P4 – proportions of Facies 1, 2, 3, and 4 respectively

Figure 4-26 Minimum misfit evolution during history matching

Figure 4-27 Five best models obtained during five history matching runs for Field Oil Production Rate (FOPR), STB/day

Figure 4-28 Five best models obtained during five history matching runs for Field Gas Production Rate (FGPR), Mscf/day

Figure 4-29 Five best models obtained during five history matching runs for Field Water Production Rate (FWPR), STB/day

Figure 4-30 Oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) total production prediction uncertainty intervals for five history matching runs

Figure 4-31 Oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) production prediction uncertainty intervals average over five history matching runs

Figure 4-32 Comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty (after NAB, average over five runs) of oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) production total
List of figures

Figure 4-33 The average (over 5 runs) Posterior Probability Distribution (PPD) of training images (TI) selected for the predictions from history matching results ........................................... 127

Figure 4-34 Comparison of the Truth case facies model and the high probability facies models, selected with NA-Bayes for reservoir prediction ................................................................. 129

Figure 5-1 The Scree plot, representing the relative eigenvalue of each dimension. The recommended number of dimensions is three on this example (from Wickelmaier, 2003) .................................................................................................................. 138

Figure 5-2 Schematic representation of calculating the dissimilarities between facies models and mapping them to the N-dimensional Euclidian space using the classic MDS approach. Only four models are shown as an example. Each colour in the Euclidian space (2D projection of the multidimensional Euclidian space) represents a different training image that was used for the simulations ....................................................................................................... 139

Figure 5-3 2D representation (dimension 1 and 2) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model's symbols). The truth case model is also represented in the metric space ........................................................................................................................................... 141

Figure 5-4 2D representation (dimension 2 and 3) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model's symbols). The truth case model is also represented in the metric space ........................................................................................................................................... 141

Figure 5-5 2D representation (dimension 3 and 4) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model's symbols). The truth case model is also represented in the metric space ........................................................................................................................................... 142

Figure 5-6 2D representation (dimension 4 and 5) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model's symbols). The truth case model is also represented in the metric space ........................................................................................................................................... 142

Figure 5-7 Comparison of production response (Field oil production total (FOPT), Field gas production total (FGPT), Field water production total (FWPT)) of three models: two of them (model 425 and 426) are close in the metric space and one (model 361) is further away in the metric space ........................................................................................................................................... 143

Figure 5-8 The Scree plot identifies the optimal number of dimensions to describe the data of MDS ........................................................................................................................................... 145

Figure 5-9 Facies classification performed from well-log data analysis, using unsupervised artificial neural networks (ANN): (a) logs data (Resistivity, Neutron-Porosity, and Density) cluster in different classes, (b) ANN classified the data and identified sedimentary facies (From Saggaf and Nebrija, 2000) ........................................................................................................... 147
List of figures

Figure 5-10 Example of classification (a) and regression (b) problem solved with supervised learning: (a) the classification algorithm identifies three classes and separates them in the x1-x2 space (From Kanevski et al., 2009), (b) the regression problem is solved by drawing a line through a cloud of points (From Kanevski et al., 2009) ........................ 148

Figure 5-11 Predicting oolitic facies from core data to uncored wells based on image logs using supervised neural networks: (A) Neural networks were trained using data from cored wells; (B) Predicted oolitic facies in uncored wells (from Qi et al., 2007) ............ 150

Figure 5-12 Example of unlabelled data effect on the classification problem results: (a) Classification problem solved without unlabelled data and (b) with unlabelled data. Unlabelled data add the information about the geometry or structure of the classification boundary between classes (From Belkin et al., 2006) .................................................................. 151

Figure 5-13 Porosity distribution maps created by using semi-supervised Support Vector Regression. Unlabelled data points are obtained from seismic data. Labelled data come from well data. Unlabelled data points give the structural information of the geobody (so-called geo-manifold); this geo-manifold is populated with data obtained from wells (From Demyanov et al., 2008) ........................................................................................................ 151

Figure 5-14 A simple sketch to explain the classification of metric space introduced into history matching: 1) we plot geological model realisations from different geological scenarios to a metric space, 2) we apply a classification technique to obtain the boundaries between different geological scenarios, 3) new point represents a new model produced during history matching. Classification result allows us to assign a new model to an appropriate geological scenario ........................................................................................................... 152

Figure 5-15 The main steps of k-means clustering classification .................................................................................................................. 154

Figure 5-16 Illustrates the choice of optimal separating hyperplane (green line) for two classes’ classification example .................................................................................................................. 156

Figure 5-17 SVM example with 2 samples sets: red and green. Bigger samples represent support vectors (SV), decision surface is maximally far away from any data point representing optimal separating hyperplane principle, one green sample is on the other side of separating hyperplane, grouped together with red points, this misclassified sample illustrates the principle of soft margin classifier, which allows us to misclassify a few training data to perform good classification results for the remaining data ................................................................. 158

Figure 5-18 Mapping of non-linear separable training data from $\mathbb{R}^2$ into $\mathbb{R}^3$ .................................................. 159

Figure 5-19 Left: Two class data set example (blue and pink points). Right: Support Vector classifier liner boundary results in poor classification on nonlinear data set (from James et al., 2014) .................................................................................................................. 160

Figure 5-20 Two class data set example (blue and pink points). Left: An example of SVM classification with a polynomial kernel of degree 3 applied to the nonlinear data set. Right: An example of SVM with a radial kernel to the same nonlinear data set. Both kernels show the capability of capturing the appropriate decision boundaries (From James et al., 2014) .................................................................................................................. 161

Figure 5-21 SVM error surface. Left: Training error surface. The error increases with kernel bandwidth for a fixed value of C. For a fixed value of kernel bandwidth, the training error decreases with C; Right: Testing error surface. Very low values of $\sigma$ result
List of figures

in overfitting, which is described by zero training and high testing error. The region of optimal \( \sigma \) and \( C \) combinations is evident (From Kanevski et al., 2009).........................164

Figure 5-22 Prediction mapping of porosity data using Gaussian RBF kernels. From left to right: the value of bandwidth decreases, the complexity of the decision boundaries increasing leading from oversmoothing to overfitting (From Kanevski et al., 2009)...165

Figure 5-23 Prediction mapping of porosity data using polynomial kernels of increasing order. From left to right: the value of bandwidth decreases, the complexity of the decision boundaries increases (from Kanevski et al., 2009)...................................................165

Figure 5-24 Prediction maps for increasing value of parameter \( C \) from left to right: the decision boundaries are too smooth and do not reflect the training data pattern (on the left map), they are close to optimal (on the central map), and they are too complex (on the right map) (From Kanevski et al., 2009)........................................................................165

Figure 5-25 A schematic example of a decision tree (Analytics Vidhya Content Team, 2016) .................................................................169

Figure 5-26 A schematic example of Random forest; each random subset corresponds to a form of Bootstrapping from which a different tree is formed..............................................170

Figure 5-27 Precision and recall schematic representation for a two-class example (From Skymind, 2017).............................................................................................................175

Figure 5-28 SVM training error surface (left) and testing error surface (right) for 12 multipliers on the eigenfunction only input data. The optimal combination of parameters is marked with a cross ..................................................................................180

Figure 5-29 SVM train error surface for 12 multipliers on the eigenfunction and 4 SNESIM geological modelling parameters (affinity and rotation) input data. The optimal combination of parameters is marked with a cross .................................................................182

Figure 5-30 Error rate against number of trees (12 multipliers on the eigenfunction input data)........................................................................................................................................................................................185

Figure 5-31 Error rate against number of trees for 12 multipliers on the eigenfunction and geological parameters input data........................................................................................................191

Figure 5-32 OOB error value for different values of mtry with the four fixed Ntree values: 500, 700, 1000, 1500.............................................................................................................193

Figure 5-33 Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=600, mtry=2)................................................................................................................................197

Figure 5-34 Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=800, mtry=8)................................................................................................................................197

Figure 5-35 Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=1000, mtry=4). ..........................................................198

x
List of figures

Figure 5-36 Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=1300, mtry=4).................................................................198

Figure 5-37 Comparison of the diagonal of confusion matrices for both input data Cases (Case 1 – multipliers on the eigenfunction only, Case 2 – multipliers and geological parameters) for k-means clustering, SVM and RF (Case 1 – Ntree =1500, mtry = 6, Case 2 – Ntree = 1300, mtry = 4) classification ...........................................................................200

Figure 6-1 Pareto fronts for two misfit components of a multi-objective history matching with SVM classification for WCA case study: (left) with SVM classification, (right) with RF classification.................................................................206

Figure 6-2 Represents the selected training images during history matching with RF and the corresponding value of misfit (<100) for one of five history matching runs. The points are coloured based on iteration number of history matching process .........................207

Figure 6-3 2D representation of SVM classification results for 11 training images: (top left) dimensions 1 and 2, (top right) dimensions 2 and 3, (bottom left) dimensions 3 and 4, (bottom right) dimensions 4 and 5. The plots highlight the predominance of training image TI 2 area in the metric space ..........................................................207

Figure 6-4 Geological modelling parameters versus misfit from history matching with SVM. Models from different training images are represented by a different colour. Affx, affy, affz – affinity x, affinity y, affinity z, Rot - rotation, P1, P2, P3, P4 – proportions of Facies 1,2,3, and 4 respectively .................................................................208

Figure 6-5 Geological modelling parameters versus misfit from history matching with RF. Models from different training images are represented by a different colour. Affx, affy, affz – affinity x, affinity y, affinity z, Rot - rotation, P1, P2, P3, P4 – proportions of Facies 1,2,3, and 4 respectively .................................................................209

Figure 6-6 Minimum misfit evolution during history matching with SVM and RF.....210

Figure 6-7 Five best models obtained during five history matching runs with SVM for Field Oil Production Rate (FOPR), STB/day.................................................................211

Figure 6-8 Five best models obtained during five history matching runs with SVM for Field Gas Production Rate (FGPR), Mscf/day.................................................................212

Figure 6-9 Five best models obtained during five history matching runs with SVM for Field Water Production Rate (FWPR), STB/day.................................................................213

Figure 6-10 Five best models obtained during five history matching runs with RF for Field Oil Production Rate (FOPR), STB/day.................................................................214

Figure 6-11 Five best models obtained during five history matching runs with RF for Field Gas Production Rate (FGPR), Mscf/day.................................................................215

Figure 6-12 Five best models obtained during five history matching runs with RF for Field Water Production Rate (FWPR), STB/day.................................................................216
List of figures

Figure 6-13 Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) total production prediction uncertainty intervals for five history matching runs with SVM ............................................................... 218

Figure 6-14 Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) total production prediction uncertainty intervals for five history matching runs with RF ............................................................... 219

Figure 6-15 Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production prediction uncertainty intervals average over five history matching runs with SVM ............................................... 220

Figure 6-16 Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) total production prediction uncertainty intervals average over five history matching runs with RF ........................................ 221

Figure 6-17 Comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty with SVM (after NAB, average over five runs) of Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production total .................................................. 222

Figure 6-18 Comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty with RF (after NAB, average over five runs) of Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production total .................................................. 223

Figure 6-19 The average (over 5 runs) Posterior Probability Distribution (PPD) of training images selected for the predictions from history matching results: (left) with SVM, (right) with RF classification ................................................................. 224

Figure 6-20 Comparison of the Truth case facies model and the high probability facies models, selected with NA-Bayes for reservoir prediction: (left) with SVM, (right) with RF ................................................................. 227

Figure 6-21 2D representation (dimension 1 and 2) of MDS for model realisations before and after history matching with SVM (all models) .................................................. 229

Figure 6-22 2D representation (dimension 1 and 2) of MDS for model realisations before and after history matching with RF (all models) .................................................. 229

Figure 6-23 2D representation (dimension 1 and 2) of MDS for model realisations before history matching and after NAB with SVM .................................................. 230

Figure 6-24 2D representation (dimension 1 and 2) of MDS for model realisations before history matching and after NAB with RF .................................................. 230

Figure 6-25 Comparison of Pareto fronts for two misfit components of a multi-objective history matching with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification .................................................. 232

Figure 6-26 Comparison of minimum misfit evolution during history matching with SVM, RF and without classification (average over 5 runs for each case) ............... 233
List of figures

Figure 6-27 Comparison of the Oil (FOPR) production rates prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification ................................................................. 234

Figure 6-28 Comparison of the Water (FWPR) production rates prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification ................................................................. 235

Figure 6-29 Comparison of the Oil (FOPT) total production prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification ................................................................. 235

Figure 6-30 Comparison of the Water (FWPT) total production prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification ................................................................. 236

Figure 6-31 Comparison of the Truth case facies model and the high probability facies models, selected with NA-Bayes for reservoir prediction: (top right) without classification, (bottom left) with SVM, and (bottom right) with RF ............................................. 237
LISTS OF TABLES

Table 2-1 Comparison between two-point and Multi-Point Statistics characteristics ....31

Table 3-1 Reservoir properties used for the flow simulation.............................................61

Table 3-2 Parameter values for the parameterization applied to the Ainsa II facies models during history matching .................................................................62

Table 3-3 Summary of parameters used for Particle Swarm Optimisation algorithm (Modified from Rojas, 2013). (*flexible PSO provides a more freely move of particles in the search space (Kathrada, 2009))v ................................................................66

Table 3-4 NAB set up for Ainsa II forecasting for Team A, Team B, and Team C models ..................................................................................................................77

Table 4-1 The average values of porosity and permeability for each sand facies ........88

Table 4-2 Summary of the differences between geological concepts (modified from Scheidt et al., 2009).....................................................................................91

Table 4-3 Summary of the of the hierarchical levels, corresponding geological uncertainties, and modelling parameters used to represent these uncertainties ........94

Table 4-4 SNESIM parameter ranges (uniform distribution) to account for geological uncertainties in West Coast of Africa reservoir ..............................................98

Table 4-5 Reservoir properties used for the flow simulation..........................................104

Table 4-6 Truth case model parameters used for facies modelling with SNESIM....106

Table 4-7 Values of the porosity and permeability for each geological facies for dynamic modelling (courtesy Stanford University, Scheidt et al., 2009).............................107

Table 4-8 Summary of the parameters used for Particle Swarm Optimisation algorithm (from Rojas, 2013) .........................................................................................114

Table 4-9 Parameter values for parametrisation applied to the West Coast of Africa facies models during history matching.................................................................115

Table 4-10 Comparison of the parameters ranges before history matching and after history matching, that provides the lowest values of misfit (<10).................................119

Table 4-11 Parameters used for NA-Bayes set up for WCA reservoir predictions under uncertainty..................................................................................................123

Table 5-1 Geological parameters used to perform facies modelling of the model 361, 425 and 426.................................................................................................144

Table 5-2 Confusion matrix of k-means classification results on testing data set received with 12 multipliers on the eigenfunction as the input data .................................177

Table 5-3 Total number of models of each training images predicted as a certain training image.......................................................................................................177
List of tables

Table 5-4 Confusion matrix of k-means classification results on testing data set received with 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data ................................................................. 178

Table 5-5 Total number of models of each training images predicted as a certain training image ............................................................................................................. 178

Table 5-6 The average values of recall and precision for k-means clustering for the test data set with two different input scenarios ................................................................................................. 179

Table 5-7 Confusion matrix of SVM classification results on testing data set received with 12 multipliers on the eigenfunction as the input data ................................................................. 181

Table 5-8 Total number of models of each training images predicted as a certain training image ............................................................................................................. 181

Table 5-9 Confusion matrix of SVM classification results on testing data set for two training images TI 2 and TI 3 only ............................................................................................................. 181

Table 5-10 Confusion matrix of SVM classification results on testing data set received with 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data ................................................................................................. 183

Table 5-11 Total number of models of each training images predicted as a certain training image ............................................................................................................. 183

Table 5-12 Confusion matrix of SVM classification results on testing data set received with 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data for two training images TI 2 and TI 3 ................................................................. 183

Table 5-13 The average values of recall and precision for SVM classification for the test data set for two different input scenarios ............................................................................................................. 184

Table 5-14 The optimal Random Forest parameter combinations and the corresponding OOB error values ............................................................................................................. 187

Table 5-15 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=500, mtry=3) ........ 188

Table 5-16 Total number of models of each training images predicted as a certain training image (ntree=500, mtry=3) ............................................................................................................. 188

Table 5-17 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=700, mtry=3) ........ 189

Table 5-18 Total number of models of each training images predicted as a certain training image (ntree=700, mtry=3) ............................................................................................................. 189

Table 5-19 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=1000, mtry=2) .... 189

Table 5-20 Total number of models of each training images predicted as a certain training image (ntree=1000, mtry=2) ............................................................................................................. 189
Table 5.21 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=1500, mtry=6)...........190

Table 5.22 Total number of models of each training images predicted as a certain training image (ntree=1500, mtry=6) ..............................................................................................................190

Table 5.23 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data for two training images TI 2 and TI 3 .............................................................................................................................................190

Table 5.24 The optimal Random Forest parameter combinations and the corresponding OOB error values ..........................................................................................................................................193

Table 5.25 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=600, mtry=2) ........................................................................................................................................194

Table 5.26 Total number of models of each training images predicted as a certain training image (ntree=600, mtry=2) ........................................................................................................................................194

Table 5.27 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=800, mtry=8) ........................................................................................................................................194

Table 5.28 Total number of models of each training images predicted as a certain training image (ntree=800, mtry=8) ........................................................................................................................................194

Table 5.29 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=1000, mtry=4) .....................................................................................................................................195

Table 5.30 Total number of models of each training images predicted as a certain training image (ntree=1000, mtry=4) .....................................................................................................................................195

Table 5.31 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=1300, mtry=4) .....................................................................................................................................195

Table 5.32 Total number of models of each training images predicted as a certain training image (ntree=1300, mtry=4) .....................................................................................................................................195

Table 5.33 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data for two training images TI2 and TI3 ......................................................................................................................................196

Table 5.34 The average values of recall and precision for RF classification for the test data set (Case 1 and Case 2 input data) .......................................................................................................................................199

Table 5.35 The best values of recall and precision for all three classification techniques: k-means clustering, SVM and RF .......................................................................................................................................200

Table 6.1 Geological parameter values for parametrisation applied to the West Coast of Africa facies models during history matching ................................................................................................................204
List of tables

Table 6-2 Other parameter values for parametrisation applied to the West Coast of Africa facies models during history matching .................................................................204

Table 6-3 Comparison of the parameters ranges before history matching and after history matching with SVM and RF, that provides the lowest values of misfit (<10) ..........210
1 CHAPTER ONE. Introduction

1.1 Thesis Motivation

Hydrocarbon exploration and production have been always associated with big investments and huge economic risks. Significant business risks arise from our incomplete knowledge of the subsurface. Decisions about investments are often based on predictions created from available analysed data. The available data might be sparse, contain errors or be difficult to analyse. It is also difficult and expensive to perform thorough data collection in very remote or harsh conditions due to the problems including deep water, severe weather, permafrost, etc. Even with advanced technologies, there are still a lot of remaining uncertainties associated with hydrocarbon reservoirs, such as interpretation of geological environment, reservoir facies description, petrophysical properties distributions, position, and geometry of faults, etc.

The value of accounting for and calculating uncertainties to aid secure decision making is recognised in the oil and gas industry. However, modelling uncertainty is difficult. It requires considering multiple sources of uncertainties, for example, geophysical survey data, geological core and log interpretational data, well measurement data and so on. These data, usually come from different scales and therefore it is challenging to include them into the geological modelling process.

Multiple geological models should be created to represent the uncertainty. With advances in computing hardware and software, it has become possible to create many reservoir models to characterise and understand the reservoir. These models can be generated by numerical simulations that represent the fluid flow in a reservoir described by geological features and distributed reservoir properties. The practice of generating multiple reservoir models is very useful to represent the associated reservoir uncertainties. Nevertheless, multiple models could be computationally expensive and even unfeasible due to the size and complexity of the model or the timeframe of the project.

There is also a challenge to propagate the uncertainties through the model calibration step. Multiple reservoir models are conventionally calibrated to the available reservoir production data. This process is called history matching. Models that match the history (production) data are considered possible representations of the distribution of geological
and physical properties of a reservoir. However, the reservoir production data can contain some measurement or interpretational errors. On top of that, history matching is known as an ill-posed inverse problem (Tarantola, 1987, 2005; Tavassoli et al., 2004). An ill-posed inverse problem means that there are many equally good solutions that match the data.

A good explanation on a Bayesian framework to solve the inverse problem is presented in Tarantola (2005): the first step is to present the prior probability distribution, and the second step is to transform the prior probability distribution into a posterior. The solution of the inverse problem is a set of models, consistent with prior information and observed data. Therefore, it is very important to account for the essential uncertainties in prior information.

However, there is a risk of underestimating subsurface uncertainties. A good example of this problem is illustrated in Vink et al., 2015 where they show that overlooked subsurface uncertainties may cause underestimated forecast uncertainty (Figure 1.1). In this example, they show that if the important characteristics of the subsurface model, for instance, alternative geological scenarios or faults are left out of the model, then this may affect the reservoir predictions. First, history matching will under-estimate uncertainty in reservoir parameters and consequently reservoir forecast will under-predict uncertainty in qualities of the interest such as fluid production rate. Figure 1.1 illustrates this problem showing the results of the forecast with a very narrow uncertainty range (in blue). This uncertainty range doesn’t represent the truth case model (in pink).

Another risk in history matching is losing the realistic geological representation of the models during the process. Many history matching papers (Gavalas et al., 1976; Gomez et al., 1999; Hajizadeh et al., 2011) have studied the tuning of reservoir properties, such as porosity, permeability, and fluid properties. These parameter tuning processes could create geologically unrealistic shapes, for example, by applying some multipliers to match the data around the wells (Figure 1.2). Such unrealistic models could decrease the reliability of the models and the resulting production response could mislead reservoir development planning.

Recently more attention has been paid to the tuning of the geological parameters, such as the geometry of geobodies (Arnold, 2008; Rojas, 2013), alteration of the structural and
stratigraphic framework (Caers and Hoffman, 2006; Park et al., 2013; Suzuki et al., 2008). These studies showed the value of geological information on the reliability of the history matched models.

Figure 1-1 A toy example of underestimated forecast uncertainty (From Vink et al., 2015)

Figure 1-2 Geologically unrealistic rectangular shapes around the wells resulted by applying porosity multipliers during history matching (From Valjak, 2008)
Chapter 1 Introduction

In this work, we propose a methodology to address the challenges of accounting for the essential geological uncertainties during history matching whilst at the same time preserving geological knowledge.

1.2 Thesis objectives

The aim of the thesis is to explore and answer three important questions related to geological uncertainties and accurate reservoir predictions (Figure 1.3):

1) How can we account for the essential geological uncertainties in facies modelling?
   Facies models greatly affect simulation results but require accounting for multiple geological uncertainties from different data sources and scales.

2) How can we propagate the facies uncertainties mentioned above and preserve their geological characteristics (geological realism) during history matching?

   In general, models are “geologically realistic” when they represent geological features observed in nature. Here, we study geological realism in facies modelling of a turbidite reservoir. We use conceptual models to provide geological description of turbidity depositional environment, which consists of channelized facies and lobes (detailed turbidite depositional environment and reservoir description is in chapter 4). In this study, we also use observed measured data (well and seismic data) to condition our facies models to real data. We are looking to obtain models reproducing the following features: 1) specific facies properties, such as channel continuity and sinuosity in channelized facies, 2) specific facies geometries, either channels or lobes, 3) specific facies proportions (defined from the well data and analogue outcrop data). These are the necessary and sufficient conditions to represent geological realism in facies modelling. We do not consider the sub-facies scale in this work for practical reasons.

3) Can geological facies uncertainties improve the reliability of the forecast?
   To evaluate the reliability of the forecast, we compare the forecast uncertainty intervals to a so-called truth case (reference) model. If the truth case model is inside the uncertainty interval, then our forecast is reliable.
To answer these questions, we propose a methodology that consists of several steps. The first step is a hierarchical approach to account for the essential geological uncertainties from different levels of geological data (depositional environment, geological scenario, geological parameters) in facies modelling (Figure 1.4). Each level of hierarchy provides different geological uncertainties, from the bigger scale – depositional environment – to the smaller scale – geological facies modelling parameters. In our case, depositional environment defines possible geological scenarios and geological parameter combinations. As the result of the hierarchical approach, we evaluate the prior range of geological uncertainties.

The second step presents how to include the prior geological uncertainties from the hierarchical approach into history matching. Figure 1.5 schematically represents the main difference between the two techniques that we propose in this work: with the first technique, we sample from the parameter space, and with the second technique, we sample from the metric and parameter space.

Sampling from the parameter space incorporates geological uncertainties in facies models (geological scenarios and geological modelling parameters) as the tuning parameters in history matching (Figure 1.5). We parametrise multiple alternative geological scenarios as categorical variables with uniform distribution and geological modelling parameters as continuous variables with a uniform distribution (Figure 1.6).
Figure 1-4 Schematic representation of the hierarchical approach to account for the geological uncertainties

Figure 1-5 Schematic view on the main difference between two proposed techniques to include geological uncertainties into history matching
Chapter 1 Introduction

Figure 1-6 Schematic representation of the first workflow proposed to include geological uncertainties from hierarchical approach into history matching process

The second technique applies the concept of metric space for modelling prior uncertainty (Suzuki and Caers, 2008). Metric space allows us to handle multiple model realisations from the alternative geological scenarios using the concept of distance between the realisations according to a certain metric (Figure 1.7). We propose to identify the geological relations between different geological scenarios by means of classification in the metric space and include them into history matching (Figure 1.5).

Without the metric space and classification, the relationship between multiple geological scenarios is arbitrary, therefore when we introduce multiple geological scenarios into history matching as unrelated categories, we impose some order which does not provide any geological meaning. Thus, if scenario 1 does not provide good values of misfit (low values), scenario 2 may be selected instead, however, scenario 2 could be very different in terms of geological description.

Classification of the metric space allows us to find geological relations between multiple scenarios. Therefore, if scenario 1 does not provide good values of misfit (low values), scenario 5, for example, may be selected instead, which has a similar geological description and provides better values of misfit. Hereby, we aim to find the appropriate combination of geological parameters and geological scenarios to improve geological realism by introducing a classification of metric space into history matching (Figure 1.8).
Chapter 1 Introduction

The final step of the methodology includes the evaluation of Posterior Probability Distribution (PPD) under Bayesian framework and estimation of the Credible Intervals (P10, P50, P90) to represent the uncertainty in the predictions (Figure 1.9).

Figure 1-7 Schematic view of the proposed metric space classification step to differentiate between different geological scenarios in metric space

Figure 1-8 Schematic view of the first workflow proposed to include geological uncertainties from hierarchical approach into history matching process

Figure 1-9 Schematic view of the final step of the proposed methodology to perform reservoir forecast under uncertainty

1.3 Thesis Overview

The novelty of this work lies in the following aspects: 1) introducing the hierarchical approach to account and to introduce geological uncertainty in facies modelling, 2) incorporating hierarchical uncertainties into history matching, 3) application of machine learning classification to evaluate the relations between geological scenarios and propagate them into history matching, and 4) improving geological realism during history matching by providing consistency of geological scenarios and related geological parameters.

This thesis consists of seven chapters:
Chapter 2 presents a state-of-the-art literature review on the problem of accounting for geological uncertainties and propagating them into history matching. In this chapter, we also discuss the proposed techniques to model and preserve geological realism during the model update process.

Chapter 3 introduces Particle Swarm Optimisation for history matching and NA-Bayes for uncertainty quantification. We show the setup of these algorithms on the Ainsa II case study. We also discuss the importance of geological knowledge, geological uncertainties and geological realism in history matching in this chapter.

Chapter 4 describes the proposed hierarchical approach to account for essential geological uncertainties in facies modelling. Also in this chapter, we show how to create multiple facies models using Single Normal Equation Simulation (SNESIM). We show how to incorporate geological uncertainties in facies representation as the tuning parameters in history matching (Figure 1.6). Finally, we show the results of history matching and uncertainty quantification. We show the application of the proposed workflow on a synthetic case study based on a real reservoir of the West Coast of Africa.

Chapter 5 explains the concept of metric space approach and multidimensional scaling. In this chapter, we introduce three machine learning techniques used to perform the classification in the metric space: k-means clustering, Support Vector Machines and Random Forest. We apply these techniques to the West Coast of Africa case and discuss the results of the classification.

Chapter 6 shows the application of the second technique when we introduce geological uncertainties based on the metric space classification results into history matching (Figure 1.8). In this chapter, we present the results of the history matching and uncertainty quantification on the West Coast of Africa example. We also compare the results of chapter 4 and chapter 6 and present the discussion.

Chapter 7 summarises the results, provide some final remarks and future work suggestions.
CHAPTER TWO. Geological Reservoir Modelling and Uncertainty Quantification

2.1 Introduction

Reservoir models are a useful tool to understand the reservoir and to predict its future performance. The reservoir’s future performance prediction should be as accurate as possible because the main economic decisions are based on it.

Reservoir models are associated with various uncertainties in geological structure and geological properties, such as the position of the faults, facies proportions and dimensions, petrophysical properties, and so on. Besides uncertainties in geological structure and properties, there are uncertainties in fluid properties, for example, the composition of oil, fluid contacts, and so forth.

Uncertainties come from a lack of data and a lack of understanding. For example, a geologist needs to build a model based on core samples and well-log data without knowing what exactly is going on between the wells. Seismic data can bring more information about the reservoir between the wells but does not necessarily reflect all the reservoir details due to the seismic resolution. This is similar to trying to complete a puzzle without having all the pieces.

Uncertainties also come from data interpretation. Specialists working with the available data use their expert knowledge to analyse and interpret the data. Such interpretations are based on specialist’s judgments and are open to bias. For example, Bond et al. (2007) showed how an individual’s training, prior knowledge, and the techniques used can result in different interpretations of the same seismic image.

Nowadays, geostatistical or object-based simulation tools (Deutsch, 2002; Journel et al., 1998) allow us to generate many geologically realistic static reservoir models to account for existing uncertainties. These multiple geological models are usually used to represent the associated uncertainties in reservoir flow dynamic response.
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

Multiple models should be calibrated to the available production data to adjust the simulated models to the observed data. However, many parameter combinations that we use to adjust the models can fit the production response and many reservoir models can be produced honouring the production data. Each of these models can produce different reservoir predictions. Therefore, reservoir predictions should be based not on one but on a set of history matched models to represent the uncertainty range.

Multiple models cannot always guarantee a representative uncertainty range (Figure 1.1). The multiple models should be diverse in order to capture the essential geological uncertainties, for example, multiple models could be based on multiple alternative geological scenarios. However, simulation of multiple models from multiple scenarios can be computationally expensive. Several methodologies have been proposed to account for multiple geological scenarios in a more efficient way (Park et al., 2013; Scheidt et al., 2009; Suzuki and Caers, 2008). We provide a review of these methodologies in this chapter.

In real life applications, history matching is usually achieved by modifying reservoir properties, such as porosity and permeability. Alteration of these parameters to achieve a good match could result in geologically unrealistic geometries, such as box-shaped or pipe-shaped geometries around the wells (Figure 1.2). An adequate match to the field data can be achieved this way, however, preserving the geological knowledge of the reservoir increases the model reliability.

In this chapter, we provide a review of the techniques and methodologies developed to create geologically realistic reservoir models and incorporate geological uncertainties into modelling processes. We also discuss the available techniques to propagate the uncertainties into history matching and to preserve realistic geology whilst adjusting the model to the history production data.

2.2 Static reservoir modelling

In this thesis, we concentrate on facies modelling to represent a depositional environment of the reservoir and describe the distribution of sedimentary facies. Facies modelling is important because it defines the modelling step that follows – petrophysical modelling – and affects the flow simulation results.
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

It is very important to create a realistic model of facies distribution and account for the associated uncertainties, although it remains quite challenging. The three main types of techniques to construct facies models are: process-based, object-based, and geostatistical techniques.

Process-based modelling aims to reproduce the physics of depositional processes of the geological system (Figure 2.1). The models simulate the genesis of geologic formations over time. Process-based models can produce very realistic geometries as they are not limited by assumptions of spatial structure (Michael et al., 2010). However, these models are difficult to condition to observational data (Anderson et al., 1997; Karssenberg et al., 2001). Therefore, process-based models have limited application in reservoir simulation and history matching studies.

Figure 2-1 Process-based facies modelling input parameters in a cross-section of a point bar. Lateral accretion surfaces and coarse channel lag deposits are indicated in the point-bar. (2) Map view of facies in the depositional environment (From Willems et al., 2017)

Object-based or Boolean modelling (Cosentino, 2001) is a process of creating models of geological elements using predefined objects, for example, using sinuous half cylinders to represent fluvial channels (Journel et al., 1998). The objects can be parametrised by the dimensions and shapes of objects based on available input data (well data, seismic interpretations, or analogue data). Object-based models can represent quite complex geological shapes, which preserve realistic reservoir connectivity representation.
Object-based models are quite difficult to condition to wells and seismic data. There are two main approaches to condition object-based facies model observations. The first approach is based on placing objects that are intersected by wells first and then distributing non-intersected objects at random locations until the desired facies proportions are met (Gundesø and Egeland, 1990). The second approach uses simulated annealing (Khachatryan et al., 1981) to measure the discrepancy between the model and the data (soft and/or hard data); then it proposes a change to the seed model (e.g. moving or changing the characteristics of an object), and accepts or rejects this change depending upon the objective function evaluation improvement (Deutsch and Tran, 2002; Skorstad et al., 1999). In both approaches, conditioning to hard data may be not feasible, especially with a dense set of hard data observations and a large number of soft data constraints (Falivene et al., 2007).

Object-based modelling could be used for geologically consistent history matching applying gradual deformation method to modify the object locations and the number of objects in a simulation (Hu, 2002). However, during these modifications, the objects suddenly appear and disappear in the model. This fact brings unwanted discontinuity in the objective function (especially for gradient-based optimisation). Le Ravalec-Dupin and Hu, 2005 reformulated the gradual deformation method to smoothly add or remove objects from a Boolean model. The smoothness helps to reduce the discontinuity of the objective function but does not fully eliminate it (Li, 2008).

