TÉCNICO LISBOA

Geostatistical History Matching coupled with Adaptive Stochastic Sampling

A zonation-based approach using Direct Sequential Simulation

Eduardo José Airoso Barrela

Thesis to obtain the Master of Science Degree in

Petroleum Engineering

Supervisor: Professor Doctor Leonardo Azevedo Guerra Raposo Pereira

Supervisor: Professor Doctor Vasily Demyanov

Examination Committee

Chairperson: Prof. Doctor Maria João Correia Colunas Pereira

Supervisor: Prof. Doctor Leonardo Azevedo Guerra Raposo Pereira

Members of the Committee: Prof. Doctor Amílcar de Oliveira Soares

October 2016

Acknowledgments

This thesis would not have been possible without the guidance and help of its contributors and their valuable assistance. In this occasion, I would like to express my sincere thanks to them.

First of all, I would like to express my deepest gratitude to my project supervisors, Doctor Leonardo Azevedo and Doctor Vasily Demyanov.

Doctor Leonardo Azevedo for his patience, guidance and absolutely unfailing support as my supervisor and mentor. This was a unique opportunity to acquire knowledge and experience and I feel fortunate for the opportunity of having him helping me along the way.

Doctor Vasily Demyanov for the supreme guidance and insights shared during the course of the project, for allowing me the chance to do my research at the Uncertainty Quantification Group at the Heriot-Watt University and for his concern and consideration for my academic needs over the period of stay. My period of stay at Heriot-Watt University was a truly unique, fantastic and memorable experience.

A word of appreciation for Doctor Amílcar Soares and Doctor Maria João Pereira, for their teachings, support and kindness and for providing me with all the necessary working conditions.

I would also like to thank Doctor Dan Arnold and Doctor Mike Christie for rendering their help, input on the project presentations and for all the joyful talks during my period of stay at Heriot-Watt. A word of thanks also goes to Jack "the magician" Talbot, whose magic tricks entertained us during the (few) periods of idleness and for his crucial assistance with the excellent Heriot-Watt in-house software, Raven.

I would like to thank the Erasmus+ Programme for partially funding my Erasmus internship period.

I would also like to thank Rock Flow Dynamics (tNavigator) for providing the necessary licenses for their incredible fluid flow simulator. Being introduced to their fluid flow simulator was another one of the great surprises I got to experience. I am a fan. Acknowledgements also goes to Schlumberger (Petrel) and MathWorks (Matlab) for providing the necessary licenses for their software.

A word of thanks also goes to my colleague Gonçalo Simões, for taking part on my international research experience. Without him as a flat mate in Edinburgh, everything would have been much more difficult. An extension of this gratitude goes towards all my course colleagues that made these past years so rewarding and fun and also for all the fantastic colleagues that I was fortunate to meet in Scotland.

Last but not least, I wish to express my gratitude to my parents, grandmother, brother and friends for their encouragement, love and support. Finally a very special mention for my girlfriend Sílvia and her parents. They were all an incredibly big part of this endeavor.

Abstract

Advances in computing technology, over the past decades, allowed the development of a number of different history matching techniques. Nevertheless, the simultaneous integration of production data under geological consistency, as part of the reservoir modeling workflow, still remains a challenge. A geologically consistent approach aims to avoid solutions that are unrealistic under the reservoir's general geological characteristics. Unrealistic history matching solutions would result in poor reservoir response forecasting. It is also essential to include only geologically realistic models for uncertainty assessment based on multiple models that are consistent with the geological data and also in order to match observed production history. Geostatistical history matching can iteratively update static reservoir model properties through conditional assimilation constrained to the production data, using geologically consistent perturbation. Multiple stochastic realizations are assimilated following a zonation approach to account for the local match quality, thus providing a way to integrate regionalized discretization of parameters with production data and engineering knowledge. The present project proposes a new history matching technique applied in uncertain reservoir conditions represented by geologically consistent reservoir zonation, based on fault presence and fluid production streamlines. The work explores the value of using a geologically consistent zonation associated with production wells in geostatistical history matching regions, coupled with adaptive stochastic sampling and Bayesian inference for uncertainty quantification and optimization of geological and engineering properties. This novel approach makes use of the *Direct Sequential Simulation* algorithm for generation of stochastic realizations and *Particle Swarm Optimization* for parameter optimization. The approach is tested in a semi–synthetic case study based on a braided–river depositional environment.

Keywords: History Matching, Geostatistics, Direct Sequential Simulation, Uncertainty Quantification, Particle Swarm Optimization, Adaptive Stochastic Sampling.

Resumo

Com os avanços na tecnologia de computação, alcançados nas últimas décadas, um número considerável de diferentes técnicas de ajuste de histórico, têm vindo a ser desenvolvidos. Apesar de tudo, a integração simultânea de dados de produção sob uma premissa de consistência geológica, como parte do fluxo de trabalho da modelação de reservatórios, continua a ser um desafio. Uma abordagem geologicamente consistente visa evitar soluções que não são realistas sob as características geológicas gerais do reservatório. Soluções de ajuste de histórico irrealistas resultam numa má previsão da resposta do reservatório. A inclusão de modelos geologicamente realistas na avaliação da incerteza, tendo como base vários modelos que são consistentes com os dados geológicos, é um aspeto essencial. O mesmo sucedendo no respeitante ao ajuste de histórico de produção observada. O ajuste de histórico geoestatístico permite atualizar iterativamente as propriedades estáticas do modelo de um reservatório, através de assimilação condicional restringida aos dados de produção, usando para isso, uma perturbação geologicamente consistente. Realizações estocásticas múltiplas são assimiladas seguindo uma abordagem de regionalização para garantir a qualidade da correspondência local da produção, permitindo a sua integração com conhecimentos de engenharia e geologia. O presente projeto propõe uma nova técnica de ajuste de histórico, aplicada em condições de incerteza nos parâmetros do reservatório, que é discretizado por uma metodologia de regionalização geologicamente consistente, com base na presença de falhas e linhas de fluxo da produção de fluidos. O trabalho explora o valor do uso de uma metodologia de regionalização geologicamente consistente, associada à produção de poços em regiões que serão ajustadas através de um algoritmo de ajuste de histórico geoestatístico. Este método será utilizado conjuntamente com um algoritmo de amostragem estocástica adaptativa e inferência Bayesiana, por forma a quantificar a incerteza e otimizar as propriedades geológicas e de engenharia do reservatório. Esta nova abordagem faz uso do algoritmo de *Simulação Sequencial Direta* para a geração de realizações estocásticas e do algoritmo de *Otimização por Enxame de Partículas*, para a otimização de parâmetros. A abordagem é testada num caso de estudo semissintético baseado num ambiente de deposição fluvial–entrelaçado.

Palavras–Chave: Ajuste de Histórico, Geoestatística, Simulação Sequencial Direta, Quantificação da Incerteza, Otimização por Enxame de Partículas, Amostragem Estocástica Adaptativa.

Table of Contents

Acknowledgments ..... **i**

Abstract ..... **iii**

Resumo ..... **v**

Table of Contents ..... **vii**

List of Figures ..... **ix**

List of Tables ..... **xi**

Acronyms ..... **xiii**

1 Introduction ..... **- 1 -**

1.1 Thesis outline ..... **- 2 -**

2 Theoretical Background ..... **- 3 -**

2.1 Geostatistical model ..... **- 3 -**

2.1.1 Random variables ..... - 3 -

2.1.2 Stationarity ..... - 4 -

2.1.3 Semivariogram modeling ..... - 6 -

2.1.4 Modeling spatial anisotropy ..... - 7 -

2.1.5 Stochastic sequential simulation ..... - 8 -

2.2 Basics on reservoir fluid flow simulation ..... **- 12 -**

2.2.1 Fluid flow concepts ..... - 12 -

2.3 History matching ..... **- 14 -**

2.3.1 Geostatistical history matching ..... - 17 -

2.4 Uncertainty Quantification ..... **- 19 -**

2.4.1 Bayesian theory ..... - 20 -

2.4.2 Adaptive Stochastic Sampling algorithms ..... - 21 -

3 Methodology Workflow ..... **- 23 -**

3.1 Dynamic evaluation ..... **- 25 -**

3.1.1 Objective function ..... - 25 -

3.1.2 Correlation coefficient ..... - 25 -

3.2 Zonation–based Geostatistical history matching ..... **- 27 -**

3.3 Coupling with Adaptive Stochastic Sampling ..... **- 33 -**

3.3.1 Parameter uncertainty ..... - 34 -

4 Case Study – "The Watt Field" ..... - 39 -

4.1 Watt field description ..... - 40 -

4.2 Base scenario selection ..... - 45 -

5 Results and Discussion ..... - 49 -

5.1 Geostatistical history matching ..... - 49 -

5.2 Coupling with Adaptive Stochastic Sampling ..... - 61 -

5.2.1 Production match analysis ..... - 61 -

5.2.2 Parameter perturbation analysis ..... - 68 -

6 Conclusions and Future Work ..... - 77 -

References ..... - 79 -

Appendix A – Best 5 Realizations (GHM coupling) ..... - A1 -

List of Figures

Figure 1 – Distributions of a collection of field data (Arnold & Demyanov, 2015). ..... - 4 -
Figure 2 –Coordinate transformation, adapted from Yaseen *et al.* (2013). ..... - 8 -
Figure 3 – Illustration of the workflow in sequential simulation, adapted from Correia (2013). ..... - 9 -
Figure 4 – Generalized workflow for reservoir modeling with history matching and forecasting (Streamsim, 2016). ..... - 15 -
Figure 5 – Generalized workflow for GHM adapted from (Unisim, 2016) ..... - 19 -
Figure 6 – GHM workflow (First stage). ..... - 24 -
Figure 7 – Coupling with Adaptive Stochastic Sampling algorithm. (Second stage) GHM is shown in detail in Figure 6 ..... - 25 -
Figure 8 – Correlation coefficient calculation on four possible cases of simulated vs observed data. - 27 -
Figure 9 – Global zonation pattern. ..... - 28 -
Figure 10 – Circular zonation pattern. ..... - 29 -
Figure 11 – Square zonation pattern. ..... - 29 -
Figure 12 – Voronoi zonation pattern. ..... - 30 -
Figure 13 – 2D top view of production streamlines from different wells. ..... - 31 -
Figure 14 – 3D view of production streamlines from different wells. ..... - 31 -
Figure 15 – Preferential flow path for water production, for well group 1 (top left), well group 9 (top right), well group 8 (bottom left) and well group 10 (bottom right). ..... - 32 -
Figure 16 – Fault and Streamline zonation–based pattern. ..... - 33 -
Figure 17 – 3D view of the Fault and Streamline zonation–based pattern. ..... - 33 -
Figure 18 – Illustration of the uncertainty in Porosity ranges. ..... - 35 -
Figure 19 – Illustration of the uncertainty in Permeability ranges. ..... - 35 -
Figure 20 – Illustration of the uncertainty in the Porosity histogram. ..... - 36 -
Figure 21 – Illustration of the uncertainty in the Permeability histogram. ..... - 37 -
Figure 22 –Illustration of the uncertainty in fault transmissibilities. ..... - 38 -
Figure 23 – Possible interpretational choices for the *Watt Field* case study (Arnold D. , 2012). ..... - 40 -
Figure 24 – Top-view of the OWC level and Top Structures for the *Watt Field* (Arnold, 2012). ..... - 41 -
Figure 25 – Facies interpretations for the *Watt Field* (Arnold D. , 2012). ..... - 41 -
Figure 26 – Well correlation panel of Wells A–F, showing the top structure well pics for TS–1, TS–2 and TS–3 (Arnold D. , 2012). ..... - 42 -
Figure 27 – Fault network definition for the Watt Field, adapted from Arnold (2012). ..... - 43 -
Figure 28 – Production history for FOPR, FWPR, FGPR and WBHP Well1 ..... - 45 -
Figure 29 – Result comparison of Relative Misfit vs Well group, over selected scenarios. ..... - 46 -
Figure 30 – Porosity field for the selected base scenario (left), Histogram for porosity (right). ..... - 47 -
Figure 31 – Permeability field for the selected base scenario (left), Histogram for permeability of log10(k) (right). ..... - 47 -
Figure 32 – Saturation regions for the selected base scenario. ..... - 48 -
Figure 33 – Best–fit Porosity for Global zonation method. ..... - 49 -

Figure 34 – Best–fit Permeability for Global zonation method. - 50 -
Figure 35 – Global Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right). - 50 -
Figure 36 – Best–fit Porosity for Square zonation method. - 51 -
Figure 37 – Best–fit Permeability for Square zonation method. - 51 -
Figure 38 – Square Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right). - 52 -
Figure 39 – Best–fit Porosity for Circular zonation method. - 53 -
Figure 40 – Best–fit Permeability for Circular zonation method. - 53 -
Figure 41 – Circular Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right). - 54 -
Figure 42 – Best–fit Porosity for Voronoi zonation method. - 55 -
Figure 43 – Best–fit Permeability for Voronoi zonation method. - 55 -
Figure 44 – Voronoi Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right). - 56 -
Figure 45 – Best–fit Porosity for Fault and Streamline zonation method. - 57 -
Figure 46 –Best–fit Permeability for Fault and Streamline zonation method. - 57 -
Figure 47 – Fault and Streamline Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right). - 58 -
Figure 48 – Comparison between different zonation methods – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right). - 59 -
Figure 49 – Misfit evolution comparison between all zonation methods. - 60 -
Figure 50 – Comparison of Best average Porosity between fault and streamline and circular zonation. (left) Total misfit values per well group and field (right). - 61 -
Figure 51 – Misfit Evolution of the zonation–based GHM algorithm, coupled with Adaptive Stochastic Sampling. - 63 -
Figure 52 – Fluid flow response for FOPR (left), FWPR (middle) and FGPR (right) for GHM coupled with Adaptive Stochastic Sampling. - 63 -
Figure 53 – Misfit evolution over different zones (x axis – Iteration, y-axis – Misfit). - 64 -
Figure 54 – Oil and Water production for WELL5 (top left, top right) and WELL7 (bottom left, bottom right) for the 5 best iterations. - 66 -
Figure 55 – Comparison of match quality by well between Base Scenario (Orange) and GHM coupling (Green), (y axis – Sum of misfits for WOPR, WWPR AND WBHP). - 67 -
Figure 56 – Fault transmissibility parameter value (y axis) vs Iteration (x axis). - 68 -
Figure 57 – Fault transmissibility parameter value (y axis) vs Misfit (x axis). - 69 -

Figure 58 – Permeability range vs Iteration (left – Horizontal range, right – Vertical range). .....- 70 -
Figure 59 – Permeability range vs Misfit (left – Horizontal range, right – Vertical range). .....- 70 -
Figure 60 – Top view of the best-fit Permeability realization (Iteration 91, Simulation 5). .....- 71 -
Figure 61 – Permeability Histogram perturbation (y axis – parameter value) vs Iteration (x axis), (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean). .....- 71 -
Figure 62 – Permeability Histogram perturbation (y axis – parameter value) vs Misfit (x axis), (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean). .....- 72 -
Figure 63 – Comparison between Permeability histogram for Base Scenario and Best Iteration. ....- 72 -
Figure 64 – Porosity range vs Iteration (left – Horizontal range, right – Vertical range). .....- 73 -
Figure 65 – Porosity range vs Misfit (left – Horizontal range, right – Vertical range). .....- 73 -
Figure 66 – Top view of the best-fit Porosity realization (Iteration 91, Simulation 5). .....- 73 -
Figure 67 – Porosity Histogram perturbation (y axis – parameter value) vs Iteration (x axis), (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean). .....- 74 -
Figure 68 – Porosity Histogram perturbation (y axis – parameter value) vs Misfit (x axis) (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean). .....- 74 -
Figure 69 – Comparison between Porosity histogram for Base Scenario and Best Iteration. .....- 75 -

List of Tables

Table 1 – Geological and engineering parameters commonly perturbed in history matching. .....- 16 -
Table 2 – Uncertainty Quantification parameters for Porosity variogram ranges. .....- 35 -
Table 3 – Uncertainty Quantification parameters for Permeability variogram ranges. .....- 35 -
Table 4 – Uncertainty Quantification parameters for Porosity histogram perturbation. .....- 36 -
Table 5 – Uncertainty Quantification parameters for Permeability histogram perturbation. .....- 37 -
Table 6 – Uncertainty Quantification parameters for Fault transmissibility perturbation. .....- 38 -
Table 7 – Operation parameters for all producer wells of the Watt Field. .....- 44 -
Table 8 – Operation parameters for all injector wells of the Watt Field. .....- 44 -
Table 9 – Model statistics for porosity. .....- 47 -
Table 10 – Model statistics for permeability. .....- 48 -
Table 11 – Summary table of comparison between zonations. .....- 59 -
Table 12 – Selection of prior distribution ranges. .....- 62 -

Acronyms

C – Covariance

cdf – Cumulative Distribution Function

Co–DSS – Direct Sequential Co–Simulation

coSGS – Sequential Gaussian Co–Simulation

coSIS – Sequential Indicator Co–Simulation

DSS – Direct Sequential Simulation

FGPR – Field Gas Production Rate

FOPR – Field Oil Production Rate

FWPR – Field Water Production Rate

GHM – Geostatistical History Matching

k – Permeability

L – Likelihood

M – Misfit Value

MCMC – Markov Chain Monte Carlo

PSO – Particle Swarm Optimization

WBHP – Well Bottomhole Pressure

WOPR – Well Oil Production Ratio

WWPR – Well Water Production Ratio

γ – Variogram

θ – Direction of anisotropy

ν – Anisotropy Ratio

σ – Sigma

φ – Porosity

1 Introduction

Reservoir modelling is a crucial step in the development and management of hydrocarbon reservoirs. Field development decisions such as location and number of production wells, production strategy and reservoir pressure maintenance depend on an accurate representation of the subsurface. An accurate reservoir model is one that honours all available data at the scale and precision at which they are available. Given the sparsity of information related with the subsurface reservoir properties and the economic gains for that information, the reservoir modeller should extract the maximum possible reservoir specific information from all available data while including the uncertainty associated with measurement errors and the lack of knowledge about the system he is trying to model.

The information available to model a given hydrocarbon reservoir is continually upgraded over the course of field development. Rock core measurements, well log data, among others, allow the estimation of the petrophysical properties in the near well-bore region as well as an indication of reservoir productivity. On the other hand, reservoir production data can be used to cross-check the quality of the models that honour the existing available data.

However, the relationship between hard data and production data in highly non-linear. Integration of both types of data into reservoir modelling is a challenging task. The process of history matching tries to address this problem by applying changes to the reservoir model parameters in order to minimize a given cost function, responsible for the quantification of the mismatch between observed production data (historical data) and the dynamic model response (simulated production data). This, however is not a single solution problem. In fact, multiple reservoir models (static or dynamic) can produce equally matched responses.

Over recent years, several approaches have been developed to address history matching of hydrocarbon reservoirs. Methods depending on data assimilation, like the Ensemble Kalman Filter (Evensen, *et al.* 2007), gradual deformation (Hu, *et al.* 2001) and probability perturbation (Caers & Hoffman, 2006) or stochastic sequential simulation and co-simulation (Mata-Lima, 2008), (Le Ravalec-Dupin & Da Veiga, 2011) have been proposed.

Other methods like Stochastic Optimisation algorithms, allow reducing computational times by sampling from the ensemble of best matched models, in order to improve matches. Algorithms such as the Neighbourhood Algorithm (Sambridge, 1999), Genetic Algorithm (Erbas & Christie, 2007), Particle Swarm Optimisation (Mohamed, 2011) and Differential Evolution (Hajidazeh, *et al.* 2009) have been successfully applied to the history matching problematic, on recent years.

The present project proposes a new history matching technique, coupling conventional Geostatistical History Matching (GHM) methodologies with the integration of Adaptive Stochastic Sampling. The proposed methodology applies a GHM technique under uncertain reservoir conditions, represented by geologically consistent reservoir zonation, based on fault presence and fluid production streamlines. The proposed methodology explores the value of using a geologically consistent zonation methodology,

associated with production wells in GHM regions, coupled with Adaptive Stochastic Sampling and Bayesian inference for uncertainty quantification and optimization of geological and engineering properties.

This novel approach makes use of the *Direct Sequential Simulation* (DSS) algorithm for generation of stochastic realizations and *Particle Swarm Optimization* (PSO) for parameter optimization. The proposed history matching approach is tested in a semi–synthetic case study based on a braided–river depositional environment.

This thesis was developed as part of the Master of Science Degree in Petroleum Engineering, promoted by the Instituto Superior Técnico, Lisbon, in a partnership between CERENA (Centro de Recursos Naturais e Ambiente), from Instituto Superior Técnico and the Uncertainty Quantification Group from the Institute of Petroleum Engineering of Heriot-Watt University, Edinburgh.

1.1 Thesis outline

The structure of this manuscript is composed by six chapters. A brief description of each chapter is presented as next:

- Chapter 1 – The first chapter introduces the topic for the work, stating its challenges and a brief description of the methodology adopted, as well as the importance of the study;
- Chapter 2 – The second chapter introduces the theoretical concepts related with the development of this work. Details about the concepts and algorithms being used for the development of this work are explained;
- Chapter 3 –In the third Chapter, the methodology used for the development of this work is explained, including the explanation of the GHM and coupling with Adaptive Stochastic Sampling algorithms being developed;
- Chapter 4 –The fourth Chapter presents the selection of the case study that was used for implementation of the algorithm, and the criteria of importance to this work that based its selection;
- Chapter 5 – The fifth Chapter is the presentation of the results obtained from the application of the developed methodology. Results for traditional GHM and GHM coupled with Adaptive Stochastic Sampling are displayed and discussed;
- Chapter 6– The sixth chapter states the conclusions to the application of the proposed methodology.

2 Theoretical Background

2.1 Geostatistical model

Geostatistics deals with modeling the spatial/temporal distribution of a given property of interest and has found a wide range of applications, from mining and oil exploration to agricultural purposes as well as image analysis. The concepts in geostatistics can be applied to a variety of problems that exhibit spatial dependence.

Geostatistics was primarily developed for application in the mining industry, during the 50s, when the mining engineer D.G. Krige (1951) and statistical H.S. Sichel (1952) developed new estimation methods for direct application to mineral reserves. Between 1957 and 1963, the French engineer G. Matheron, based on the observations of D.G. Krige, developed the theory of regionalized variables (Matheron, Principles of geostatistics, 1963). Since 1968, geostatistics has been used for hydrocarbon reserve estimation. Between 1968 and 1970, the theory of Universal Kriging (name introduced by Matheron in honor of Krige) was developed for application to underwater mapping. In 1973, Matheron created the intrinsic theory of Order K, for application to meteorology problems (Matheron, 1973). Between 1972 and 1973, the principle of Convex Analysis emerged, aimed at maximizing the recoverable reserves of underground reservoirs. In 1974 the theory of recovery functions and, based on it, the concepts of applied nonlinear geostatistics (Agterberg, 1974), were developed for application on the selection of recoverable reserves.

The purpose of geostatistics, applied to the oil industry, is to improve prediction of the spatial distribution of the petrophysical properties of interest, by building more realistic and heterogeneous models of a given reservoir.

Geostatistical stochastic simulation procedures are often aimed at the generation of multiple plausible results. Contrary to a deterministic approach, where you have a unique model of the reality, multiple and equiprobable stochastic or geostatistical realizations, allow a quantification of the spatial uncertainty associated to the models. The advantage of using many equiprobable geostatistical realizations for modeling a hydrocarbon reservoir, requires multiple reservoir flow simulations and consequent time and computational costs, but the advantages of having such a tool outweigh its inherent disadvantages.