Geostatistical sequential simulation is widely spread in petroleum industry nowadays. One group of geostatistical methods uses variogram-based geostatistical algorithms (Figure 2.2), such as the Sequential Indicator Simulation (SIS) (Deutsch and Journel, 1998; Journel, 1982), Truncated Gaussian Simulation (TGS) (Ravenne and Beucher, 1988), and Pluri-Gaussian Simulation (PGS) (Le Loc’h and Galli, 1997). The main limitation of the variogram-based geostatistical methods is that they cannot capture complex curvilinear shapes, such as meandering channels.

Xu, 1996 proposed to adapt the local indicator kriging systems to varying anisotropy using a distribution of local angles. The resulting models were able to reproduce curvilinear structures, yet being discontinuous. Discontinuous structures could affect the accuracy of the flow simulations and reservoir flow predictions.
The other group of geostatistical methods uses Multi-Point Statistics (MPS), accounting for correlation between three or more locations at the same time (Table 2.1). MPS uses a so-called training image (Guardiano and Srivastava, 1993) (Figure 2.2) instead of a variogram (SIS, SGS, PGS) to capture the correlation structure (Van den Boogaart, 2006). Training image is a conceptual representation of sedimentary objects honouring geological knowledge without any specific well or seismic data to be tied to at that stage. The training image can be created based on outcrop data, photos of the modern depositional environment, analogue data. Training images represent an essential database of geological patterns. Once the required patterns are retrieved from the training images, they need to be tied to the available subsurface data (well, seismic or/and production data).

<table>
<thead>
<tr>
<th>Comparison criteria</th>
<th>Two-point statistics</th>
<th>Multi-Point Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>Statistical relation of two points</td>
<td>Structures and pattern beyond two-point</td>
</tr>
<tr>
<td></td>
<td>$z(u)$ and $z(u + h)$</td>
<td>correlation</td>
</tr>
<tr>
<td>Conditional probability</td>
<td>Variogram model</td>
<td>Training image</td>
</tr>
<tr>
<td>Parametrization</td>
<td>Sill, range, nugget, shape of variogram</td>
<td>Proportions, scale, anisotropy</td>
</tr>
</tbody>
</table>

*Table 2.1 Comparison between two-point and Multi-Point Statistics characteristics*
Figure 2-2 An example illustrating the Sequential Indicator Simulation using a variogram (Deutsch and Journel, 1998) and Multiple-point simulation, performed with the method IMPALA (Straubhaar et al., 2011) using a training image. The simulated values come from a prior model based on the known cell configuration in the neighbourhood of the cell to simulate (From Rongier et al., 2016).

Figure 2.3 helps illustrate how MPS works on a simple model example (SNESIM algorithm). A training image is used for the calculation of the conditional probability of an unknown central cell surrounded by the cells where the data is already known. This group of cells is called “data event”. In this example, the “data event” consists of two sand cells (yellow) and one shale sand (shale). The probability is calculated by scanning a training image for “replicates” of a given “data event”. Three such events are found of which one gives a central sand value, therefore the probability of having sand in this cell is 1/3 (Caers, 2005). A random value is drawn and combined with the calculated
probability to determine the facies value in this cell. The process is repeated until the grid is full. The modelling process results in a realisation with a pattern of spatial continuity similar to a training image. A different model realisation can be made by varying the random numbers and/or the order in which all the cells are visited.

MPS algorithms, as well as most geostatistical methodologies, rely on some form of stationarity to provide the spatial variability information. For example, Single Normal Equation Simulation (SNESIM) (Strebelle, 2002) requires a stationary training image to retrieve the patterns and calculate the conditional probability.

In many cases, the geological features are not stationary over the entire model. Some non-stationarity should be incorporated into a modelling process to represent the realistic geology (Honarkhah and Caers, 2012). To introduce non-stationarity, models are often decomposed into a trend component and a stationary stochastic component. In MPS, the training image is used to extract stationary higher-order statistics, such as conditional probabilities (Strebelle, 2002), higher-order cumulants (Dimitrakopoulos et al., 2010), patterns (Arpat and Caers, 2005), and data events, (Mariethoz et al., 2010). After extraction, they are combined in various ways to create non-stationary models. For example, in SNESIM algorithm (used in this work) local and global parameters of rotation and affinity can be used to introduce non-stationarity into the model. SNESIM algorithm and modelling parameters will be discussed in more detail in chapter 4.

The advantage of MPS is that it can be easily conditioned to well and seismic data. For example, for a conditional simulation (as in SNESIM), the hard data are honoured by assigning the data to the closest nodes of the current grid level before simulating it; seismic data can be converted into probability cubes (Caers et al., 2006) and the realisations are conditioned to these seismic probability cubes.

An issue with MPS is that it uses the short relative correlation length of the search neighbourhood, in comparison with the long correlation length of geological structures such as channels. This fact could lead to a discontinuity of the modelled sand bodies. To preserve the continuity, as this feature has direct impact on flow simulation results, some improvements (Strebelle and Remy, 2005) and new modelling methods were proposed, such as simulating all the cells within a neighbourhood at once from the training image instead of a single cell (Arpat and Caers, 2005, Chatterjee and Dimitrakopoulos, 2012,
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

Tahmasebi et al., 2012). However, these improvements and methods have conditioning difficulties.

Mariethoz et al. (2010) proposed Direct Sampling (DS) as an alternative Multi-Point Statistics technique to construct continuous variables, such as sand channels. This method is statistically equivalent to SNESIM MPS implementations (chapter 4) but instead of scanning and storing the obtained conditional probabilities in a database, it samples the training images directly, making the database unnecessary. The main advantages of Direct Sampling are: it is computationally fast, applicable with various cases of nonstationarity, and straightforward to implement.

Zhang et al., 2006 showed the application of SNESIM and FILTERSIM (a filter-based implementation of MPS, Zhang, 2006) to model both categorical and continuous variables of a carbonate reservoir in Kazakhstan. In this example, SNESIM was used to model the two main reservoir regions – platform and slope; FITERSIM was applied to model the 3D porosity distribution by capturing the porosity patterns from the training images and anchoring them into the reservoir by conditioning to well data.

SNESIM provided satisfactory results; however, FILTERSIM resulted with some issues. FILTERSIM results showed that in the case when the available data are dense and contain
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

continuous variables, the training images did not match the real data patterns because training images contained “idealised” patterns. Therefore, training-image patterns can be considered as soft information to constrain the models so that the tradeoff between training patterns and well-data conditioning is achieved.

Recently, Rongier et al., 2016 proposed a systematic method to compare static connectivity of the geological realisations. An incorrectly modelled static connectivity can bias the results of the flow simulation (Journel et al., 1990), therefore it is important to control the model’s connectivity. However, the differences in static connectivity cannot always be recognised by two-point statistics analysis. Rongier et al. (2016) presented a set of indicators to quantify and analyse the static connectivity of the stochastic simulations. The indicator analysis is unbiased and can handle a great number of realisations, compared with a visual analysis. This methodology helps to rank the model realisations obtained by different simulation methods.

In this thesis, we use Sequential Normal Equations Simulation (SNESIM) in the open-source Stanford Geostatistical Modelling Software (SGeMS) to perform geological facies modelling. We chose SNESIM because (1) it allows us to condition model realisations to available well and seismic data; (2) SNESIM modelling parameters allow us to introduce hierarchical uncertainties (Level 3) for each geological scenario (training image).

2.3 Dynamic reservoir modelling

The main goal of the dynamic reservoir modelling is to predict reservoir performance and hydrocarbon recovery for a given field development scenarios. Dynamic modelling also helps to evaluate the effect of different development scenarios and conditions on hydrocarbon recovery and compare the economics of different recovery scenarios.

Reservoir simulation is a complex and thorough approach to predict reservoir performance. Reservoir simulation is based on a conservation of mass (Dake, 1983) and Darcy’s law for multiphase flow (Young et al., 1983). The reservoir is represented as multiple grid cells, where each cell is assigned a reservoir properties value (porosity, permeability, saturation, etc.) (Figure 2.4). These properties are usually incorporated from a previously created geological model. The reservoir model is used to calculate a
numerical approximation of the field fluid flow by calculating flow between the model grid cells.

The mathematical model in reservoir simulations represents the physics of reservoir fluid behaviour between grid cells in the reservoir model. It captures the physical properties of fluids, such as viscosity, density and compressibility. To create a mathematical model of reservoir physics we need to solve a set of differential equations with a set of boundary equations (Peaceman, 1977).

![Figure 2-4 3D 3 phase flow simulation of a volatile oil system with gas and water Injectors (50x30x7 Reservoir Grid System) (From the (“Marathon Center of Excellence for Reservoir Studies, Showcase,” 2014) ![Figure 2-4 3D 3 phase flow simulation of a volatile oil system with gas and water Injectors (50x30x7 Reservoir Grid System) (From the (“Marathon Center of Excellence for Reservoir Studies, Showcase,” 2014)](image)

The flow of a single fluid (single phase) in one dimension (1D) is described with Darcy equation as follows:

\[ q = -\frac{kA \Delta P}{\mu L} \]  

(Eq. 2.1)

where \( q \) is flow rate (cm\(^3\)/s), \( k \) is intrinsic permeability (Darcy), \( A \) is a cross-sectional area (cm\(^2\)), \( \mu \) is the viscosity of the fluid (cp), \( \Delta P \) is the pressure difference (atm), and \( L \) is the distance (cm).
A partial differential form of Darcy equation is used for the reservoir simulation, where $u$ is the Darcy velocity:

$$u = -\frac{k \partial P}{\mu \partial x}$$  \hspace{1cm} (Eq. 2.2)

The mathematical model also considers the interaction between reservoir fluids (oil, gas, water). Whilst oil and water are immiscible fluids (i.e. they do not mix), some gases, such as nitrogen ($N_2$), carbon dioxide ($CO_2$) are miscible (i.e. the gases mix with the oil).

The interaction of immiscible fluids can be modelled by relative permeability ($k_r$) curves. These fluid relationships are included into a flow simulator to predict the flow accounting for different fluid interaction.

Two types of fluid description can be used in numerical reservoir simulation: black oil model and compositional model. The black oil model treats “gas” and “oil” components as “gas” and “oil” phases. It considers the fluid's volumetric (formation volume factor and solution gas-oil ratio) and flow characteristics (viscosity) as a single-value function of pressure (Lyons and Plisga, 2011). In complex fluid mixtures, the volumetric and flow characteristics are a function of pressure and composition. These fluids are represented as N-component hydrocarbon mixture, where N represents the components (e.g. C$_1$, C$_2$, C$_3$ and so on). For these fluids, a compositional model applies an equation of state to describe the volumetric and flow characteristics of fluids. An equation of state characterises a fluid in terms of the fundamental physical properties of its components (methane, ethane, heptanes-plus, and so on). These fundamental physical properties (critical pressure, critical temperature, critical volume, and so on) are unique for each compositional fluid description derived for a simulation study.

Generally, the mathematical model for reservoir simulation is a set of mass balance calculations (phase dependent) and a set of Darcy based flow equations that should be solved. They can be solved through reservoir spatial and temporal discretization via grid blocks and time steps (discrete points in time). The mathematical model is very often discretized using a finite difference scheme. Full finite difference reservoir simulation is computationally expensive, especially for big and complex geological models.
An alternative method, called streamline simulation (SLS) emerged in the early 1990s and since then has advanced significantly. The main differences of streamline-based flow simulation from conventional reservoir simulations are: 1) it does not conserve mass exactly, and 2) the fluid is transported along the streamlines (Figure 2.5) and not from cell-to-cell (Thiele et al., 2010).

Streamline-based flow simulation is especially effective in solving large, geologically complex and heterogeneous systems, where fluid flow is defined by well positions and rates, rock properties (e.g. permeability, porosity, fault distributions), fluid mobility (phase relative permeabilities and viscosities), and gravity (Thiele, 2001). For streamline simulation to work well, flow rates should be high so that capillary and gravity forces play only a small role in the displacement.

A streamline simulator uses two separate time scales in the process of field scale fluid flow through porous media: the time scale of pressure diffusion and the time scale of saturation advection. The fact that the time scale of pressure diffusion is much smaller than the time scale of saturation advection, allows us to use the quasi-stationarity hypothesis. This provides the decoupling of the pressure and saturation dynamic problems. After that, the full 2D/3D problem transforms into multiple 1D problems solved along streamlines, using the “time of flight” parameter along streamlines.

The “time of flight” along the streamlines is defined as follows (Batycky et al., 1997):

$$\tau = \int_0^S \frac{\varphi}{u_t(\zeta)} d(\zeta)$$  \hspace{1cm} (Eq. 2.3)

which gives us the time required to reach a point $S$ on the streamline-based on the total velocity $u_t(\zeta)$, along the streamline.
The main advantages of streamline are speed and the ability to simulate complex geology without losing essential heterogeneities (Datta-Gupta and others, 2000). Nevertheless, the streamline simulator could be considered less accurate as the pressure equation is updated at a limited number of time steps, compared to the full finite difference reservoir simulation, where the pressure is updated much more frequently (Datta-Gupta and others, 2000). Datta-Gupta (2000) highlights the streamline simulation should not be considered as a replacement for the conventional grid-based simulation but as an alternative tool to solve geologically complex problems faster, because the conventional reservoir simulator conserves mass to machine accuracy, whilst streamline simulator does not.

Dynamic simulations play important role in history matching and uncertainty quantification. They allow obtaining an accurate reservoir dynamic response to match the production history data and evaluate the uncertainty in reservoir future production performance (section 2.4 and 2.5).
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

To perform dynamic simulations in this thesis, we are going to use “3DSL” streamline simulator (“Streamsim Technologies, User manual – v2015.0911”) to perform large number of simulations in reasonable time.

2.4 Bayesian Approach to Model Uncertainty

Uncertainty is generally present in all steps of the modelling process. Challenor and Tokmakian (2011) divided the model uncertainty into two types: aleatoric and epistemic. Aleatoric uncertainty represents a random irreducible uncertainty. For example, prices of crude oil and gas products (Lin et al., 2012; Thunnissen, 2003) is irreducible uncertainty because the market prices of crude hydrocarbons are continuously changing and individual companies have little influence on the price fluctuations. Epistemic uncertainty, on the contrary, can be reduced with gathering more data or knowledge.

Bayesian statistics is used to describe epistemic uncertainty using probabilities. Tarantola, 2005 provides a good explanation of Bayesian framework for modelling uncertainty as a consistent and repeatable mathematical framework. Bayes Theorem is written as follows:

\[ p(m|O) = \frac{p(O|m)p(m)}{\int p(O|m)p(m)dm} \]  

(Eq. 2.4)

where \( m \) is our model parameter that we aim to estimate based on some observation \( O \), \( p(m) \) is the prior probability distribution and \( p(O|m) \) is the so-called likelihood function.

The integral in the denominator is the normalisation constant and Bayes Theorem could be simplified as:

\[ \text{posterior} \propto \text{prior} \times \text{likelihood} \]  

(Eq. 2.5)

Taking logarithms, we obtain the following expression:

\[ \log(\text{posterior}) = \log(\text{prior}) + \log(\text{likelihood}) \]  

(Eq. 2.6)

The prior distribution combined with likelihood provides a posterior probability distribution (updated epistemological uncertainty). The posterior probability represents the uncertainty of the model taking into account the data and the assumptions we made.
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

Bayes Theorem applied to petroleum industry problems helps to predict reservoir model parameters by updating these parameters probabilities based on available field observation data, such as reservoir pressure, production rates, etc.

As the choice of prior distribution directly affects the resulting posterior distribution, it is important. A realistic geological prior distribution can be based on data from cores and seismic measurements of the reservoir, outcrop and analogue reservoir data, and/or scientific publications and expert’s knowledge.

The likelihood of a reservoir model represents the probability of observation data being equal to simulated responses from a given reservoir model. In this thesis, the likelihood will be calculated using the misfit values between the historical production data of the reservoir (Obs) and simulated production response obtained from the model (Sim). Likelihood at a time step $t$ is defined as follows:

$$p(O_t | m) = \left( \frac{1}{\sigma \sqrt{2\pi}} \right) \exp \left[ -\frac{1}{2} \frac{(Obs - Sim)^2_t}{\sigma^2} \right]$$  (Eq. 2.7)

where $\sigma$ is the standard deviation in the measurement field data (chapter 3), $t$ is the time step, $Obs$ is the historical production data, and $Sim$ is the simulated production response.

We assume that the measurement errors (in historical production data) are independent in each time step and normally distributed. Therefore, the likelihood is calculated as the product of the probabilities of individual measurements at all time steps ($t$):

$$p(O | m) = \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^N \prod_{t=1}^{N} \exp \left[ -\frac{1}{2} \frac{(Obs - Sim)^2_t}{\sigma^2} \right]$$  (Eq. 2.8)

where $\sigma$ is the measurement error (chapter 3), $t$ is the time step, $N$ is the number of data points, $Obs$ is the historical production data, and $Sim$ is the simulated production response.

The Bayesian statistics is going to be used to perform geologically consistent history matching and uncertainty quantification.
2.5 History matching approaches

History matching is a process of adjusting input parameters of a reservoir model so the simulated production agrees with the observed behaviour. The difference between the observed and simulated data is called an objective function (Glimm and Sharp, 1999), and we aim to minimise it by tuning the unknown model parameters.

A numerical value to represent the discrepancy between the observed and simulated data is called misfit. Misfit ($M$) is defined as follows:

$$M = -\log(\text{posterior})$$  \hspace{1cm} (Eq. 2.9)

The most common objective function for history matching obtains the misfit by the least squares formula:

$$M = \frac{1}{2} \left( \sum_{i=0}^{N} \frac{(\text{Obs}_i - \text{Sim}_i)^2}{\sigma_i^2} \right)$$  \hspace{1cm} (Eq. 2.10)

where $\text{Obs}_i$ is the observed data, $\text{Sim}_i$ is the simulated result at time $i$, $N$ is the number of data points, and $\sigma_i^2$ is the measurement error.

The least squares method for calculating the misfit (Eq. 2.10) is the most common approach due to its simplicity, however it does not account for a time-dependent variance that may exist in the data (least squares uses a single value for $\sigma^2$ at each time step of the simulation (O’sullivan and Christie, 2005). A more complete, time-dependent description of misfit uses a covariance matrix to represent the measurement errors (Christie et al., 2005).

Another adjustment of the misfit calculation is to include an estimate of the solution error which is caused by using a model grid resolution that does not capture the full detailed characteristic of the reservoir.

O’Sullivan and Christie (2005) showed that using the error model to estimate reservoir model parameters can overcome the problems of bias and inappropriate measures of variance. The bias, which results from using a coarse grid model in simulations, can be
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

eliminated by introducing mean error into the misfit calculations. This results in a more accurate representation of the true value by the maximum likelihood estimate.

Replacing the single value for a variance with the time-dependent covariance means the spread of the error, as it changes with time, is adequately represented in the misfit function. As the covariance is calculated using a number of realisations, one can obtain a level of confidence in parameter estimates and oil recovery predictions (O’Sullivan and Christie, 2005).

Throughout this thesis, the least squares misfit is used due to its simplicity to implement on reservoir studies.

Equation 2.11 shows the relationship between misfit and likelihood. In this thesis, we calculate the likelihood values directly from the least square misfits (Eq. 2.10) of history matched model.

\[
p(\mathcal{O}_t | m) \propto e^{-M}
\]

(Eq. 2.11)

Many history matching studies (Gavalas et al., 1976; Gomez et al., 1999; Hajizadeh et al., 2011) used reservoir properties, such as porosity, permeability, and fluid properties as tuning parameters in history matching. However, changes in these parameters during tuning can affect the geological realism in model representation, for example resulting in unrealistic geobody shapes. These unrealistic models affect the reliability of resulting production response and mislead reservoir development planning.

Recently more attention has been paid to the tuning of different geological parameters. For example, Arnold (2008) used channel width and thickness as the parameters that define the geometry of geobodies. He applied the empirical equation to control the realistic width and thickness parameter combinations during history matching.

Another example of geological parametrisation has been presented by Suzuki et al. (2008) where they considered structural uncertainty as a parameter for history matching. They proposed to use a set of structural realisations that encapsulates different horizon and fault positions to perform a history matching that honours geological/geophysical constraints.
In addition to parameterization, the main challenges of history matching are:

- ill-posedness, which means that are many equally good solutions that match the data (resulted in low values of misfit), and there is no way to differentiate between them (Tavassoli et al., 2004);
- a large number of model parameters, which complicates the history matching process,
- and time-consuming forward simulations, which arise from the previous two challenges.

Various approaches have been proposed to overcome these challenges (Oliver and Chen, 2011). They could be divided into three groups: gradient-based, stochastic and assimilation methods.

Gradient-based methods were the first optimisation method used in automated history matching (Slater et al., 1971). They have been widely used due to their fast convergence (Oliver et al., 1996; Van Ditzhuijzen et al., 2001). Gradient-based methods require the calculation of the derivative of the objective function with respect to the model parameters. The gradients are obtained by changing each parameter individually and performing the reservoir simulation to evaluate how each parameter affects the objective function. The adjustment of each model parameter sets the limitation on the number of parameters that are practical to use.

An alternative solution, such as the adjoint method was proposed by Li et al., 2001. The advantage of this method is that the number of matrix solutions needed to compute the sensitivity coefficients (sensitivity of production data to the model parameters) is independent of the number of reservoir model parameters to be estimated (Li et al., 2001). Calculating these sensitivity coefficients allows computation of a posterior covariance matrix and to estimate the reduction in uncertainty (obtained by conditioning to production data). After these sensitivities are computed, the Levenberg-Marquardt or Gauss-Newton method methods can be used to perform history matching.

Levenberg-Marquardt and Gauss-Newton are the conventional optimisation methods used for gradient-based methods. The main advantage of Gauss-Newton and Levenberg-Marquardt methods is the quadratic convergence: when the number of steps is doubled, the
error goes down by an order of 4. There have been multiple studies of history matching geological features with these methods. For example, Caers, 2002 used Gauss-Newton gradient approach to history match production rates of Multi-Point Statistical models, and Bi et al., 1999 used Levenberg-Marquardt algorithm to condition stochastic channel models to pressure data.

These methods are very efficient in converging to a local minimum in the objective function, but there is no guarantee that this is the global minimum. The nonlinearity in static and dynamic data result in the solution getting stuck in a local minimum (Gómez-Hernández et al., 1997; RamaRao et al., 1995).

Gradient-based methods are limited for uncertainty quantification because they quantify uncertainty in respect to a single local minimum. Gomez et al. (1999) combined a gradient-based method with a global optimisation technique called the “tunnelling” (Figure 2.6). This method allowed multiple solutions to be found during history matching. Though there was not much difference between these multiple solutions in the history match, the forecast of each solution gave a different prediction. This study demonstrated the importance of obtaining multiple history match solution for better uncertainty evaluation.

Stochastic optimisation algorithms, such as Genetic Algorithm (GA) (Goldberg, 1989; Holland, 1975) and Neighbourhood Algorithm (NA) (Sambridge, 1999a) can overcome the problem of getting stuck in a local minimum, however, they are less computationally efficient than gradient-based methods.

Genetic algorithm (GA) is based on natural processes (Darwin’s theory of evolution) and has been used widely in history matching to solve optimisation problems with an evolutionary process. The method simply starts with a population of individuals (solutions) who produce new population using the principles of reproduction, crossover and mutation based on their objective function values. GA was used to predict permeability alteration in formation damage (Erbas et al., 2001), well placement and trajectory optimisation (Montes et al., 2001), and history matching (Romero et al., 2000; Williams et al., 2004).
Neighbourhood Algorithm (NA) is a derivative-free method that aims at finding an ensemble of acceptable models rather than seeking for a single solution. NA uses the properties of Voronoi cells in high dimensions to achieve multiple history matched models. The key approximation in NA is that the misfit surface is constant in the Voronoi cell surrounding a sample point in parameter space. NA has been used in a number of reservoir history matching studies (Christie et al., 2002; Erbas et al., 2007; Rotondi et al., 2006; Subbey et al., 2003, 2004).

Particle Swarm Optimisation (PSO) is a more recent generation of stochastic optimizers. PSO explores the parameter space by a population of particles that search for an optimum solution (Eberhart et al., 2001). These particles replicate the behaviour of birds flocking or fish schooling looking for the best place for food in nature. In PSO each “bird” (or particle) is a single solution in the search space, and “food” is the global minima. The example of PSO application to reservoir engineering problems was presented in Kathrada (2009) where he combined PSO with a hierarchical clustering algorithm to obtain history-matched models. Fernandez et al. (2009) applied PSO to perform seismic history matching and production optimisation.

Mohamed et al. (2010a) compared the efficiency of PSO for history matching with Neighbourhood Algorithm (NA). She concluded that PSO obtains good history matched models with lower value of misfits from fewer simulations than NA. Mouser and Dunn
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

(2005) demonstrated that PSO provides better history matching results (lower values of misfit function) in less number of iterations than Genetic Algorithm.

These stochastic methods find a solution honouring the dynamic data, however, they are not originally designed to preserve the realistic geological model representation. Therefore, some other stochastic techniques were developed to address this issue.

Other stochastic techniques, such as Gradual Deformation method (GDM) (Hu, 2000; Hu et al., 2001, Le Ravalec-Dupin and Nø etinger, 2002; Roggero et al., 1998) and Probability Perturbation method (PPM) (Jef Caers, 2003; Caers and Hoffman, 2006) to account for not only dynamic production data but also static geological data as model constraints.

GDM gradually transforms an initial geological model preserving the geological continuity of the model, until it matches the history data (Caers, 2003). The GDM proposed by Hu et al. (2001) is based on a perturbation of the random numbers of a sequential simulation algorithm used to generate a geostatistical realisation. Any type of sequential simulation can be used with GDM. The Gradual Deformation method uses the fact that certain linear combinations of independent Gaussian fields preserve second order statistics. When the linear combination is formed, so-called deformation parameters are used to weight each individual realisation. Then these weights are adjusted to perturb the resulting realisation to achieve a history match.

During history matching, GDM is essentially a geostatistical parametrization, which can reduce the number of independent variables in history matching. However, this parametrisation can result in a slowdown of the convergence rate by being too restrictive and consequently narrowing down the search space.

The Gradual Deformation method is limited to work with geological data that can be represented by a Gaussian distribution. Therefore, some form of transformation of the adjusted parameter is required if it is not Gaussian. In practice, GDM is not suitable for reservoirs with complex geology that cannot be obtained from a Gaussian distribution (Caers, 2007).
Probability Perturbation method is similar to Gradual Deformation method; however, it does not rely on any assumptions about the statistics of the adjusted property and therefore is more flexible and suitable for complex geology. PPM uses the local constrained probability functions from the sequential simulation. The local probability function of an undefined value in a grid block is constructed from the data derived from the neighbouring grid blocks. PPM perturbs the probability of a given event on the grid block using a conditional probability function. The main difference in perturbation process between GDM and PPM is that Gradual Deformation perturbs physical properties at the grid block directly, whilst Probability Perturbation perturbs probabilities at the grid block which results in a perturbation of the physical properties after a sequential simulation.

The local conditional probability functions in PPM can be derived from a training image representing a realistic geological scenario, assuming that the training image is known (Caers, 2003). However, in real field applications, the geological scenario can be uncertain. Caumon et al. (2004) presented an example where the choice of the geological scenario has a stronger influence on the evaluation of reservoir uncertainty than the stochastic variation of local properties distribution. To account for uncertainty in geological scenarios, Suzuki et al. (2006) proposed an automatic history matching method to model reservoir properties by jointly perturbing the geological scenario and the local spatial distribution of reservoir properties.

They used the distance measure (section 2.6) to search for the history matched models efficiently from a large predefined set of model realisations from multiple alternative training images (“prior uncertainty space”). In this study, they demonstrated the benefits of accounting for multiple geological scenarios on a 2D synthetic case study (consisting of meandering channel system with two facies).

Both PPM and GDM require a large number of forward simulations which may be very computationally expensive. Also, they produce only one posterior model, therefore a large number of optimisations with different initial models need to be done to estimate the uncertainty.

Several approaches were proposed to both decrease the computational time and to minimise the loss of realistic geological representation. Rojas (2013) used Machine Learning techniques to, first, model realistic relationships between geometrical variables
(width, thickness, etc.) of geobodies during facies modelling, and second, to reject models with an unrealistic relationship between geobody parameters during history matching. This approach allowed to ensure that only geologically realistic models were history matched and used for the reservoir future performance prediction. It also resulted in decreasing the computational time because the rejected unrealistic models were not simulated.

An alternative way to perform history matching is by assimilation methods. Ensemble Kalman Filter (EnKF) is a particular example of ensemble assimilation methods that has been widely applied to history matching problems (Evensen et al., 2007; Gu et al., 2005; Haugen et al., 2008). EnKF is a Monte Carlo version of the classic Kalman filter. EnKF uses an ensemble of samples to represent necessary statistics, such as covariance of model parameters (e.g., porosity, permeability) and the correlations between model parameters and observations (Lin et al., 2017). EnKF runs the simulation model one time step at a time, updates the model parameters and only then moves to the next time step (Evensen, 2007). Consequently, the EnKF is suitable for real-time data assimilation allowing updating of the ensemble continuously when new data are available (Lin et al., 2017).

Sarma et al. (2009) stated that the main drawback of the standard EnKF is that EnKF only uses the covariance and cross-covariance between the random fields (e.g., permeability) and observations, thereby only preserving two-point statistics. However, some geological structures (e.g., channels) are depicted more realistically with Multi-Point Statistics. There is a concern that EnKF is unable to maintain geological structures. Therefore, Sarma et al. (2009) proposed a generalised EnKF using kernels, which is capable of representing non-Gaussian random fields (e.g., permeability) characterised by Multi-Point geostatistics.

Recently, Chang et al. (2016) proposed a workflow that combines new facies parametrisation and adaptive Gaussian mixture (AGM) filter to preserve realism in facies modelling during history matching. They also used dummy wells to update facies-log data and regenerate facies model realisations honouring the updated geological information. AGM assumes a more flexible Gaussian mixture model of the forecast ensemble comparing to EnKF. They applied this workflow to the Brugge field and reported a satisfactory data match and maintenance of geological facies representation.
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

Lange et al. (2012) addressed the problem of preserving geological realism during history matching with the Frequency Matching (FM) method. The FM method uses a dissimilarity measure to compare the solution models of the inverse problem to the training image that acts as an expected model. The solution models may be penalised depending on how much the frequency distribution deviates from the expected models. In other words, the FM method is used to evaluate how geologically reasonable is a solution model compare it to the training image which represents the available prior information. However, this technique requires many forward simulations.

To minimise the number of forward simulations Melnikova et al. (2015) proposed a smooth formulation of Multi-Point Statistic coupled with gradient-based optimisation. The solution of the inverse problem is found by minimising the sum of two mismatches: the mismatch with the history data and the mismatch with the Multi-Point Statistics of a training image. However, the convergence of the gradient-based optimisation depends on the starting guess, therefore, additional data, for example, seismic data could be beneficial.

Different reparameterization techniques were proposed to reduce the number of parameters for history matching. For example, zonation method is proposed to divide the reservoir into a number of zones, in each of which the unknown properties are either treated as uniform or modified by multipliers (Carter et al., 1974).

Another method, known as Karhunen-Loeve (K-L) expansion (Huang et al., 2001) has been introduced to approximate geological models as a linear combination of orthogonal basis functions. As a linear K-L approximation, Principle Component Analysis methods (Chen et al., 2012) have been applied to reduce the number of parameters for reservoir characterisation, yet persevering the main characteristic of prior geological distribution. PCA follows the similar idea of Eigenvalue decomposition to approximate a covariance matrix from a given training ensemble of realisations. Eigenvalue decomposition of the covariance matrix, required for each parameter update, is usually prohibitively expensive for high-resolution models (Bhark et al., 2011).

To overcome the expense of Eigenvalue decomposition, Li and Cirpka (2006) assumed periodicity of a random fluctuation in the property heterogeneity used for model calibration, since the eigenfunctions of its corresponding periodic covariance are known
analytically. The corresponding eigenvalues, required for reconstruction of the random function, are then computed by fast Fourier transformation (Bergland, 1969).

In this work, we use combine Particle Swarm Optimisation algorithm combined with new hierarchical approach and metric space classification to provide geologically realistic history matching.

2.6 Similarity distance and metric space

In mathematics, a metric space represents a set where a distance (called a metric) is determined between elements of the set (Fenwick and Batycky, 2011). Figure 2.7 represents a conceptual illustration of metric space, where model realisations are the elements of the set. Model realisations are represented as points in a metric space, where the distance between any two points is defined by a distance function and is called “similarity distance”.

This distance function indicates how close (or far apart) a pair of models is to each other in the metric space. Simultaneously, the similarity distance represents how similar (or dissimilar) a pair of models is: i.e. if the models are close to each then they are more similar (for example, in terms of visual facies representation) than the ones that are far from each other. The critical assumption of this approach is that a similarity in spatial properties between the realisations represents the similarity in production response of the realisations (Suzuki et al., 2006).

Suzuki and Caers (2008) proposed to use a static distance between the models as a dissimilarity measure in a metric space. They showed that there is a correlation between the static model distances and their flow response: models close to each other produced a more similar dynamic response. Suzuki and Caers (2008) used Hausdorff distance (Huttenlocher et al., 1993) to measure the similarity of geometry between any two stratigraphic reservoir models. Hausdorff distance measures the greatest of all the distances from a point in one set to the closest point in the other set.

Suzuki and Caers (2008) chose Hausdorff distance because it is a well-established distance in pattern recognition and shape matching of objects problems. In this study, Hausdorff distance defined the parameter space for history matching where the well
matched models were searched by a stochastic search method (Neighbourhood algorithm). This method was capable of handling several alternative structural model interpretations during history matching, which are the major source of structural uncertainty.