2.1.1 Random variables

In classical statistics, a random variable is considered to be a variable that can be sampled repeatedly (at least, in a theoretical point of view) and that each sample is independent from the previous (Soares, 2006).

By taking into account the example of a core extracted from a rock, a core sampled at location x is not random. It is in fact unique and physically determined, although with a measurement error associated. This way, there is no possibility of getting a repeated core sample from exactly the same location, which is something that violates the repeatability condition previously assumed. Nevertheless, a new sample can be taken within a proximity of location x, which in other words means that the condition of repeatability can be approximately met. However, the sample independence condition is violated, as the two cores will not be independent from each other, since the second core sample is likely to be similar to the first sample. This is what is behind the concept of spatial continuity, meaning that independence between samples cannot be assumed and that in fact, samples are spatially correlated among each other and that this correlation depends directly on the distance separating the samples.

Figure 1 depicts the end result of the collection of data of a given property (i.e. porosity or permeability), where the concept of repeatability and non–dependency of the samples is demonstrated. Spatial continuity between the different regions of the plane is shown by the similitude between the distributions of neighboring cells, as opposed to the dissimilitude of the distributions between far away cells.

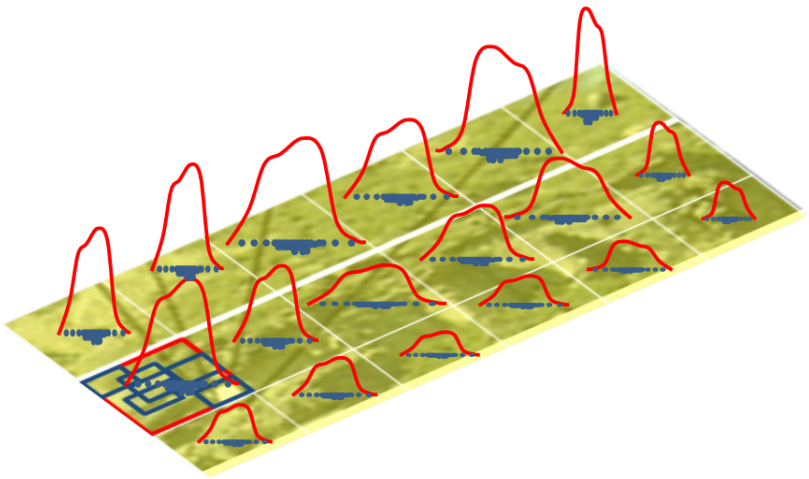


Figure 1 – Distributions of a collection of field data (Arnold & Demyanov, 2015).

2.1.2 Stationarity

As previously seen, sampling data is considered as been originated by a random process (hence the name random variable), this however, does not equal to say that the phenomena that generated the sample data, also originates from a random process. In fact, random variables are what compose the random function and from this random function, only the sample data is known.

Geostatistics assumes several degrees of stationarity of the random function. According to the assumption that the correlation between any two random variables is dependent on their distance (h), when applied to the covariance of a random function, stationarity is stipulated to be independent from:

- Distribution of the variables (meaning strict stationary);
- Mean m(x), which is assumed to be constant;
- Variance σ^2(x), also assumed to be constant;
- Covariance, C(x, h).

As such, by considering a property z which is assumed to be a spatially distributed variable (for instance porosity or permeability), z(x) can then be treated as a realization of the stochastic variable Z(x), with x indicating that z is a regionalized variable. This way, the expected value of Z(x) is given as

$$E\{Z(x)\} \;=\; m(x) \tag{1}$$

and the covariance is given by

$$C(x,h) \;=\; E\{[Z(x)-m(x)][Z(x\;+m)-m]\} \tag{2}$$

If the mean of the property in question is independent of position, that is m(x) = m, Z is denoted as first order stationary. If also the covariance is independent of the position, then Z is denoted as second order stationary. Within these assumptions, the variogram can be defined as:

$$2\gamma(h) \;=\; E\{[Z(x)-Z(x\;+h)]^2\} \tag{3}$$

The intrinsic hypothesis of geostatistics states the independency of the variogram in terms of the position, only depending on the lag, i.e. the distance h, between the two points (Goovaerts, 1997). Mathematically stated the intrinsic hypothesis says:

$$\gamma(x,h) \;=\; \gamma(h) \tag{4}$$

If the stochastic variable Z is second order stationary, then the intrinsic hypothesis is valid. Under 2$^{\text{nd}}$ order stationarity the semivariogram is given as:

$$\gamma(h) \;=\; C(0)-C(h) \tag{5}$$

The validity of the previous equation can be shown by the following proof:

The variance is defined as

$$\sigma^2 \;=\; E\{[Z(h)-m]^2\} \;=\; E\{Z^2(h)\}-m^2 \tag{6}$$

and the covariance is defined as

$$\begin{aligned} C(h) \;&=\; E\{[Z(x)-m][Z(x\;+h)-m]\} \\ &=\; E\{Z(x)\;\cdot\;Z(x\;+h)\}-mE\{Z(x)\}-mE\{Z(x\;+h)\}+m^2 \\ &=\; E\{Z(x)\;\cdot\;Z(x\;+h)\}-m^2 \end{aligned} \tag{7}$$

The variogram is defined as

$$\begin{aligned} 2\gamma(h) &= E\{[Z(x) - Z(x\ +h)]^2\} \\ &= E\{Z^2(x)\} + E\{Z^2(x\ +h)\} - 2E\{Z(x)\ \cdot\ Z(x\ +h)\} \end{aligned} \tag{8}$$

By combining the previous equations, the expression becomes

$$\begin{aligned} 2\gamma(h) &= 2(\sigma^2\ +m^2) - 2(C(h) + m^2) \\ \gamma(h) &= \sigma^2\ - C(h) \\ &= C(0) - C(h) \end{aligned} \tag{9}$$

From a number of observations the experimental semivariogram is easily calculated by

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{k=1}^{N(h)} \big(Z(x_i) - Z(x_i\ +h)\big)^2 \tag{10}$$

where N(h) is the number of pairs of observations with a lag distance of h.

2.1.3 Semivariogram modeling

A number of empirical semivariogram models can be used to describe the spatial continuity pattern of z within a given area, e.g. the area comprising an oil reservoir. The following equations show the most common variogram models used for semivariogram modeling. Note that the variogram model selected to model the experimental variogram should be positive.

Spherical model

$$\gamma^*(h) = \begin{cases} 0 & h\ =\ 0 \\ C_0\ + C_1\ \left(\frac{2}{3}\frac{h}{R} - \frac{1}{2}\frac{h^3}{R^3}\right) & 0\ <\ h\ <\ R \\ C_0\ + C_1 & h\ \geq\ R \end{cases} \tag{11}$$

Exponential model

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0\ + C_1\left(1 - e^{\left(-\frac{3h}{R}\right)}\right) & h > 0 \end{cases} \tag{12}$$

Gaussian model

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0\ + C_1\left(1 - e^{\left(-\frac{3h^2}{R^2}\right)}\right) & h > 0 \end{cases} \tag{13}$$

The parameter C_0 is denoted as the *nugget effect* and acts as a measure of measurement error and discontinuity in the vicinity of h $=$ 0. The quantity $C_0\ + C_1$ is designated by *sill* which is the value of the semivariogram as the lag tends to infinity, i.e. the total variance of the existing experimental samples. R is the *range of influence* and measures the extent to which a measurement has an impact on a new estimation, i.e. the spatial correlation of a given sample in space.

2.1.4 Modeling spatial anisotropy

The previous sections introduced the concepts related with a statistically homogeneous two–dimensional field. Such a field is called *isotropic,* indicating that the variability of the property in question (e.g. permeability) does not depend on the direction of the separation vector h – only the distance (or length of the vector) is responsible for any variability. In many applications such an assumption is invalid. In reservoir modeling and characterization a model may be systematically build up by areas with unique properties according to some underlying geologic process.

Hohn (1999) defines *geometric anisotropy* as the situation where the *sill* is independent of the direction of h and only the *range of influence* depends on the direction.

Consider the two spherical semivariograms which account for horizontal variability and vertical variability, respectively:

$$\gamma_1(h) = \frac{3h}{2R_1} - \frac{h^3}{2R_1^3} \tag{14}$$

$$\gamma_2(h) = \frac{3h}{2R_2} - \frac{h^3}{2R_2^3} \tag{15}$$

The *anisotropy* ratio (ν) is defined as

$$\nu = \frac{R_1}{R_2} \tag{16}$$

Inserting νh in the first semivariogram gives

$$\gamma_1(\nu h) = \frac{3R_1 h}{2R_2 R_1} - \frac{R_1^3 h^3}{2R_2^3 R_1^3} = \frac{3h}{2R_2} - \frac{h^3}{2R_2^3} = \gamma_2(h) \tag{17}$$

Thus, the semivariogram in the vertical direction can be calculated by the semivariogram for the horizontal direction if the separation distance is multiplied with the *anisotropy* ratio. Such an approach, however, only enables the user to determine the semivariogram in two directions – the horizontal and the vertical directions. If the direction lies between these, a change of coordinate system must be applied. Let a linear coordinate transformation be given as

$$h'_h = a_{11} h_h + a_{12} h_v \tag{18}$$

$$h'_v = a_{21} h_h + a_{22} h_v \tag{19}$$

where $\mathrm{h} = [\mathrm{h_h}\ \mathrm{h_v}]^{\mathrm{T}}$ and $\mathrm{h}' = [\mathrm{h'_h}\ \mathrm{h'_v}]^{\mathrm{T}}$.

Figure 2 shows an ellipsoid with axes corresponding to the directions of influences associated with the *ranges of influence* $\mathrm{R_1}$ and $\mathrm{R_2}$.

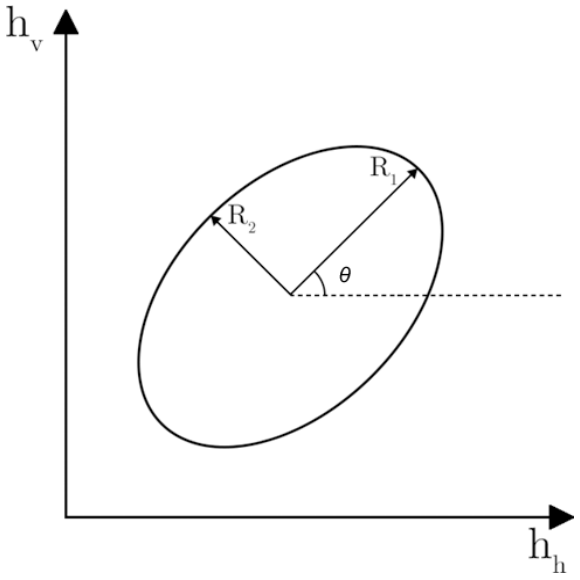


Figure 2 –Coordinate transformation, adapted from Yaseen *et al.* (2013).

In the following discussion θ will denote the angle between the principal axis of the ellipsoid and the horizontal axis as depicted in Figure 2.

Hohn (1999) gives the transformation

$$h' = \begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix} \begin{bmatrix} cos\,\theta & sin\,\theta \\ -sin\,\theta & cos\,\theta \end{bmatrix} h = SRh \tag{20}$$

This transformation aligns the kriging coordinate system with the axes of influence.

The semivariogram value can now be determined by the use of the transformed coordinates:

$$\gamma(h) = \frac{3\|h'\|}{2R_1} - \frac{\|h'\|^3}{2R_1^3} \tag{21}$$

where

$$\|h'\| = \sqrt{h_h'^2 + h_v'^2} \tag{22}$$

2.1.5 Stochastic sequential simulation

In Geostatistics, kriging methods can be generalized to be part of the so–called estimation methods. These methods are characteristic in the sense that the estimated field is unique, and deterministic because the estimation variance is minimized. This means that whenever a number of sample values are present and the variability (the semi–variogram) is determined, only *one* model will fulfill the

requirement of minimum estimation variance. The so–called stochastic sequential simulation techniques are different with respect to this property.