Caers and Scheidt (2010) proposed the concept of metric space for modelling uncertainty. The first step of modelling in metric space requires the calculation of the distance to build a metric space for the initial set of models. The second step represents the projection of high-dimensional metric space to a low-dimensional metric space. The resulting metric space could be used for multiple purposes.

![Figure 2-7 Conceptual illustration of similarity distance in metric space between model realisations](image)

Caers (2008) and Scheidt and Caers (2008) used metric space to generate new models to solve a pre-image problem (which involves reconstructing the model obtained in high-dimensional feature space back to lower-dimensional input space). Scheidt et al. (2009) and Caers and Scheidt (2011) applied the clustering techniques in metric space to select a few representative models for uncertainty quantification. Scheidt and Caers (2009) and Caers and Scheidt (2011) performed sensitivity analysis and uncertainty assessment for models in metric space. Caers and Park (2008) and Park et al. (2008a), Park et al. (2008b) demonstrated updating of models in metric space when constraining to dynamic data by applying a metric EnKF.
Park et al. (2013) used metric space to determine which training images are inconsistent with the production data. These inconsistent training images were rejected and only consistent are were history matched using Probability Perturbation Method (PPM). However, the reality could be that no specific training image but rather a region between several training images in the metric space represents the reservoir best of all.

Rojas (2013) combined a metric space approach with machine learning classification to include three different training images into history matching and provide more flexibility to find model realisations that better match the history data. The classification algorithm identified the borders between three different geological concepts. The classified metric space (Figure 2.8) was used for sampling during history matching. The history matching results showed that the lowest values of misfit between simulated and observed data were obtained not from one specific training image but from a region represented by all three different training images.

In this thesis, we expand the idea of accounting for geological uncertainties and incorporating multiple training images into history matching. We propose a hierarchical approach to introduce geological uncertainties from different levels of geological data (depositional environment, geological scenario, geological parameters) in facies modelling.

Figure 2-8 Classification results of metric space: each class represents a different geological scenario (training image). Red square represents the region with low misfit models identified during history matching (from Rojas, 2013)
Chapter 2 Geological Reservoir Modelling and Uncertainty Quantification

We use a metric space approach to account for multiple geological scenarios and multiple geological parameters. We define the geological similarities and relationships between different geological scenarios by means of classification in the metric space (k-means clustering, Support Vector Machines, and Random Forest).

Classification of the metric space allows us to find geological relations between multiple scenarios. After classification, geological scenarios are no longer introduced into history matching as arbitrary categorical variables. Instead, we account for the geological relations between multiple scenarios found in the metric space and introduce them into history matching. Overall, we aim to find the appropriate combination of geological parameters and geological scenarios to improve geological realism by introducing a classification of metric space into history matching.

In the following chapters, we explain the proposed approaches in more details and present the results and discussion.
3 CHAPTER THREE. The Impact of Geological Parametrisation and Interpretation on History Matching and Uncertainty Quantification

3.1 Introduction
In this chapter, we discuss the influence of geological parametrisation and interpretation on the history matching and forecast results. Geological parametrisation and interpretation inform prior probability distribution in Bayesian framework history matching.

As mentioned earlier (Section 2.4) the prior distribution affects the resultant posterior distribution, because there is almost always insufficient data to dominate the prior. Therefore, it is important to provide a realistic geological prior distribution by accounting for the available geological information. A realistic geological prior distribution can be based on data from different sources: cores and seismic measurements of the reservoir or/and outcrop, analogue reservoir data, scientific publications, reports, and expert knowledge.

Prior information can vary from non-informative to highly-informative with a range in between. Non-informative priors are usually represented by a uniform distribution where the same probability is assigned to all the values in a defined range. Use of non-informative priors is usually justified in statistics to eliminate possible bias (Jaynes, 2003). However, when we define a parameter range, we already impose some information about the parameters and the reservoir and therefore, introduce some geological knowledge.

Informative priors can be moderately informative or highly informative. Moderately informative priors are used when the information about the parameter is limited. Highly informative priors are used when fairly information about the parameter is available.

Informative priors can improve the results of history matching and reservoir predictions. For example, Arnold (2008) used informative priors in history matching by modelling the prior information of fluvial channel geometry with the equation that provides the relation
between channel width and thickness. He generated a 2D region that represented only realistic combinations of channel width and thickness. Arnold used this region to select only realistic channel width and thickness combinations for history matching and rejected the unrealistic ones. This allowed the simulation time to be decreased (by rejecting and not simulating the unrealistic ones) and to improve the reliability of the prediction by accounting only for realistic models.

Rojas (2013) explored the similar idea to Arnold (2008). He used Machine Learning techniques to model realistic relationships between geometrical variables (width, thickness, etc.) to form models of geobodies. These models were used as informative or “intelligent” geological priors. Only models with realistic parameters combinations were selected for history matching. The results showed that in the case of informative priors, the lowest value of misfit was reached twice as fast compared to the non-informative, and the simulation time was 1.5 times faster with informative than non-informative.

Rojas (2013) also used multiple geological scenarios (training images) as informative priors that were included into the history matching framework. This approach provided more flexibility in history matching and allowed for geological uncertainties.

In this chapter, we are going to demonstrate the influence of non-informative geological priors on reservoir predictions on the Ainsa II turbidite channel case study. These priors will be used to perform history matching and uncertainty quantification under the Bayesian framework.

We also introduce the Particle Swarm Optimisation algorithm for history matching and NA-Bayes (NAB) technique for uncertainty quantification in this chapter.

A case study with informative priors will be demonstrated in Chapter 4 (West Coast of Africa case study).
3.2 Ainsa II Case Study

3.2.1 Outcrop interpretation

Ainsa II outcrop was chosen for this exercise because it has been studied before and detailed descriptions of it is available in the literature (Clark, 1995 and Thurmond et al., 2007). The most commonly held interpretation is that of Clark (1995). The outcrop is described as a deep-water channelized system; it consists of a number of successive stacked erosive channels. A detailed 2D simulation model of the outcrop (at a resolution of 2500 by 200 cells) was produced by the Genetic Units Project (GUP) project at Heriot-Watt University (Figure 3.1), capturing some specific features of the outcrop (for example, the offsets in the layering at the boundaries between the different channels). The simulation results of this model were used as synthetic historical production data and as the reference case to validate the forecast results.

A group of Reservoir Evaluation and Management (REM) MSc students from Heriot-Watt carried out an exercise designed to illustrate the influence of interpretation and parametrization on the history matching and forecasting results during a field trip in the Northern Spain in 2014. The students were divided into 3 subgroups (Team A, Team B, and Team C) to perform a description of the same outcrop – Ainsa II turbidite channel complex (Figure 3.1, top). The purpose of this exercises is to demonstrate that there are usually multiple possible interpretations of the same data (outcrop in this case) due to multiple uncertainties (in facies interpretation, facies proportion, petrophysical properties, etc.).

Each team was asked to describe the outcrop from a distance. Students needed to determine the depositional environment and to sketch the facies distribution model into a 50 x 10 grid (Figure 3.1). Each team also assigned the ranges of parameter distributions (porosity, permeability, etc.) for each facies to perform flow simulation of their models.

Each team delivered a different model (Figure 3.1) as the result of their interpretation of the same outcrop which was based on a limited amount of data they can gather from the visual observation from a distance.
Figure 3-1 Comparison of Ainsa II outcrop photo, high-resolution model, and facies interpretation models produced by Team A, Team B, and Team C. Each 2D model has an injector (I) and a producer (P) well at each side of the model (Courtesy of A. Gardiner, Heriot-Watt University)
The main differences between the three models are in:

1) depositional environment: Team C described the outcrop as fluvial, whilst Team A and B – as turbidite;

2) distribution and number of facies: Team A described the outcrop with 7 facies, Team B – with 4 facies, and Team C described with 5 facies.

These differences highlight the fact that the interpretations are often highly subjective. At this stage, the students had very limited data to base their interpretations on. Therefore, there were a lot of uncertainties, even in determining the depositional environment. Some depositional environments can be confused with others. For example, a turbidity environment is more likely to be confused with fluvial due to some geomorphic similarities: fluvial meandering channels look similar to deep-marine channels, however, their origins are very different (Figure 3.2). Nevertheless, it is very unlikely that fluvial environment may be confused with the lacustrine environment (Figure 3.3), due to distinguishing differences in bedding types, common rock type and rock structure, common fossils and so on. Availability of more data could help to reduce the uncertainty in a type of a depositional environment.

![Figure 3-2 Visual comparison between fluvial meandering and deep-marine channels. A: Meandering channel, Chubut River, Argentina (from Foix et al., 2012); B: Deep-marine channel, De Soto Canyon, Gulf of Mexico (from Posamentier and Kolla, 2003)](image-url)
Despite the differences identified in students’ interpretations, some common features are recognised in all interpretations, for example, all Teams identified the location of a massive sand facies (bright yellow) approximately at the same location.

To simplify the models for the dynamic simulation all shale facies were lumped together, resulting in 6 facies in total instead of 7 for Team A.

The facies models were assigned homogeneous petrophysical properties and were used to perform history matching and uncertainty quantification.

In this chapter, we discuss the uncertainties in outcrop interpretation and the influence of these uncertainties on history matching and reservoir predictions. However, the same ideas can be applied to seismic interpretation.

3.2.2 Simulation models for history matching

For this study, a 2D model with the 50 x 10 grid blocks was created. The reservoir model size is 950m long x 40m high, resulting in individual block sizes of 19m x 4m. The top reservoir was set at a depth of 2240m.

One vertical injector was put on the left side of the model and one vertical producer well was put on the opposite side (Figure 3.1).
Chapter 3 The impact of Geological Parametrisation and Interpretation on History Matching and Uncertainty Quantification

Waterflooding was performed using the total liquid rate control (preferred control) for both injector and producer (17 m³/day at reservoir conditions). The injector bottom hole pressure (BHP) was constrained to be no higher than 1500 psi. If injecting the specified rate caused BHP to go over the limit, the rate was reduced to meet the constraints. The producer BHP was left at the default value (atmospheric pressure).

Other properties used during the simulation are summarised in Table 3.1.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water-oil contact (km)</td>
<td>3000</td>
</tr>
<tr>
<td>Initial pressure (psi)</td>
<td>500</td>
</tr>
<tr>
<td>Fluid density (kg/m³)</td>
<td>Oil: 0.876</td>
</tr>
<tr>
<td></td>
<td>Water: 1</td>
</tr>
<tr>
<td>Fluid viscosity (cp)</td>
<td>Oil: 2</td>
</tr>
<tr>
<td>(at reference pressure 500 psi)</td>
<td>Water: 1</td>
</tr>
<tr>
<td>Formation volume factor</td>
<td>Oil: 1.15</td>
</tr>
<tr>
<td>(at reference pressure 500 psi)</td>
<td>Water: 1.01</td>
</tr>
</tbody>
</table>

Table 3-1 Reservoir properties used for the flow simulation

The simulation model was run for 820 days in total. The history matching was performed for the first 420 days, and the future reservoir prediction was carried out for the remaining 400 days.

Each model created by the students was history matched to a synthetic production data (oil rate and water rate) from a detailed high-resolution 2D model produced by GUP project mentioned earlier (Figure 3.1). The full simulation model (820 days) production data were used as a reference case, so-called “Truth case” to validate the model predictions.
3.3 History Matching

3.3.1 Case study parametrisation

To perform history matching of facies models, the following parametrization (Table 3.2) was used: uniform distributions across specified ranges of porosity and permeability for each facies, global vertical permeability, and relative permeability defined by the Corey exponents (Dake, 1983).

All other properties were assumed to be known and fixed during the simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global porosity</td>
<td>(0.05-0.3)</td>
</tr>
<tr>
<td>Global permeability, mD</td>
<td>(10-1000)</td>
</tr>
<tr>
<td>Global vertical permeability, mD</td>
<td>(0 - 1)</td>
</tr>
<tr>
<td>Global Corey coefficients</td>
<td>(1.5 - 6)</td>
</tr>
</tbody>
</table>

*Table 3.2 Parameter values for the parameterization applied to the Ainsa II facies models during history matching*

3.3.2 History matching algorithm

**Particle Swarm Optimisation** (PSO) was chosen for this study as a more recent generation of stochastic optimizers that demonstrates better history matching results (lower values of misfit function) in fewer number of iterations in previous studies, comparing to Genetic Algorithm (Mouser and Dunn, 2005) and Neighbourhood Algorithms (Mohamed et al., 2010a). PSO has a small number of parameters to adjust which makes it simple in both formulation and computer implementation (Mohamed, 2011).

PSO explores the parameter space by a population of particles that search for an optimum solution (Eberhart et al., 2001). These particles replicate the behaviour of birds flocking or fish schooling looking for the best place for food in nature. In PSO each “bird” (or particle) is a single solution in the search space, and “food” is the global minima.

Each particle has a random position in the search space at the beginning. Then the position of each particle is changed according to the velocity of the particle $v_i$ value and the memory of the previous best position $pbest_i$. 
The *pbest* vector of the particle with the best fitness in the neighbourhood is denoted as *gbest*.

The particle’s velocity $v_i$ is updated as follows:

$$v_i^{k+1} = \omega v_i^k + c_1 r_{1i} \times (pbest_i^k - x_i^k) + c_2 r_{2i} \times (gbest^k - x_i^k) \quad (Eq. 3.1)$$

where $x_i^k$ is the current position of the particle $i$ at iteration $k$, $\omega$ is inertial weight, $c_1$ and $c_2$ are user-defined parameters that define particle’s attraction to its own known best position $pbest_i^k$ and the global best position $gbest^k$ respectively up to iteration $k$, $r_1$ and $r_2$ are two random vectors with components, $r_{1i}$ and $r_{2i} \in (0,1)$. These random vectors are introduced to represent the stochastic unpredictable component of swarms’ behaviour in nature.

The particular implementation of PSO used in this thesis employs a technique called “indirect acceleration” (A. Reynolds, private communication). This means that the random numbers $r_1$ and $r_2$ vary by parameter. The consequence of this choice is that PSO is no longer constrained to the space spanned by the initial random samples, and can explore variations in other dimensions as well, thereby enhancing exploratory performance.

The velocity update consists of three main components (Engelbrecht, 2006):

1) Inertia – defines the particle’s tendency to move in the same direction;

2) Memory – linear attraction towards the global best position;

3) Social knowledge - linear attraction towards the particle’s best position.
The particle’s position is added to the particle’s velocity to obtain the new position.

\[ x_i^{k+1} = x_i^k + v_i^{k+1} \quad \text{(Eq. 3.2)} \]

where \( x_i^{k+1} \) is the new position of the particle \( i \) at iteration \( k+1 \), \( x_i^k \) is the current position of the particle \( i \) at iteration \( k \), and \( v_i^{k+1} \) is the velocity update.

The particle’s position is then updated regardless of the progress with the objective function.

Figure 3.4 illustrates the velocity and the particle’s position update process.

Equation 3.3 represents the update equation of the personal best position \( p_{\text{best}}^i \).

\[
p_{\text{best}}^{i+1} = \begin{cases} p_{\text{best}}^i & \text{if } f(x_i^{k+1}) \geq f(p_{\text{best}}^i) \\ x_i^{k+1} & \text{if } f(x_i^{k+1}) < f(p_{\text{best}}^i) \end{cases} \quad \text{(Eq.3.3)}
\]

where \( f \) is the objective function (that is being minimised), and \( k \) is the iteration (generation) number.
Chapter 3 The impact of Geological Parametrisation and Interpretation on History Matching and Uncertainty Quantification

The PSO workflow is performed by the following steps (Mohamed, 2011):

1. Initialise the initial swarm models \((N_{\text{init}})\) by assigning the particles at random locations in parameter space. Each particle is also assigned a plausible random velocity.

2. Solve the forward problem (dynamic simulation) and obtain the relevant objective function value for each model (particle).

3. Update the position and value of \(p_{\text{best}}\) for each particle. If the current objective function value of one particle is better than its \(p_{\text{best}}\) value, then its \(p_{\text{best}}\) value and the position are replaced by the current objective function value and position, respectively as in Eq. (3.3).

4. Find the global best objective function value and the corresponding best position \(g_{\text{best}}\) among all swarm's \(p_{\text{best}}\). Update if appropriate.

5. Update the velocities and positions of all the particles using Eqs. (3.1) and (3.2).

6. Repeat steps 2 – 5 until a stopping criterion is met (e.g. the maximum number of iterations is reached or a sufficiently good objective function value is obtained).

The main parameters used to set up PSO for this work are summarised in Table 3.3. The values of the parameters are chosen based on studies done by (Shi and Eberhart, 1998, Mohammed, 2011, and Rojas, 2013).
Table 3-3 Summary of parameters used for Particle Swarm Optimisation algorithm (Modified from Rojas, 2013). (*flexible PSO provides a more freely move of particles in the search space (Kathrada, 2009))

<table>
<thead>
<tr>
<th>PSO parameter</th>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of particles</td>
<td>Number of models used in the optimisation</td>
<td>20</td>
</tr>
<tr>
<td>Group size</td>
<td>Number of particles for each group of particles</td>
<td>4</td>
</tr>
<tr>
<td>Initial inertia</td>
<td>Tendency of the particles to move in the same direction</td>
<td>0.9</td>
</tr>
<tr>
<td>Initial decay</td>
<td>Weight used to reduce the initial inertia in every step</td>
<td>0.9</td>
</tr>
<tr>
<td>Cognitive component</td>
<td>Linear attraction towards the best position ever found by the particle</td>
<td>1.333</td>
</tr>
<tr>
<td>Group component</td>
<td>Linear attraction towards the best position found by a group of particles</td>
<td>1.333</td>
</tr>
<tr>
<td>Social component</td>
<td>Linear attraction towards the best position found by any particle</td>
<td>1.333</td>
</tr>
<tr>
<td>Energy retention</td>
<td>Allows the particles to retain the strategy used in the previous steps</td>
<td>0.8</td>
</tr>
<tr>
<td>Particle behaviour</td>
<td>Select between conventional and flexible behaviour of the particles</td>
<td>Flexible*</td>
</tr>
</tbody>
</table>

**3.3.3 History matching results**

The automated history matching was performed for 420 days using PSO for three different models provided by three student teams (Team A, Team B, and Team C). A least squares objective function (Eq.3.4) was used to minimise the discrepancy between the simulated and historical production data (Field Oil Production Rate (FOPR) and Field Water Production Rate (FWPR) at surface conditions).

\[
M = \frac{1}{2} \left( \sum_{i=0}^{N} \frac{(Obs_i - Sim_i)^2}{\sigma_i^2} \right) \quad \text{(Eq. 3.4)}
\]

where \(Obs_i\) is the observed data, is the \(Sim_i\) simulated result at time \(i\), \(N\) is the number of data points, and \(\sigma_i^2\) is the measurement error.
The constant value of $\sigma_i^2$ is equal to 1 m$^3$/day. This $\sigma_i^2$ value was chosen to represent the measurement error for Ainsa II study.

Figures 3.5 – 3.7 show an example of parameter evolution (porosity for each facies) for different models (Team A, Team B, Team C) through history matching. We could see that there is a reduction in uncertainty during history matching; the reduction is greater for some facies than for the others, for example, porosity range for facies 5 (po5, Team A) is collapsing from [0.05 – 0.3] to [0.25 – 0.3], where the majority of the models are gathered (Figure 3.5). We could also observe that porosity for some facies is converging better than for the others, for instance, porosity for facies 3 and 4 (po3 and po4, Team B) converges much better than porosity for the facies 1 and 2 (Figure 3.6). We also can see that the converged porosity range is different for different facies, for example, it is [0.13 – 0.17] for facies 1 (po1, Team C), [0.21 – 0.26] for facies 2 (po2, Team C), and [0.15 – 0.25] for facies 3 (po3, Team C) (Figure 3.7).

Figure 3-5 Example of porosity variation for each facies during history matching (Team A). Other parameters varied are not shown
These results show that different model interpretation (Team A, Team B, Team C) lead to different history matching parameter combinations that deliver good matches of compatible quality (Figure 3.8). Figure 3.8 shows five “best” model with the lowest values of misfit for oil and water production rates after 500 iterations. The model responses are very similar, therefore, only one model is visible in Figure 3.8 for each team.
The error bar (Figure 3.8) represents the accepted variation of the simulation results. In this case, it is equal to 1m$^3$/day and goes to both sides of the history data values ($\pm 1$). We understand that the value of FWPR cannot be negative but the error bar goes to $-1$ when the observed FWPR equals to zero. We could have used a different definition, for example, a log-normal distribution, but it would have made no difference to these results.

The models were history matched on historic surface rates (FOPR and FWPR), therefore, a pressure effect was introduced to the misfit. Formation volume factor (Bo) is used to calculate surface rate values (surface rate = reservoir rate/formation volume factor). Formation volume factor is pressure dependent and 1/Bo is linear with pressure (in 3DSL streamline simulator). The models that reached the pressure limit (or early breakthrough) had a high value of misfit (misfit > 10): 159 models for Team A, 175 models for Team B, and 78 models for Team C.

Models that have no water production and have a pressure within the limits were further discriminated because of a slight pressure effect on the surface rates (Figure 3.8).

The majority of the models selected by NAB for the forecasting did not have water breakthrough and stayed within the pressure limits. The misfit values of these models are very small: from 0 to 0.1. However, several models reached the pressure limit: 11 models for Team A, 2 models for team B and 7 models for Team C. The values of misfit for these models are from 7 to 162.

**3.3.4 Geological realism during history matching**

In this case study, we introduced geological realism in terms of the observed correlation between porosity and permeability values. Figure 3.9 (left) shows that there is a monotonic relation between porosity and permeability values prior to history matching in each team’s model: when the value of porosity increases the value of permeability increases too.
Chapter 3 The impact of Geological Parametrisation and Interpretation on History Matching and Uncertainty Quantification

Figure 3-8 Results of automated history matching (five best models, models are very similar, therefore only one model is visible) for oil and water production rates of three different model interpretations from Team A, Team B, and Team C. These models have no water production and have pressure within the limits, therefore were further discriminated because of a slight pressure effect on the surface rates.

However, during the history matching the model parameters are tuned to minimise the misfit function and the monotonic relation between porosity and permeability is lost. Figure 3.9 (right) shows no monotonic correlation between the porosity and permeability values for all three models. These results illustrate the problem of losing geological realism in petrophysical properties during history matching. Though there is no apparent effect of this on history matching results, geologically unrealistic history matched models may affect the reliability of the reservoir forecast. Model predictions will be discussed in the next sections.
Figure 3-9 (left) Linear correlation between porosity and permeability values before history matching; (right) No linear correlation between porosity and permeability values after history matching

3.4 Uncertainty Quantification in Reservoir Predictions

3.4.1 Posterior probability distribution

Posterior Probability Distribution (PPD) as mentioned in Section 2.4 represents the updated probability distribution based on prior knowledge and observation data. The process of calculating PPD requires multiple history matched models, usually referred as an ensemble of reservoir models.

There are several ways to determine PPD:

1) Local categorisation of PPD around the Maximum Likelihood (ML) or Maximum a posteriori (MAP). Maximum Likelihood (ML) is the best history matched model
from the ensemble. It is called Maximum a posteriori (MAP) if the prior term is incorporated. If the objective function is multi-modal, it is possible to use more than one equally good solutions. In this case, multiple ML/MP solutions can be found.

2) Using the subset of an ensemble of history-matched models: Randomised Maximum Likelihood (RML) (Oliver et al., 1996) and Pilot Point (PP) (Tjelmeland, 1997). In RML, a sample is drawn from the prior reservoir model, and another sample is drawn from the observed data. Then the pairs of reservoir model samples observed data samples are history matched individually and the objective function value is obtained. The objective function includes misfits from both of these history matches. The pilot point is an approximation of RML, in which model parameters are varied only in specific locations (pilot points) in the parameter space. Thus, RML is sensitive to the pilot point selection.

3) Using the complete ensemble of history-matched models: Rejection Sampling (RS) and Markov Chain Monte Carlo (MCMC) (Oliver et al., 1997). In Rejection Sampling (RS), reservoir models are generated independently from a simple distribution. Then they are accepted or rejected based on a decision function (Casella et al., 2004). Markov Chain Monte Carlo (MCMC) generates a model at each step of a chain by perturbing a model at the current state. The model is accepted or rejected based on the acceptance criteria (e.g. Metropolis-Hasting). The acceptance criteria for both RS and MCMC is defined in proportion to the likelihood of the models. It means that the PPD is obtained as a result of a long chain, which usually requires a huge number of iterations to perform.

More details and examples of these methods could be found in (Liu et al., 2001). Liu et al. (2001) compared the ability of these methods to approximate PPD. They showed that only RS and MCMC type methods should be considered suitable for sampling the PPD as they sample the PPD correctly, whilst the others provide only an approximation of the PPD.
RS and MCMC are computationally expensive as they do not have an optimisation step. Therefore, in this thesis, the methodology called NA-Bayes (NAB) (subsection 3.4.2) that is based on MCMC and have the optimisation step.

Figure 3.10 represents the Bayesian framework used in this thesis. It consists of two parts: 1) Sampling the parameter space, 2) Posterior inference.

For the sampling part, we use PSO (Section 3.3.2) as an optimisation technique to generate multiple history matched models.

The second part allows us to determine the PPD of the predictions by an MCMC based posterior sampling algorithm. As mentioned above, we use NAB to estimate the PPD of reservoir predictions. NAB determines the PPD from the complete ensemble without running further reservoir simulations, which makes it computationally efficient.

3.4.2 NA-Bayes algorithm

The NAB algorithm proposed by Sambridge (1999b) differs from MCMC by a resampling option. NAB selects from an existing ensemble generated by sampling in parameter space using a search algorithm (e.g. PSO). NAB infers the information from the complete ensemble rather than a best-fitting model with the lowest value of misfit.

NAB uses Voronoi cells (Aurenhammer, 1991) to represent the model space and to interpolate the PPD of unknown points in the search space (Hajizadeh, 2011). The misfit surface is assumed to be constant over each Voronoi cell surrounding a particle. NAB converts the posterior probability densities to a posterior probability for the model at the centre of the Voronoi cell (Sambridge, 1999b). Consequently, the forward simulations are only required for the models that are selected by NAB.

NAB uses a Gibbs sampler which is a special case of the Metropolis-Hasting algorithm (Robert and Casella, 2005) to construct and approximate the PPD, where steps are made in a particular pattern (one parameter at a time). NAB uses interpolated likelihood surface represented by discontinuous Voronoi tessellation in high-dimensional parameter space.
Figure 3.11 illustrates the steps required to perform NAB using the Gibbs sampler (Sambridge, 1999b):

1) Choose the initial start point (B) (typically the lowest misfit or maximum likelihood model) and perform a series of random walks (Kaye, 2008) along each parameter axis in turn.

2) Produce a conditional probability distribution function (PDF) by calculating the intersection points of each Voronoi cells with each cut line for each axis (for example, the XX’ or YY’ cut line on Figure 3.9). The probability is determined as the product of the PPD value and the width of the intersection.

3) Each walk is performed on the selected axis by generating a uniform random deviation from the conditional PPD along the axis. The proposed step $x_i^p$ is accepted or rejected based on following equation:

$$ r \leq \frac{P_{NA}(x_i^p|x_{-i})}{P_{NA}(x_i^{max}|x_{-i})} $$

(Eq. 3.5)

where $P_{NA}(x_i^{max}|x_{-i})$ is the maximum value of the conditional PDF along the selected axis, and $r$ is a second random variable, $r \in (0, 1)$. 

74
4) If step 3 is rejected, then the process is repeated until a step is accepted.

5) The Gibbs sampler goes on by generating next steps and cycles through each parameter axis in turn. An iteration is completed when all dimensions have been cycled throughout once.

6) After many independent walks starting from different locations, the constructed conditional PDF is considered to be a good approximation to the true posterior distribution.

The output of NAB is a list of Voronoi cells visited, along with a visit frequency that equals to the posterior probability of the model at the centre of the cell. These models are then used to generate the forecast curves.

The accuracy of NAB depends on the complexity of the misfit surface and the size of samples. If there is a complex misfit surface (e.g. large change of misfit value within a short distance) or a limited number of samples (e.g. large Voronoi cells), then the NAB approach should be used carefully.

Figure 3.12 illustrates how the NAB results are used to quantify uncertainty in the reservoir forecast. NAB selects a subset of models from the history matched samples; this subset consists of the models visited during the NAB Gibbs sampling, and each model has an associated probability. We use all of those models in the forecast stage. Reservoir simulations of selected models are performed to receive production forecast. Each of these simulated models is assigned with the PPD value. Then we sort the models by PPD and construct a cumulative distribution at each time step. From the cumulative distribution, we obtain the values of production response corresponded to 10%, 50%, and 90% probability. These production values are used to build the credible intervals (P10-P50-P90) to represent the uncertainty range in the forecast.

More details and examples of NAB application could be found in Demyanov et al., 2008, 2004; Erbas et al., 2007; Hajizadeh et al., 2009; Mohamed et al., 2010a; Subbey et al., 2004.
3.4.3 NA-Bayes prediction results

A NAB chain for 10,000 random walk steps was run for the history matched model ensembles of Team A, Team B, and Team C. A burn-in period, a specified number of iterations at the beginning of a random walk that is discarded to ensure the independence of the resulting distribution from the starting point, was 5,000. It means that the first 5,000 models in the chain were discarded. NAB set up parameters are summarised in Table 3.4.
Chapter 3 The impact of Geological Parametrisation and Interpretation on History Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of chains</td>
<td>Quantity of chains</td>
<td>1</td>
</tr>
<tr>
<td>Chain length</td>
<td>Number of random walks</td>
<td>10,000</td>
</tr>
<tr>
<td>Burn-in period</td>
<td>Number of discharged iterations</td>
<td>5,000</td>
</tr>
</tbody>
</table>

Table 3-4 NAB set up for Ainsa II forecasting for Team A, Team B, and Team C models

The reservoir forecast was performed for 820 days. Figure 3.13 shows the uncertainty intervals (P10-P50-P90) (probability of exceedance) for the oil and water production rates for all three models (Team A, Team B, and Team C). We could see that even all three models had very similar results during history matching (Figure 3.8), they result in very different forecasts.

The high-resolution model mentioned before (produced by GUP, Figure 3.1) was used as “Truth case” to validate the predictions. Although we could observe that each of these forecast ranges more or less encapsulate the “truth case” model, the spread of the ranges (P10-P50-P90) of each model varies significantly. For example, Team A model predicts water breakthrough after 29 days, whilst Team B and Team C models after 580 days.

The risk, in this case, is to base the decision making on one of the model’s forecast results as there is no guarantee that the particular prediction represents the realistic flow behaviour.

This simple example highlights that we cannot evaluate how good a model interpretation is based only on the results of history match. There is also a risk to lose geologically realistic relationships between the parameters during history matching, which may affect the credibility of the forecast results.

This example also represents the importance of considering multiple geological interpretations to account for essential geological uncertainties and produce a more reliable forecast.
Overall, this study leads us to the key research question: whether imposing geological realism would improve the quality of forecast or not? Next chapters of this thesis are dedicated to this question.

Figure 3.13 Forecast results (probability of exceedance) for the oil and water production rates for all three models (Team A, Team B, and Team C)

3.5 Summary
In this chapter, we showed an impact of geological interpretation and parametrization on the history matching and uncertainty quantification results. We used three different Ainsa II outcrop interpretation to perform history matching with Particle Swarm Optimisation (PSO) algorithm. All three different models (based on different geological interpretation) resulted in a good history match for oil and water production performance.
During the history matching parameters tuning, the geologically realistic relations between the parameters were lost. Though there was no apparent effect on history matching results, it is possible that geologically unrealistic history matched models may affect the reliability of the reservoir forecast. This leads to a research question examined in this thesis.

We used NA-Bayes (NAB) algorithm to determine the Posterior Probability Distribution (PPD) of the reservoir forecast. NAB resamples from the history matched ensemble of models and results in a fewer number of models than the history matched ensemble. This fact makes NAB a computationally efficient technique to calculate production forecast.

We used the NAB results to calculate the forecast uncertainty (P10-P50-P90). All three models resulted in very different forecasts. To validate the uncertainty intervals, we used the “Truth case” model of Ainsa II outcrop. All three forecast uncertainty intervals more or less encapsulate the “truth case” model; however, the spread of these intervals varies significantly.

The main conclusions of this case study are:

1) Multiple different model interpretations can result in equally good history match. Therefore, we cannot judge the quality of the interpretation based on the results of history match only.

2) A good history match doesn’t guarantee a reliable forecast. There is a great risk in decision making on a single model forecast result.

3) There is a risk to lose geologically realistic relationships between the parameters during history matching. It may affect the credibility of the forecast results.

4) Considering multiple geological interpretations during history matching may help to account for essential geological uncertainties and improve the reliability of the forecast.
CHAPTER FOUR Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

4.1 Introduction

In previous chapters we discussed the following problems: 1) underestimation of geological uncertainties in static modelling, 2) preservation of geological uncertainties and geological realism during history matching, and 3) different forecast results from different history matched model ensembles for the same reservoir.

In this chapter, we propose a hierarchical approach (Figure 1.4, chapter 1) to account for the essential geological uncertainties in facies modelling. The approach is hierarchical because it combines different geological uncertainties from different levels of geological data (depositional environment, geological scenario, geological parameters). The hierarchy helps us to account for relationships between the uncertainties from different scales, for example, a geological scenario represents a specific facies architecture, which should be characterised by a certain combination of geobody parameters to provide geological realism. We establish these relationships to create a hierarchy based on the available geological data and data interpretations.

Since the hierarchy establishes relationships between concepts at different scales, it allows us to enforce a greater degree of geological realism than other approaches on its own, such as Object modelling and Multipoint statistics, by accounting for multiple geological concepts and multiple facies modelling parameter combinations.

In this chapter, we develop the hierarchical approach on a West Coast of Africa (WCA, section 4.2) example. We use the available geological data and description to create a hierarchy of the geological uncertainties (subsection 4.3.1) and perform static geological modelling with MPS SNESIM (subsection 4.3.2 – 4.3.6). We use one of the model realisations as the synthetic production data (section 4.5) to perform history matching (section 4.6). This synthetic production data is also used as the Truth case model to validate the forecast results (section 4.7).
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

After applying the hierarchical approach, we can calculate the prior range of geological uncertainties in geological parameters. We incorporate this range into history matching to calibrate the models to historical production data.

In this chapter, we parametrize multiple alternative geological scenarios (Level 2 of the hierarchy) as categorical variables with uniform prior distribution for history matching. We also include geological modelling parameters (Level 3 of the hierarchy) as continuous variables with uniform distribution into history matching.

We sample from the created parameter space (Figure 1.5, chapter 1) to incorporate geological uncertainties in facies models (geological scenarios and geological modelling parameters) into history matching and uncertainty quantification of the reservoir.

4.2 West Coast of Africa reservoir

The proposed workflow can in principle be applied to any geological depositional environment or even expanded to multiple depositional environments. However, in this research, we consider the turbidite depositional environment case study.

4.2.1 Turbidite depositional environment overview

A turbidite channel system is a common type of sandstone deposit on the continental slope and has been recognised one of the most common types of hydrocarbon reservoirs in the deepwater settings (Weimar, 2000). Despite their prevalence and many years of study, both in industry and academia, characterization and predictability of this reservoir type are very challenging due to the three-dimensional complexity and diversity of channel systems (McHargue et al., 2011). There is an interest in improving characterization of turbidite channel systems, because of the high cost and high risks of discovering and production of the hydrocarbons.

Turbiditic systems are generally classified based on two main parameters (Reading and Richards, 1994; Richards et al., 1998):

1) The number of system feeders and their characteristics.
2) The size of the sediment particles, e.g., gravel-rich, sand-rich or mud-rich systems.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

The system feeders are the current suppliers for the system. They can be divided into three categories (Figure 4.1):

1) One point source, forming submarine fans (West Coast of Africa reservoir).
2) Multiple point sources, forming submarine ramps.
3) Multiple linear sources, forming slope aprons.

Submarine fans and submarine ramps rely on the stable feeders, such as rivers or deltas and tend to have an organised architecture. Slope aprons come from more ephemeral feeders, such as a slope failure and they do not have a clear organisation, which reduces their potential as reservoirs.

The following sections will focus on submarine fans because West Coast of Africa belongs to this category.

The system architecture of turbidites results from the alternation of erosion and deposition events from the shore to the oceanic basin. Sediment gravity flows (Middleton and Hampton, 1973) are the most important deep-sea processes in developing submarine fans. Channelized turbidity currents and associated debris flows are the two dominant types of sediment gravity flow responsible for transporting and depositing submarine fans (Shanmugam and Moiola, 1988).

Submarine-fan sequences mainly consist of turbidites that occur as channel-fill, lobes, and sheet sands (Figure 4.2). Submarine-fan channels can be recognised by their sedimentological and geophysical characteristics, for examples, by thinning- and fining-upward cycles (Mufti and Ricci-Lucchi, 1972; Ricci-Lucchi, 1975). Sometimes, the turbidity current overspills from the channel margins, due to the low difference in density between the turbidity current and the surrounding water (Figure 4.3). Such overspills are called levees or overbanks.

Submarine lobes result from the accumulation of sediments at a channel way out. Lobes can be at the mouth of the channels and are called terminal lobes. They are the most distal element of a turbiditic system. Avulsion lobes occur when an overbank breaches, which is a first step to a formation of a new channel path (Droz et al., 2007).
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

Generally, a lobe has an ellipsoidal and convex-up morphology. Lobes are usually several meters thick and tens of kilometres wide (Shanmugam and Moiola, 1988). The horizontal shape is more or less elongated depending on the confinement.

Sheet sands in turbidites can develop in broad channels and in fan-fringe areas. Sheet sands are common in the lower-fan and fan-fringe regions of large modern fans such as the Amazon, Bengal, Indus, and Mississippi (Shanmugam and Moiola, 1988). Depositional lobes often consist of sheet sands (Figure 4.2).

Next sections describe West Coast of Africa reservoir geological settings, its geological interpretation, and related uncertainties. In this work, we aim to capture channel and lobe morphology of a turbidite submarine fan to create geologically realistic models.

4.2.2 West Coast of Africa reservoir description

The West Coast of Africa (WCA) field is a heterogeneous clastic deepwater turbidite reservoir deposited on the steepest part of the slope in a slope-valley system (Figure 4.4). The reservoir is offshore in 1600 ft of water and 4600 ft below sea level (Park et al., 2013). WCA is deepening slightly from east to west (Caers et al., 2006).

Generally, petroleum reservoirs found in such high-energy locations (steepest part of the slope) tend to be more heterogeneous than the ones found lower down the slope (Figure 4.4). These reservoirs are associated with large drilling costs and significant risks.

The reservoir is formed by amalgamated and aggradational channel complex that fills the canyon cut (Hoffman et al., 2006). The channels at the top of the reservoir are more aggradational with defined meandering pattern that had been recognised from seismic data (Figure 4.5) (Hoffman et al., 2006).

The WCA reservoir is divided into four structural blocks by a number of faults: block 1, block 2, block 3, and block 4 (Figure 5.3). Block 1, 2 and 3 are in communication with one another, based on the production data; however, the faults reduce the transmissibility. Block 4 is not in communication with the other blocks (Park et al., 2013).
Figure 4-1 Examples of three turbidite categories based on the type of the feeder (Modified from Stow and Mayall, 2000)
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

Figure 4-2 Schematic view of a slope mini-basin and mud rich fine-grained submarine fans depicting main turbidite structural elements: channel, leveed channels, lobes and sheet lobes (modified from DeVay et al., 2000)

Figure 4-3 Conceptual model showing a turbidity current within a channel (from Shanmugam et al., 1993)
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

An aquifer exists in the east of the reservoir (Figure 4.6). A fault separates it from the rest of the reservoir (Park et al., 2013).

Water support comes from below the reservoir. The oil-water contact varies for the different blocks, but it has been described at around 5440 ft (Caers et al., 2006).

The reservoir dimensions are approximately 1 mile long, 0.5 of a mile wide, and 800 feet thick. The reservoir has 20 production (red) and 8 injection (blue) wells (Figure 4.3).

Figure 4.4: Schematic description of geological zones of slope-valley turbidite. This scheme also describes the location of the WCA (from Hoffman et al., 2006)
Four depositional facies were described from the well logs: channel sand (Facies 1), two poorer quality sand facies (Facies 2 and 3) and shale (Facies 4) (Scheidt et al., 2009). The total amount of sand is approximately 55% of total volume, where Facies 1 is
approximately 30% of the reservoir, Facies 2 and Facies 3 are 10% and 15% of the gross reservoir volume respectively.

The three sand facies were classified based on their petrophysical properties, measured in the wells. The average values of porosity and permeability for each facies is summarised in Table 4.1.

<table>
<thead>
<tr>
<th>Facies number</th>
<th>Porosity</th>
<th>Permeability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3</td>
<td>2000 mD</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>400 mD</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>100 mD</td>
</tr>
</tbody>
</table>

*Table 4-1* The average values of porosity and permeability for each sand facies

4.2.3 West Coast of Africa reservoir uncertainties

The main uncertainties of the reservoir are facies depositional setting (how the facies are distributed within the reservoir), facies architecture (geometry of geobodies), and facies proportions (what is the amount of each of the facies).

Multiple facies models realisations conditioned to data could be created to capture the uncertainties in facies depositional settings. Each realisation represents a different facies distribution within the reservoir.

The key uncertainties in facies architecture are related to channel facies sinuosity, channel thickness, and poorer quality sand facies (Facies 2 and Facies 3) architecture: these facies could be interpreted as channels or as lobes. Eleven geological scenarios were created by experienced geologists (courtesy of Stanford University and Chevron, Scheidt et al., 2009) to represent these uncertainties in facies architecture.

Figure 4.7 shows a 2D top view of 3D conceptual models, where four different facies are represented by a different colour: dark red (brown) – channel sand (Facies 1), light blue and yellow - poorer quality sand (Facies 2 and Facies 3), dark blue – shale (Facies 4).

These conceptual models also differ by the channel sinuosity. Channel sinuosity shows how much the course of a channel deviated from the shortest possible path (Gordon et al.,
Models 4, 5, 6, 7, 10 and 11 have high sinuosity channels, whilst models 1, 2, 3, 8 and 9 have low sinuosity channels.

Channel thickness is another criteria to differentiate between these conceptual models. Models 2, 3, 4, 5, 8, 9, and 11 have visually thicker channels than models 1, 6, 7, and 10.

All conceptual models can be divided into two big groups based on poorer quality sand (Facies 2 and Facies 3) architecture. The first group consists of models, where Facies 2 and Facies 3 are interpreted as meandering channels (model 1, 2, 3, 5, 7 and 11). The second group consists of models, where Facies 2 and Facies 3 are interpreted as lobes (model 4, 6, 8, 9 and 10).

Table 4.2 summarises the differences between conceptual models.

To account for the uncertainties in facies proportions, multiple models (conditioned to data) with different amount of each facies can be created.

In terms of geological realism within these uncertainties, we aim to obtain models reproducing the following features: 1) specific facies properties, such as channel continuity and sinuosity in channelized facies, 2) specific facies geometries, either channels or lobes, 3) specific facies proportions (defined from the well data and analogue outcrop data). These are the necessary and sufficient conditions to represent geological realism in facies modelling.
Figure 4-7 Geological scenarios of West Coast of Africa reservoir generated by geologists (courtesy of Stanford University and Chevron). Each colour represents different facies: dark red (brown) – channel sand (Facies 1), light blue and yellow - poorer quality sand (Facies 2 and Facies 3), dark blue – shale (Facies 4)
4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>Conceptual model</th>
<th>Facies Architecture</th>
<th>Channel Thickness</th>
<th>Channel Sinuosity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>channels + channels</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>2</td>
<td>channels + channels</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>3</td>
<td>channels + channels</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>4</td>
<td>channels + lobes</td>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>5</td>
<td>channels + channels</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td>6</td>
<td>channels + lobes</td>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>7</td>
<td>channels + channels</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td>8</td>
<td>channels + lobes</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>9</td>
<td>channels + lobes</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>10</td>
<td>channels + lobes</td>
<td>low</td>
<td>high</td>
</tr>
<tr>
<td>11</td>
<td>channels + channels</td>
<td>high</td>
<td>high</td>
</tr>
</tbody>
</table>

Table 4-2 Summary of the differences between geological concepts (modified from Scheidt et al., 2009)

4.3 Hierarchical Facies Modelling with MPS

4.3.1 Hierarchical uncertainties

As mentioned in the previous section, the description of the WCA reservoir facies is subject to uncertainty. The new hierarchical approach is proposed to account for the facies uncertainties and introduce them into history matching. In this work, the hierarchy consists of 3 levels of geological data from the different scale (from the bigger scale to the smaller scale) (Figure 4.8):

1) the largest scale is represented by a depositional environment, analysed from the available geological data;

2) the medium scale has multiple alternative geological conceptual models, called training images in Multi-point statistics;

3) the smallest scale in this hierarchy accounts for facies architecture, proportions, and orientation.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

The first level – depositional environment – is a deepwater turbidite. Geological features of turbidite environment are propagated into geological concept models.

Eleven geological concept models were created by experienced geologists (courtesy of Stanford University, Scheidt et al., 2009) to account for geological facies uncertainties in WCA reservoir (Figure 4.7). The concept models are based on interpretation of the available geological data, outcrop descriptions, analogue data, and modern satellites images. WCA conceptual models have been described earlier (subsection 4.2.3) and the summary of the differences between them is provided in Table 4.2.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

The last level of hierarchy (Level 3) account for the uncertainties in facies architecture, proportions, and orientation within each geological concept. These uncertainties are expressed by geological facies modelling parameters and are based on analogue outcrop data. An analogue for WCA is the Ainsa outcrop (a weakly confined channel system) in the South-Central Pyrenees (Spain) (Clark, 1995; Clark and Pickering, 1996; Falivene et al., 2006; Larue, 2004; Thurmond et al., 2007). The description of this outcrop was used to define facies architecture through geobody geometries (channel width and thickness) and facies proportions from facies net-to-gross data.

The summary of the hierarchical levels, corresponding geological uncertainties, and modelling parameters used to represent these uncertainties is presented in Table 4.3.

4.3.2 SNESIM overview

The facies modelling was carried out with the Sequential Normal Equations Simulation (SNESIM) implementation of Multi Point Statistics (MPS) (Strebelle, 2001) to perform geological facies modelling. We chose SNESIM because (1) it allows us to condition model realisations to available well and seismic data; and (2) SNESIM modelling parameters allow us to introduce hierarchical uncertainties (Level 3) for each geological scenario (training image).

SNESIM is a sequential simulation algorithm, which relies on the idea of simulating each grid cell property (facies or petrophysical) sequentially along a random path; the simulation of cells later in the process is constrained by cells earlier simulated, along with well and/or seismic data (Caers and Zhang, 2004). SNESIM scans a training image using a pre-defined data template called “data event” to find the “replicates” of this “data event” on the training image. Then SNESIM retrieves the corresponding histogram of the central value of the “data events”. Once “data events” and their associated central values are extracted from the training image, SNESIM stores them in a data structure called a “search tree” (Knuth, 1997).
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>Level of hierarchy</th>
<th>Uncertainties</th>
<th>Modelling parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Depositional environment</td>
<td>No modelling parameters</td>
</tr>
<tr>
<td>2</td>
<td>Geological concepts (training images)</td>
<td>Training image number</td>
</tr>
<tr>
<td>3</td>
<td>a) geometry of geobody</td>
<td>a) channel width and thickness</td>
</tr>
<tr>
<td></td>
<td>b) direction of paleoflow</td>
<td>b) facies orientation</td>
</tr>
<tr>
<td></td>
<td>c) facies net-to-gross</td>
<td>c) facies proportions</td>
</tr>
</tbody>
</table>

Table 4-3 Summary of the of the hierarchical levels, corresponding geological uncertainties, and modelling parameters used to represent these uncertainties

SNESIM reproduces the structure of the training image (Figure 4.9) by reproducing its conditional proportions and, at the same time, SNESIM honours any available well or seismic data (Arpat, 2005).

Figure 4-9 An example of unconditional SNESIM facies realisation (right) based on the training image (left) (From Arpat, 2005)
4.3.3 SNESIM data conditioning

SNESIM realisations can be conditioned to well data, usually referred as “hard data”, and seismic data, referred as “soft data”.

Hard data conditioning usually refers to the exact reproduction of the original well data values. Geological continuity should be also respected around these well values to reproduce the realistic geology in the models. MPS methods (including SNESIM) achieve data conditioning by selecting whole training image events that are consistent with the available data. Although this conditioning approach gives satisfactory results, it could also result in geologically unrealistic realisations. For example, the SNESIM algorithm searches for an exact match in the search tree, dropping data nodes if necessary until such a match is found. Such dropping of data nodes leads to a loss of the large-scale information which may result in the generation of artificial discontinuities in the realisations (Arpat, 2005).

Soft data are typically used as a “guide” to modelling and are not reproduced exactly. Soft data are usually obtained from indirect measurements (e.g. seismic data). The soft data is coded into prior probabilities using well-calibration or clustering methods (Avseth, 2000; Caers et al., 2003; Pairazian, 1998). Once such soft-data-based prior probabilities are calculated, they are combined with the hard-data-based probabilities as reading from the training image (Journel, 2002; Strébelle and Journel, 2000).

Figures 4.10 and 4.11 represent WCA hard (well data) and soft conditioning data (seismic data). A seismic inversion cube was created and then transformed into a facies probability cube calibrated to the well data (Hoffman et al., 2005).
Figure 4-10 WCA well data color-coded by facies and used for hard data conditioning with SNESIM (courtesy of Stanford University)

Figure 4-11 WCA facies probability cubes for each facies used for soft data conditioning with SNESIM (courtesy of Stanford University)
4.3.4 SNESIM affinity and rotation

Affinity and rotation SNESIM modelling parameters allow us to create a variety of model realisations based on one training image. Changing the value of rotation parameter, we can obtain the different orientation of the facies. Figure 4.12 shows an example of affinity and rotation transformation of a training image.

Affinity parameter works as a scaling coefficient or a multiplier to modify the dimensions of geobodies in 2 (2D model) or 3 (3D model) directions. Affinity parameter could stretch (if the affinity value > 1) or shrink (if the affinity value < 1) geobody architecture. (Liu, 2006).

In this thesis, we use global affinity and rotation transformation to introduce geological uncertainty from Level 3 of hierarchical approach (Figure 4.6) for each WCA training image.

![Figure 4-12 Example of multiple different model realisations based on one training image: top left – training image, top right – new realisation with global rotation angle = 45°, bottom left - new realisation with global affinity x = 1 and y = 2, bottom right - new realisation with global affinity x = 1, y = 3 and global rotation angle = 30° (modified from Caers and Zhang, 2004)]
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

4.3.5 SNESIM set up
To account for the facies uncertainties of WCA discussed before (subsection 4.3.1, Table 4.3), we use MPS SNESIM to create multiple facies realisations (60 realisations) for each geological scenario (11 training images).

We modify facies width, thickness, channel sinuosity using affinity multipliers (Table 4.4). We intend to preserve realistic channel facies width-to-thickness ratio, based on analogue data of several turbidite reservoirs and outcrops, found in the literature (Abreu et al., 2003; Wonham et al., 2000; Wynn et al., 2007). In this study, the width and thickness of channel facies of training images were manually calculated. Then it was multiplied by affinity parameter value. The result values do not exceed the analogue turbidite ranges from the literature.

Global rotation parameter was used to account for uncertainties in paleoflow direction. We used the interpretation of the seismic map of the top layer of the WCA reservoir (Figure 4.5) to introduce the variation in facies orientation (Table 4.4).

The uncertainty in facies proportions was based on Ainsa outcrop analogue (Clark, 1995; Clark and Pickering, 1996; Thurmond et al., 2007). We use the values of low and high net-to-gross in sand and shale facies to create possible ranges for WCA facies proportions (Table 4.4). For each realisation, the sum of four facies proportions should be equal to 100%.

The random number parameter was not changed during the facies modelling.

<table>
<thead>
<tr>
<th>Geometry of geobodies</th>
<th>Facies orientation</th>
<th>Facies proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td>affinity x: (0.5; 3)</td>
<td>Global: (0° - 45°)</td>
<td>Facies 1: (25% – 34%)</td>
</tr>
<tr>
<td>affinity y: (0.5; 3)</td>
<td></td>
<td>Facies 2:  (9% - 14%)</td>
</tr>
<tr>
<td>affinity z: (0.5; 9)</td>
<td></td>
<td>Facies 3:  (6% - 12%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Facies 4:  (40% – 60%)</td>
</tr>
</tbody>
</table>

*Table 4-4 SNESIM parameter ranges (uniform distribution) to account for geological uncertainties in West Coast of Africa reservoir*
4.3.6 SNESIM results

SNESIM was used to create 60 3D facies realisations for 11 training images, which resulted in 660 realisations in total. Figure 4.13 shows some examples of facies realisations (top view) for each training image: the first realisation has no affinity or rotation parameters applied, and the second realisation has different values of affinity. The value of rotation is 0°, and facies proportions are Facies 1: 52%, Facies 2: 11%, Facies 3: 9%, Facies 4: 28% for these examples (Figure 4.13).

In this example, the scale of the model heterogeneiity relative to the flow means that the parameters have a greater influence than the random seed on the simulation results. Therefore, we have not performed modelling with different random seed in this work.

4.3.7 SNESIM discussion of the results

From the visual examination, we can distinguish between realisations that have Facies 2 and Facies 3 as channels and the ones that have them as lobes. We also can observe that some model realisations reproduce a given training image better than the others. For example, both realisations of training image 3 (TI 3) reproduce channelised architecture for all three sand facies and channels connectivity rather well, whilst realisations of training image 5 (TI 5) do not fully reproduce channelised architecture and channels connectivity for Facies 2 and Facies 3 (light blue and yellow).

As we modify the value of affinity, we receive model realisation with different channel width, so models from initially different training images look quite alike. For example, training image 5 (TI 5) has much wider channels than training image 1 (TI 1). However, when we apply affinity $y = 0.5$ for TI 5, we get a model realisation with narrower channels that looks more like a realisation from TI 1. This fact can affect our ability to distinguish between model realisations from different training images.

Not all SNESIM model realisations were able to preserve channel facies continuity, which could affect the flow simulation results.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>Training image</th>
<th>Realisations</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="TI 1" /></td>
<td><img src="image2" alt="Realisation 1" /> <img src="image3" alt="Realisation 2" /> <img src="image4" alt="Realisation 3" /></td>
</tr>
<tr>
<td>TI 1</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
<tr>
<td><img src="image5" alt="TI 2" /></td>
<td><img src="image6" alt="Realisation 1" /> <img src="image7" alt="Realisation 2" /> <img src="image8" alt="Realisation 3" /></td>
</tr>
<tr>
<td>TI 2</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
<tr>
<td><img src="image9" alt="TI 3" /></td>
<td><img src="image10" alt="Realisation 1" /> <img src="image11" alt="Realisation 2" /> <img src="image12" alt="Realisation 3" /></td>
</tr>
<tr>
<td>TI 3</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
<tr>
<td><img src="image13" alt="TI 4" /></td>
<td><img src="image14" alt="Realisation 1" /> <img src="image15" alt="Realisation 2" /> <img src="image16" alt="Realisation 3" /></td>
</tr>
<tr>
<td>TI 4</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
</tbody>
</table>
### Chapter 4. Hierarchical Geological Realism in Multiple Training Images History

Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>TI 5</th>
<th>TI 6</th>
<th>TI 7</th>
<th>TI 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image](TI 5)</td>
<td>![Image](TI 6)</td>
<td>![Image](TI 7)</td>
<td>![Image](TI 8)</td>
</tr>
<tr>
<td>affinity x,y,z:</td>
<td>1, 1, 1</td>
<td>1, 0.5, 1</td>
<td></td>
</tr>
<tr>
<td>affinity x,y,z and rotation:</td>
<td>1, 1, 1</td>
<td>1.25, 1, 1</td>
<td></td>
</tr>
<tr>
<td>affinity x,y,z:</td>
<td>1, 1, 1</td>
<td>2, 1, 2</td>
<td></td>
</tr>
<tr>
<td>affinity x,y,z:</td>
<td>1, 1, 1</td>
<td>0.5, 0.5, 0.5</td>
<td></td>
</tr>
</tbody>
</table>
Some of the model realisations resulted in visually unrealistic facies model representation (Figure 4.14). These geobody shapes might be caused by inappropriate affinity parameter combinations. Affinity parameters change width, thickness, amplitude and wavelength in the channelized geobodies. However, it is difficult to control and change these features of geobodies separately. That is why some of the models have the unrealistic facies representation.

However, even visually unrealistic facies model representations have continuous channels, which will allow fluid to flow, for example, training image 1 (TI 1) model.

<table>
<thead>
<tr>
<th>Training Image</th>
<th>Affinity Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI 9</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
<tr>
<td>TI 10</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
<tr>
<td>TI 11</td>
<td>affinity x,y,z: 1, 1, 1</td>
</tr>
</tbody>
</table>

Figure 4-13 Examples of facies model realisation with SNESIM for 11 training images: the first realisation has no affinity or rotation parameters applied, and the second realisation has different values of affinity.
realisation (Figure 4.14, in the middle). Therefore, we use SNESIM and the current parametrization for history matching, allowing the optimisation process tune the parameters to get the acceptable model match results.

![Figure 4.14 Examples of geologically unrealistic visual facies representations created with MPS SNESIM](image)

**4.4 Dynamic simulations**

Corner-point geological models were populated with homogenous petrophysical properties (Table 4.6) to perform dynamic simulations. The structure of the model is provided by the geological description (Section 4.2.2): reservoir consists of four blocks, separated by 3 faults; there is also an aquifer at the east side of the reservoir (Figure 4.6) (Hoffman et al., 2006). However, only aquifer is included in the simulations but not the faults.

The reservoir model has the dimensions of 78 × 59 × 116 cartesian grid-blocks. The dimensions of each grid-block are 75 m × 75 m × 8 ft. There are approximately 100,000 active grid-blocks in the model. The active grid-blocks are displayed in Figure 4.6.

The initial fluid contacts, pressure and fluid properties vary between different blocks (Figure 4.6), but the oil-water contact depth is approximately 5.440 feet, the initial pressure is approximately 2,300 psi, and the oil gravity is approximately 24 ° API. Densities for oil and water are 57 and 63 lbm/ft³, respectively (Park et al., 2013). The relative permeability of the reservoir is provided in Figure 4.15.

Reservoir properties used for the flow simulation are summarised in Table 4.5.
As was mentioned before, there are 20 production wells and 8 injection wells. Well positions are taken from the actual field (Figure 4.6). Waterflooding was performed using bottom hole pressure (BHP) control mode at producers and injectors.

The forward modelling was performed using 3DSL streamline simulator (Streamsim Technologies).

<table>
<thead>
<tr>
<th>Property</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>OWC</td>
<td>5440 ft (TVD)</td>
</tr>
<tr>
<td>Initial pressure</td>
<td>2300 psi</td>
</tr>
<tr>
<td>Fluids density (lb/ft³)</td>
<td>Oil: 57</td>
</tr>
<tr>
<td>(at reference pressure 2200 psi)</td>
<td>Gas: 0.05</td>
</tr>
<tr>
<td></td>
<td>Water: 63</td>
</tr>
<tr>
<td>Fluids viscosity (cp)</td>
<td>Oil: 3.03</td>
</tr>
<tr>
<td>(at reference pressure 2200 psi)</td>
<td>Gas: 0.01692</td>
</tr>
<tr>
<td></td>
<td>Water: 0.42</td>
</tr>
<tr>
<td>Formation volume factor</td>
<td>Oil: 1.1647</td>
</tr>
<tr>
<td>(at reference pressure 2200 psi)</td>
<td>Gas: 1.2222</td>
</tr>
<tr>
<td></td>
<td>Water: 1.02</td>
</tr>
</tbody>
</table>

Table 4-5 Reservoir properties used for the flow simulation
4.5 Synthetic Truth Case

In this thesis, we do not use the production data of the real reservoir. Instead, we create a so-called Truth case model to generate production data from it by forward simulation. The results of the forward simulation are treated as historical field production data. The advantage of this is (1) the data can be published without an issue of disclosing sensitive information, and (2) we can use Truth case model simulation response to evaluate the reliability of the forecast results.

Truth case 3D geological model was generated from a randomly picked training image (TI 9) with MPS SNESIM implementation. Figure 4.16 shows the top view of the Truth case model, where channelized sand is represented by dark-red/brown, lobes deposits are in light-blue and yellow, and shale is in dark-blue. Table 4.6 summarises SNESIM model parameters that were used to create Truth case model.

The forward modelling was run for 1216 days in total. Production data for 820 days were used as historical observed data during history matching; production data for the rest 396 days were used as the reference case to evaluate the quality of the reservoir predictions.

Figure 4-16 2D geological model of the Truth Case (top view), built from training image TI 9 with the following parameters: affx=1, affy=0.75, affz=1.25, rot=0°, facies proportions: shale=58% (blue), poorer quality sand 1=11% (light blue), poorer quality sand 2=9% (yellow), channel sand=28% (dark red/brown), conditioned to wells and seismic data
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

The reservoir model characteristics are the same as described in section 4.4. Reservoir properties used for the flow simulation are summarised in Table 4.5. The relative permeability of the reservoir is provided in Figure 4.15.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training image</td>
<td>TI 9</td>
</tr>
<tr>
<td>Facies architecture</td>
<td>Channels + lobes</td>
</tr>
<tr>
<td>Affinity x, y, z</td>
<td>1, 0.75, 1.25</td>
</tr>
<tr>
<td>Global rotation</td>
<td>0°</td>
</tr>
<tr>
<td>Facies proportions</td>
<td>(Facies 1, Facies 2, Facies 3, Facies 4)</td>
</tr>
<tr>
<td></td>
<td>28%, 11%, 9%, 58%</td>
</tr>
</tbody>
</table>

Table 4-6 Truth case model parameters used for facies modelling with SNESIM

Figure 4.17 represents the 2D top view of porosity and permeability of the Truth case model. Homogeneous values of porosity and permeability for each facies is shown in Table 4.7.

The reservoir simulation was performed with 3DSL streamline simulator (Streamsim Technologies). Figure 4.18 shows the field fluids production response for 1216 days.

Figure 4.19 shows how the Truth case model floods out at the end of the simulation (saturation of water at the final time step).

4.6 Assisted History Matching

4.6.1 Multi-Objective History Matching

Conventionally, a single-objective history matching operates on a single number (the misfit between observed and simulated values) which is obtained by summing the misfits of all the quantity of interest to form a global objective function (Christie et al., 2013). This objective function is then minimised to obtain best possible overall matches for different targets in well and field scale (Hajizadeh et al., 2011).
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

Figure 4-17 Porosity and permeability of the Truth case model, 2D top view

<table>
<thead>
<tr>
<th>Facies</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facies 1</td>
<td>Porosity: 0.3</td>
</tr>
<tr>
<td></td>
<td>Permeability: 2000mD</td>
</tr>
<tr>
<td>Facies 2</td>
<td>Porosity: 0.2</td>
</tr>
<tr>
<td></td>
<td>Permeability: 400mD</td>
</tr>
<tr>
<td>Facies 3</td>
<td>Porosity: 0.15</td>
</tr>
<tr>
<td></td>
<td>Permeability: 100mD</td>
</tr>
<tr>
<td>Facies 4</td>
<td>Porosity: 0.0001</td>
</tr>
<tr>
<td></td>
<td>Permeability: 0.0001mD</td>
</tr>
</tbody>
</table>

Table 4-7 Values of the porosity and permeability for each geological facies for dynamic modelling (courtesy Stanford University, Scheidt et al., 2009)
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

Figure 4-18 Truth case fluid production profile

Figure 4-19 Saturation of water of the Truth case model at the end of the simulation
History matching is essentially a multi-objective problem. Often, there are multiple match criteria in individual well and field in general, that can even be in conflict with each other. For example, when we try to improve the oil rate match in a particular well, we may deteriorate the gas rate match in another well (Hajizadeh et al., 2011). In multi-objective history matching, the objectives are divided into several match quality component. Each of this quality components is history matched individually. The overall history match misfits are obtained by summing all the misfit components at the end.

In this thesis, we used the multi-objective optimisation. The optimiser works using only the three objectives. In order to get the forecast, we need to compute the negative log posterior from Bayes theorem (Eq. 2.4). In that term, the negative log likelihood, or misfit, calculated as:

\[
\text{Misfit} = Obj_1 + Obj_2 + Obj_3
\]  

(Eq. 4.1)

The multi-objective approach provides two main improvements to history matching over a single-objective approach: (1) it encourages diversity in history matching, which should lead to improved forecasting, and (2) an increase in history matching speed (Christie et al., 2013).

Multi-objective optimisation evaluates the solutions in a so-called objective function space. Figure 4.20 represents the decision variable space, where we search the parameter values and the objective function space for a two-objective optimisation problem with three decision variables.
4.6.2 Pareto Dominance Concept

In multi-objective optimisation, we use the Pareto dominance concept (Pareto, 1971) to evaluate and compare the solutions generated during history matching.

The concept of dominance can be expressed by two conditions as follows (for minimising):

\[
\begin{cases}
  f_m(A) \leq f_m(B), \text{for all } m = 1, 2, \ldots, M \\
  f_m(A) < f_m(B), \text{for at least one } m = 1, 2, \ldots, M
\end{cases}
\]  

(Eq. 4.2)

where \( m \) is the number of objectives.

Equation 4.2 means that, if there exists a solution (A) that is better than a solution (B) for at least one of the objective functions, and this solution (A) is at least equal to (B) in terms of all objective functions, than solution (A) dominates solution (B).

Figure 4.21 shows a graphical representation of the Pareto dominance concept for a two-objective problem where both objectives are minimised. Six solutions for this two-objective problem have been obtained. Solution 1 is better (minimised) than the other solutions for both objectives; therefore it satisfies both Pareto dominance conditions (Eq. 4.2) which means that solution 1 dominates all other solutions. Solution 4 is equal to
solution 5 for objective 2, and it is better than solution 5 for objective 1, thus solution 4 dominates solution 5.

![Figure 4.21 An example of two-objective optimisation with six solutions illustrating the Pareto dominance concept (modified from Hajizadeh et al., 2011)](image)

A set of Pareto optimal solutions (those that are not dominated by any other feasible solutions) is called a Pareto front (Figure 4.22).

![Figure 4.22 Schematic view of a Pareto front for a two-objective optimisation](image)
Whilst solving any multi-objective optimisation, we aim to:

1) Obtain an ensemble of solutions as close as possible to the true Pareto front,
2) Obtain an ensemble of solutions as diverse as possible on the Pareto front.

The diverse Pareto front has been shown to lead to improved forecasting (Hutahaean et al., 2016).

Generally, the choice of the objectives is based on engineering knowledge and specifics of the reservoir. In this thesis we used three objectives to match each fluid production rate separately: 1) Objective 1 – Field Oil Production Rate (FOPR), 2) Objective 2 – Field Gas Production Rate (FGPR), and 3) Objective 3 – Field Water Production Rate (FWPR).

### 4.6.3 AHM set up

In this thesis assisted history matching and uncertainty quantification was performed with the software Raven (Epistemy Ltd.). The Particle Swarm Optimisation (PSO) algorithm (Section 3.3.2) was used for sampling. Raven provides various default choices of optimal parameters, and the ones we used are summarised in Table 4.8.