Such *simulation techniques* require a two stage process. First, there is a need to estimate the expected value at a given location by kriging along with the corresponding kriging variance. These estimates are carried out on the available sample data along with any other previously simulated values. From these estimates a local distribution function is created and a value is simulated by using *Monte Carlo* simulation. The simulation techniques visit each unknown point sequentially following a random path that visits all the grid cells. The way the local probability function is modelled varies from method to method. Figure 3 illustrates the workflow in sequential simulation.

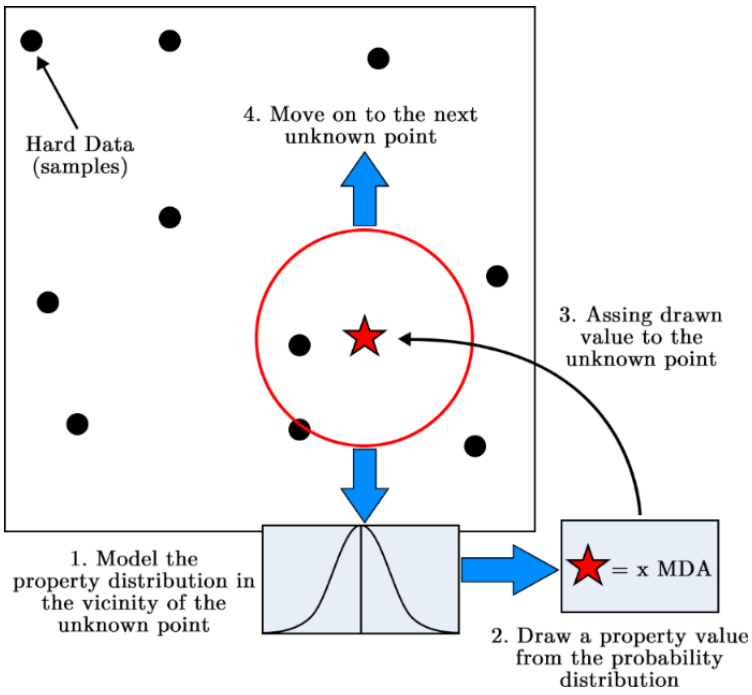


Figure 3 – Illustration of the workflow in sequential simulation, adapted from Correia (2013).

Stochastic simulation models are intended to reproduce equiprobable images of the reality, considering what is known of the reality, meaning, the variability of the available samples, including the variable probability distribution and its spatial continuity, revealed by the variogram or covariance study of the sample (Caragea & Smith, 2006).

Making a large number of simulations of a given reality, makes it possible to characterize the uncertainty of a given phenomenon. Most recent simulation methods are based on a sequential simulation approach, where an extension of the methods of stochastic simulation based on Bayes rule on the conditional probability kriging.

$$P(A|B) = \frac{P(A,B)}{P(B)} \tag{23}$$

2.1.5.1 Direct sequential simulation

DSS performs simulation without transformation of the original variables, thus not requiring any prior indicator coding or multi–Gaussian assumptions. The covariance model is reproduced as long as the conditional distributions identify the local kriging means and variances. Any distribution type (Gaussian, log–normal, uniform) can be used to build the local conditional distribution (Journel, 1994).

However, histogram reproduction becomes an issue when using the DSS algorithm, because DSS does not have the equivalent of the normal score back transform of the often used Sequential Gaussian Simulation (SGS) algorithm. A solution proposed by Soares (2001) consists at sampling the global target histogram at each simulation node, by means of simple kriging, in such a way as to identify the local kriging mean and variance. This method provides an approximation of the target histogram, while honoring the covariance model, resulting in a re–sampling of the global cdf $F_Z(Z)$ in order to obtain a new distribution $F'_Z(Z)$ with its intervals centered on the local mean and with a spread proportional to the conditional local variance (Soares, 2001). This process will determine the shape of the local cdf, at all locations along the path, such that the marginal distribution is approximated at the end of each realization.

DSS can be summarized by the following algorithm (Soares 2001):

- *Definition of a random path that encompasses all nodes to simulate* x_u *where* , $u = 1, \ldots, N$ *and* N *equals to the number of nodes to simulate;*
- *Estimation of the assumed local mean and variance of* $z(x_u)$*, respectively, the simple kriging estimate* $z(x_u)^*$ *and estimation variance* $\sigma^2_{SK}(x_u)$ *conditioned to the experimental data* $z(x_i)$ *and previous simulated values* $z^S(x_i)$*;*
- *Definition of the interval of the global cdf* Fz (z) *to be sampled, by using the algorithm described in Soares (2001);*
- *Draw a simulated value* $z_S(x_u)$ *from the selected interval of the global cdf* $F_z(z)$*;*
- *Loop until all* N *nodes have been visited and simulated.*

2.1.5.2 Direct Sequential co–Simulation

Direct sequential co–simulation (coDSS) is an extension of DSS, for application to two or more spatially depended variables (Soares, 2001). As such, simulation must reproduce the correlation, the individual distributions and the variograms. Specifically, the generated values must be reproduced based on a joint simulation or co–simulation. One of the biggest advantages of coDSS instead of sequential Gaussian co–simulation (coSGS) or sequential indicator co–simulation (coSIS) is its ability to better reproduce the variograms and covariograms of the simulated variables.

This way, the process of the joint distribution simulation

$$F\left(Z_1, Z_2, \ldots, Z_{N_v}\right) = F(Z_1)F(Z_2|Z_1)\ldots F(Z_{N_v}|\ Z_1, Z_2, \ldots, Z_{N_v-1}) \tag{24}$$

assumes N_v variables $Z_1(x), \ldots, Z_{N_v}(x)$ that one intends to simulate. For instance, when considering 2 variables and being $Z_1(x)$ the secondary variable means that this one has the major spatial continuity and is the first to be individually simulated. Subsequently, $Z_2(x)$ is conditionally simulated to the previous simulated values of $Z_1(x)$. Thus, in the spatial location of x_0, the $Z_2(x_0)$ values are generated based on the following conditional distribution law

$$F\left(Z_2(x_0)|Z_2(x_\alpha)\right) = z_2(x_\alpha), Z_1(x_i) = z_1^s(x_i), i = 1, N) \tag{25}$$

whereas $z_1^s(x_i)$ represents the previous simulated values from $Z_1(x)$, $z_2(x_0)$ are the experimental data and the eventually simulated values from $Z_2(x)$ around x_0.

After having the images of $Z_1(x)$, obtained by DSS, the same algorithm is applied to the other variable $Z_2(x)$. One of the images of $Z_1(x)$ is assumed as secondary and to have a positive spatial correlation between $Z_1(x)$ and $Z_2(x)$, characterized by the correlogram $\rho_{1,2}(h)$. This means that both variables have the same spatial pattern statistics, e.g. variogram and global histogram. The $Z_2(x)$ values are generated, in any spatial location x_0, from the conditional distribution laws of the previous simulated $Z_2(x)$ and $Z_1(x)$. The conditional distribution laws are defined, for each location x_0, from the local means $[z_2(x_0)]^*$ and variance $\sigma^2_{SK}(x_0)$, obtained by co–estimation of $z_2(x_0)$ based on the values $z_2(x_\alpha)$, $z_1^s(x_i)$, $i = 1, N$.

Given the fact of the conditioning variable $Z_1(x)$ being plentiful, co–located co–kriging is used for the calculation of local mean and variance, which is a linear combination of the selected values from $z_2(x_\alpha)$ around x_0 and the know value $z_1^s(x_0)$ from the previous simulated image

$$[z_2(x_0)]^*_{csk} = \sum_{\alpha=1}^{N} \lambda_\alpha[z_2(x_\alpha) - m_2] + \lambda_\beta[z_1^s(x_0) - m_1] + m_2 \tag{26}$$

The process of coDSS can then be summarized by the following algorithm (Soares 2001):

1. Definition of a random path encompassing all nodes of x_0 of a given regular grid to be co–simulated;
2. Calculation of the co–estimated local mean and variance of x_0, which are respectively assumed to be, the collocated simple cokriging estimator $[z_2(x_0)]^*_{csk}$ and the estimation variance $\sigma^2_{csk}(x_0)$, conditioned to the experimental values $z_2(x_\alpha)$, the previously simulated values $z_2^s(x_\alpha)$ and to the previously known simulated value $z_1(x_0)$;
3. Definition of the interval of $F_{Z2}(z)$ to be resampled, based on a multi–Gaussian distribution given by:
4. $G\left([y(x_0)]^*, \sigma^2_{csk}(x_0)\right)$, where $[y(x_0)]^* = \varphi([z_2(x_0)]^*)$;

5. Calculation of a simulated value $z_2^s(x_0)$ from the previously defined distribution $F_{Z2}(z)$, by doing:
   - Drawing of a value p from a uniform distribution U(0,1);
   - Drawing of a value y^s of $G\left([y(x_0)]^*, \sigma^2_{csk}(x_0)\right) : y^s = G^{-1}(p)$;
   - Calculation of the simulated value $z_2^s(x_0) = \varphi^{-1}(y^s)$;
6. Return to 1 and loop until all N nodes have been visited and co–simulated.

2.2 Basics on reservoir fluid flow simulation

Fluid flow simulation of hydrocarbon reservoir models is one of the main tasks of a reservoir engineer. Reservoir dynamic simulation resorts to mathematical modeling to study the hydrodynamic behavior of fluids within the oil reservoir. An adequate simulation model is one that represents the main features of the system, accounting for their behavior, yet remaining simple enough so that its calculation time (flow simulation) is appropriate.

The usage of numerical simulation allows the study of a given reservoir in detail, by using individualized cells, populated with relevant properties that enable representation of a realistic version of the system. The reservoir model is assembled by defining and populating the mesh of cells that fit the scale, dimension and spatial distribution of the properties to be represented (i.e. geological or engineering parameters). The end result is a representation of the reality in a discretized way, where rock and fluid properties are defined.

Results for fluid flow simulation are obtained by calculating a set of flow equations of the fluids being transmitted from cell to cell. These flow equations are governed by the principle of mass conservation, combined with Darcy's law and the state equation. This system of partial differential equations describes the motion of all the phases of the fluid within the reservoir.

2.2.1 Fluid flow concepts

As previously mentioned, fluid flow equations, as used in reservoir engineering, are obtained from the combination of three basic equations, respectively, the continuity equation, Darcy's law and the state equation.

The continuity equation stems from the application of the principle of mass conservation of a known volume, through which occurs the flow of a fluid. The mass balance can be translated with the following equation:

$$m_{inflow} - m_{outflow} = \Delta m \tag{27}$$

The terms m_{inflow} e $m_{outflow}$ correspond, respectively, to the mass that goes in or out of the known volume and Δm refers to the mass variation in the volume element.

Darcy's law states that the flow velocity of a uniform fluid in a porous medium is proportional to a pressure drop over a given distance and inversely proportional to the fluid viscosity.

$$u = \frac{q}{A} = -\frac{k}{\mu}\frac{dp}{dl} \tag{28}$$

The fluid velocity (u) is expressed in cm/s, the flow rate (Q) in cm^3/s, the cross–sectional area of the rock (A) in cm^2, permeability (k) in Darcy (D), the fluid viscosity (μ) in cP (centipoise) and the pressure gradient (dp/dL) in atm/cm.

Darcy's law applies only in the case of laminar flow (i.e. when in presence of a low order Reynolds number). Flow within the reservoir is assumed to be laminar, hence the assumption of validity of Darcy's law, with the exception for specific situations such as very high injection and production flows in the vicinity of a well (Silva, 1997).

State equation is a ratio used to represent the mass of the specific behavior of a compressible fluid as a function of pressure and temperature. The choice of the appropriate state equation is dependent on the type of fluid. Generally, as the flow is assumed to be isothermal, the formulation of the state equation will depend only on the pressure.