To perform history matching, the following parametrization was used: uniform distribution of the specified ranges of SNESIM modelling parameters, such as affinity parameter, global rotation, and facies proportions (Table 4.8). The number of a training image was used as a categorical value to introduce multiple training images into history matching (Table 4.9).

The automated history matching was performed for 820 days. The least squares objective function (Eq. 4.3 – 4.5) was used to minimise the discrepancy between the simulated and historical production data for three objectives.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

\[ M_1 = \sum_i^N \frac{(FOPR_i - \text{Obs}_{FOPR}(i))^2}{2\sigma^2} \]  

(Eq. 4.3)

\[ M_2 = \sum_i^N \frac{(FGPR_i - \text{Obs}_{FGPR}(i))^2}{2\sigma^2} \]  

(Eq. 4.4)

\[ M_3 = \sum_i^N \frac{(FWPR_i - \text{Obs}_{FWPR}(i))^2}{2\sigma^2} \]  

(Eq. 4.5)

where \( M_1, M_2, M_3 \) are three misfit functions, \( FOPR_i, FGPR_i, FWPR_i \) are the simulated field oil, gas and water production rate at time \( i \), \( \text{Obs}_{FOPR}(i), \text{Obs}_{FGPR}(i), \text{Obs}_{FWPR}(i) \) are the observed oil, gas and water production rate data at time \( i \), \( N \) is the number of data points, and \( \sigma^2 \) is the measurement error.

4.6.7 Empirical Bayes approach to obtain the measurement error

When we do not have an estimate of measurement error \( \sigma \), we can use Bayes Theorem to obtain one (Carlin and Louis, 2000). First, we recap (section 2.4) that taking the logarithm of the simplified Bayes Theorem (Eq. 2.5) we obtain the following expression:

\[ \log p(m|O) = \log p(O|m) + \log p(m) \]  

(Eq. 4.6)

Then, we multiply both sides by -1 and substituting for the likelihood, we obtain the following expression for the misfit:

\[ M = -\log p(m|O) = \sum \frac{(o_i - s_i)^2}{2\sigma^2} - \log p(m) \]  

(Eq. 4.7)

To this misfit expression obtained above, we add the term containing \( \sigma \) from the normalising constant for the Gaussian errors. We do this as \( \sigma \) is no longer considered constant, since we will need to vary it to determine the value:

\[ M(m, \sigma) = \sum \frac{(o_i - s_i(m))^2}{2\sigma^2} - \log p(m) - n \log \sigma \]  

(Eq. 4.8)

There are various approaches in the Bayesian literature to handle this case, of which the most robust is to marginalise over the unknown value by integrating it out. However, we
can obtain a simple and relatively reliable approach by looking for the maximum likelihood value of $\sigma$.

We imagine that we have obtained a set of model parameters that give a good fit. (If we have not, we can adopt an iterative procedure where we guess a value of $\sigma$, generate some models that match, take the best fitting model and use that to update $\sigma$).

We then differentiate the misfit expression to obtain the following:

$$\frac{\partial M}{\partial \sigma} = -\frac{1}{\sigma^2} \sum (o_i - s_i(m))^2 + \frac{n}{\sigma}$$  \hspace{1cm} (Eq. 4.9)

Then we set the derivative equal to zero to obtain the following:

$$\sigma^2 = \frac{1}{n} \sum (o_i - s_i(m))^2$$  \hspace{1cm} (Eq. 4.10)

where the parameters $m$ are those associated with the best fitting model.

In this thesis, we use this approach to obtain $\sigma$ (constant for each time step), as we do not have an estimate of the measurement error.

<table>
<thead>
<tr>
<th>PSO parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of particles</td>
<td>40</td>
</tr>
<tr>
<td>Group size</td>
<td>4</td>
</tr>
<tr>
<td>Initial inertia</td>
<td>0.9</td>
</tr>
<tr>
<td>Initial decay</td>
<td>0.9</td>
</tr>
<tr>
<td>Cognitive component</td>
<td>1.333</td>
</tr>
<tr>
<td>Group component</td>
<td>1.333</td>
</tr>
<tr>
<td>Social component</td>
<td>1.333</td>
</tr>
<tr>
<td>Energy retention</td>
<td>0.8</td>
</tr>
<tr>
<td>Particle behaviour</td>
<td>Flexible*</td>
</tr>
</tbody>
</table>

*Table 4-8 Summary of the parameters used for Particle Swarm Optimisation algorithm (from Rojas, 2013)
### 4.6.4 Results of history matching

In this section, we show the results of 5 history matching runs with multi-objective PSO for WCA case study.

Figure 4.23 shows the 2D projection of the Pareto front of multi-objective history matching for 5 runs. Pareto models are color-coded in blue and the rest of the models are green. Blue dotted line represents the Pareto front for a set of Pareto optimal solutions. All Pareto fronts are relatively diverse, which will benefit in improved forecasting.

Figure 4.24 illustrates misfit values obtained during history matching for different training images. From this plot, we can observe which training images resulted in lower values of misfit. For example, five lowest values of misfit were obtained with models from training images TI 2, TI 4, and TI 5.

The misfit values for each of the geological modelling parameters obtained during history matching are represented in Figure 4.25. Training images are represented by different colours on these plots. We can observe that the lowest value of misfit was obtained by a model from training image TI 5 with the following geological modelling parameters: affinity x = 1.5, affinity y = 1.8, affinity z = 2.4, global rotation = 3°, proportion of Facies 1 = 0.3, proportion of Facies 2 = 0.09, proportion of Facies 3 = 0.07, and proportion of Facies 4 = 0.54.

---

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affinity x</td>
<td>(0.5, 3)</td>
</tr>
<tr>
<td>Affinity y</td>
<td>(0.5, 3)</td>
</tr>
<tr>
<td>Affinity z</td>
<td>(0.5, 9)</td>
</tr>
<tr>
<td>Rotation</td>
<td>(0, 45)</td>
</tr>
<tr>
<td>Facies 1 (channel sand)</td>
<td>(0.25, 0.34)</td>
</tr>
<tr>
<td>Facies 2 (poorer quality sand)</td>
<td>(0.09, 0.14)</td>
</tr>
<tr>
<td>Facies 3 (poorer quality sand)</td>
<td>(0.06, 0.12)</td>
</tr>
<tr>
<td>Facies 4 (shale)</td>
<td>1 – (Facies 1 + Facies 2 + Facies 3)</td>
</tr>
<tr>
<td>Training image</td>
<td>(1, 11)</td>
</tr>
</tbody>
</table>

*Table 4-9 Parameter values for parametrisation applied to the West Coast of Africa facies models during history matching*
We can compare the original range of the parameters with the range which provide the models with the lowest value of misfit (< 10) in Table 4.10. The lowest misfit ranges are slightly lower for most of the parameters, however, they are significantly lower for affinity z and for the proportion of Facies 4. It means that for these parameters (aff z and P4) the most reduction in uncertainty was achieved during history matching.

Figure 4-23 Pareto fronts for two misfit components of a multi-objective history matching with PSO for WCA case study
Figure 4-24 Represents the selected training images during history matching and the corresponding value of misfit (<100) for one of five history matching runs. The points are coloured based on iteration number of history matching process.
Figure 4.25 Geological modelling parameters versus misfit. Models from different training images are represented by a different colour. Affx, affy, affz – affinity x, affinity y, affinity z, Rot – rotation, P1, P2, P3, P4 – proportions of Facies 1,2,3, and 4 respectively.
**Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior ranges</th>
<th>Ranges with the lowest misfit (&lt;10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>affinity x</td>
<td>(0.5; 3)</td>
<td>(0.6; 2.56)</td>
</tr>
<tr>
<td>affinity y</td>
<td>(0.5; 3)</td>
<td>(0.56; 2.8)</td>
</tr>
<tr>
<td>affinity z</td>
<td>(0.5; 9)</td>
<td>(0.57; 5.25)</td>
</tr>
<tr>
<td>rotation</td>
<td>(0° - 45°)</td>
<td>(2° - 42°)</td>
</tr>
<tr>
<td>Facies 1</td>
<td>(25% – 34%)</td>
<td>(25% – 33%)</td>
</tr>
<tr>
<td>Facies 2</td>
<td>(9% - 14%)</td>
<td>(9.3% - 14%)</td>
</tr>
<tr>
<td>Facies 3</td>
<td>(6% - 12%)</td>
<td>(6.5% - 11.7%)</td>
</tr>
<tr>
<td>Facies 4</td>
<td>(40% – 60%)</td>
<td>(46% – 55%)</td>
</tr>
</tbody>
</table>

Table 4-10 Comparison of the parameters ranges before history matching and after history matching, that provides the lowest values of misfit (<10)

Figure 4.26 shows the minimum misfit evolution during history matching. The minimum value of misfit drops from 71 to 11 after 10 iterations (where iteration means the repetition of the process but with a different combination of parameters).

![Figure 4-26 Minimum misfit evolution during history matching](image-url)
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

Five best models (with the lowest value of misfit) obtained during history matching for Field Oil, Gas, and Water Production Rate (FOPR, FGPR, FWPR) for 5 runs are presented in Figures 4.27 – 4.29.

All five models match the history very well for Oil and Gas Production Rate is very similar for all five runs (Figure 4.27 and 4.28). However, not all of the five models are matching the history for Water Production Rate after 500 days of production: only one (Figure 4.29) and two models (Figure 4.29, top left) are matching the data after 500 days.

![Figure 4-27 Five best models obtained during five history matching runs for Field Oil Production Rate (FOPR), STB/day](image-url)

Figure 4-27 Five best models obtained during five history matching runs for Field Oil Production Rate (FOPR), STB/day
Figure 4-28 Five best models obtained during five history matching runs for Field Gas Production Rate (FGPR), Mscf/day
In this thesis, we use the weighted sum approach (Eq. 4.1) to scalarize three objective functions into one using the same weight factor = 1 for all the objective functions.

### 4.7 Uncertainty Quantification in Reservoir Prediction

History matching results were used to perform reservoir predictions under uncertainty with the NA-Bayes (NAB) algorithm (subsection 3.4.2). A NAB chain for 10,000 random walk steps was run each of five history matching runs separately. The burn-in period was 5,000. NAB set up parameters are summarised in Table 4.11.

The reservoir forecast was performed for 1216 days. Figure 4.30 shows the uncertainty intervals (P10-P50-P90) for the oil, gas and water production rates (on the left) and total production (on the right) for all five runs. Figure 4.30 represents the average uncertainty intervals over five runs.
The Truth case model (section 4.5) was used to validate the prediction results. All intervals (Figure 4.30) cover the Truth case model very well. However, we can observe that intervals for water rate and total water production are much wider than for oil and gas. From the average plots (Figure 4.31) it is noticeable that the Truth case model is closer to P90 curve after 800 days of production for oil and gas rates. However, it is close to P50 for total oil and gas production. For water rate and total production, the Truth case model is close to P10.

The results show good history matches and Truth case model is inside the credible interval, even though training image TI 9, which was used to create a Truth case model, was not selected at all. TI 5 and TI 9 are initially different: Facies 2 and Facies 3 are represented as channels in TI 5 and as lobes in TI 9. However, after SNESIM simulations we get that some models from TI 5 look similar to models from TI 9: Facies 2 and Facies 3 are represented as lobes instead of channels (Figure 4.13).

Figure 4.32 shows the comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty (after NAB, average over five runs) of oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) production total. We could see that the posterior forecast uncertainties are much smaller than the prior uncertainties before history matching.

### Table 4-11 Parameters used for NA-Bayes set up for WCA reservoir predictions under uncertainty

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of chains</td>
<td>1</td>
</tr>
<tr>
<td>Chain length</td>
<td>10000</td>
</tr>
<tr>
<td>Burning period</td>
<td>5000</td>
</tr>
<tr>
<td>Distance refresh frequency</td>
<td>1000</td>
</tr>
</tbody>
</table>

The Truth case model (section 4.5) was used to validate the prediction results. All intervals (Figure 4.30) cover the Truth case model very well. However, we can observe that intervals for water rate and total water production are much wider than for oil and gas. From the average plots (Figure 4.31) it is noticeable that the Truth case model is closer to P90 curve after 800 days of production for oil and gas rates. However, it is close to P50 for total oil and gas production. For water rate and total production, the Truth case model is close to P10.

The results show good history matches and Truth case model is inside the credible interval, even though training image TI 9, which was used to create a Truth case model, was not selected at all. TI 5 and TI 9 are initially different: Facies 2 and Facies 3 are represented as channels in TI 5 and as lobes in TI 9. However, after SNESIM simulations we get that some models from TI 5 look similar to models from TI 9: Facies 2 and Facies 3 are represented as lobes instead of channels (Figure 4.13).

Figure 4.32 shows the comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty (after NAB, average over five runs) of oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) production total. We could see that the posterior forecast uncertainties are much smaller than the prior uncertainties before history matching.
Figure 4.30 Oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) total production prediction uncertainty intervals for five history matching runs.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>Rate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="FOPR" /></td>
<td><img src="image2" alt="FOPT" /></td>
</tr>
<tr>
<td><img src="image3" alt="FGPR" /></td>
<td><img src="image4" alt="FGPT" /></td>
</tr>
<tr>
<td><img src="image5" alt="FWPR" /></td>
<td><img src="image6" alt="FWPT" /></td>
</tr>
</tbody>
</table>

Figure 4.31 Oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPT) production prediction uncertainty intervals average over five history matching runs.
Figure 4.32 Comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty (after NAB, average over five runs) of oil (FOPR), gas (FGPR) and water (FWPR) production rates and Oil (FOPT), gas (FGPT) and water (FWPRT) production total.

Figure 4.33 shows the Posterior Probability Distribution (PPD) of training images selected for the predictions from history matching results (averaged over five runs). It was produced by summing up the posterior probability of each selected training image (from 5 runs) and divided by the number of the training images selected. For example, training image TI 1 was selected two times with the corresponding probabilities = 0.02 and 0.08; then the average PPD will be (0.02+0.08)/2 = 0.05. Training image TI 5 has the highest Probability, whilst training images TI 8, TI 9, and TI 11 were not selected at all.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

4.8 Geological Realism in Predictions

In this work, we compare geological models selected for reservoir predictions (models with high probability after NAB) to the Truth case geological model by visual examination. Figure 4.34 shows a top view of geological models and geological parameters used for the modelling.

The main differences between selected models and the Truth case model are:

- channel sand facies (dark red/brown) is not sinusoidal in the models used for prediction;

- the value of global facies rotation parameter is higher than the Truth case rotation value for most of the models;

- the continuity of the channel sand facies (dark red/brown) that we can observe in the Truth case model is not preserved in most of the models.

However, the values of facies proportion are quite similar. From this, we can conclude that affinity and rotation parameters have the biggest influence on 3D geological model realisation in this case study.
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

<table>
<thead>
<tr>
<th>TI 9;</th>
<th>affx</th>
<th>1</th>
<th>channel sand</th>
<th>28%</th>
</tr>
</thead>
<tbody>
<tr>
<td>affy</td>
<td>0.75</td>
<td>poorer quality sand 1</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>affz</td>
<td>1.25</td>
<td>poorer quality sand 2</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td>rot</td>
<td>0°</td>
<td>shale</td>
<td>52%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TI 5;</th>
<th>affx</th>
<th>1.46</th>
<th>channel sand</th>
<th>30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>affy</td>
<td>1.74</td>
<td>sand 1</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>affz</td>
<td>2.34</td>
<td>sand 2</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>rot</td>
<td>3°</td>
<td>shale</td>
<td>52%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TI 5;</th>
<th>affx</th>
<th>1.43</th>
<th>channel sand</th>
<th>28%</th>
</tr>
</thead>
<tbody>
<tr>
<td>affy</td>
<td>1.3</td>
<td>sand 1</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>affz</td>
<td>5.18</td>
<td>sand 2</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>rot</td>
<td>15°</td>
<td>shale</td>
<td>53%</td>
<td></td>
</tr>
</tbody>
</table>
4.9 Discussion and Conclusions

4.9.1 Facies modelling

In this chapter, we showed the application of the hierarchical approach to introducing geological uncertainties into facies modelling. The hierarchical approach allowed us to consider multiple geological concepts and multiple geological facies modelling parameters for each concept. However, from a geological point of view, the original
Chapter 4. Hierarchical Geological Realism in Multiple Training Images History Matching and Uncertainty Quantification

concepts are quite similar, that is we selected the same parameter ranges for all the concepts. We believe, that it might be very interesting to create a set of more diverse training images with different parameter ranges to explore the broader range of geological uncertainties.

We used SNESIM implementation of MPS to create geological facies realisation with hierarchical uncertainties. We choose SNESIM over object modelling as it allowed to condition model realisations to the available well and seismic data, and, also to introduce hierarchical uncertainties (Level 3) for each geological scenario (Level 2).

Geological facies modelling parameters (Level 3) allowed as to introduce the variability into each of geological concept: we created very different facies realisations within each of the training images. However, this variability led to a difficulty to visually distinguish between model realisations from different concepts: resulting models from different images look more alike than the initial concepts. This fact complicates the process of distinguishing between model realisations from different training images.

Modification of the affinity parameters, which has the main effect of the resulting realisation, as they defined the geometry of geobodies, was difficult to control. Some affinity parameter combination resulted in unrealistic facies representation. The facies geometry has a great impact on flow simulation results. Unrealistic models could result in unrealistic flow predictions.

From the visual examination of the resulting models, we observed that some model realisations were able to reproduce a given training image better than the others by capturing channelised architecture and continuity for sand facies.

However, not all SNESIM model realisations were able to preserve channelised architecture and channel facies continuity, which could affect the flow simulation results.

Therefore, we would recommend considering how we could improve the facies modelling by controlling the realism in facies geometry and facies connectivity, yet accounting for the uncertainties.
4.9.2 History matching and forecasting

We introduced hierarchical geological uncertainties into history matching and uncertainty quantification. We were able to achieve good history matching results for oil and gas rates but not for the water rate. The resulting credible intervals in forecasting has the Truth case model inside the uncertainty interval, however, the interval for water rate and water total production is much wider than for oil and gas.

The visual examination of the geological models selected for reservoir predictions (models with high probability after NAB) showed that these models did not preserve the geologically realistic facies architecture and connectivity.

Therefore, in the next chapters, we propose and show the results of the metric space classification integrated into history matching to preserve the geologic representation of facies models and to compare the effect on the history matching and forecasting results.
CHAPTER FIVE Metric Space Approach and Machine Learning Classification

5.1 Introduction

5.1.1 Metric space approach

Metric space approaches have been used for history matching and uncertainty modelling in multiple studies (Park et al., 2013; Scheidt and Caers, 2009, n.d.; Suzuki et al., 2006; Suzuki and Caers, 2008). In these studies, model realisations are represented in the metric space to evaluate a distance function (similarity distance) between any two realisations. The evaluated distance function represents how similar (or different) a pair of models is: the models that are close to each are more similar than the ones that are far from each other.

The main purpose of the metric space approach is to reduce the simulation time by identifying the similar models in a metric space because the main assumption is that the similar models in the metric space produce the similar simulation response.

Depending on the applications, any type of metric may be defined as long as the metric represents the difference in responses of interest and is calculated efficiently (Park et al., 2011). For example, Suzuki and Caers (2008) used a static metric to measure the similarity of geometry between two reservoir models. The geometry of a model was represented as a point set, which consisted of the corner points of the stratigraphic grid belonging to the top and bottom horizon surfaces of the structure. Suzuki and Caers (2008) used a Hausdorff distance (Huttenlocher et al., 1993) to measure the similarity of this geometry between any two stratigraphic reservoir models. The Hausdorff distance is a well-established distance in pattern recognition and shape matching of objects problems. Suzuki and Caers (2008) showed there is a correlation between the static model distances and their flow response (models close to each other produced a more similar dynamic response), which is an important assumption for this technique to work well.
Scheidt et al. (2015) proposed to use a static metric to perform global pattern comparison between forward synthetic seismic responses and observed seismic data, which are analysed using discrete wavelet transforms (DWTs) and associated wavelet coefficient histograms. The authors used a Kullback-Leibler distance (Kullback, 1997) to distinguish between two different images. The Kullback-Leibler distance measures how one probability distribution diverges from another expected probability distribution. This technique (Scheidt et al., 2015) allowed to evaluate consistency between geological assumptions and seismic data, to update the distribution of the uncertainty parameters, and to eliminate some of the parameters that are not consistent with geological data.

In an earlier work, Scheidt and Caers (2009) used a dynamic metric to measure the similarities between cumulative oil and cumulative water production of the field. The authors used the metric approach to select a subset of reservoir models reflecting the same uncertainty in flow response as the full set to reduce the computation time by reducing the number of models. The distances between any two realisations were calculated using a streamline simulator (Batycky et al., 1997) to receive a fast simulation response. Scheidt and Caers (2009) stated that a streamline simulator provided a good measure of the difference in the degree of connectivity for this particular case study (channelized, connected sand/shale systems, with water injected for pressure support). In this study, the authors used a specific flow-based distance, therefore the conclusions they got are only specific to this case study. The authors recommend using static-based distance to obtain more general results.

Rongier et al. (2016) proposed to use a static metric to compare the connectivity of the sedimentary deposits between model realisations. The authors used a set of indicators (e.g. shape indicators, skeleton indicators, global indicators) based on the connected components of the realisations to evaluate the similarities between any two realisations. Rongier et al. (2016) used a combined heterogeneous Euclidian/Jensen-Shannon (Wilson and Martinez, 1997) distance to compute the indicator dissimilarities because the indicators are heterogeneous (they are either based on histograms or on continuous values). This approach was able to provide the quality control of multiple geological model realisations quality control by assessing their static connectivity. However, there is no evidence of a correlation between these static connectivity indicators and dynamic model response.
Chapter 5. Metric Space Approach and Machine Learning Classification

In this work, we propose to use a static metric to measure the similarities between multiple facies model realisations from different geological scenarios. We chose Euclidian distance due to its simplicity and as a starting (reference) point to which more complex and specific distance measures could be compared (e.g. Hausdorff, Frechet, Minkowski distance and so on). The major assumption of this approach is that there is a correlation between similarity in facies models in static metric space and the similarity in production response of these facies models.

The full description of the metric space and multidimensional scaling are presented in this chapter. We apply the metric space and multidimensional scaling on a West Coast of Africa (WCA, section 4.2) example. The results are presented in this chapter as well.

5.1.2 Machine learning metric space classification

The main idea behind metric space classification is to identify the geological relations between different geological scenarios in the metric space and include them into history matching (Section 1.3). The main challenge is to find the relationship in data in a high-dimensional space. Therefore, we propose to use machine learning classification as it is known to perform well in high-dimensional space (Vapnik, 1995).

Classification of the metric space allows us to find geological relations between multiple scenarios and map the boundaries in the metric space. Therefore, classification helps to find better models with the appropriate combination of geological parameters and geological scenarios during history matching. This would lead to an improve geological realism in facies models.

In this chapter, we provide the description of three different classification algorithms: k-means clustering (MacQueen et al., 1967), Support Vector Machines (SVM) (Vapnik, 1995) and Random Forest (RF) (Breiman, 2001) along with the application and results. We also present the comparison of the results of all three classifiers. Classifiers that perform better than the others will be included into history matching and the results of this study will be presented in chapter 6.
Chapter 5. Metric Space Approach and Machine Learning Classification

5.2 Multidimensional scaling

5.2.1 Principle and method used

Multidimensional scaling (MDS) is a technique applied in a metric space to analyse similarity or dissimilarity of a set of objects. MDS is used to translate the dissimilarity matrix into a configuration of points in n-dimensional (reduced dimensional) Euclidian space (Borg and Groenen, 2005). The points in the metric space are arranged in such a way that their Euclidian distances match as much as possible to the dissimilarities of the objects. The main advantage of this technique is to graphically represent the structure of the data and to facilitate the understanding and analysis of this data.

There are several Multidimensional scaling methods available; the difference between them lies in the type of dissimilarities and in the way to represent these dissimilarities with the distances. The usual MDS method called the classical scaling (Gower, 1966; Scheidt et al., 2009; Torgerson, 1952, 1958) assumes that the dissimilarities are already Euclidean distances.

The classical MDS is based on the fact that the coordinate matrix X can be derived by eigenvalue decomposition from a Gram matrix B, which is obtained by converting the dissimilarity matrix D into a scalar product (Scheidt et al., 2009).

The following steps summarize the classical MDS process:

1) Construct a matrix A with function elements of the dissimilarity distance between any two models \( \delta_{ij} \):

\[
a_{ij} = -\frac{1}{2} \delta_{ij}^2
\]  
(Eq. 5.1)

2) Construct a matrix B by centring A: \( B = HAH \) using the symmetric matrix \( H = I - \frac{1}{n} 1^T 1 \)

3) Extract the \( p \) largest positive eigenvalues \( \lambda_1, \ldots, \lambda_p \) of B and the corresponding \( p \) eigenvectors \( e_1, \ldots, e_p \)
4) A $p$-dimensional spatial configuration of the $N_R$ objects is derived from the coordinate matrix $X = E_p A_p^{1/2}$ where $E_p$ is the matrix of $p$ eigenvectors and $A_p$ is the diagonal matrix of $p$ eigenvalues of $B$, respectively.

The map obtained by MDS is derived by the dissimilarity distances in the matrix, the absolute location of the points is irrelevant.

One of the most studied distances in MDS is the Euclidean distance which is defined as:

$$d_{ij}(X) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2}$$  \hspace{1cm} (Eq. 5.2)

where $x_i$ and $x_j$ represents a pair of models. A single model can be represented by a vector, which contains properties of the model on a grid.

Equation 5.2 can be also generalised to $m$-dimensions as follows:

$$d_{ij}(X) = \left[\sum_{a=1}^{m} (x_{ia} - x_{ja})^2\right]^{1/2}$$  \hspace{1cm} (Eq. 5.3)

Another MDS method is called Scaling by Majorizing a Complicated Function (SMACOF) (De Leeuw and Heiser, 1977, 1980). SMACOF uses a stress majorization optimisation of an objective function, where for a set of $n$ $m$-dimensional data, a configuration of $X$ of $n$ points in $r$ ($<<m$)-dimensional space is sought that minimizes the so-called stress function $\sigma(X)$. This function $\sigma$ is a loss function that measures the squared differences between ideal $m$-dimensional distances and actual distances in $r$-dimensional space. It is defined as follows:

$$\sigma(X) = \sum_{i<j=n} w_{ij} (d_{ij}(X) - \delta_{ij})^2$$  \hspace{1cm} (Eq. 5.4)

where $w_{ij} > 0$ is is a weight for the measurement between a pair of points $(i, j)$, $d_{ij}(X)$ is the Euclidean distance between $i$ and $j$, and $\delta_{ij}$ is the ideal distance between the points in the $m$-dimensional data space. The weight can be used to specify a degree of confidence in the similarity between points.
Chapter 5. Metric Space Approach and Machine Learning Classification

A configuration $X$ which minimizes $\sigma(X)$ gives a plot in which points that are close together correspond to points that are also close together in the original $m$-dimensional data space.

There are many ways that $\sigma(X)$ could be minimized (e.g. an iterative steepest descent approach (Kruskal, 1964). However, the iterative majorization method (De Leeuw and Heiser, 1977) has been recognised as the one that gives significantly better results (in terms of guarantees and rate of convergence). This method minimizes a simple convex function which both bounds $\sigma$ from above and touches the surface of $\sigma$ at a point $Z$, called the supporting point at each step. In convex analysis such a function is called a majorizing function. That is why this iterative majorization process is also referred to as the Scaling by MAjorizing a COmplicated Function (SMACOF) algorithm.

In this work, we use the classical MDS and Euclidean distance to measure the dissimilarities between multiple facies model realizations from different geological scenarios in a static metric space. The classical MDS assumes that dissimilarities are Euclidean. It is not always true and other distances such as Hausdorff (Huttenlocher et al., 1993), Frechet (Eiter and Mannila, 1994), Minkowski (Singh et al., 2013) are available. However, they are not compared in this work because it is beyond the scope of this research.

5.2.2 Validation of the number of dimensions: the scree plot

MDS is used as a dimensionality reduction technique. MDS allows us to explore visually the dissimilarities between given objects by producing a configuration of the objects as optimal as possible in a lower dimension. MDS computes a matrix of pairwise distances and uses distance scaling to obtain a representation of high-dimensional data in lower dimensions as close as possible (Buja et al., 2008).

A Scree plot (Borg and Groenen, 2005) is a method for determining the optimal number of dimensions to describe the data of MDS. The scree plot is a histogram that shows the eigenvalues of each dimension. The relative eigenvalues express the ratio of each eigenvalue to the sum of the eigenvalues. The relative eigenvalue of a dimension gives the proportion of the data variance explained by this dimension.
Chapter 5. Metric Space Approach and Machine Learning Classification

The Scree plot provides a graphical way to evaluate the number of dimensions required to capture most information contained in the data. In a scree plot, the relative eigenvalues decrease when the dimension number increases. To define the optimal number of dimensions to describe the data, one should look for the ‘elbow’ (Figure 5.1). The dimension value right after the point with the highest flexion is generally enough for a decent representation (e.g. 3 dimensions are enough on Figure 5.1).

![Scree plot](image)

*Figure 5-1 The Scree plot, representing the relative eigenvalue of each dimension. The recommended number of dimensions is three on this example (from Wickelmaier, 2003)*

Other methods, such as the Shepard diagram (Borg and Groenen, 2005) and point position confidence (Rongier, 2016) are available to validate the number of dimensions for representation of the dissimilarities.

In this work, we use the Scree plot approach, which is sufficient enough on its own, to define the number of dimensions required to capture most of the information from the data.

5.2.3 Multidimensional scaling results

As mentioned before, we used static metric space and Euclidian distance to evaluate the dissimilarities between facies models obtained from different training images with SNESIM algorithm. Figure 5.2 represents a schematic view of the process (just for four models, as an example). Each of four facies was defined as a number (0, 1, 2, 3) and a dissimilarity distance matrix was constructed.
Chapter 5. Metric Space Approach and Machine Learning Classification

The dissimilarity matrix contains the distance between any two facies models, which describes how similar (or different) these two facies models are. The distance matrix is then used to map all the realisations into an N-dimensional Euclidian space (number of dimensions is equal to a number of models) using the classic MDS approach (Figure 5.2 represents a 2D projection of multi-dimensional Euclidian space). Each model is colour-coded according to the training image that was used for the simulation.

![Dissimilarity Matrix and MDS Projection](image)

*Figure 5-2 Schematic representation of calculating the dissimilarities between facies models and mapping them to the N-dimensional Euclidian space using the classic MDS approach. Only four models are shown as an example. Each colour in the Euclidian space (2D projection of the multidimensional Euclidian space) represents a different training image that was used for the simulations.*

The results of the MDS for WCA are presented in Figures 5.3 – 5.6; 660 models from 11 training images were represented in 660-dimensional Euclidean space. In 660 dimensions, all models are almost equally apart. Figures 5.3 – 5.6 represent 2D projections (dimensions 1 and 2, 2 and 3, 3 and 4, 4 and 5) of the MDS in the metric space. Static facies models that are close to each in the metric space are more similar than models that are further apart.
Chapter 5. Metric Space Approach and Machine Learning Classification

We could observe that different training images are close to each other in the metric space. Despite the described facies property variation of training images (described in subchapter 4.2.3), the training images are quite similar to each other, which is quite realistic as all of them represent a description of the same depositional environment. We could also see that the training images are close to some of the model realisation in the metric space.

Facies models are widely spread in the metric space. A lot of realisations are quite far from the training images. This could be explained by the variation of geological modelling parameters (facies proportions, global rotation, and global affinity) applied to account for hierarchical geological uncertainties in the model realisations.

The models in Figures 5.3 – 5.6 are color-coded according to the number of a training image. These plots represent a 2D projection of a high-dimensional space where the models from different training images are mixed. This presents a challenging task to classify these models in the metric space. Therefore, we suggest using machine learning that can perform classification in a high-dimensional space to distinguish between different training images.

Facies models from different training images are very mixed in the metric space (2D projections). This could be due to several reasons: 1) the applied hierarchical uncertainties resulted in some facies models from initially different training images being created more similar as being to each other, 2) the initial set of training images is rather similar and the resulting models are quite alike, 3) SNESIM approach scrambles the initial differences between training images and creates the realisations that are more similar to each other than to their respective training image, or 4) the choice of the static metric and Euclidian distance might not be appropriate for this data to reproduce the differences between the realisations.
Figure 5-3 2D representation (dimension 1 and 2) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model’s symbols). The truth case model is also represented in the metric space.

Figure 5-4 2D representation (dimension 2 and 3) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model’s symbols). The truth case model is also represented in the metric space.
Chapter 5. Metric Space Approach and Machine Learning Classification

Figure 5-5 2D representation (dimension 3 and 4) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model’s symbols). The truth case model is also represented in the metric space.

Figure 5-6 2D representation (dimension 4 and 5) of MDS in Euclidian space for 660 facies models from 11 training images. Models are colour coded according to different training images. Training images are also mapped into the metric space (symbols are bigger than the model’s symbols). The truth case model is also represented in the metric space.
Chapter 5. Metric Space Approach and Machine Learning Classification

To identify the reason for these results a further thorough investigation is recommended, which is beyond the scope of this research.

Figure 5.7 shows the comparison of production response of the models that are close to each other in metric space (model 425 and model 426) and the model that is further away (model 361). As we can see, models 425 and 426 has more similar production response, whilst model 361 has more different production response for Field oil and gas production total (FOPT, FGPT). However, Field water production total (FWPT) results differ significantly for all three models.