Since the compressibility of a fluid is given by $C = -\frac{1}{V} \times \frac{\partial V}{\partial P}$ and its density by $\rho = \frac{m}{V}$, the formulation of fluid compressibility can be made according to the following:

$$C = -\frac{\rho}{m} \times \frac{\partial\left(\frac{m}{\rho}\right)}{\partial P} = \frac{1}{\rho}\frac{\partial \rho}{\partial P} \tag{29}$$

Considering the three previously established equations, it is then possible to formulate the basic differential equation, formulated to describe fluid flow in the vicinity of a well. The analytical solutions of the equation can be obtained for various initial conditions and reservoir boundary conditions are applied to the description of the movement of fluids into the well.

Considering the flow of fluid through a volume of thickness dr, located at distance r from the center (radial flow) and applying the principle of mass conservation, the following expression is obtained:

$$q\rho_{inflow} - q\rho_{outflow} = 2\pi r h \emptyset \; dr \frac{\partial P}{\partial t} \tag{30}$$

Where, inflow and outflow correspond respectively to the sections defined by r + dr and r. The volume of the element of width dr is given by 2πrhØ dr.

A simplification of the previous equation is given by:

$$\frac{\partial(q\rho)}{\partial r} = 2\pi r h \emptyset \frac{\partial P}{\partial t} \tag{31}$$

By applying Darcy's law for a radial and horizontal flow, the flow rate can then be substituted by:

$$q = \frac{2\pi khr}{\mu}\frac{\partial P}{\partial r} \tag{32}$$

Which will give us:

$$\frac{\partial}{\partial r}\left(\frac{2\pi Khr}{\mu}\rho\frac{\partial P}{\partial r}\right) = 2\pi rh\emptyset\frac{\partial P}{\partial t} \tag{33}$$

Or,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{K\rho}{\mu}r\frac{\partial P}{\partial r}\right) = \emptyset\frac{\partial P}{\partial t} \tag{34}$$

Considering the previously defined state equation, the fundamental differential equation that regulates radial flow through a porous medium of a monophasic fluid can be obtained by the following expression:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{K\rho}{\mu}r\frac{\partial P}{\partial r}\right) = \emptyset C\rho\frac{\partial P}{\partial t} \tag{35}$$

This expression is nonlinear as it depends implicitly on the density, compressibility and viscosity of the fluid, as shown in the coefficients $\frac{K\rho}{\mu}$ and $\emptyset C\rho$.

2.3 History matching

History matching is summarized as the process of calibrating a hydrocarbon model through improvements on the reservoir characterization based on the mismatch between the response given by the simulated model and the observed historical production. This task is an intrinsic part of the job of any reservoir engineer and/or field management professional.

History matching's importance is mainly attributed to the acknowledgment of:

1. Uncertainty in reservoir characterizations;
2. Economic reasons and their deep connection to the quality and quantity of the information used to describe a given reservoir;
3. Field management decisions, based on predictions in terms of field production;
4. Loss of information in the transition between the geological scale to the simulation scale;
5. Production forecasting.

Typically, history matching is performed at later stages of a reservoir's life. Being a data driven process (i.e. based on historical field production data), the more available data there is, the stronger is the base for the history matching process to work on. This way, history matching can be considered as a method of continuous improvement and validation of the reservoir characterization, as production information is more and more available throughout the reservoir production life.

History matching also entails the problematics behind inverse problems. In short, there is a need to find a set of conditions that is responsible for the replication of a particular result. The problem behind this lies in the fact that inverse problems are almost never well posed. In the particular case of reservoir engineering, an infinite number of equally good solutions can exist, i.e. several different earth models originate the same production response (Carrera, *et al.* 2005). In fact, obtaining a multitude of equally good solutions (history matched models), allows the assessment of the posterior uncertainty associated with a given model. Many combinations of parameters can provide equally good history matching results and be located in different regions of the model parameter space. Such regions are termed local minima and may not be the global minima (optimal minima), or be located far from the global minima. On the other hand, the global minima, which is a representation of reality, can be the best solution for a given model but may not represent the truth (Arnold D. , 2008).

Although numerical representations of reservoir models can be considered an approximation of the real subsurface geology, a multitude of parameters are involved in their description. This increases the difficulty in the history matching process and introduces the problem of nonlinearity, as the relationships between all the parameters involved in the description of the reservoir model, are complex. Nonlinearity of inverse engineering problems, and in particular in the case of reservoir engineering, tries to be addressed by the implementation of history matching algorithms and methodologies that make use of tools and concepts capable of achieving a calibration of the reservoir model in a fast, yet thorough approach.

In a typical history matching workflow, the model (or a set of possible models) generated by geologists is passed on to the reservoir engineering team, which then applies changes directly to the static and/or dynamic assumptions in order for the fluid flow simulation to be able to replicate the real dynamic response. In this process of perturbation, a variety of parameters can be used. They can both be of geologic nature (porosity, permeability, fault transmissibilities, etc.) or engineering parameters (relative permeabilities, saturations, etc.). A generalized workflow for reservoir modeling with history matching and forecasting is shown in Figure 4:

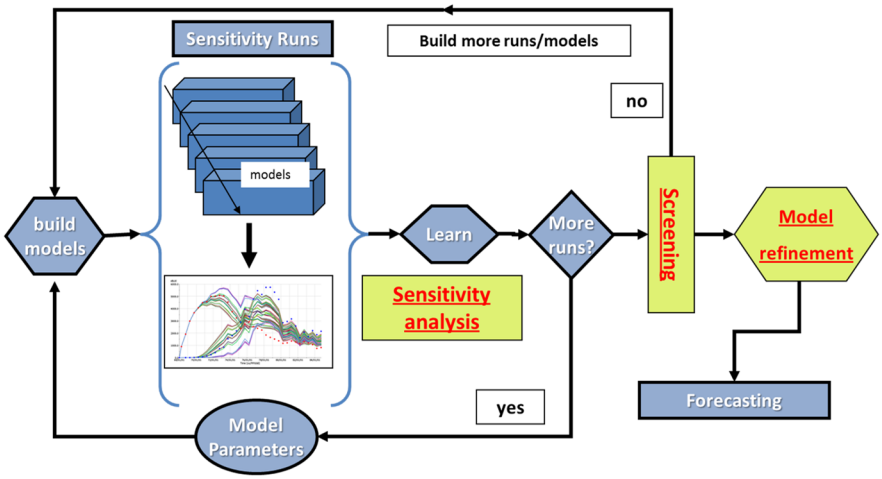


Figure 4 – Generalized workflow for reservoir modeling with history matching and forecasting (Streamsim, 2016).

Knowledge of the reservoir's heterogeneity is required for its full characterization and essential for predicting its behavior. The characterization step should include as much information as possible such as cores, logs, seismic, analogs, outcrops. After building the reservoir model(s) that describe, as best as possible, the known reality, a sensitivity analysis is carried out, using a numerical fluid flow simulator to obtain the dynamic responses of the static models. These dynamic responses will then be submitted to evaluation through the use of an objective function. This objective function should reflect the mismatch between the observed and the simulated production data. It can be summarized as a difference between the observed data and the simulated data. After the calculation of the objective function for the dynamic responses, a decision is made on whether to screen out the calculated models for refinement or if the models need further tuning (or perturbation) of the geological and engineering parameters. This process will run consecutively, in a loop, until a final set of history matched models is chosen and applied in the reservoir's production forecasting.

A list of commonly perturbed geological and engineering parameters is provided by (Williams, *et al.* 1998).

Table 1 – Geological and engineering parameters commonly perturbed in history matching.

Data Type	Data
Pore Volumes	Volume Net–to–Gross (NTG) Porosity
Permeability or Transmissibilities	Horizontal Vertical
Relative Permeabilities	Curves End–Points
Capillary Pressures	
Contacts	Gas–Oil Contact (GOC) Oil–Water Contact (OWC)
Compartments	Gas–Oil Contact (GOC) Oil–Water Contact (OWC)
Fault	Location Transmissibility
Rock	Compressibility
Saturations	Irreducible Water Saturation Oil Residual Saturations
Aquifer	Size (affects Pore Volume) Strength (affects Permeability)

History matching can be performed either manually or in an assisted manner. Manual adjustment is characterized by being a time consuming process of trial and error. Any type of parameter perturbation or combination of perturbations have to be tested by running a dynamic simulation, which is then followed by an interpretation of the results, followed by the drawing of conclusions on what needs to be additionally changed in order to further improve the match, or on what changes are dependent or influent on another. Due to this, manual history matching should be a task reserved only to professionals with deep knowledge on all of the involved geological and engineering concepts in order to preserve the consistency of the perturbed model with what is observable and possible to occur in reality.

Alternatively to manual methods, assisted history matching is the process used to automate the manual adjusting of reservoir parameters, as previously described. It is a process that makes use of an optimization algorithm that tries to find the best parameter combination responsible for the minimization of an objective function. The calculation of the objective function and the result of this calculation is the base on which any algorithm decides if a given response is better (has a lower misfit value) than another. This way, it is able to decide if a given type of perturbation is responsible for a better dynamic response.

The main advantage of assisted history matching is its ability to greatly reduce or minimize time–consumption by leaving professionals free to address other tasks or being able to subdivide the process in stages where human intervention is unnecessary. Another equally big advantage of assisted history matching is having the possibility to explore a much wider spectrum of parameter perturbations.

2.3.1 Geostatistical history matching

In today's industry, the processes of history matching and geostatistical modeling are still widely considered as two separate stages of the reservoir modeling workflow.

Over the past decades, significant developments have been taking place, regarding computer technology. On the other hand, recent years have shown great advances in the development of new history matching techniques. Latest efforts have been focusing on combining both processes.

If fact, recent years have been proliferous in terms of research being made on this subject, with many authors addressing the issue of combining history matching with geostatistical modeling, making it a subject of growing importance in the industry.

Bennett & Graft (2002) proposed a methodology to integrate a history matching tool based on gradient calculation with the combining of different geological settings. One of the principles of the method was the selection or rejection of static models according to the difference between the simulated and observed data. The process consisted in the generation of multiple realizations that were used to perform adjustments in the reservoir model, by using deterministic reservoir parameters, such as relative permeability.

Tureyen & Caers (2003) proposed a methodology for optimization on two levels, by considering one of the levels on a geological scale and another on the simulation scale.

Gross *et al.* (2004) presented a method based on streamline simulation, combined with geostatistical sequential simulation. A similar method of using streamline simulation combined with models being generated by geostatistical simulation was also approached by Le Ravalec–Dupin and Fenwick (2002).

Caers (2002) addressed the combination of history matching with geostatistical modeling through the use of Markov Chain techniques. Later on, Caers (2003) showed a gradual deformation method also combined with streamline simulation. In works published in 2005 and 2006, Caers and Hoffman (2005) (2006) explored methods of history matching by applying regional perturbation of reservoir models with the use of geostatistical modeling techniques. Suzuki and Caers (2006) submitted a stochastic search method, designed to search for reservoir models that matched the historical production data considering a "similarity measure" between likely structural model realizations. In 2007, Hoffman & Caers (2007) submitted a study where history matching was achieved by jointly perturbing local facies proportions and their spatial distribution.

Maschio *et. al* (2008) developed a framework to integrate geostatistical modeling and history matching, where geostatistical images were treated as matching parameters, using genetic algorithm and direct search methods.

Mata–Lima (2008) proposed a methodology to reduce the uncertainty in reservoir characterization by a process of iterative updating of permeability realizations using DSS and co–DSS.

Le Ravalec–Dupin & Da Veiga (2011) proposed an algorithm for continuously modifying realizations representing porosity or permeability fields using co–simulation. The realizations resulting from the co–simulation process would depend on linear correlation coefficients between considered random functions, hence resulting in induced variations in the realizations by perturbing these correlation coefficients.

Soares *et. al* (2014) proposed a geostatistical methodology for history matching conditioned to seismic data. In this approach, the stochastic models simultaneously match the production data for a given reservoir, while producing a synthetic seismic reflection data that matches the observed one.

Although not addressing the issue, Kashib & Srinivasan (2006) argued that the influence of the variogram modeling in the coupling of geostatistical simulation with history matching, is an important research topic.

Figure 5 shows the general workflow adopted in GHM studies.