![Comparison of production response](image)

*Figure 5-7 Comparison of production response (Field oil production total (FOPT), Field gas production total (FGPT), Field water production total (FWPT)) of three models: two of them (model 425 and 426) are close in the metric space and one (model 361) is further away in the metric space*

Table 5.1 summarizes geological parameters used to perform facies modelling of these models.
Chapter 5. Metric Space Approach and Machine Learning Classification

<table>
<thead>
<tr>
<th>Model 361</th>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Affinity</td>
<td>0.5, 1, 2</td>
<td></td>
</tr>
<tr>
<td>Rotation</td>
<td>45°</td>
<td></td>
</tr>
<tr>
<td>Facies proportions</td>
<td>52, 11, 9, 28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 425</th>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Affinity</td>
<td>0.5, 1.5, 1</td>
<td></td>
</tr>
<tr>
<td>Rotation</td>
<td>0°</td>
<td></td>
</tr>
<tr>
<td>Facies proportions</td>
<td>52, 11, 9, 28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model 426</th>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Affinity</td>
<td>0.5, 1.5, 3</td>
<td></td>
</tr>
<tr>
<td>Rotation</td>
<td>0°</td>
<td></td>
</tr>
<tr>
<td>Facies proportions</td>
<td>52, 11, 9, 28</td>
<td></td>
</tr>
</tbody>
</table>

Table 5-1 Geological parameters used to perform facies modelling of the model 361, 425 and 426

5.2.3 Identifying key number of dimensions

To identify the necessary number of dimensions that represent the similarities between the models, we plot the relative eigenvalues against the number of dimensions to get the scree-plot (Figure 5.8). From Figure 5.8 we can see that the relative eigenvalues decrease when the dimension number increases. As the relative eigenvalue of a dimension gives the proportion of the data variance explained by this dimension, we chose eight dimensions as the relative eigenvalue after eight dimensions is varying insignificantly (<0.02). All the following dimensions would introduce only a minor improvement to represent the similarities between models.
Figure 5-8 The Scree plot identifies the optimal number of dimensions to describe the data of MDS

5.3 Machine Learning – Learning from Data

5.3.1 Motivation and challenges
Identifying various geological relationships, for example, facies geometry, sediment grain-size, bed thickness, etc. is a very complicated nonlinear and multivariate problem. This problem is usually complicated by the fact that geological data come from multiple sources and different scales (outcrops, seismic interpretation, satellite images, well data, analogous oil fields and so on).

There are many statistical techniques trying to capture and evaluate these complicated relationships but there are still a few challenges remaining (Rojas, 2013):

- The number of observations is commonly low to capture the true underlying relations;

- Statistical relationships between the geological variables are non-linear, although many authors try to use linear relationships to explain in a simple way the interaction between some geological variables (Bridge and Mackey, 1993; Crane, 1982; Leeder, 1973).
- All the data usually contain noise because the sampling process, preparation and analysis are highly associated with errors and noise.

- Geological information is often high-dimensional since it includes multiple process and factors that are influencing the generation of a specific rock or the distribution of different rock properties;

The high dimensional data are usually associated with the increasing difficulty to find the dependencies from high-dimensional data. The common problem associated with high-dimensions is called a curse of dimensionality (Bellman, 2015), which is defined as the emptiness in space for increasing dimensions. It means that when the dimensionality increases, the volume of the space increases so fast that the available data becomes sparse. This sparsity is problematic for any method that requires statistical significance. In order to obtain a statistically sound and reliable result, the amount of data needed to support the result often grows exponentially with the dimensionality.

Machine Learning Techniques (MLT) have become very popular as they are able to address the challenges mentioned above: MLT performs very well in high-dimensional spaces, works with noisy data and identifies nonlinear relationships between variables (Vapnik, 1995).

Machine learning provides the opportunity to learn from data by eliciting complex and hidden statistical dependencies from some examples. MLT tools are universal, adaptive to nonlinear data, robust and efficient (Kanevski et al., 2009). MLT can be applied for the various type of problems: regression, classification, and density modelling problem in high-dimensional spaces.

5.3.2 Types of learning process
In this section, we describe three types of the learning process in machine learning: 1) unsupervised learning – uses only input data to perform, 2) semi-supervised learning – uses incomplete input and output data sets, and 3) supervised learning – uses full output with limited input data. It is important to choose an appropriate type of learning for the particular problem based on data availability.
Unsupervised learning

As mentioned above, unsupervised learning performs learning using only input data, when no output data are available. It means that the goal of an algorithm is to explore the data and find some structure within. Unsupervised learning is used to find patterns, clusters, and structures in a set of N input data samples.

Popular techniques include self-organizing maps (Haykin, 1994), nearest-neighbour mapping (Altman, 1992), k-means clustering (MacQueen et al., 1967), and singular value decomposition (Stewart, 1993). Some examples of these algorithms are in text topics segmentation, items recommendation, and identification of data outliers.

Clustering of the data has been a very popular application of unsupervised machine learning in geoscience and petroleum engineering studies. Romero et al. (2016) showed the results of unsupervised learning for remote sensing image classification. (Lary et al., 2016) applied Self-Organizing Maps (SOMs) to perform a classification task of a large number of surface types from multi-spectral satellite data. Saggaf and Nebrija (2000) used unsupervised neural networks to segregate a well into distinct facies classes based on the logs behaviour (Figure 5.9). The obtained results showed that the neural network method yielded closely similar results to those obtained manually by an experienced geologist.

![Figure 5-9 Facies classification performed from well-log data analysis, using unsupervised artificial neural networks (ANN): (a) logs data ( Resistivity, Neutron-Porosity, and Density) cluster in different classes, (b) ANN classified the data and identified sedimentary facies (From Saggaf and Nebrija, 2000)](image-url)
Supervised learning

Supervised learning is trained using the data with known output. In other words, the learning algorithm receives a set of inputs along with the corresponding outputs, and then the algorithm learns by comparing its produced output with correct output to find errors. Based on this comparison, the algorithm then modifies the model accordingly. Supervised learning uses patterns to predict the new values of the data. Supervised learning could perform classification (Figure 5.10 (a)), regression (Figure 5.10 (b)), prediction and gradient boosting.

Qi et al. (2007) demonstrated an example where they used for the supervised neural network to predict oolitic facies from image logs without any core data (Figure 5.11). Neural network model was trained from digital well logs based on 1000 half-foot intervals of cored wells. Core data were used to calibrate core depth. The trained neural network model was then used to predict lithofacies in wells where the core was absent.

Figure 5-10 Example of classification (a) and regression (b) problem solved with supervised learning: (a) the classification algorithm identifies three classes and separates them in the x1-x2 space (From Kanevski et al., 2009), (b) the regression problem is solved by drawing a line through a cloud of points (From Kanevski et al., 2009)

Semi-supervised learning

Semi-supervised learning is applied to the data analysis cases when the sets of inputs and outputs are incomplete. It uses both known (labelled) and unknown (unlabelled) data for training, usually a small amount of labelled data with a large amount of unlabeled data. The reason behind this is that unlabelled data is typically less expensive to get and less effort to acquire. Therefore, semi-supervised learning is useful when it is too costly to
acquire the full set of labelled data. This type of learning can be used for classification, regression, and prediction tasks.

The information obtained from the unlabelled part of the dataset affects the learning results. For instance, Figure 5.12 shows two examples of a classification problem results: (a) without unlabelled data and (b) with unlabelled data. We could see that the differences in the obtained classification boundary geometry are defined by the structure given by the unlabelled data. Whilst unlabelled data do not have output values, these data affect the structure that links samples from different populations.

Demyanov et al. (2008) applied semi-supervised Support Vector Regression (SVR) (after Belkin et al., 2006) to a synthetic fluvial reservoir to model petrophysical properties taking into account the prior geological knowledge for more geologically realistic results. In this paper, the petrophysical properties were modelled by adapting the concept of geo-manifold, which assumes that in the high-dimensional data space there exists a low-dimensional manifold which implies useful dependencies for the data-driven models. The authors used porosity and permeability data from wells and unlabelled data (channel geo-manifold) from seismic in order to generate permeability distribution maps in a fluvial reservoir (Figure 5.13). As a result, semi-supervised Support Vector Regression model of reservoir properties took into account the natural similarities in space and data relation to reproduce fluvial geological structure and anisotropy of the fluvial system in a more realistic way.
Chapter 5. Metric Space Approach and Machine Learning Classification

Figure 5-11 Predicting oolitic facies from core data to uncored wells based on image logs using supervised neural networks: (A) Neural networks were trained using data from cored wells; (B) Predicted oolitic facies in uncored wells (from Qi et al., 2007)
Chapter 5. Metric Space Approach and Machine Learning Classification

Figure 5-12 Example of unlabelled data effect on the classification problem results: (a) Classification problem solved without unlabelled data and (b) with unlabelled data. Unlabelled data add the information about the geometry or structure of the classification boundary between classes (From Belkin et al., 2006)

Figure 5-13 Porosity distribution maps created by using semi-supervised Support Vector Regression. Unlabelled data points are obtained from seismic data. Labelled data come from well data. Unlabelled data points give the structural information of the geobody (so-called geo-manifold); this geo-manifold is populated with data obtained from wells (From Demyanov et al., 2008)

To sum up, for a successful result of the machine learning study, it is important to choose the type of learning process based on the specifics of the problem and the available data.

In this work, we used supervised machine learning because we had a set of input data with the corresponding outputs. During supervised learning, the algorithm learnt from the data by comparing the newly produced output with known output to find errors.
5.4 Solving Classification Problem with Machine Learning

In this thesis, we propose to use machine learning classification to account for essential geological uncertainties and improve geological realism in reservoir models during history matching. Machine learning classification allows us to find geological relationships between multiple scenarios and create the boundaries in metric space representing areas of different geological scenarios. Knowing where these boundaries are, it allows us to control what geological scenario and which appropriate geological parameters are selected during history matching to improve geological realism in history matched models.

A simple sketch is introduced below (Figure 5.14) to illustrate the following workflow:

1) generate a number of geological model realisations from different geological scenarios with SNESIM;
2) plot these models in a chosen metric space;
3) apply a classification technique to define the boundaries between model realisations from different geological scenarios (number of classes equals to the number of geological scenarios);
4) perform history matching;
5) new models are generated during history matching process. Each of these models can be mapped to the metric space. Based on the position of the model in the metric space we identify which geological scenario it belongs to.

![Figure 5.14 A simple sketch to explain the classification of metric space introduced into history matching: 1) we plot geological model realisations from different geological scenarios to a metric space, 2) we apply a classification technique to obtain the boundaries between different geological scenarios, 3) new point represents a new model produced during history matching. Classification result allows us to assign a new model to an appropriate geological scenario](image-url)
In this thesis, three different algorithms were used to perform classification in the metric space:

- \textit{k}-means clustering (MacQueen et al., 1967);
- Support Vector Machines (SVM) (Vapnik, 1995);
- Random Forest (Breiman, 2001).

All three algorithms satisfy the requirements of the studied classification problem: they all aim to learn from the known outputs to define the boundaries between different classes in multi-dimensional metric space. However, the learning process and the resulting boundaries are very different between the methods. In this thesis, we aim to apply them and compare their performance to a specific problem.

Here we provide the description of these three classification algorithms used.

\subsection*{5.4.1 \textit{k}-means clustering}

\textit{k}-means clustering is one of the most popular and simplest method of classification. The aim of the algorithm is to partition \(N\) number of observations into \(k\) clusters. As the result, each observation belongs to a cluster with the nearest mean, so-called centroid. The number of clusters could be predefined a priori or be provided by the algorithm.

\(k\)-means clustering aims to minimise an objective function known as squared error function:

\[ J(v) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} (||x_i - v_j||)^2 \]  

(Eq. 5.5)

where \(||x_i - v_j||\) is the Euclidean distance between \(x_i\) and \(v_j\), \(c_i\) is the number of data points in the \(i^{th}\) cluster, and \(c\) is the number of cluster centres.

Figure 5.15 represents the main steps of \(k\)-means:

1) Define the number of clusters;
Chapter 5. Metric Space Approach and Machine Learning Classification

2) Determine the initial centroids for each cluster (the Forgy method of initialization is recommended (Hamerly and Elkan, 2002); the Forgy method randomly chooses $k$ observations from the data set and uses these as the initial means);

3) Calculate the distance between the data points and centroids (Euclidian distance is used as the standard distance);

4) Group data points according to the calculated distance into clusters (classes);

5) Recalculate the position of centroids for each cluster;

6) Repeat step 3 and 4 until the centroids no longer move (the algorithm is converged).

$k$-means algorithm is very sensitive to the initial randomly selected cluster centres, so-called centroids. Therefore, as was mentioned above it is recommended to run the algorithm multiple times to reduce this effect.

Figure 5-15 The main steps of $k$-means clustering classification
Advantages of k-means
The main advantages of k-means clustering are listed below:

1) Simple to understand and interpret;
2) Fast for low dimensional data;
3) Gives the best result when the data set is distinct or well separated from each other.

Disadvantages of k-means
A list of disadvantages of k-means clustering includes:

1) Needs to be performed multiple times, otherwise, provides different result with each run (depends on initial random centroids assignments);
2) If there are two highly overlapping data subsets, then k-means will struggle to resolve that there are two clusters;
3) Not suitable for finding clusters with nonconvex shapes;
4) Unable to handle noisy data and outliers.
5) The algorithm performs poorly for the nonlinear data set.

5.4.2 Support Vector Machines
Support Vector Machines (SVM) has become a very popular technique because it helps to solve difficult classification problems in a wide range of application domains. The main advantages of SVM are: it produces very accurate classifiers, less overfitting, and robust to noise.

The foundations of Support Vector Machines (SVM) was first invented by Vapnik and Chervonenkis in 1963. The current standard incarnation was developed by Cortes and Vapnik in 1993 and published in 1995 (Cortes and Vapnik, 1995).

Here we explain the main principles and terminology of SVM.

Classification using a separating hyperplane
To understand how SVM works, we first need to introduce the main principles on a simple example. The simplest way to solve the classification problem is to distinguish the samples of two different classes with a linear decision surface that separates data into two
classes. Figure 5.16 illustrates that there are many possible linear classifiers (blue lines) that can separate the data into two classes. However, there is only one linear classifier (green line) that maximises the margin, i.e. maximises the distance between the classifier and the nearest data point of each class. This linear classifier is called the *optimal separating hyperplane* (James et al., 2014).

![Figure 5-16 Illustrates the choice of optimal separating hyperplane (green line) for two classes’ classification example](image)

Generally, a hyperplane is defined as a flat subspace of dimension \( p - 1 \) (James et al., 2014). For example, in 2-dimensional space it is a line (Figure 5.16), in 3 dimensions – it is a plane. The mathematical definition of a hyperplane in \( p \)-dimensions is represented by the following equation:

\[
\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p = 0
\]  
*(Eq. 5.6)*

where \( \beta_0, \beta_1, \) and \( \beta_p \) are the parameters and a point \( X = (X_1, X_2, \ldots, X_p)^T \) in \( p \)-dimensional space lies on the hyperplane.

When a point \( X \) does not lie on the hyperplane, it represents that a point \( X \) lies to the one or to the other side of the hyperplane respectively (Equation 5.7 and 5.8).

We can determine on which side the point lies by calculating the sign of the left-hand side of the Equation 5.6.
Chapter 5. Metric Space Approach and Machine Learning Classification

\[ \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p > 0 \quad \text{(Eq. 5.7)} \]

or

\[ \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p < 0 \quad \text{(Eq. 5.8)} \]

Intuitively, we would expect the *optimal separating hyperplane* (Figure 5.16) to generalise well as opposed to the other possible boundaries. By generalisation, we mean the ability of learning algorithm to describe the unknown underlying process from the data. The common problem for learning algorithm is that the amount of data is usually small and that the function chosen by algorithm describes only the dependencies in the data but not the underlying unknown process. This problem is even more common for high-dimensional cases.

**Support vectors and soft margin classifier**

Only the points at the margin determine or “support” the solution of the classifier (Figure 5.17). If all the other data samples except these points would be removed from the data set, this will not affect the solution. Since mathematically each point is treated as a vector, these critical patterns are referred to as "support vectors" (Figure 5.17).

In the previous simple example (Figure 5.16), we consider that the separation is possible without misclassification. This is called the *linearly separable data case*. The algorithm that finds the optimal solution for the linearly separable data is called *large margin classifier* (James et al., 2014).

Generally, real data are rarely linearly separable. That is why we are looking for a linear classifier that allows for some errors; this type of classifier is called a *soft margin classifier* (James et al., 2014). Soft margin classifier allows misclassifying a few training observations (from the training data that are used for algorithm learning process) in order to perform good classification results for the remaining data. (Figure 5.17).
Figure 5.17 SVM example with 2 samples sets: red and green. Bigger samples represent support vectors (SV), decision surface is maximally far away from any data point representing optimal separating hyperplane principle, one green sample is on the other side of separating hyperplane, grouped together with red points, this misclassified sample illustrates the principle of soft margin classifier, which allows us to misclassify a few training data to perform good classification results for the remaining data.

**Kernels function**

Support Vector Machines belong to the class of Kernel Methods. As all kernel-based learning algorithms, they are composed of a general-purpose learning machine (in the case of SVM a linear machine) and a problem specific kernel function (Hofmann, 2006). Since the linear machine can only classify the data in a linear separable feature space, the role of the kernel-function is to induce such a feature space by implicitly mapping the training data into a higher dimensional space where the data is linearly separable. This process is called “Kernel Trick”.

Figure 5.18 illustrates a simple example of the “Kernel Trick”, where we are mapping a nonlinear function \( \Phi: I = R^2 \rightarrow F = R^3 \) from the 2-dimensional input space I into the 3-dimensional feature space F, which is defined in the following way:

\[
\Phi(\vec{x}) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T
\]  

(Eq. 5.9)
Taking the equation for a separating hyperplane into account (Eq. 5.10):

\[ \mathbf{w}^T \mathbf{x} + b = 0 \]  

(Eq. 5.10)

where \( \mathbf{w} \) is the normal to the hyperplane and \( b \) is the perpendicular distance of the hyperplane to the origin, we get a linear function in \( \mathbb{R}^3 \):

\[ \mathbf{w}^T \Phi(\mathbf{x}) = w_1 x_1^2 + w_2 \sqrt{2} x_1 x_2 + w_3 x_2^2 = 0 \]  

(Eq. 5.11)

Eq. 5.11 is an elliptic function when set to a constant \( c \) and evaluated in \( \mathbb{R}^2 \). Hence, with an appropriate mapping function, we can use our linear classifier in \( F \) on a transformed version of the data to get a nonlinear classifier in \( I \) with no effort. After mapping our nonlinear separable data into a higher dimensional space, we can find a linear separating hyperplane (Figure 5.18).

A kernel function can be interpreted as a kind of similarity measure between the input objects, which we define to perform data classification. In practice, several kernels are available in Support Vector Machines. The most commonly used are linear, polynomial and Gaussian Radial Basis kernels.
Chapter 5. Metric Space Approach and Machine Learning Classification

Equation 5.12 is known as a linear kernel because the support vector classifier is linear in the features:

\[ K(x_i, x_{i'}) = \sum_{j=1}^{p} x_{ij} x_{i'j} \]  \hspace{1cm} \text{(Eq. 5.12)}

where \( K \) is a kernel function, \( K(x_i, x_{i'}) \) a generalisation of the inner product of 2 vectors, and \( p \) is the number of parameters.

The linear kernel quantifies the similarity of a pair of observations using Pearson correlation. Figure 5.19 shows an example of the linear kernel applied to a nonlinear data set (toy example). Different colours of the points represent two different classes that need to be separated by the classifier. We could observe the linear kernel obviously fails to separate the points into two different classes appropriately.

![Figure 5.19 Left: Two class data set example (blue and pink points). Right: Support Vector classifier liner boundary results in poor classification on nonlinear data set (from James et al., 2014)](image)

Equation 5.13 represents another type of kernel – a polynomial kernel:

\[ K(x_i, x_{i'}) = (1 + \sum_{j=1}^{p} x_{ij} x_{i'j})^d \]  \hspace{1cm} \text{(Eq. 5.13)}

where \( d \) is an order of polynomial kernel, which defines the smoothness of the function.
Chapter 5. Metric Space Approach and Machine Learning Classification

The polynomial kernel provides much more flexible decision boundaries, which is especially useful in a higher-dimensional space involving polynomials of degree $d$, than in the original feature space.

Figure 5.20 (left) shows an example of an SVM classification with a polynomial kernel applied to nonlinear data. This example shows a significant improvement of the classification comparing to the linear kernel (Figure 5.19).

![Figure 5.20](image)

*Figure 5-20 Two class data set example (blue and pink points). Left: An example of SVM classification with a polynomial kernel of degree 3 applied to the nonlinear data set. Right: An example of SVM with a radial kernel to the same nonlinear data set. Both kernels show the capability of capturing the appropriate decision boundaries (From James et al., 2014)*

Another popular choice is Gaussian Radial Basis (RBF) kernel (Figure 5.20, right). Equation 5.14 represents Gaussian Radial Basis kernel as follows:

$$K(x_i, x_i') = \exp \left( -\gamma \sum_{j=1}^{P} (x_{ij} - x_{i'j})^2 \right)$$  \hspace{1cm} (Eq. 5.14)

where $\gamma = \frac{1}{2\sigma^2}$, and $\sigma$ is the kernel width.
Chapter 5. Metric Space Approach and Machine Learning Classification

The radial kernel has very local behaviour, defined by the fact that only nearby training observations will have an effect on the class label of a test observation. For example, if a test observation $x' = (x'_1 \ldots x'_p)^T$ is far from a training observation $x_i$, defined by Euclidian distance, then $\sum_{j=1}^{p}(x'_j - x_{ij})^2$ will be large, and $K(x', x_i) = \exp (-\gamma \sum_{j=1}^{p}(x'_j - x_{ij})^2)$ will be very small. In other words, training observations that are far from $x'$ will not influence the prediction of a class label for $x'$.

The main advantage of using kernels is computational, as one only needs to compute $K(x_i, x_{i'})$ for all given pairs $i, i'$ (James et al., 2014).

In this thesis, a Gaussian RBF kernel was applied as it has a capability of capturing appropriate decision boundaries for our data in the metric space.

Here we provide the description of the Gaussian RBF kernel parameters and their tuning process.

**Gaussian RBF parameters and their tuning**

The main parameters to tune for RBF kernel are: $\sigma$ - the kernel width, and $C$ – the cost or complexity factor (determines the number and severity of the misclassification that can be tolerated).

The choice of appropriate parameter values is very important. To define the optimal parameter values, a comprehensive search based on computing training and testing error surfaces of kernel width ($\sigma$) and cost parameter ($C$) is recommended.

The following examples are based on real data (Kanevski et al., 2002). The original data (both coordinates $(x,y)$ and porosity values) were rescaled into a $[0,1]$ interval. Original continuous data of porosity were transformed into “low” and “high” values to create a binary problem: porosity higher than 0.5 were coded as $+1$ and less than 0.5 as $-1$. The data were split into training data (200 samples) and testing data (94 samples).

Training data is used to understand the hidden relationships among the data points. Testing data is used to validate the found relationships. The testing set should be big enough and representative to perform the algorithm validation.
We perform a tuning ($K$-fold cross-validation is recommended to speed up the computations (Kanevski and Maignan, 2004)) to select the appropriate values for SVM parameters (kernel width ($\sigma$) and cost parameter ($C$)). In this example, the simple grid search is demonstrated. Figure 5.21 represents the error surface of training (left) and testing error (right) for different SVM parameters combination, for example for $\sigma = 0.1$ and $\log(C) = 4$ the training error is 0, which means that all the training data were classified correctly (from Kanevski et al., 2009).

From this example, we could analyse the general behaviour of the error surfaces (Kanevski et al., 2009):

1) The training error is small and even zero or even equal to zero in the region with the small kernel widths (Figure 5.21, left). This region is the overfitting region. The generalisation of the model in this region is bad and the testing error (Figure 5.21, right) is high.

2) In the region of high values of kernel width, the training error is high (Figure 5.21, left). The testing error is increasing after reaching a minimum (at intermediate bandwidth) (Figure 5.21, right). The effect in this region is called oversmoothing.

3) An optimal region is reached at intermediate values of kernel width and $C$ parameter (Figure 5.21, left and right).

Generally, the training error of the optimal model is likely to be low if the data are not too noisy. At the same time, it should not be too low as to lead to the overfitting. The value of the testing error is as important as the stability of the model in the neighbourhood of the selected parameters. The most stable model is likely to be the correct one. So, the training and the testing errors of the model change smoothly in the optimal region of low testing error in Figure 5.21 (right). The minimum of the testing error is quite well defined.
Chapter 5. Metric Space Approach and Machine Learning Classification

![Figure 5-21 SVM error surface. Left: Training error surface. The error increases with kernel bandwidth for a fixed value of C. For a fixed value of kernel bandwidth, the training error decreases with C; Right: Testing error surface. Very low values of σ result in overfitting, which is described by zero training and high testing error. The region of optimal σ and C combinations is evident (From Kanevski et al., 2009)](image)

The choice of the parameters affects the prediction results based on the classification solution. Figures 5.22 – 5.24 illustrate the prediction results of porosity data using different kernels and different values of parameters (example from Kanevski et al., 2009).

In Figure 5.22 we see the prediction results of porosity data using the Gaussian RBF kernel with a fixed value of C and different values of bandwidths. Large values of bandwidth give simple decision boundaries and result in oversmoothing (Figure 5.22, left). Small values of bandwidth lead to very complex decision boundaries and overfitting (Figure 5.22, right).

Figure 5.23 shows the same tendency for the polynomial kernels of increasing order: large values of bandwidth lead to oversmoothing (Figure 5.23, left) and small values of bandwidth lead to overfitting (Figure 5.23, right).

Figure 5.24 illustrates the dependence between the prediction results and the value of C parameter. The smooth decision boundaries do not represent the training data pattern for small values of C (Figure 5.24, left), whilst the big value of C leads to overfitting (Figure 5.24, right).
Chapter 5. Metric Space Approach and Machine Learning Classification

Figure 5-22 Prediction mapping of porosity data using Gaussian RBF kernels. From left to right: the value of bandwidth decreases, the complexity of the decision boundaries increasing leading from oversmoothing to overfitting (From Kanevski et al., 2009)

Figure 5-23 Prediction mapping of porosity data using polynomial kernels of increasing order. From left to right: the value of bandwidth decreases, the complexity of the decision boundaries increases (from Kanevski et al., 2009)

Figure 5-24 Prediction maps for increasing value of parameter C from left to right: the decision boundaries are too smooth and do not reflect the training data pattern (on the left map), they are close to optimal (on the central map), and they are too complex (on the right map) (From Kanevski et al., 2009)
Multi-class classification with SVM

Originally, SVM was developed as a binary classifier. Over the years, a number of proposals have been made to extend SVM to classify data with more than two classes. Such extension is called multi-class classification. Multi-class classification is used in this thesis because we classify 11 training images into 11 classes. The most popular approaches to perform multi-class classification are called one-versus-one and one-versus-all approaches.

One-versus-all classification approach

In the one-versus-all approach, we train $K$ SVMs, where $K$ is the number of classes, each time comparing one of the $K$ classes to the remaining $K - 1$ classes (James et al., 2014). This strategy involves training a single classifier per class, with the samples of that class as positive samples and all other samples as negatives.

Let $\beta_{0k}, \beta_{1k}, \ldots, \beta_{pk}$ denote the parameters that result from fitting and SVM comparing the $k$th class (coded as +1) to the others (coded as -1). Let $x^*$ define a test observation. We assign the observation to the class for which $\beta_{0k} + \beta_{1k}x_1^* + \beta_{2k}x_2^* + \cdots + \beta_{pk}x_p^*$ is the largest, as this amounts to a high level of confidence that the test observation belong to the $k$th class rather than to any other of the classes (James et al., 2014).

Although this strategy is popular, it has the following problem: even if the class distribution is balanced in the training set, the binary classification learners see unbalanced distributions because typically the set of negatives the learners see is much larger than the set of positives (Bishop, 2006).

One-versus-one classification approach

In this approach, a binary SVM classifier is constructed for each pair of classes and build an ensemble. For example, if the multi-class problem has $n$ classes, the one-versus one ensemble of classifiers will be composed of $n(n-1)/2$ classifiers. So, for three classes A, B and C the ensemble will be composed of 3 (= 3 * (3 - 1) / 2) binary classifiers. The first classifier will discriminate A from B, the second A from C, and the third B from C. The label assigning stage is then performed by majority voting. Say you have three classes, A, B, and C. The OVO ensemble will be composed of 3 (= 3 * (3 - 1) / 2) binary classifiers.
Chapter 5. Metric Space Approach and Machine Learning Classification

Now, if $x$ is to be classified, $x$ is presented to each binary classifier of the ensemble to create a vector of individual classifications, e.g. (A, B, B). The final step of this approach is to assign a label to $x$ based on which class were defined the most times in a vector of individual classification (so-called majority voting). In this example, $x$ would belong to class B.

One-versus-one is more computational demanding than one-versus-all. However, one-versus-one has been reported to produce more robust results (Hsu and Lin, 2002). Therefore, in this thesis, we use the one-versus-one approach in R interface to libsvm in package e1071 (Chang and Lin, 2011).

Advantages of SVM

The main advantages of SVM are the following:

1) Performs well for high dimensional problems;
2) Delivers unique solution;
3) Provides good generalisation performance;
4) Does not require extra knowledge about kernels to perform the transformation (“Kernel trick”).

Disadvantages of SVM

Nevertheless, it has some disadvantages as well:

1) It could be rather slow when performing tuning for the large data set,
2) It could be quite tricky to avoid data overfitting,
3) It lacks transparency in the results.

5.4.3 Random Forest

The original algorithm for random decision forests was created by Ho (Ho, 1995) and an extension of this algorithm – Random Forest – was developed by Breiman and Cutler (Breiman, 2001, Liaw, 2013). Breiman and Cutler combined the original random selection of features (Ho, 1995) with the “bagging” feature (see below for more details on bagging feature).
Random Forest has been successfully applied for high dimensional problems with a large number of training examples in different areas, for example, in gaming - Kinect for Xbox 360 (Shotton et al., 2013), in vision – scene classification (Bosch et al., 2007), and in medicine – digital anatomy (Criminisi et al., 2013). However, Random Forest is fairly new for the petroleum industry (Aulia et al., 2014). Aulia et al., 2014 is recommended for the Random Forest introduction reading.

Random forest is called an ensemble learning method because it consists of several decision trees combined into an ensemble. As a result of the learning process, random forest outputs the class that is the mode of the classes of the individual trees for classification problems and the class that is the mean prediction of the individual trees for the regression tasks (Ho, 1995).) Every decision tree is made by randomly selecting data from the available data set and then amalgamating together to get an accurate and stable decision (Ali et al., 2012). The main advantage of random forest over a single decision tree is that it corrects for decision trees’ habit of overfitting to their training set (Friedman et al., 2001).

Each decision tree combined in an ensemble in random forest consists of the following parts (Figure 5.25):

- **Root Node** – it represents the entire data set; further it is divided into two or more homogeneous sets. The dividing is done by splitting a node into two or more sub-nodes.

- **Decision Node** – when a sub-node splits into more sub-nodes, then it is called a decision node.

- **Terminal Node (Leaf)** – nodes that do not split are called Terminal or Leaf node.

- **Branch / Sub-Tree** – a sub section of the entire tree is called branch or sub-tree.

- **Splitting** – is defined by a user specified parameter during tuning and it depends on the number of the input variables (more details in “Tuning parameters of
Chapter 5. Metric Space Approach and Machine Learning Classification

Random Forest” below). As a result of splitting a decision or a terminal node is formed.

In Random Forest, to classify a new object based on the input attributes, each tree predicts (outputs) a class number (the class that is the mode of the classes of the individual trees) and then the Forest chooses the final classification results based on the most frequently outputted class from all the trees in the Forest (Figure 5.26).

In this research, we interrogate the multipliers on the eigenfunction on each node. The nodes are formed automatically within the algorithm after splitting. The algorithm chooses the order of the input attributes to split at random (Breiman, 2001). We end up with 11 classes. The number of output classes is specified at the beginning of the classification as RF is a supervised algorithm which means that it learns from the known output. The final classification prediction (number of a training image) is derived from each tree predictions as follows: for example, from 100 trees 20 predicted as training image 1, and all other trees predicted the rest of the training images (8 trees – TI 2, 8 trees – TI 3 and so on), then the final prediction will be training image 1 as the majority of predicted (Breiman, 2001).

Figure 5-25 A schematic example of a decision tree (Analytics Vidhya Content Team, 2016)
Bagging feature

Bagging or bootstrap aggregation is an important feature of Random Forest. Bagging repeatedly selects a random sample with replacement of the training set and fits trees to these samples. Random Forest uses a so-called "feature bagging", that selects at each candidate split a random subset of the features. This procedure attempts to reduce the correlation between estimators in an ensemble by training them on random samples of features instead of the entire feature set. The reason for "feature bagging" instead of simple bagging is to avoid the correlation between the trees: if one or a few features are very strong predictors for the response variable, these features will be selected in many of the trees, causing them to become correlated, which could affect the accuracy of the classification (Breiman, 1996). Bootstrapping and ensemble scheme allows Random Forest to overcome the overfitting problem (Ali et al., 2012).

In Random Forest, each tree is grown as follows (Breiman, 2002):

1. If the number of cases in the training set is \( N \), a sample of these \( N \) cases is taken randomly and with replacement ("feature bagging") from the original data. This sample will be used as the training set for growing the tree.
Chapter 5. Metric Space Approach and Machine Learning Classification

2. If there are \( M \) input variables, a number \( m \) is selected such that \( m < M \) at each node; \( m \) variables are selected at random out of the \( M \). The best split on these \( m \) is used for splitting the node. The value of \( m \) is kept constant during the forest growing.

3. Each tree is grown to the largest possible extent without pruning.

4. New data are predicted by aggregating the predictions of all trees (majority votes for classification, the average for regression).

In each bootstrap training set, about one-third of the data is left out. This data is called out-of-bag. This data is used to estimate the training error of the classifier. This is a Random Forest cross validation method. It is very similar to leave one out validation technique, but much faster. This method simply tags every observation used in different trees and then it finds out a maximum vote score for every observation based on only trees which did not use this particular observation to train itself.