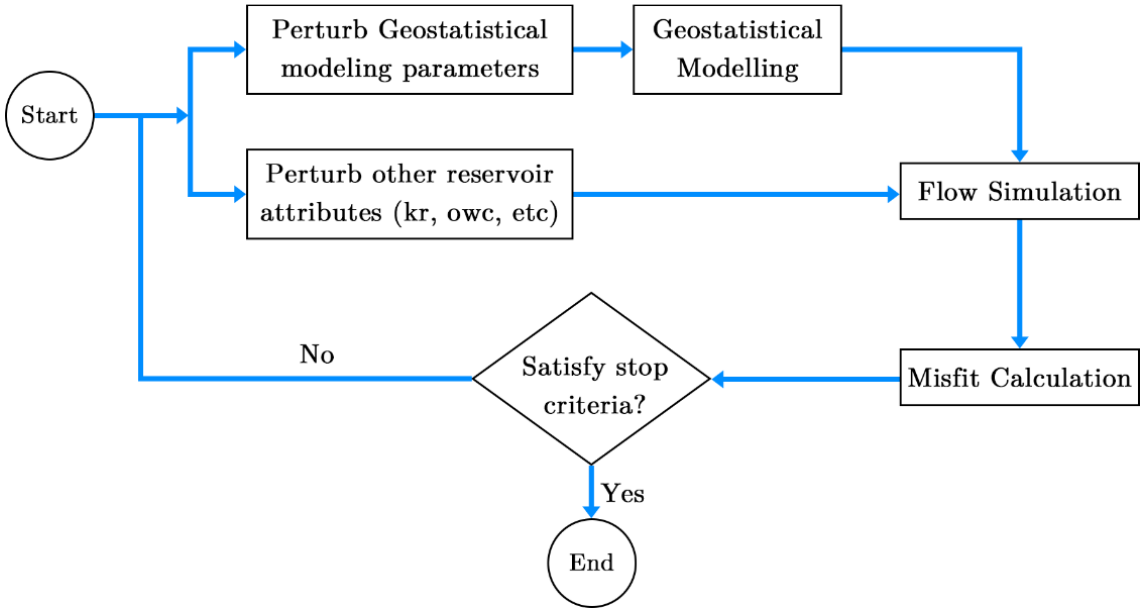


Figure 5 – Generalized workflow for GHM adapted from (Unisim, 2016).

2.4 Uncertainty Quantification

Engineering problems often rely on the integration of real world variability and probabilistic behavior on the finding of solutions. The purpose of uncertainty quantification is to provide means of quantifying, characterizing, investigating and managing the uncertainty associated to these problems.

For the specific case of reservoir simulation, whether static or dynamic, uncertainty comes primarily from the lack of data of the reservoir, measurement errors for the input data and complexity of the static and dynamic phenomena. In order to assess the impact of reservoir static and dynamic uncertainties in reservoir history matching, uncertainty in geologic and engineering parameters needs to be explored under a Bayesian framework (Abdollazdeh, *et al.* 2012).

Reservoir uncertainty quantification is an essential part of reservoir management in which, history matched reservoir models are used for production optimization and forecasting, incorporated with estimation of the uncertainty. By more deeply understanding and quantifying these sources of uncertainty, uncertainty uuantification provides the ability to make decisions with known levels of confidence (Christe *et al.* 2006).

2.4.1 Bayesian theory

The process of calculating the input parameters that match a given observed data is referred to as an inverse problem (Carrera, *et al.* 2005). In a Bayesian framework, if p(m) is the prior distribution of the input parameters (m), the goal is to update this knowledge and find the distribution of parameters that best fit the observed data (o). A rigorous way to incorporate the new information is to use the Bayes rule,

$$p(m|o) = \frac{(p(o|m)p(m))}{p(o)} \tag{36}$$

where p(m|o) is the posterior distribution conditioned on o, and $p(o|m)$ is the likelihood function, which is the probability of the data, given the model m is true.

In complex representations of the reality, such as is the case of reservoir models, analytical formulation of the posterior from the prior is unfeasible as the likelihood implies complex forward simulations of the processes. Most importantly, the denominator in the Bayes rule equation is generally unknown.

These issues are resolved by Monte Carlo Markov Chains (MCMC) techniques. MCMC techniques makes use of ratios of densities, allowing them to sample from unnormalized distributions and obtaining samples from the posterior distribution (Robert & Casella, 2004).

The choice of prior distribution is a critical step. Rojas *et al.* (2009) have shown that, in a multi–model framework, the posterior plausibility of the various conceptual models heavily depends on the prior model probabilities. Ideally, extremely large datasets could be acquired to eliminate the influence of the choice of prior. This is unrealistic for oil reservoir applications due to the large number of uncertain parameters and to the sparsity of data. A possible strategy is to obtain expert opinion on how to identify a reliable prior.

The likelihood function is defined by the error model selected for the analysis (Kavetski, *et al.* 2006). Yustres *et al.* (2012) remark that the correct construction of the likelihood function and the identification of the sources of uncertainty are essential to carry out a meaningful analysis. As previously mentioned, a misfit value is calculated to quantify the discrepancy between simulated and observed data. The most common choice is to assume that the misfit dominates other sources of errors and that the errors are of Gaussian nature.

This way, by considering that the likelihood L is given by:

$$L = p(o|m) \propto e^{-M} \tag{37}$$

with M being the misfit value. M can be obtained by the following expression.

$$M = -log\big(p(o|m)\big) = -log(L) \tag{38}$$

Apart from the problems related to the choice of prior and likelihood distributions, the main limitation of Monte Carlo approaches comes from the computational cost of simulating the flow responses, particularly when in presence of complex physical processes. Dynamic response evaluation is necessary at each iteration (i.e. for every step of perturbations) resulting in an intractable computational cost.

2.4.2 Adaptive Stochastic Sampling algorithms

Adaptive Stochastic Sampling algorithms are a branch of optimization algorithms commonly used in the oil industry for solving inverse problems, such as is the case of history matching. Several of these algorithms were inspired by natural phenomena. Some of the most commonly used Adaptive Stochastic Sampling algorithms are:

- Genetic algorithms
- Simulated Annealing
- Evolutionary algorithms
- Swarm algorithms

Evolutionary and swarm algorithms are among the most recent algorithms. The main difference between them is their ability to find multiple local minima, which is an essential characteristic for achieving a realistic quantification of uncertainty prediction. Swarm algorithms take advantage of the tracing of the actuating particles for achieving a balance between exploration and exploitation of the search space, in the search for an optimal solution.

The present work makes use of a particular case of the family of swarm algorithms, namely the PSO algorithm, which is explained in the following chapter.

2.4.2.1 Particle Swarm Optimization

PSO is a population based evolutionary, developed by Kennedy and Eberhart in 1995 (Kennedy & Eberhart , 1995). It is a stochastic optimization technique inspired by the social behavior of bird flocking or fish schooling.

The principle behind the PSO algorithm implies considering a search space where a swarm of particles is randomly scattered. Apart from a position in the search space domain, all particles have a specific velocity at which they update their position.

Every particle is a candidate to the solution of the problem and retains its previous best position. All of the positions of all of the particles are simultaneously compared at every step of the algorithm by a fitness function evaluation. From this, a global best position is obtained, which is then responsible (along with all the best positions of all different particles) to converge the solution of the problem to a minimum value.

The PSO workflow can be summarized in the following steps (Mohamed, 2011):

1. *Initialization of the swarm by assigning a random position in the search space to each particle with sensible random velocity.*
2. *Evaluation of the fitness function for each particle.*
3. *For each particle, update the position and value of* p_{best} *– the best solution the particle has seen. If current fitness value of one particle is better than its* p_{best} *value, then its* p_{best} *value and the corresponding position are replaced by the current fitness value and position, respectively.*
4. *Update the global best fitness value and the corresponding best position.*
5. *Update the velocities and positions of all the particles.*
6. *Repeat steps 2–5 until a stopping criterion is met.*

Other common stochastic sampling algorithms used for uncertainty quantification are the Hamiltonian Monte Carlo algorithm, the Neighborhood Algorithm and the Multi–Objective PSO algorithm, which is an extension of the PSO algorithm for the case of multi–objective optimization (more than one objective function),.

For further details on these algorithms, the reader is encouraged to consult the works made by Mohamed (2011) and Mohamed *et al.* (2011).

3 Methodology Workflow

Within the scope of this work a new algorithm for GHM is proposed, following the works on regionalization–based history matching algorithms done by Hoffman & Caers (2005) , Mata–Lima (2008) and Le Ravalec-Dupin & Da Veiga (2011). The proposed methodology couples Adaptive Stochastic Sampling with zonation–based geostatistical history matching.

The proposed algorithm encompasses a two–stage optimization process. The first stage, or the inner loop of the algorithm, is the stage where the geological trends of the reservoir, in terms of porosity and permeability, are perturbed locally, under a regionalization model using DSS with multi–distribution functions and spatial continuity patterns (Nunes, *et al.* 2016). After each iteration, perturbation on the porosity and permeability models are assimilated through geostatistical conditioning, by resorting to best porosity and permeability cubes, generated at the end of each iteration. These volumes will serve as secondary variable (*soft data*) along with a corresponding correlation coefficient volume that is set according to the dynamic evaluation of the data. The calculation of these correlation coefficients are based on the mismatch between observed and simulated values given by the calculation of an objective function.

The second loop of the proposed algorithm encompasses the first loop and is responsible for exploring the uncertainty in the large scale geological parameter space: variogram parameters, porosity and permeability histograms and fault transmissibilities that interconnect the different zones of the reservoir. For this, a workflow using Adaptive Stochastic Sampling and the PSO algorithm is used. At each iteration, the PSO algorithm optimizes the uncertain large scale parameters which are then used as input for the inner GHM loop. The process iterates until a final number of iterations is reached, or when convergence of the parameter perturbation and misfit value is above a pre–defined threshold.

The proposed methodology is synthetized by the following workflow:

First stage (Geostatistical history matching with zonation; Figure 6):

1. Regionalization of the reservoir area according to a given fault model and area of influence of the wells resulting in a cube with a zone being assigned for each well or for a group of wells;
2. Simulation of a set of permeability and porosity realizations through DSS, honouring the well data, histograms and spatial distribution revealed by the variogram;
3. Evaluation of the dynamic responses for each of the realizations and calculation of the mismatch between the dynamic response and real production data, using an objective function;
4. Calculation of a correlation coefficient between each dynamic response and generation of a cube composed of different correlation coefficients per zone;
5. Creation of a composite cube of porosity and permeability, with each region being populated by the corresponding realization with the lowest mismatch, calculated in step 3;

6. Return to step 2, using co–DSS and the cubes calculated in steps 4 and 5 as local correlation coefficient and soft data, respectively. The algorithm is expected to run up to a maximum number of iterations, or until a pre–defined mismatch value is reached.

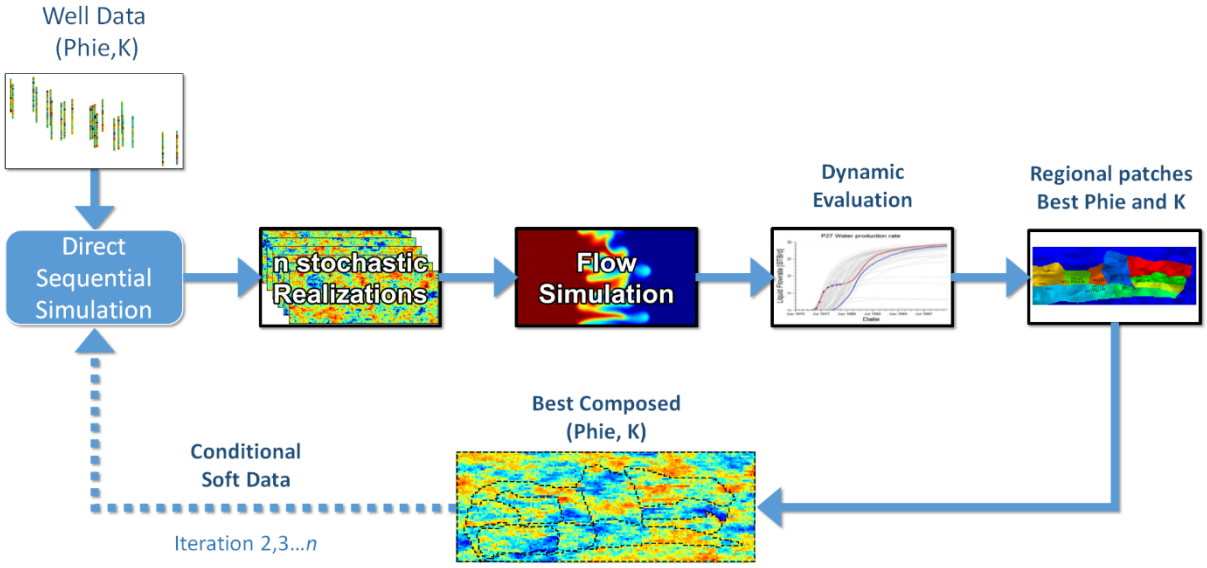


Figure 6 – GHM workflow (First stage).

Second stage (Coupling with Adaptive Stochastic Sampling):

The second stage can be considered as an outer loop, encompassing the results obtained by the first stage. At each iteration of the zonation–based GHM loop, a best porosity and permeability cube is obtained, along with their respective correlation coefficient cubes. A process of matching relevant engineering and geological parameters will then occur, using Bayesian inference and Adaptive Stochastic Sampling. The second phase runs for a maximum number of iterations.

For this work, the parameters considered for perturbation were the variogram parameters being used to generate the geostatistical realizations on the inner GHM loop, histogram perturbation and fault transmissibilities. Other geological or engineering parameters can be used with the same proposed procedure.

The flowchart on Figure 7 illustrates both stages of the proposed algorithm methodology.

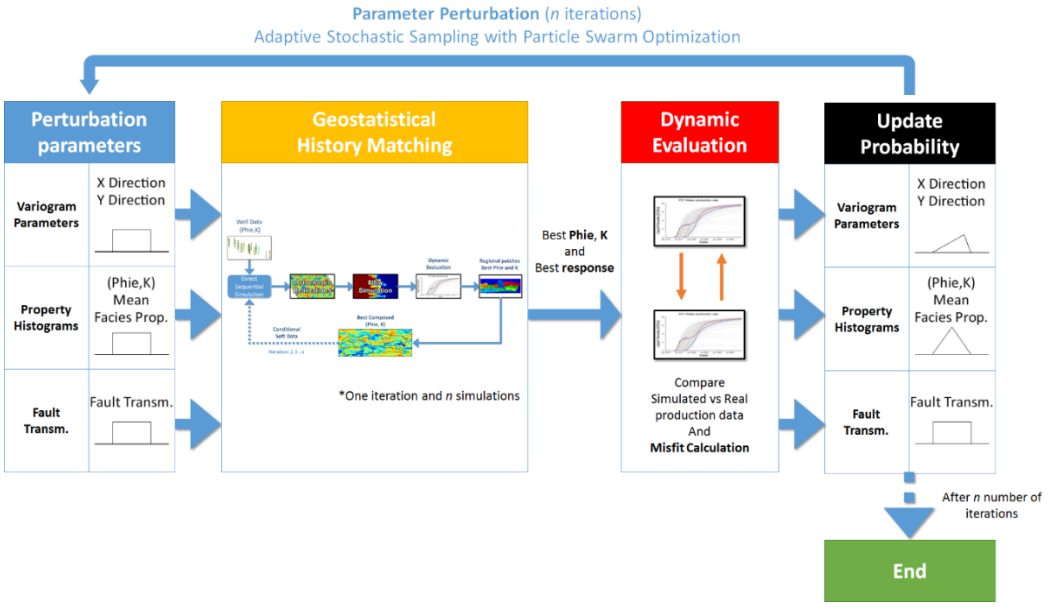


Figure 7 – Coupling with Adaptive Stochastic Sampling algorithm. (Second stage) GHM is shown in detail in Figure 6

3.1 Dynamic evaluation

3.1.1 Objective function

The objective function is formulated as a mathematical expression that measures how close a problem solution (simulated data) is towards an optimal value (observed data). The definition of such a metric is critical in order to achieve convergence in the iterative procedure.

The most commonly used objective function for history matching is the least square norm, which calculates a measure of the discrepancy between the simulated and observed value, returning a "mismatch" value, M. The least squares norm formula to be used in the present work is defined as:

$$M = min \sum_{t=i}^{n} \frac{(R_{t,obs} - R_{t,sim})^2}{2\sigma^2} \tag{39}$$

Where $R_{t,obs}$ is the observed, or historic value of a given variable at timestep t, $R_{t,sim}$ is the simulated value of a given variable at timestep t, σ^2 is the error, or standard deviation, associated to the measurement of a given variable at timestep t.

3.1.2 Correlation coefficient

In GHM (first stage of the proposed algorithm) convergence is assimilated through geostatistical soft conditioning. This means that the perturbation being made by the many realizations generated by the

algorithm, will be carried on to the following iterations by means of conditioning, according to its match quality and the degree to which they are close to the real observed data. The measure adopted to depict this degree of closeness to the observed data is done through a correlation coefficient applied to every simulated response.

Due to the importance of the correlation coefficient in the proposed GHM workflow, it is crucial to make sure its value mimics the misfit value obtained by the objective function. Meaning, a process that integrates calculations of misfit and correlation coefficients should guarantee that, when calculated, these values change symmetrically and in the same proportion, under different configurations of observation data curves.

This issue was also tackled within the scope of this work, with the development of a new approach to calculate a correlation coefficient based on the misfit value. The following step–by–step list describes the calculation of the correlation coefficient being used by the GHM algorithm.

1. For all timesteps and for the production variable to be matched, calculate the response difference, or a deviation from the response, towards the observed data and the corresponding error threshold, respectively Δ_{sim} and Δ_{sigma}, according to the following:

$$\Delta_{sim} = \left|R_{(t,obs)} - R_{(t,sim)}\right| \tag{40}$$

$$\Delta_{sigma} = \sigma_t^2 - \Delta_{sim} \tag{41}$$

2. The next step is the calculation of the final deviation towards observed data for a given timestep, Δ_{obs} , according to the following:

$$\Delta_{obs} = \begin{cases} R_{t,obs} - (\Delta_{sim} + \Delta_{sigma}) & R_{t,obs} \geq \Delta_{sim} + \Delta_{sigma} \\ 0 & R_{t,obs} < \Delta_{sim} + \Delta_{sigma} \end{cases} \tag{42}$$

3. After getting the final deviation value of the simulated response, there is a need to normalize this value in order to get a value from 0 to 1. The calculation of each timestep correlation coefficient, x_t ,is done by the following calculation:

$$x_t = \begin{cases} 1 - \dfrac{\Delta_{sim}}{(\Delta_{sim} + \Delta_{sigma} + \Delta_{obs})} & \Delta_{sim} \leq \Delta_{sigma} + \Delta_{obs} \\ 0 & \Delta_{sim} > \Delta_{sigma} + \Delta_{obs} \end{cases} \tag{43}$$

4. The final correlation coefficient for a given response is then obtained by averaging all x_t from all timesteps according to the following:

$$\overline{x} = \frac{\sum_{t=1}^{N} x_t}{N} \tag{44}$$

Where $R_{t,obs}$ is the observed, or historic value of a given variable at timestep t, $R_{t,sim}$ is the simulated value of a given variable at timestep t, σ^2 is the error of the measurement at timestep t.

This method ensures a degree of compatibility between the misfit value resulting from the objective function and the correlation coefficient to be used in the co–simulation of a new set of models. It also takes into account different configurations of obtaining simulated values and attempts to make the calculation of the resulting correlation coefficients consistent, regarding the relative position between observed and simulated values as well as their respective error threshold. For further explanation, Figure 8 depicts a timestep of four possible cases of simulated data value when compared to the real data.

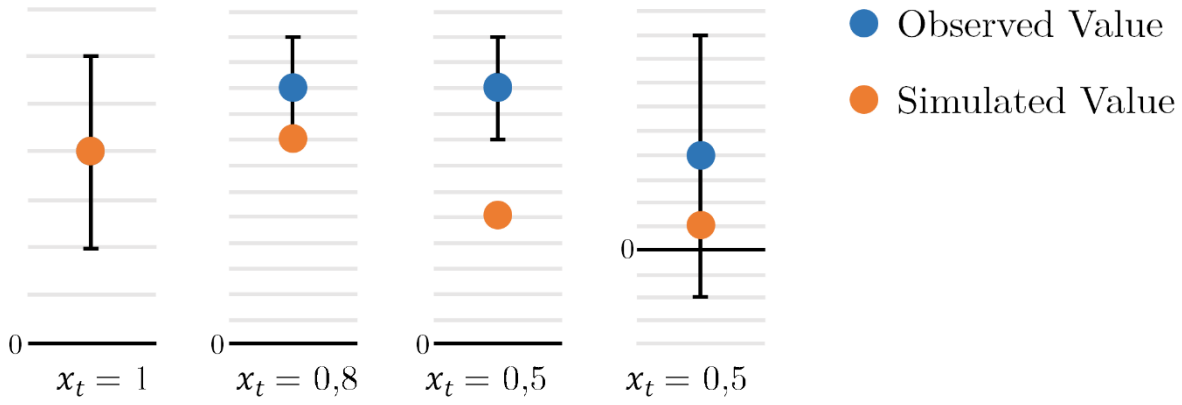


Figure 8 – Correlation coefficient calculation on four possible cases of simulated vs observed data.

Figure 8 shows, from the first through the third plot, the generic case for simulated value over observed value at a given timestep, regardless of the sigma interval. The fourth plot shows a case where the simulated value is inside the sigma interval, but the sigma value goes below 0. In this case, since the simulated response is half way close to the observed value, the correlation coefficient is calculated as 0.5. This is particularly important for specific cases where observed values are close to zero at several timesteps, such as is the case of Water Production near the beginning of the simulation (before water breakthroughs). By looking at the fourth plot, it is possible to observe that if this case were not to be accounted for, the correlation coefficient given by the normalization of the simulated value over observed value would return a much lower x_t value (in some cases even almost zero or division by zero), when in reality the simulated value is located at about half the lower bar of the accepted error threshold, for the specific timestep. In this case, the value for the timestep correlation coefficient is calculated as 0.5.

3.2 Zonation–based Geostatistical history matching

Zonation methodologies (also known as regionalization or compartmentalization) are a common approach in reservoir engineering and history matching, in order to reduce the number of parameters of a given reservoir description. The zonation methodology, as adopted in petroleum reservoir models, consists on grouping grid blocks of a given reservoir's static or dynamic model in order to create sub–regions with property values assumed to be either constant, correlated or homogeneous. The calibration of these properties is then obtained by the process of history matching, which in this case, is the

adjustment of the parameters on these zones, by applying perturbation aimed towards reproducing the observed production data of the reservoir.

One of the objectives of this work was to explore different regionalization methodologies and their influence on the history matching quality and convergence. The following sub–chapters are a presentation of the explored zonation methodologies.