**Tuning parameters of Random Forest**

Random Forest is quite fast and easy to tune. Two main tuning parameters of Random Forest are: \( ntree \) (\( R \) interface nomenclature) – number of trees, and \( mtry \) (\( R \) interface nomenclature) – number of variables available for splitting at each tree node.

To choose the appropriate number of trees (\( ntree \)) to build, one should have in mind that a higher number of trees give better but slower performance. At some point, the increased number of trees doesn’t perform any significant improvements in the performance. Therefore, it is recommended to perform validation for a different number of trees to find out the optimal number.

It is a bit more complicated to choose the number of variables for splitting (\( mtry \)) parameter value. The default value differs depending on which software (or \( R \) package) is used. For example, in `randomForest` (\( R \) package) the default of \( mtry \) is the square root of the number of predictor variables (for classification models), and it is the number of predictor variables divided by 3 (for regression models). In `party` (\( R \) package) the default is always 5.
Chapter 5. Metric Space Approach and Machine Learning Classification

There is an extensive discussion about the influence of \( mtry \) on the output reported in the literature. Cutler et al. (2007) showed that different values of \( mtry \) did not affect the correct classification rates of their model. They also mentioned that other performance metrics (sensitivity, specificity, etc.) were stable under different values of \( mtry \). Contrariwise, Strobl et al. (2008) reported that \( mtry \) had a strong influence on predictor variable importance estimates.

Due to such conflicting evidence reported in the literature, it is recommended to start with the default value of \( mtry \) and perform validation for different \( mtry \) values to find the one that works better for the specific model.

**Variable Importance**

As mentioned before, when we create a bagged ensemble of trees, we improve prediction accuracy over prediction achieved using a single decision tree (James et al., 2014). However, the ensemble of bagged trees is much more difficult to interpret, for example, it is no longer clear which variables are most important to the process.

Several measures are available to estimate variable importance in Random Forest:

1) **Gini Importance** or **Mean Decrease in Impurity (MDI)** computes each variable importance as the sum over the number of splits (across all trees) that include the variable, proportionally to the number of samples it splits. Each time a particular variable is used to split a node, the Gini coefficient for the new node is calculated and compared to the original node. A low Gini (i.e. high decrease in Gini) means that a particular variable plays an important role in partitioning data into classes.

2) **Permutation Importance** or **Mean Decrease in Accuracy (MDA)** is evaluated for each variable by removing the relation between that variable and the target. This is achieved by randomly displacing the values of the variable and measuring the resulting increase in error. The mean decrease in accuracy is obtained during the out-of-bag error calculation. The more the accuracy of Random Forest decreases due to exclusion of a single variable, the more important the variable is. Therefore, variables with a large mean decrease in accuracy are more important for classification.
Chapter 5. Metric Space Approach and Machine Learning Classification

These measures can be estimated for each class, which is very useful for detecting the variables that might degrade performance for a specific class whilst acting positively on average.

**Advantages of Random Forest**
Random Forest has a very significant list of advantages (Ali et al., 2012):

1) It is fast and efficient;
2) It is suitable for high dimensional data modelling because it can handle continuous, categorical, and binary data as well as missing values;
3) It can handle thousands of input variables. Also, it could identify the most significant variables and work as dimensionality reduction method;
4) It provides the option to use the out-of-bag error estimate instead of setting aside a test set;
5) It produces accurate predictions and good generalisation which are achieved due to bootstrapping and ensemble scheme;

**Disadvantages of Random Forest**
Random Forest has a couple of disadvantages:

1) It may overfit for some datasets with noise classification task;
2) It can feel like a black box approach for some statistical modellers because there is a very little control on what the model does. All that the modeller can do is to try different parameters and random seeds.

5.5 Assessing Multi-class Classification Performance Quality
To evaluate the multi-class classification performance, we first analyse the individual performance of each class and then average the values to get the overall estimation. The common way to evaluate the results of classification problem is to calculate the accuracy, which is the proportion or percentage of correctly predicted labels over all predictions. Nevertheless, to evaluate multi-class classification problem results, accuracy is not very representative. Accuracy alone is sometimes quite misleading, for example, in a multi-class problem a model with relatively 'high' accuracy might be predicting the 'not so
important’ class labels fairly accurately but making mistakes on the classes that are actually critical to the application.

Therefore, to analyse the individual performance of each class of classification, we calculate precision and recall for each class label. After that, we average the values of precision and recall getting the overall estimation.

**Precision and Recall**

*Precision* (also called positive predictive value) is the fraction of retrieved instances that are relevant. In other words, precision shows how many instances were correctly predicted or how useful the search results are.

*Recall* (also known as sensitivity) is the fraction of relevant instances that are retrieved. In other words, how many instances were correctly captured or how complete the search results are.

Figure 5.27 illustrates what precision and recall represent on a two-class example. For classification tasks, the terms – true positives, true negatives, false positives, and false negatives – are used to compare the results of the classifier to the test data. The terms positive and negative relate to the classifier's prediction (sometimes known as the expectation), and the terms true and false relate to whether that prediction corresponds or does not correspond to the observation.

Thereby, we could calculate precision and recall for two-class classification problem as follows:

\[
\text{Precision of first class} = \frac{\text{true positive of first class}}{\text{total predicted of first class}}
\]

\[
\text{Recall of first class} = \frac{\text{true positive of first class}}{\text{total of first class}}
\]
Chapter 5. Metric Space Approach and Machine Learning Classification

 Whilst it is straightforward to compute precision and recall for a binary classification problem, it can be quite confusing for a multi-class classification problem. The following steps are recommended to calculate precision and recall for multi-class classification problem:

1) Create a confusion matrix for all the classes (usually most of the machine learning packages generate it automatically); A confusion matrix, also known as an error matrix (Stehman, 1997), is a specific table layout that provides a visualisation of the performance of a classification algorithm. Each column of the matrix represents the instances in an actual class, while each row represents the instances in a predicted class. The diagonal of the confusion matrix shows how many instances were classified correctly. The confusion matrix is also useful to track if something goes wrong during the classification process.

2) Calculate precision and recall as described before for the first class.

3) Repeat the same calculations for each class.

4) Average the values of precision and recall to get the overall estimation.
5.6 Geological Metric Space Classification Results

To perform training, we divided the whole data set into training (70%) and testing (30%) data. We used this partitioning for $k$-means clustering and SVM classification approaches. This partitioning allows us to perform the validation of our classification results. RF does not require separation of data into training and testing sets as it is done automatically when RF is learning from data. In Random Forest, only 2/3 of the total data is used for growing each tree (training data), whilst the remaining 1/3 is left out (out-of-bag) (testing data). This 1/3 then is used to perform the validation. Out-of-bag (OOB) error or out-of-bag estimate is the mean prediction error that represents the misclassification rate of the Random Forest classification model. It is calculated using predictions from the trees that do not contain the training observations.

Based on the results of Multidimensional scaling showing that the model realisations from different training images are mixed together (Figures 5.3 – 5.6) we propose to use additional input parameters (geological parameters) hoping to improve the learning outcome. Therefore, we use two different input data sets for the machine learning classification: 1) where we use 12 multipliers on the eigenfunction only, and 2) where we use 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation). Even though multipliers on eigenfunction already represent a combination of a training image and the geological parameters, we still can use geological parameters as the additional input data because machine learning classifiers can automatically handle the information redundancy in the inputs. Redundancy, according to Phatak (1999), can even improve the generalisation ability of the classifier, avoiding the overfitting.

5.6.1 $k$-means clustering

Set up

The main aim of $k$-means clustering is to partition the metric space into the number of clusters. We chose to partition the metric space into 11 clusters, which is equal to the number of training images. By partitioning the metric space into 11 clusters, we ensure that we include all training images into the following history matching process. $k$-means clustering classification was performed with clue library in R (Chang et al., 2016).
Chapter 5. Metric Space Approach and Machine Learning Classification

Classification results – Case 1
(12 multipliers on the eigenfunction only)

We tested the classifier on a test data set and show the results in the form of a confusion matrix.

Table 5.2 shows the confusion matrix of $k$-means classification results on the test data set with 12 multipliers on the eigenfunction as the input data. The diagonal of the matrix shows how many models of each training images were classified correctly (percentage of the total number). We can see that the values in the diagonal of confusion matrix are very low. There is only one training image – TI 8 – that is 50% classified as TI 8. Training images TI 5 and TI 11 got 20%. All the rest of training images are misclassified.

Table 5.3 shows how many models were predicted as a certain training image. For example, the biggest number of models were predicted as training images TI 5 and TI 11. Very few models were predicted as training images TI 1, TI 2, TI 7, and TI 10. No models were predicted as training images TI 4 and TI 9.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>TI1</th>
<th>TI2</th>
<th>TI3</th>
<th>TI4</th>
<th>TI5</th>
<th>TI6</th>
<th>TI7</th>
<th>TI8</th>
<th>TI9</th>
<th>TI10</th>
<th>TI11</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI1</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI3</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI5</td>
<td>20</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI6</td>
<td>40</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>TI7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI8</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI10</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TI11</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>90</td>
<td>50</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 5-2 Confusion matrix of $k$-means classification results on testing data set received with 12 multipliers on the eigenfunction as the input data

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>14</td>
<td>0</td>
<td>24</td>
<td>15</td>
<td>3</td>
<td>13</td>
<td>0</td>
<td>7</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 5-3 Total number of models of each training images predicted as a certain training image
Chapter 5. Metric Space Approach and Machine Learning Classification

Classification results – Case 2
(12 multipliers on the eigenfunction and geological parameters)

Table 5.4 shows the confusion matrix of k-means classification results on the test data set received with 12 multipliers on the eigenfunction and SNESIM MPS geological modelling parameters (affinity and rotation) as the input data. In this case, we also have few models that were classified correctly (in the diagonal of confusion matrix). There is again only one training image TI 3 that is 50% classified as TI3. Training images TI1, TI8, and TI11 got 20%. All the rest of the training images are misclassified.

In Table 5.5 we calculate how many models were predicted as a certain training image. For example, the biggest number of the models were predicted as training image TI8. Very few models were predicted as training images TI2, TI5, TI7, and TI9. No models were predicted as training images TI4.

To evaluate the total multi-class classification results we calculate recall and precision values (Section 5.5) for each class. The highest value of recall and precision identifies better results of the classification. The average results of recall and precision of k-means

### Table 5.4: Confusion matrix of k-means classification results on testing data set received with 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data

<table>
<thead>
<tr>
<th>Predicted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>50</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>40</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40</td>
<td>40</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>30</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>60</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>40</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>20</td>
<td>40</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

### Table 5.5: Total number of models of each training images predicted as a certain training image

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>17</td>
<td>1</td>
<td>15</td>
<td>0</td>
<td>3</td>
<td>11</td>
<td>1</td>
<td>28</td>
<td>1</td>
<td>14</td>
<td>19</td>
</tr>
</tbody>
</table>

Multi-class classification evaluation

To evaluate the total multi-class classification results we calculate recall and precision values (Section 5.5) for each class. The highest value of recall and precision identifies better results of the classification. The average results of recall and precision of k-means
clustering are presented in Table 5.6. From the Table 5.6, we could see that inclusion of SNESIM geological modelling parameters as input data slightly improves the classification results.

<table>
<thead>
<tr>
<th>k-means clustering</th>
<th>Multipliers only</th>
<th>Multipliers and geology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>0.08</td>
<td>0.1</td>
</tr>
<tr>
<td>Precision</td>
<td>0.008</td>
<td>0.04</td>
</tr>
</tbody>
</table>

*Table 5-6 The average values of recall and precision for k-means clustering for the test data set with two different input scenarios*

Overall, k-means clustering performed poorly on this data.

### 5.6.2 Support Vector Machines

**Set up**

Support Vector Machines (SVM) use a kernel function to perform classification in high-dimensional space. For this case, we used Gaussian Radial Basis (RBF) kernel to perform classification. Two main parameters of RBF: $\gamma$ - the positive constant ($\gamma = \frac{1}{2\sigma^2}$, where $\sigma$ is the kernel width), and $C$ – the cost (a parameter that allows the misclassification). We tuned these parameters to get the optimal combination for the best performance of the classifier.

SVM one-versus-one multi-class classification was performed with *libsvm* in *e1071* library in *R* (Chang and Lin, 2011).

**SVM tune – Case 1**

*(12 multipliers on the eigenfunction only)*

The tuning process was performed on training and testing data to get the error surfaces. Figure 5.28 represents the training and testing error surface. We compared the training and testing error surfaces to choose the optimal parameter combinations that give us the minimum error value but do not overfit.

Training error surface has a much bigger area corresponding to the small error values comparing to the test error. Also, the minimum values of the training error are smaller than testing ($0 < 5$). The choice of the parameters from the region of minimum training
error values ($\gamma = 1.4991$, $\text{C}= 20$, Error $= 0.1527$, represented with a cross) would lead to an overfitting and bad generalisation of the model. We chose the optimal combination of parameters that gives us the minimum error on testing error surface: $\gamma = 1.2121$, $\text{C}= 4$, Error $= 0.1527$, represented with a cross.

![Figure 5-28 SVM training error surface (left) and testing error surface (right) for 12 multipliers on the eigenfunction only input data. The optimal combination of parameters is marked with a cross](image)

**Classification results – Case 1**

**(12 multipliers on the eigenfunction only)**

Table 5.7 shows the confusion matrix of SVM classification results on testing data set. The diagonal of the matrix shows how many models of each training images were classified correctly (percentage of the total number). There is a noticeable improvement in the number of the training images that were classified correctly compared with $k$-means clustering results (Table 5.2).

If we sum up all the numbers in each row and divide them by 100, we get how many models were predicted as a certain training image (Table 5.8). For example, the biggest number of models were predicted as training images TI 2 and TI 9. Fewer models were predicted as training images TI 1, TI 3, TI 6, TI 7 and TI 8.

SVM confusion matrix also shows that half of the models from training image TI 3 were predicted as the other training image TI 2 (Table 5.7). In this case, we suggest performing a separate classification just between TI 2 and TI 3. Table 5.9 shows that the additional
Chapter 5. Metric Space Approach and Machine Learning Classification

classification between two training images (TI 2 and TI 3) could significantly improve the classification results when conducted separately.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>20</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>10</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>40</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>90</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 5-7 Confusion matrix of SVM classification results on testing data set received with 12 multipliers on the eigenfunction as the input data

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>8</td>
<td>14</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>14</td>
<td>13</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 5-8 Total number of models of each training images predicted as a certain training image

<table>
<thead>
<tr>
<th>Predicted</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>70</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 5-9 Confusion matrix of SVM classification results on testing data set for two training images TI 2 and TI 3 only

**SVM tune – Case 2**

(12 multipliers on the eigenfunction and geological parameters)

Figure 5.29 represents the training and testing error surface for 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data. We can see that the training error surface with small error values is much bigger (Figure 5.29, left) compared with the small error values area for the testing surface (Figure 5.29, right). The minimum values of the training error are smaller than testing errors (0 < 6.2). The combination of tuning parameters from the region of minimum training error values ($\gamma = 1.0201$, $C = 21$, Error = 0.0969, represented with a cross) would
lead to an overfitting and bad generalisation of the SVM model. Instead we chose the optimal combination of parameters that gives us the minimum error (Error = 6.21) on testing error surface: \( \gamma = 0.7681, C = 11 \), represented with a cross.

![Figure 5-29 SVM train error surface for 12 multipliers on the eigenfunction and 4 SNESIM geological modelling parameters (affinity and rotation) input data. The optimal combination of parameters is marked with a cross.](image)

**Classification results – Case 2**

*(12 multipliers on the eigenfunction and geological parameters)*

Table 5.10 shows the confusion matrix of SVM classification results on testing data set. The diagonal of the confusion matrix shows the correctly classified models (percentage of the total number). The values in this diagonal of the matrix are slightly worse than the values from the previous results of 12 multipliers on the eigenfunction only (Table 5.7).

Table 5.11 represents how many models were predicted as a certain training image. For example, the biggest number of models were predicted as training images TI 3 and TI 10. Fewer models were predicted as training images TI 1, TI 2, TI 5, TI 6, and TI 8.

SVM confusion matrix also shows that some of the models from training image TI 3 were predicted as training image TI 2, and 50% of the models from training image TI 2 were predicted as TI 3 (Table 5.10). In this case, we performed a separate classification just between TI 2 and TI 3. Table 5.12 shows that the additional classification between two training images significantly improved the classification results when conducted separately. There is no difference in the results of SVM classification for two training
Chapter 5. Metric Space Approach and Machine Learning Classification

images TI 2 and TI 3 performed with two different input parameters sets (Table 5.9 and Table 5.12).

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.10 Confusion matrix of SVM classification results on testing data set received with 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5</td>
<td>8</td>
<td>15</td>
<td>10</td>
<td>9</td>
<td>6</td>
<td>11</td>
<td>8</td>
<td>13</td>
<td>15</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 5.11 Total number of models of each training images predicted as a certain training image

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 5.12 Confusion matrix of SVM classification results on testing data set received with 12 multipliers on the eigenfunction and SNESIM geological modelling parameters (affinity and rotation) as the input data for two training images TI 2 and TI 3

Multi-class classification evaluation – Case 1 and Case 2

To evaluate the total multi-class classification results we calculate recall and precision values (Section 5.5) for each class. The highest value of recall and precision identifies better results of the classification. The average recall and precision results of SVM classification for both input data cases are presented in Table 5.13. From the Table 5.13 we could see that SNESIM geological modelling parameters as input data did not improve the classification results, on the contrary, they worsened both recall and precision values by approximately 16%.
5.6.3 Random Forest

Random Forest multi-class classification was performed in RandomForest library in R (Liaw, 2013).

Random Forest tune – Case 1

As was mentioned before, two main parameters to tune of Random Forest are: ntree (R interface nomenclature) – number of trees and mtry (R interface nomenclature) – number of variables available for splitting at each tree node.

The forest error rate depends on:

1. The correlation between any two trees in the forest: the higher the correlation, the bigger the error rate.
2. The strength of each individual tree: the tree with a low error rate is considered as a strong classifier. It means that by increasing the strength of the individual trees we can decrease the total forest error rate.

Number of trees – Case 1

(12 multipliers on the eigenfunction only)

Generally, more trees lead to higher accuracy of the classifier. However, more trees also mean an increase in computational time. That’s why it is recommended to find the optimal number of trees with stabilised value of the out-of-bag error, after which the improvement in the error rate is negligible.

To do that, we set the other parameter – mtry – to the default value, and built Random Forest with a different number of trees (100, 200, …, 2500) several times (10 times). We recorded the OOB error rate to find where the OOB rate stabilises and reaches the minimum.

<table>
<thead>
<tr>
<th>SVM</th>
<th>Multipliers only</th>
<th>Multipliers and geology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>0.6909</td>
<td>0.5272</td>
</tr>
<tr>
<td>Precision</td>
<td>0.7258</td>
<td>0.5634</td>
</tr>
</tbody>
</table>

Table 5-13 The average values of recall and precision for SVM classification for the test data set for two different input scenarios
Figure 5.30 represents the OOB error rate (dark green line in the middle of the plot) and the individual error rates for each class (training image). From the figure, we could see that the error rate is stabilising faster for some training images than for the others, for example, it is faster for training image TI 6 than for TI 1. Most importantly, we observe that the OOB rate is quite stable after 300 trees, with the minimum values with 500, 700, 1000 and 1500 trees.

![Graph showing error rate against number of trees](image)

**Figure 5.30 Error rate against number of trees (12 multipliers on the eigenfunction input data)**

**Number of variables for splitting – Case 1**

*(12 multipliers on the eigenfunction only)*

Breiman’s recommendation (Breiman, 2002) for the number of variables for splitting (\(mtry\)) is to take a square root of a total number of predictors. This is the default value in the `RandomForest` library in R.

Nevertheless, there is a function `tuneRF` (in R) available for optimising this parameter. This function starts from the default value of \(mtry\) (the square root of the number of predictor variables = 3.5 in our case) and searches for optimal value to minimise the OOB error. In `tuneRF` one also needs to define a stepFactor. The stepFactor value defines how
mtry is inflated (or deflated) at each iteration. For example, on our case the stepFactor = 1.5. The search is going in both directions: first left, then right. To the left new mtry = default value/stepFactor. In our case, new mtry to the left equals 3.5/1.5 = 2.33, rounded up to the be an integer, which gives 3. To the right new mtry = default value*stepFactor. In our case, new mtry to the right equals 3.5*1.5 = 5.25, rounded up to the be an integer, which gives 6.

Then it goes to the left again by dividing the default mtry by 3 (1.5+1.5), which equals 3.5/3 = 1.16, rounded up to 2. The stopping criteria are defined by the value specified by improve, which measures the (relative) improvement in OOB error. The OOB improvement must be by the specified value (which is 0.05 in our case) for the search to continue. In our case, the was no OOB improvement after the third iteration and the search had stopped.

Figure 5.31 represents the results of mtry tuning for a different number of trees value: 500, 700, 1000 and 1500.

Table 5.14 represents the summary of the several considered optimal parameter combinations and the corresponding OOB error values. The smallest OOB error is achieved with 1500 trees and 6 variables for splitting. However, the OOB error value is slightly different every time we perform Random Forest for each combination of parameters due to bootstrap sampling. That is why we perform prediction for all 4 combinations of parameters to compare which one delivers better classification results.
Figure 5-31 OOB error value for different values of mtry with the four fixed Ntree values: 500, 700, 1000, 1500

<table>
<thead>
<tr>
<th>Parameter combination</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ntree</td>
<td>500</td>
<td>700</td>
<td>1000</td>
<td>1500</td>
</tr>
<tr>
<td>mtry</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>OOB error</td>
<td>46.91</td>
<td>46.73</td>
<td>46.36</td>
<td>46.55</td>
</tr>
</tbody>
</table>

Table 5-14 The optimal Random Forest parameter combinations and the corresponding OOB error values

Classification results – Case 1

(12 multipliers on the eigenfunction only)

Tables 5.15, 5.17, 5.19, and 5.21 show the confusion matrix of RF classification results on testing data set for four different parameter combinations respectively (see Table 5.14).
Chapter 5. Metric Space Approach and Machine Learning Classification

The diagonal of the confusion matrixes shows how many models of each training images were classified correctly (percentage of the total number). Models from training images TI 4, TI 5, TI 8, TI 9, TI 10, and TI 11 were classified much better than the rest of the training images for all parameter combinations.

Tables 5.16, 5.18, 5.20, and 5.22 represent the total number of predicted models as a certain training image (sum of all the values in a row/100). The highest number of models were predicted as training images TI 9. Fewer models were predicted as training images TI 1, TI 2, TI 3, TI 6, and TI 7.

From the RF confusion matrix, we could also see that from 30% to 40% of the models from training image TI 3 were classified as training image TI 2, and from 20% to 40% of the models from training image TI 2 were classified as TI 3 (Tables 5.15, 5.17, 5.19, and 5.21). To improve the results, we performed a separate classification just between training images TI 2 and TI 3. Table 5.23 shows that the additional classification between two training images significantly improves the classification results when performed separately.

<table>
<thead>
<tr>
<th>Predicted Actual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>30</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>40</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>70</td>
<td>10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5-15 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=500, mtry=3)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>11</td>
<td>12</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>14</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 5-16 Total number of models of each training images predicted as a certain training image (ntree=500, mtry=3)
Chapter 5. Metric Space Approach and Machine Learning Classification

<table>
<thead>
<tr>
<th>Predicted Actual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>40</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>30</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>80</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 5-17 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=700, mtry=3)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>11</td>
<td>11</td>
<td>7</td>
<td>6</td>
<td>10</td>
<td>16</td>
<td>14</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 5-18 Total number of models of each training images predicted as a certain training image (ntree=700, mtry=3)

<table>
<thead>
<tr>
<th>Predicted Actual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>50</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>20</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>80</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 5-19 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=1000, mtry=2)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>7</td>
<td>12</td>
<td>6</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>10</td>
<td>17</td>
<td>14</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 5-20 Total number of models of each training images predicted as a certain training image (ntree=1000, mtry=2)
Chapter 5. Metric Space Approach and Machine Learning Classification

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>10</td>
<td>30</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0</td>
<td>40</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5-21 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data (ntree=1500, mtry=6)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>15</td>
<td>14</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 5-22 Total number of models of each training images predicted as a certain training image (ntree=1500, mtry=6)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 5-23 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction only as the input data for two training images TI 2 and TI 3

**Number of trees – Case 2**

*(12 multipliers on the eigenfunction and geological parameters)*

We performed the same search approach for the optimal value of the number of trees for the classification with a new input set as we did for the Case 1. We started with the number of variables for the splitting parameter (mtry) equals to the default value, and built Random Forest with a different number of trees (100, 200, …, 2500) several times (10 times). We recorded the OOB error rate to find where the OOB rate stabilises and reaches the minimum.

Figure 5.32 represents the OOB error rate (bright green line in the middle of the plot) and the individual error rates for each class (training image). From the figure, we could see that the error rate stabilises faster for some training images than for the others, for
example, it is faster for training image TI 4 than for TI 3. We can also observe that for
some training images error rate is: first, decreasing, then stabilising, and then increasing
again when the number of trees is increasing. For example, we could see this effect for
training image TI 7: first, error rate decreases, then when the number of trees is around
1700, it stabilises, but when the number of trees is 2000 and greater, the error rate goes
up again. However, we observe that the OOB rate is quite stable after 300 trees, with the
minimum values with 600, 800, 1000 and 1300 trees.

![Error rate against number of trees for 12 multipliers on the eigenfunction and
geological parameters input data](image)

Figure 5-31 Error rate against number of trees for 12 multipliers on the eigenfunction and
geological parameters input data

**Number of variables for splitting – Case 2**

*(12 multipliers on the eigenfunction and geological parameters)*

We use the same function `tuneRF` (in R) for optimising the number of variables for
splitting as we used for the Case 1. `tuneRF` function starts from the default value of `mtry`
(the square root of the number of predictor variables = 5 in our case) and searches for
optimal value to minimise the OOB error.
Chapter 5. Metric Space Approach and Machine Learning Classification

The stepFactor value, in this Case, is also 1.5. The search is going in both directions: first left, then right. To the left new mtry = default value/stepFactor = 5/1.5 = 3.33, rounded up to the be an integer, which gives 4. To the right new mtry = default value*stepFactor = 5*1.5 = 7.5, rounded up to the be an integer, which gives 8. Then it goes to the left again by dividing the default mtry by 3 (1.5+1.5), which equals 5/3 = 1.6, rounded up to 2. The stopping criteria are defined by the value specified by improve, which measures the (relative) improvement in OOB error. The OOB improvement must be by the specified value (which is 0.05 in our case) for the search to continue. In our case, the was no OOB improvement after the third iteration and the search had stopped.

Figure 5.33 represents the results of mtry tuning for a different number of trees value: 600, 800, 1000 and 1300.

Table 5.24 represents the summary of the several considered optimal parameter combinations and the corresponding OOB error values. The smallest OOB error is achieved with 800 trees and 8 variables for splitting. However, as we mentioned before, the OOB error value is slightly different every time we perform Random Forest for each combination of parameters due to bootstrap sampling. That is why we perform classification for all 4 combinations of parameters to compare which one delivers better classification results.

Classification results – Case 2
(12 multipliers on the eigenfunction and geological parameters)

Tables 5.25, 5.27, 5.29 and 5.31 show the confusion matrix of RF classification results on testing data set (Case 2) for four different parameter combinations respectively (see Table 5.24).

The diagonal of the confusion matrixes shows how many models of each training images were classified correctly (percentage of the total number). Models from training images TI 4, TI 5, TI 8, TI 9, and TI 10 were classified much better than the rest of the training images for all four parameter combinations.
Chapter 5. Metric Space Approach and Machine Learning Classification

We show the total number of predicted models as a certain training image (sum of all the values in a row/100) in Tables 5.26, 5.28, 5.30, and 5.32. The greatest number of models were predicted as training images TI 9. Fewer models were predicted as training images TI 1, TI 3, TI 6, and TI 11.

Figure 5.32 OOB error value for different values of mtry with the four fixed Ntree values: 500, 700, 1000, 1500

<table>
<thead>
<tr>
<th>Parameter combination</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ntree</td>
<td>600</td>
<td>800</td>
<td>1000</td>
<td>1300</td>
</tr>
<tr>
<td>mtry</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>OOB error</td>
<td>45.09</td>
<td>43.27</td>
<td>45.27</td>
<td>44.18</td>
</tr>
</tbody>
</table>

Table 5.24 The optimal Random Forest parameter combinations and the corresponding OOB error values
Table 5-25 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=600, mtry=2)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>50</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>20</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>70</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>70</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
</tr>
</tbody>
</table>

Table 5-26 Total number of models of each training images predicted as a certain training image (ntree=600, mtry=2)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>9</td>
<td>11</td>
<td>6</td>
<td>11</td>
<td>13</td>
<td>7</td>
<td>7</td>
<td>10</td>
<td>15</td>
<td>13</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 5-27 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=800, mtry=8)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>50</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>20</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 5-28 Total number of models of each training images predicted as a certain training image (ntree=800, mtry=8)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>13</td>
<td>9</td>
</tr>
</tbody>
</table>
### Chapter 5. Metric Space Approach and Machine Learning Classification

#### Table 5-29 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=1000, mtry=4)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>50</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>20</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
</tr>
</tbody>
</table>

#### Table 5-30 Total number of models of each training images predicted as a certain training image (ntree=1000, mtry=4)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>14</td>
<td>13</td>
<td>9</td>
</tr>
</tbody>
</table>

#### Table 5-31 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data (ntree=1300, mtry=4)

<table>
<thead>
<tr>
<th>Predicted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>50</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>20</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>80</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
</tr>
</tbody>
</table>

#### Table 5-32 Total number of models of each training images predicted as a certain training image (ntree=1300, mtry=4)

<table>
<thead>
<tr>
<th>TI</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5</td>
<td>12</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>16</td>
<td>13</td>
<td>7</td>
</tr>
</tbody>
</table>
Chapter 5. Metric Space Approach and Machine Learning Classification

From the RF confusion matrixes, we could also see that 40% of the models from training image TI 3 were predicted as training image TI 2, and 20% of the models from training image TI 2 were predicted as TI 3 (Tables 5.25, 5.27, 5.29 and 5.31). To improve the results of the classification, we performed a separate classification just between TI 2 and TI 3. Table 5.33 shows that the additional classification between two training images significantly improves the classification results for these two classes.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>70</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>30</td>
<td>80</td>
</tr>
</tbody>
</table>

*Table 5.33 Confusion matrix of RF classification results on testing data set received with 12 multipliers on the eigenfunction and geological parameters as the input data for two training images TI2 and TI3*

**Variable Importance**

With Random Forest, we can estimate which variables are more important to the classification process. However, as we performed MDS beforehand, our multipliers on the eigenfunction are already sorted by their relative importance. That is why we perform variable importance estimation only for the Case 2 where we added the geological parameters.

We performed the estimation of two measures for variable importance: **Mean Decrease in Accuracy (MDA)** and **Mean Decrease Gini (MDG)** for four different Random Forest parameter combinations (Table 5.24) for Case 2 input data set (Figures 5.33 – 5.36). Multipliers on the eigenfunction 1 and 2 obviously have the highest decrease in both Accuracy and Gini, which indicates they are the most important variables for the classification, whilst the least important are the geological parameters (affinity and rotation).
Chapter 5. Metric Space Approach and Machine Learning Classification

**Figure 5-33** Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=600, mtry=2)

**Figure 5-34** Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=800, mtry=8)
Chapter 5. Metric Space Approach and Machine Learning Classification

Figure 5-35 Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=1000, mtry=4)

Figure 5-36 Mean Decrease in Accuracy (MDA) and Mean Decrease Gini (MDG) for 12 multipliers on the eigenfunction and geological parameters input data (ntree=1300, mtry=4)
Chapter 5. Metric Space Approach and Machine Learning Classification

**Multi-class classification evaluation – Case 1 and Case 2**

To evaluate the multi-class classification results we calculate recall and precision values (Section 5.5) for each class. The highest value of recall and precision identifies better results of the classification. The average recall and precision results of RF classification for Case 1 and Case 2 are presented in Table 5.29. From the Table 5.34, we could see that including SNESIM MPS geological modelling parameters as input data can slightly improve recall value for most parameter combinations (3 first parameter combinations), but not for the fourth one. Precision value is also high with geological parameters, with the maximum for the eighth parameter combination (Ntree = 1300, mtry = 4). We use the eighth parameter combination as the best one for the further comparison analysis between all three classification techniques.

<table>
<thead>
<tr>
<th>RF</th>
<th>Multiplication only</th>
<th>Multiplication only</th>
<th>Multiplication only</th>
<th>Multiplication + Multiplication only</th>
<th>Multiplication + Multiplication only</th>
<th>Multiplication + Multiplication only</th>
<th>Multiplication + Multiplication only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ntree, mtry</td>
<td>500 3</td>
<td>700 3</td>
<td>1000 2</td>
<td>1500 6</td>
<td>600 2</td>
<td>800 8</td>
<td>1000 4</td>
</tr>
<tr>
<td>Recall</td>
<td>0.6272</td>
<td>0.6454</td>
<td>0.6454</td>
<td>0.6727</td>
<td>0.6364</td>
<td>0.6636</td>
<td>0.6636</td>
</tr>
<tr>
<td>Precision</td>
<td>0.6578</td>
<td>0.6634</td>
<td>0.6767</td>
<td>0.6465</td>
<td>0.6579</td>
<td>0.6886</td>
<td>0.6886</td>
</tr>
</tbody>
</table>

*Table 5.34: The average values of recall and precision for RF classification for the test data set (Case 1 and Case 2 input data)*

### 5.6.4 Classification results comparison between different algorithms

Figure 5.37 compares the diagonals of the confusion matrices for both input data Cases (Case 1 – multipliers on the eigenfunction only, Case 2 – multipliers and geological parameters) for k-means clustering, SVM and RF (Case 1 – Ntree = 1500, mtry = 6, Case 2 – Ntree = 1300, mtry = 4) classification. From this comparison, we can see how many models (percentage of the total number) were classified correctly for each training image. K-means clustering performed poorly on this data, whilst SVM and RF resulted in higher values of correctly classified models for each training image.

Table 5.35 represents the comparison of the recall and precision values for k-means clustering, SVM and RF. We compare the best classification results (the highest values of recall and precision) achieved with each classifier. From Table 5.35 we can see that both SVM and RF resulted in high values of recall and precision, while k-means clustering performed very poor for this case. SVM with 12 multipliers on the eigenfunction as input
Chapter 5. Metric Space Approach and Machine Learning Classification

performed slightly better than RF with 12 multipliers on the eigenfunction and geological parameters input. However, SVM and RF results are quite close, therefore, we introduce both SVM and RF classification models into history matching to improve geological realism during history.

![Figure 5.37 Comparison of the diagonal of confusion matrices for both input data Cases (Case 1 – multipliers on the eigenfunction only, Case 2 – multipliers and geological parameters) for k-means clustering, SVM and RF (Case 1 – Ntree = 1500, mtry = 6, Case 2 – Ntree = 1300, mtry = 4) classification](image)

<table>
<thead>
<tr>
<th>Parameter combination</th>
<th>k-means clustering</th>
<th>SVM</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>Multipl + geology</td>
<td>0.1</td>
<td>0.6909</td>
</tr>
<tr>
<td>Precision</td>
<td>Multipl only</td>
<td>0.04</td>
<td>0.7258</td>
</tr>
</tbody>
</table>

Table 5.35 The best values of recall and precision for all three classification techniques: k-means clustering, SVM and RF

5.7 Conclusions

Results of multidimensional scaling in static metric space using Euclidian distance showed that the model realisations are widely spread in space, they are quite far away from the training images and also the realisations from different geological scenarios are mixed in 2D projections of high-dimensional space. We explain this with several reasons: 1) the applied hierarchical uncertainties resulted in some facies models from initially different training images becoming more similar to each, 2) the initial set of training images is rather similar and the resulting models are quite alike, 3) SNESIM approach scrambles the initial differences between training images and creates the realisations that
Chapter 5. Metric Space Approach and Machine Learning Classification

are more similar to each other than to their respective training image, or 4) selected metric could not appropriately reproduce the differences between the realisations.

These results in the metric space make the classification very difficult and affect the quality of results: $k$-means clustering performed very poor on this data, SVM and RF resulted only in satisfactory classification.

The idea to introduce geological parameters as additional input parameters for classification showed no improvement of the classification results with SVM and minor improvement in precision with RF because in this example the values of geological parameters applied for geostatistical modelling are similar for all geological scenarios. We believe that in the case of more diverse geological scenarios and geological parameters the effect could be much more significant.
6 CHAPTER SIX Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification

6.1 Introduction

In the previous chapter, we discussed the metric space approach and machine learning classification to account for multiple geological uncertainties and to identify geological relations between different geological scenarios.

In this chapter, we propose to use the results of Support Vector Machines and Random Forest classification (the resulted classifications of the metric space) during history matching (Figure 1.5). Based on the position of the new model created during history matching in the metric space, each model created is assigned to an appropriate geological scenario (based on the predefined classification model results) and linked with appropriate geological parameters. This should lead to an improved geological realism in model realisations during history matching.

We develop the metric space classification history matching workflow on a West Coast of Africa (WCA, section 4.2) example. The reservoir description and uncertainties are described in subsection 4.2.2 and 4.2.3 respectively. Reservoir dynamic model description is introduced in section 4.4. The characteristics of a Truth case model are presented in section 4.5.

6.2 Assisted History Matching (AHM)

6.2.1 AHM set up

In this chapter, as in chapter 4, assisted history matching and uncertainty quantification was performed with the software Raven (Epistemy Ltd.). Multi-objective Particle Swarm Optimisation (MOPSO) (Section 3.3.2) was used for sampling with the same values of parameters as in chapter 4 (Table 4.8).

To perform history matching, the following parametrization was used: uniform distribution of the specified ranges of SNESIM modelling parameters, such as affinity
parameter, global rotation, and facies proportions (Table 6.1). The uniform distribution of multipliers on the eigenfunction that were used to identify the number of the geological scenario (training image) based on the classification are summarised in Table 6.2.

In this chapter we also used three objectives (as in chapter 4) to match each fluid production rate separately: 1) Objective 1 – Field Oil Production Rate (FOPR), 2) Objective 2 – Field Gas Production Rate (FGPR), and 3) Objective 3 – Field Water Production Rate (FWPR).

The automated history matching was performed for 820 days. A least squares objective function (Eq. 6.1 – 6.3) was used to minimise the discrepancy between the simulated and historical production data for three objectives.

\[
M_1 = \sum_{i}^{N} \frac{(FOPR_i - Obs_{FOPR}(i))^2}{2\sigma^2}
\]  
\[
M_2 = \sum_{i}^{N} \frac{(FGPR_i - Obs_{FGPR}(i))^2}{2\sigma^2}
\]  
\[
M_3 = \sum_{i}^{N} \frac{(FWPR_i - Obs_{FWPR}(i))^2}{2\sigma^2}
\]

where \(M_1, M_2, M_3\) are three misfit functions, \(FOPR_i, FGPR_i, FWPR_i\) are the simulated field oil, gas and water production rate at time \(i\), \(Obs_{FOPR}(i), Obs_{FGPR}(i), Obs_{FWPR}(i)\) are the observed oil, gas and water production rate data at time \(i\), \(N\) is the number of data points, and \(\sigma^2\) is the measurement error defined by the Empirical Bayes approach (subsection 4.6.7).

### 6.2.2 Results of history matching

In this section, we show the results of 5 history matching runs with SVM classification (Case 1) and 5 history matching runs with RF classification (Case 2). Both history matching cases are with multi-objective PSO for WCA case study.
Table 6-1 Geological parameter values for parametrisation applied to the West Coast of Africa facies models during history matching

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affinity x</td>
<td>(0.5, 3)</td>
</tr>
<tr>
<td>Affinity y</td>
<td>(0.5, 3)</td>
</tr>
<tr>
<td>Affinity z</td>
<td>(0.5, 9)</td>
</tr>
<tr>
<td>Rotation</td>
<td>(0, 45)</td>
</tr>
<tr>
<td>Facies 1 (channel sand)</td>
<td>(0.25, 0.34)</td>
</tr>
<tr>
<td>Facies 2 (poorer quality sand)</td>
<td>(0.09, 0.14)</td>
</tr>
<tr>
<td>Facies 3 (poorer quality sand)</td>
<td>(0.06, 0.12)</td>
</tr>
<tr>
<td>Facies 4 (shale)</td>
<td>1 – (Facies 1 + Facies 2 + Facies 3)</td>
</tr>
</tbody>
</table>

Table 6-2 Other parameter values for parametrisation applied to the West Coast of Africa facies models during history matching

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiplier on eigenfunction 1</td>
<td>[-214.0, 425.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 2</td>
<td>[-236.0, 243.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 3</td>
<td>[-109.0, 627.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 4</td>
<td>[-90.0, 918.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 5</td>
<td>[-547.0, 642.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 6</td>
<td>[-211.0, 232.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 7</td>
<td>[-351.0, 481.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 8</td>
<td>[-182.0, 602.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 9</td>
<td>[-378.0, 768.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 10</td>
<td>[-257.0, 724.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 11</td>
<td>[-291.0, 526.0]</td>
</tr>
<tr>
<td>Multiplier on eigenfunction 12</td>
<td>[-142.0, 519.0]</td>
</tr>
</tbody>
</table>

Figure 6.1 shows the 2D projection of the Pareto front of multi-objective history matching for 5 runs for SVM (on the left) and for RF (on the right). Pareto models are colour-coded in blue and the rest of the models are green. Blue dotted line represents the Pareto front for a set of Pareto optimal solutions. All Pareto fronts are relatively diverse, which will benefit in improved forecasting (Hutahaean et al., 2016). However, we could observe that in comparison SVM Pareto fronts are more diverse than RF Pareto fronts. However, the
misfit values of the Pareto models are very similar for history matching with SVM and RF.

Figure 6.2 illustrates misfit values obtained during history matching for different training images with RF. From this plot, we could see that only training images TI 1, TI 3, TI 5, TI 6, TI 8 and TI 11 were selected during history matching. The lowest values of misfit were obtained with models from training images TI 6 and TI 8.

Only models from training image TI 2 were selected during history matching with SVM. To explain these results, we refer to the SVM classification results (Figure 6.3). From Figure 6.3 we could clearly see that training image TI 2 dominates in the metric space.

Despite that only one training image TI 2 was selected during history matching with SVM classification, there is still diversity in the resulting models due to the variation of geological modelling parameters: affinity, rotation, and facies proportion (introduced with hierarchical approach).

Figures 6.4 and 6.5 represent misfit values obtained during history matching for each of the parameters with SVM and with RF respectively. The combination of multipliers on the eigenfunction parameters was used to define the position of the model in the metric space. Based on this position in the metric space, the number of the training image was obtained according to the existing classification boundaries. Obtained training images are represented by different colours on the following plots for every geological modelling parameter versus corresponding misfit values (Figure 6.4 and 6.5). When we introduce a stochastic element in the generation of the geological models, the consequence is that the range of parameters where we get a good match is higher. This also explains the of training images.

From the plots, we can observe that the affinity ranges (*aff*, *affy*, *affz*) corresponding to the minimum misfit values are wider and more diverse for history matching with RF classification. The ranges of rotation (Rot), proportion of shale (P1), poorer quality sand 1 (P2), poorer quality sand 2 (P3), and channel sand (P4) parameters are very similar for history matching with SVM and RF.
Figure 6-1 Pareto fronts for two misfit components of a multi-objective history matching with SVM classification for WCA case study: (left) with SVM classification, (right) with RF classification.
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Figure 6-2 Represents the selected training images during history matching with RF and the corresponding value of misfit (<100) for one of five history matching runs. The points are coloured based on iteration number of history matching process.

Figure 6-3 2D representation of SVM classification results for 11 training images: (top left) dimensions 1 and 2, (top right) dimensions 2 and 3, (bottom left) dimensions 3 and 4, (bottom right) dimensions 4 and 5. The plots highlight the predominance of training image TI 2 area in the metric space.
Figure 6-4 Geological modelling parameters versus misfit from history matching with SVM. Models from different training images are represented by a different colour. Affx, affy, affz – affinity x, affinity y, affinity z, Rot - rotation, P1, P2, P3, P4 – proportions of Facies 1, 2, 3, and 4 respectively
Figure 6.5 Geological modelling parameters versus misfit from history matching with RF. Models from different training images are represented by a different colour. Affx, affy, affz – affinity x, affinity y, affinity z, Rot - rotation, P1, P2, P3, P4 – proportions of Facies 1,2,3, and 4 respectively.

We can compare the original range of the parameters with the range which provides the models with the lowest value of misfit (< 10) in Table 6.3. The lowest misfit (< 10) ranges for affinity is slightly lower with SVM than RF. However, these ranges for rotation is slightly lower with RF than SVM. The lowest misfit ranges for proportions is very similar with SVM and RF. The lowest ranges of the parameters are associated with the reduction in uncertainty during history matching.
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior ranges</th>
<th>Ranges with misfit &lt;10 (SVM)</th>
<th>Ranges with misfit &lt;10 (RF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>affinity x</td>
<td>(0.5; 3)</td>
<td>(1.1; 1.9)</td>
<td>(0.5; 2.4)</td>
</tr>
<tr>
<td>affinity y</td>
<td>(0.5; 3)</td>
<td>(1; 1.8)</td>
<td>(0.5; 2.25)</td>
</tr>
<tr>
<td>affinity z</td>
<td>(0.5; 9)</td>
<td>(0.5; 3)</td>
<td>(0.5; 5.9)</td>
</tr>
<tr>
<td>rotation</td>
<td>(0° - 45°)</td>
<td>(5° - 43°)</td>
<td>(5° - 28°)</td>
</tr>
<tr>
<td>Facies 1</td>
<td>(25% – 34%)</td>
<td>(25% – 34%)</td>
<td>(25% – 34%)</td>
</tr>
<tr>
<td>Facies 2</td>
<td>(9% - 14%)</td>
<td>(9.2% - 13%)</td>
<td>(9.8% - 12%)</td>
</tr>
<tr>
<td>Facies 3</td>
<td>(6% - 12%)</td>
<td>(6.5% - 11.7%)</td>
<td>(6.2% - 10.1%)</td>
</tr>
<tr>
<td>Facies 4</td>
<td>(40% – 60%)</td>
<td>(44% – 57%)</td>
<td>(48% – 56%)</td>
</tr>
</tbody>
</table>

Table 6-3 Comparison of the parameters ranges before history matching and after history matching with SVM and RF, that provides the lowest values of misfit (<10)

Figure 6.6 shows the minimum misfit evolution during history matching with SVM (orange) and RF (green). The minimum value of misfit drops from 75 to 11 with SVM and from 38 to 9 with RF after 50 iterations (where iteration means the repetition of the process but with a different combination of parameters). The difference between misfit evolution with SVM and RF is insignificant.

![Figure 6-6 Minimum misfit evolution during history matching with SVM and RF](image)
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Five best models (with the lowest value of misfit) were obtained during history matching for Field Oil, Gas, and Water Production Rate (FOPR, FGPR, FWPR) for both cases: SVM and RF (5 runs of each case are presented in Figures 6.7 – 6.12). The error bars on these Figures represent the measurement error defined by the Empirical Bayes approach (subsection 4.6.7).

All five models from both cases (SVM and RF) match the history very well for Oil, Gas and Water Production Rate. However, the models’ response of Oil and Gas is very similar for all five runs (Figures 6.7, 6.8, 6.10 and 6.11), whilst the models’ response for Water is more different (Figures 6.9 and 6.12).

In this thesis, we used the weighted sum approach (Eq. 4.1) to scalarize three objective functions into one using the same weight factor = 1 for all the objective functions.

![Figure 6-7 Five best models obtained during five history matching runs with SVM for Field Oil Production Rate (FOPR), STB/day](image_url)
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Figure 6-8 Five best models obtained during five history matching runs with SVM for Field Gas Production Rate (FGPR), Mscf/day
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Figure 6-9 Five best models obtained during five history matching runs with SVM for Field Water Production Rate (FWPR), STB/day
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Figure 6-10 Five best models obtained during five history matching runs with RF for Field Oil Production Rate (FOPR), STB/day
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Figure 6-11 Five best models obtained during five history matching runs with RF for Field Gas Production Rate (FGPR), Mscf/day
6.3 Uncertainty Quantification in Reservoir Prediction

History matching results were used to perform reservoir predictions under uncertainty with the NA-Bayes (NAB) algorithm (subsection 3.4.2). NAB algorithm set up parameters are the same as were used in chapter 4 (Table 4.11).

The reservoir forecast was performed for 1216 days. Figures 6.13 and 6.14 show the uncertainty intervals (P10-P50-P90) for the Oil, Gas and Water production rates (on the left) and Oil, Gas and Water total production (on the right) for all five runs with SVM and RF respectively. Figures 6.15 and 6.16 represent the average uncertainty intervals over five runs with SVM and RF.

The Truth case model (section 4.5) was used to validate the prediction results. All intervals (Figures 6.13 - 6.14) mainly cover the Truth case model. However, we can observe that the truth case model is closer to the lowest P90 curves for oil and gas.
production rates and totals. The truth case model is slightly outside the intervals with SVM between 300 – 400 days, 610 – 630 days and 900 – 1100 days for oil and gas rates and between 950 – 1214 days for oil and gas totals (Figure 6.15). The truth case model is closer to P50 curve for water rate and to P10 for water total with SVM (Figure 6.15).

The results with RF are similar to the results with SVM for oil, gas and water rates and also water totals. However, for the oil and gas totals, the results with RF are slightly better than with SVM, as the truth case model is slightly outside the uncertainty intervals only between 1100 – 1214 days (Figure 6.16).

The results show reasonably good history matches and Truth case model is inside the credible interval with both SVM and RF classifications, even though training image TI 9, which was used to create a Truth case model, was not selected at all. TI 6 and TI 8 (mostly selected with RF) are very similar to TI 9: Facies 2 and Facies 3 are represented as lobes for all these training images, and the differences in channel thickness and channel sinuosity are reduced by the selected affinity parameters during history matching.

TI 2 (selected with SVM) is initially different to TI 9: Facies 2 and Facies 3 are represented as channels in TI 2 and as lobes in TI 9. However, after SNESIM simulations we observe that some models from TI 2 look similar to models from TI 9: Facies 2 and Facies 3 don’t have the channel continuity and look more as lobes instead of channels (Figure 4.13).

Figures 6.17 and 6.18 show the comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty (after NAB, average over five runs) of Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production total with SVM and RF respectively. We could see that the posterior forecast uncertainties are much smaller than the prior uncertainties before history matching for both cases.
Figure 6-13 Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) total production prediction uncertainty intervals for five history matching runs with SVM
**Figure 6-14** Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) total production prediction uncertainty intervals for five history matching runs with RF
Figure 6-15 Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production prediction uncertainty intervals average over five history matching runs with SVM
## Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

<table>
<thead>
<tr>
<th>Rate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FOPR</strong></td>
<td><strong>FOPT</strong></td>
</tr>
<tr>
<td><img src="image1" alt="Rate FOPR" /></td>
<td><img src="image2" alt="Total FOPT" /></td>
</tr>
<tr>
<td><strong>FGPR</strong></td>
<td><strong>FGPT</strong></td>
</tr>
<tr>
<td><img src="image3" alt="Rate FGPR" /></td>
<td><img src="image4" alt="Total FGPT" /></td>
</tr>
<tr>
<td><strong>FWPR</strong></td>
<td><strong>FWPT</strong></td>
</tr>
<tr>
<td><img src="image5" alt="Rate FWPR" /></td>
<td><img src="image6" alt="Total FWPT" /></td>
</tr>
</tbody>
</table>

**Figure 6-16** Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) total production prediction uncertainty intervals average over five history matching runs with RF.
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

Figure 6-17 Comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty with SVM (after NAB, average over five runs) of Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production total
Figure 6.18 Comparison of the prior forecast uncertainty (40 non-history matched models) and posterior forecast uncertainty with RF (after NAB, average over five runs) of Oil (FOPR), Gas (FGPR) and Water (FWPR) production rates and Oil (FOPT), Gas (FGPT) and Water (FWPT) production total.

Figure 6.19 shows the Posterior Probability Distribution (PPD) of training images selected for the predictions from history matching results (averaged over five runs) with SVM (left) and RF (right). It was produced by summing up the posterior probability of each selected training image (from 5 runs) and divided by the number of the training images selected. For example, training image TI 5 was selected four times with the corresponding probabilities = 0.02415, 0.0124, 0.0321 and 0.03055; then the average PPD will be (0.02415+0.0124+0.0321+0.03055)/5 = 0.01984. Only training image TI 2 was selected with SVM, therefore, the PPD equals to 1. Training image TI 6 has the
highest Probability with RF, whilst training images TI 2, TI 3, TI 4, TI 7, TI 9 and TI 10 were not selected at all.

![Figure 6-19 The average (over 5 runs) Posterior Probability Distribution (PPD) of training images selected for the predictions from history matching results: (left) with SVM, (right) with RF classification](image)

### 6.4 Geological Realism in Predictions

In this work, we compare geological models selected for reservoir predictions with SVM and RF classification (models with high probability after NAB) to the Truth case geological model by visual examination. Figure 6.20 shows a top view of geological models and the corresponding geological parameters used for the modelling.

The main differences between selected models and the Truth case model are:

- channel sand facies (dark red/brown) is wider for most of the models, except models 4 and 5 with RF;

- some channel sand facies (dark red/brown) are not sinusoidal: models 2 with both SVM and RF;

- the value of global facies rotation parameter is higher than the Truth case rotation value, especially for the models 4 and 5 with SVM and the models 1 and 5 with RF;

- the continuity of the channel sand facies (dark red/brown) that we can observe in the Truth case model is not preserved in the models 2 with SVM and RF;
Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

- Facies 2 and Facies 3 (light blue and yellow) are much wider for most of the models, except models 3 and 5 with SVM.

However, the values of facies proportion are quite similar to the Truth case for both cases. Therefore, we conclude that affinity and rotation parameters have the biggest influence on 3D geological model realisation in this case study.
### Chapter 6. Hierarchical Geological Realism and Classification in History Matching and Uncertainty Quantification – West Coast of Africa Case Study

<table>
<thead>
<tr>
<th></th>
<th>Rotation</th>
<th>Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>rot</strong></td>
<td>10°</td>
<td>shale</td>
<td>53%</td>
</tr>
<tr>
<td><strong>rot</strong></td>
<td>40°</td>
<td>shale</td>
<td>51%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>channel sand</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>affx</strong></td>
<td>1.39</td>
<td>28%</td>
<td></td>
</tr>
<tr>
<td><strong>affy</strong></td>
<td>1.60</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td><strong>affz</strong></td>
<td>0.19</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td><strong>rot</strong></td>
<td>11°</td>
<td>57%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>channel sand</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>affx</strong></td>
<td>2.16</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td><strong>affy</strong></td>
<td>1.70</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td><strong>affz</strong></td>
<td>4.81</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td><strong>rot</strong></td>
<td>15°</td>
<td>47%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>channel sand</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>affx</strong></td>
<td>1.74</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td><strong>affy</strong></td>
<td>1.60</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td><strong>affz</strong></td>
<td>0.13</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td><strong>rot</strong></td>
<td>9°</td>
<td>55%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>channel sand</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>affx</strong></td>
<td>2.08</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td><strong>affy</strong></td>
<td>1.08</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td><strong>affz</strong></td>
<td>2.67</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td><strong>rot</strong></td>
<td>12°</td>
<td>52%</td>
<td></td>
</tr>
</tbody>
</table>

226
6.5 Metric Space and Multidimensional Scaling Comparison

In chapter 5 (Section 5.2), we introduced and presented the results of metric space approach and Multidimensional scaling (MDS) on the prior set of models for WCA reservoir (Figures 5.3 – 5.6).
In this chapter, we add the models after history matching and the models selected by NAB with SVM and RF to the same metric to compare their position in metric space over the prior model position (Figures 6.21 – 6.24).

From 2D MDS projections we observe the different behaviour of the models after history matching with SVM and RF (Figures 6.21 and 6.22). All models after history matching are shown here; they are widely spread with SVM and evenly distributed in the metric space covering the areas without prior models (Figure 6.21). Models with RF are also widely spread; however, they are more densely distributed in the middle of the metric space mainly covering the same area as the prior models (Figure 6.22).

Models selected by NAB with SVM are spread out in the centre of the metric to the right from the Truth case model and the area which is densely populated with the prior models (Figure 6.23). High probability models are situated rather far from the Truth case model. Models selected by NAB with RF are also spread out in the centre of the metric space, however, they are clustered much closer to the Truth case model and the area which is densely populated with the prior model. High probability models are very close to the Truth case model and some medium probability models are surrounding the Truth case model very closely.
Figure 6-21 2D representation (dimension 1 and 2) of MDS for model realisations before and after history matching with SVM (all models).

Figure 6-22 2D representation (dimension 1 and 2) of MDS for model realisations before and after history matching with RF (all models).
Figure 6-23 2D representation (dimension 1 and 2) of MDS for model realisations before history matching and after NAB with SVM

Figure 6-24 2D representation (dimension 1 and 2) of MDS for model realisations before history matching and after NAB with RF
6.6 Discussion

The results from this study show that the proposed hierarchical approach combined with the metric space classification allows us to introduce geological uncertainty from different levels (geological scenario, model parameters) into history matching and uncertainty quantification. We can handle multiple geological scenarios and multiple parameter combinations more efficiently by comparing the static facies models in metric space introduced into history matching. Partitioning the metric space with a classification approach allows us to navigate in the metric space during history matching.

The results of Multidimensional scaling in metric space affect the classification results as we can observe that the models from different training images are mixed in the metric space (2D projections of the high-dimensional space). As was mentioned in chapter 5, there several reasons that could be the possible explanations of these results (section 5.7).

We used three different machine learning classification techniques to perform classification in high-dimensional metric space: \( k \)-means clustering, Support Vector Machines (SVM) and Random Forest (RF). SVM and RF significantly outperformed \( k \)-means clustering, therefore we used only SVM and RF for the history matching in this chapter.

Here we compare the results of history matching and forecasting with classification (both SVM and RF) and without classification (discussed in section 4.6.3) on the following characteristics: Pareto front models’ plots, minimum misfit evolution, reservoir forecast, and geological facies models.

Figure 6.25 shows the comparison of Pareto models plots (top left – SVM classification, top right – RF classification, and at the bottom – no classification). We could observe that history matching with SVM classification and without classification resulted in more diverse Pareto models plots than RF. Also, we could see that the Pareto models’ misfit values with SVM and RF are lower than without classification.

Figure 6.26 represents the comparison of minimum misfit evolutions for all three cases (average over 5 runs for each case). The minimum value of misfit drops from 75 to 11 with SVM, from 38 to 9 with RF, and from 38 to 6 without classification after 50
iterations. The lowest average values of misfit (3.6) are obtained without classification. However, the lowest values of misfit with SVM and RF are 6.8 and 7, which is not significantly higher than without classification.

Figure 6-25 Comparison of Pareto fronts for two misfit components of a multi-objective history matching with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification
Figures 6.27 – 6.30 compare the reservoir production predictions in the following values (for all 3 cases):

- oil production rate;
- water production rate;
- total oil production;
- total water production.

Forecast uncertainty range is wider with RF than with the other approaches for oil production rate and total oil production (Figures 6.27 and 6.29). P10 – P90 range is sufficiently wider with SVM and RF than without classification for water production rate and total water production (Figures 6.28 and 6.30). Truth case model goes along the P50 curve without classification for total oil production, whilst it is slightly outside the P10-P90 range with both classification approaches (Figure 6.29).

The uncertainty range for total water production without classification is very narrow and the Truth case model is outside the uncertainty range, whilst it is inside the range of both classification approaches (Figure 6.30).
The differences in average uncertainty range (P10 – P90) for total oil and water production at the last day of the forecast (1214 days) for all 3 cases are as follows:

- approximately 14mm STB with SVM, 28mm STB with RF, and 16mm STB without classification for oil total production;
- approximately 24mm STB with SVM and RF, and 13 without classification for water total production.

![Comparison of the Oil (FOPR) production rates prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification](image)

*Figure 6-27 Comparison of the Oil (FOPR) production rates prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification*
Figure 6-28 Comparison of the Water (FWPR) production rates prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification

Figure 6-29 Comparison of the Oil (FOPT) total production prediction uncertainty intervals average over five history matching runs with: (top left) SVM classification, (top right) RF classification, and (bottom) no classification
By introducing hierarchical approach and classification into history matching, we aimed to account for the essential geological uncertainties and improve realistic facies model representation whilst history matching. However, we can observe that the parameter ranges that we used for geological modelling allowed a great degree of variability resulting in models that differ from the truth case and partially lose the geological realistic facies representation. Nevertheless, if we compare the high probability models with and without classification, we could observe that the geological model representation with classification is improved: there are no unexpected artificial shapes in the facies models and all models look more realistic after history matching with classification than without classification.
Figure 6-31 Comparison of the Truth case facies model and the high probability facies models, selected with NA-Bayes for reservoir prediction: (top right) without classification, (bottom left) with SVM, and (bottom right) with RF

We evaluated the resulting models mainly by the visual investigation. Whilst it works for facies examination, a deeper investigation would be beneficial, especially if the heterogeneous petrophysical properties distribution would be used.

6.7 Conclusions

In this chapter, we showed the application of the hierarchical approach and metric space classification introduced into history matching and uncertainty quantification.

The results show that we could achieve good history matching results with both classification techniques (SVM and RF). There is no significant difference observed in
forecasting with SVM and RF. However, the Truth case model is better represented inside the P10-P90 interval without classification.

High probability facies models selected with NAB for the predictions look more realistic with classification (capturing the channelized architecture and continuity for sand facies) than without classification.

However, the training image TI 9 which was used to create the Truth case model was never picked during history matching neither with classification nor without it. This could be caused either by the great degree of variability introduced with the hierarchical approach during geostatistical simulations. This variability led to a difficulty to visualise and distinguish between model realisations from different concepts. The choice of the metric space could also have affected the results of the classification because different metric space would produce a different pattern of the model realisations in the metric space.
7 CHAPTER SEVEN Conclusions and Future Work

Suggestions

7.1 Overview
This chapter provides a summary of the main results presented in this work and makes suggestions for the future work. In the thesis, we suggested a methodology to account for geological uncertainties and improve geological realism in facies modelling during history matching and uncertainty quantification:

1) We proposed a hierarchical approach to account for essential geological uncertainties from different levels of geological data (depositional environment, geological scenario, geological parameters) in facies modelling;

2) We suggested that Multidimensional scaling and geological metric space classification could be introduced into history matching to improve geological realism of facies models;

3) We investigated how use of hierarchical approach, Multidimensional scaling and geological metric space classification affected the reliability of reservoir predictions.

The suggested methodology was applied to a synthetic case study based on a real reservoir of the West Coast of Africa.

7.2 Thesis Conclusions
The thesis conclusions are as follows:

1) The visual examination of the facies models with hierarchical approach and MPS SNESIM showed that not all the models reproduced a given training image with the same level of quality: some facies models were able to reproduce the training images better than the others. For example, realisations of training image 3 (TI 3) reproduced channelised architecture for all three sand facies and channels connectivity rather well, whilst realisations of training image 5 (TI 5) do not fully
reproduce channelised architecture and channels connectivity for Facies 2 and Facies 3 (light blue and yellow) (Figure 4.13).

Some of the model realisations resulted in visually unrealistic facies model representation. Some of these geologically unrealistic models had connected facies, which were able to produce a reasonable simulation response. However, other unrealistic models did not preserve the channel facies connectivity, which resulted in a poor simulation response.

2) Multidimensional scaling resulted in models from different training images being scrambled together in the metric space. This result introduced an additional difficulty into metric space classification. This result could be due to several reasons: 1) the applied hierarchical uncertainties resulted in some facies models from initially different training images being more alike, 2) the initial set of training images was rather similar and the resulted in very similar models, 3) SNESIM approach scrambles the initial differences between training images and creates realisations that are more similar to each other than to their respective training image, or 4) the selected metric could not appropriately reproduce the differences between the realisations.

3) Classification results showed that k-means clustering performed poorly on these data. However, both SVM and RF were able to perform a satisfactory classification, resulting in high values (66 – 72%) of both precision (shows how many models were correctly predicted) and recall (how many models were correctly captured).

Geological parameters as additional input data for the classification showed no improvement of the classification results with SVM and minor improvement in precision with RF. These results might be due to similar values of the geological parameters for all training images. In the case of more diverse training images and geological parameters the effect of introducing additional geological parameters as input data could be more significant.
Chapter 7. Conclusions

4) When we history matched without classification, good results were achieved for oil and gas rates but not for the water rate. The resulting credible intervals in forecasting are much narrower compared to the prior uncertainty range. The Truth case model is inside the uncertainty intervals for oil and gas production, whilst it is slightly outside the uncertainty range for water production.

The visual examination of the geological models selected for reservoir predictions (models with high probability after NAB) showed that these models did not preserve the geologically realistic facies architecture and connectivity.

5) Good history matching results were achieved for oil, gas rates and water rate with both SVM and RF classification. The resulting posterior uncertainty ranges are also much narrower than the prior ranges. The uncertainty ranges (P10 – P90) are sufficiently wider with both classification approaches (SVM and RF) than without classification for water production rate and total water production prediction.

The Truth case model is it is slightly outside the uncertainty range for oil and gas production with both classification approaches, whilst it is inside the range for water production.

The visual examination of the geological models selected for reservoir predictions (with NAB) with classification showed an improved geological facies representation than the models selected without classification:

a) The geobodies shape of the selected models is closer to the Truth case with classification than without;
b) The sinuosity of the channel sand facies is represented better in the models with RF classification than with SVM and without classification;
c) The global rotation value is closer to the Truth case in the models with classification.

In summary, we can conclude the proposed approach were able to improve the geologically realistic facies model representation during history matching and uncertainty quantification.
Chapter 7. Conclusions

The further work on the identified issues would allow investigation of the highlighted problems and improve the suggested solutions.

7.3 Future Work Suggestions

The following suggestions can be applied to enhance the achieved results:

- Some additional controls on facies architecture and facies continuity modelling could be introduced to improve the quality of the facies realisations; for example, a set of indicators to measure the static and dynamic connectivity of the models (Rongier et al., 2016, 2017).

- A more diverse set of training images would allow verifying the methodology with more diverse initial settings (different geological parameters). Such a set of training images will involve a separate parametrisation and might significantly affect Multidimensional scaling and classification results.

- The results of the metric space classification can be used to select representative models of each class and perform uncertainty quantification on a smaller number of models.

- A different type of metric space (static or dynamic) and different distance measures (Hausdorff, Manhattan, etc.) could be applied to examine the influence of a specific type of metric space and a specific method on the methodology results.

- An experiment can be made checking that if we choose the input parameters (training image and geological facies modelling parameters) of a model based on its position in the metric space and then perform an MPS simulation, would we get the resulting realisation located at the same place where it was sampled? This experiment can examine the influence the stochasticity of the MPS modelling.

- The heterogeneous petrophysical properties could be simulated to add more geological realism in model realisations.
References


Analytics Vidhya Content Team, 2016.


References


References


References


References


References


References


References


References


James, G., Witten, D., Hastie, T., 2014. An Introduction to Statistical Learning: With Applications in R.

References


References


References


Marathon Center of Excellence for Reservoir Studies, Showcase, 2014.


References


References


References


References


References


260
References


References


References


References