3.2.1.1 Global zonation

As a way of testing the influence of having a no zonation based approach, a global zonation methodology was adopted. In this zonation approach, there is only a global region that encompasses the entire reservoir area. The global zonation is therefore used as a form of depicting the type of conditional assimilation applied by the iterative geostatistical conditioning process, under a global correlation coefficient. Since all reservoir wells are encompassed by this global region, a global correlation coefficient is used for the whole reservoir and calculated according to the match quality of simulated FIELD data over observed FIELD data (Field Oil Production Rates, Field Water Production Rates and Field Gas Production Rates). Figure 9 illustrates a top view of the Global zonation pattern:

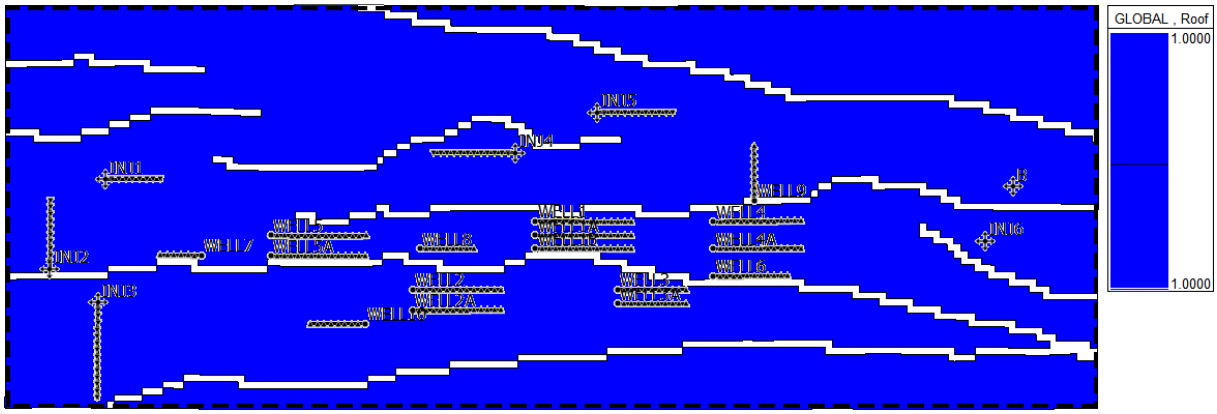


Figure 9 – Global zonation pattern.

3.2.1.2 Circular zonation

A circular area of influence around each well is also considered an important aspect when carrying out regionalization approaches. Within the scope of this work a zonation method exploring such pattern was also considered. In this regionalization method, eleven different zones are considered. Ten of them corresponding to the different well groups of the reservoir and one reserved for the area outside. Permeability and porosity are updated inside each individual area according to the mismatch between the observed data and simulated production data of the corresponding wells or well groups (Well Oil Production Rates, Well Water Production Rates and Well Bottom Hole Pressure). The same going for other regionally based methods.

To avoid overlapping areas, the area of influence considered for each well, or groups of wells, was defined by a distance of 500 meters from every well completion node (roughly 4 to 5 grid blocks). This resulted in the multilayered spherical shaped pattern shown in Figure 10.

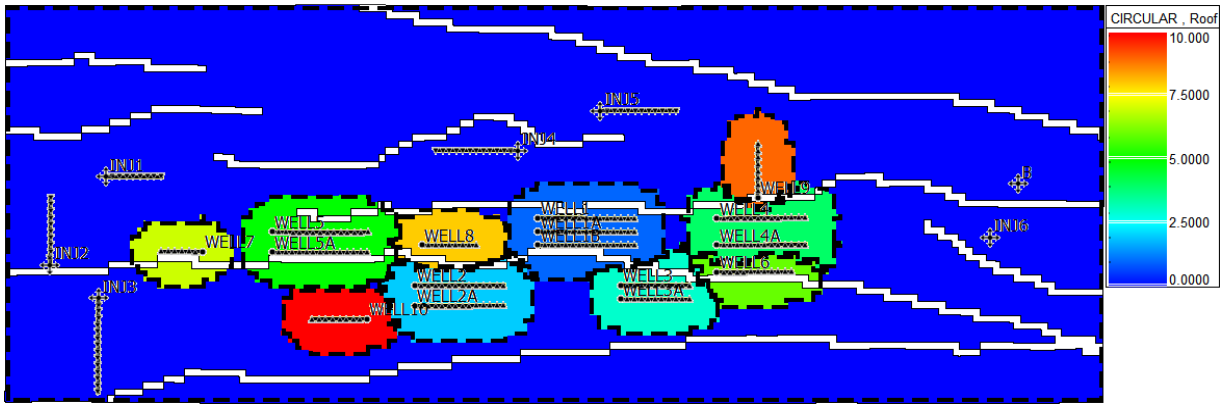


Figure 10 – Circular zonation pattern.

3.2.1.3 Square zonation

In order to compare results observed on different zonation criteria, there was a need to set up a regionalization pattern that did not take into account a specific zonation criteria, such as circular areas of influence, geological consistency or fluid flow pattern.

The idea behind using a squared zonation was to observe the contrast between adopting a non–geological consistent zonation methodology and a geological consistent zonation methodology. In this case, the adopted sub–regions are square shaped, disregarding not only any geological consistency of parameter perturbation but also well production areas of influence and relevant geological features such as fault presence. Different well groups are also encompassed by the same regions, meaning that the update of petrophysical models constrained to production data is disturbed. This resulted in the multilayered square shaped zonation, shown in Figure 11:

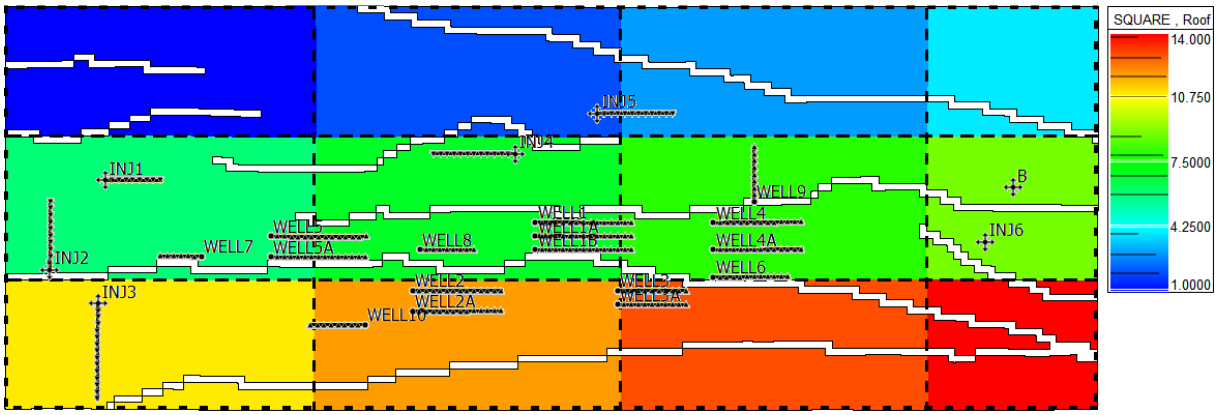


Figure 11 – Square zonation pattern.

3.2.1.4 Voronoi zonation

Many GHM approaches use Voronoi polygons to address zonation. The main advantage behind Voronoi zonation is its ease of use, its ready–to–use implementation on several available software packages and the ability to efficiently partition 2D or 3D space between a collection of points. The outcome of a Voronoi zonation method is such that each subdivided region (also known as Voronoi cell or domain) consists of a grouping of all points in the space that are closer to a specific site than any other, in this case well locations.

As in the circular zonation methodology, the regionalization by Voronoi polygons is also a type of representation of the area of influence of each well or group of wells. Figure 12 shows the zonation resulting from the Voronoi partition and shows the existence of a zone for every group of well (zone 1 to 10). Some regions had to be tweaked in order to include all wells in a specific group (i.e. WELL4 and WELL4A) or to associate injector wells to a given region (i.e. INJ6 and B). The final obtained pattern is the multilayered Voronoi type regionalization shown in the Figure 12.

Figure 12 – Voronoi zonation pattern.

3.2.1.5 Fault and Streamline–based zonation

After interpreting the results obtained from the running of several GHM loops using the regionalization patterns previously described, a *hybrid-like* regionalization pattern was adopted in order to capture the accuracy obtained by the different regionalization methods in each specific case. For this, a regionalization method containing aspects of geological consistency, fluid flow paths along with their area of influence and fault zonation, was proposed. This regionalization method was primarily conceived in order to improve mismatch results on the GHM stage, while maintaining a degree of geologic consistency during parameter perturbation. As for the second stage, concerning the integration with Adaptive Stochastic Sampling, fault transmissibility perturbation is addressed, along with other geological parameters such as variogram *ranges of influence* and property histograms. This stresses the point of having a regionalization methodology which is able to provide a support, which is compatible

with the type of sampling perturbation being addressed. Considering a geologically consistent zonation method to serve as a strong base for geostatistical perturbation, by addressing preferential flow paths, the area of influence for each well and fault zonation, is expected to provide better and more efficient match qualities under Adaptive Stochastic Sampling perturbation.

In terms of areas of influence, the assumption of a radius of influence, made in the circular zonation method, was replaced by the area consisting of the most common flow paths for production, given by streamlines. For this, a number of flow simulations were ran over several petrophysical models with a global regionalization approach matched towards the observed production data (in an attempt to reach an unbiased common flow pattern, in terms of any regionalization assumptions).

Figure 13 illustrates a top view example of the of the well production streamlines for a given realization.

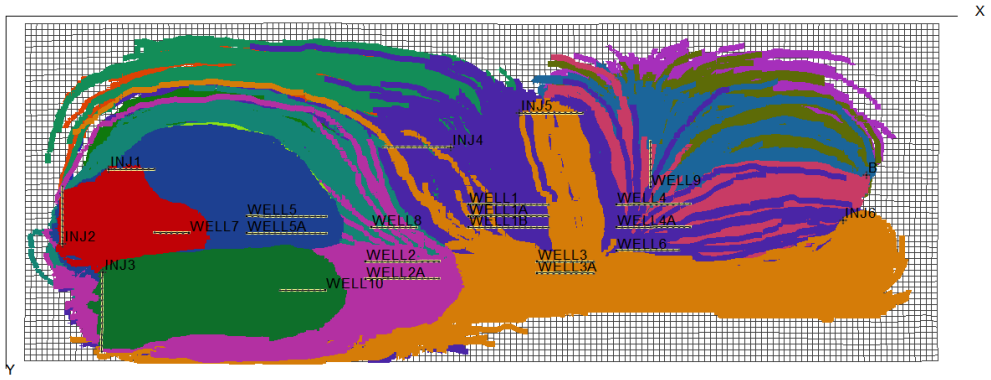


Figure 13 – 2D top view of production streamlines from different wells.

Figure 14 illustrates the 3D view of the example given above. Note the characteristic shape of the production streamlines associated with each well and the way they overlap.

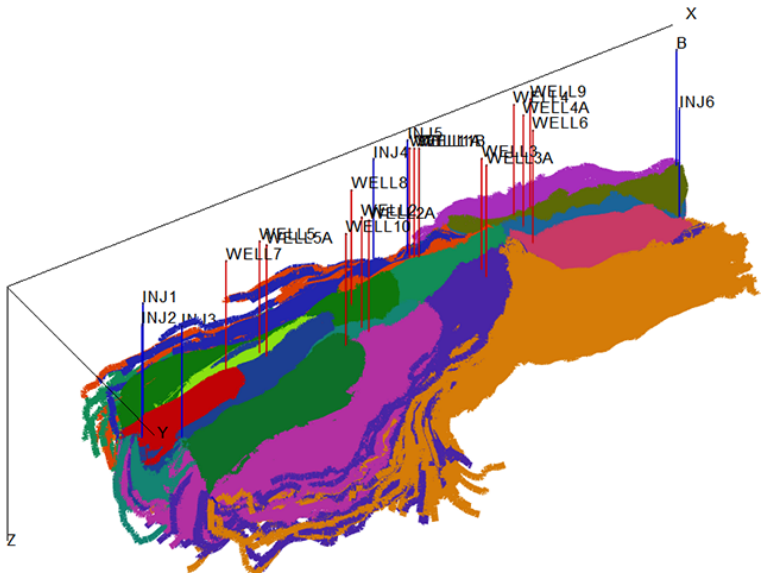


Figure 14 – 3D view of production streamlines from different wells.

In order to arrive at a fault and streamline–based zonation, fault locations were considered and integrated with production streamline data. In this way, it is possible to observe preferential flow paths and their interference with existing faults in the model. Water production streamlines were selected, since they are the hardest to match in the present case study.

Figure 15 is an example of the preferential flow path for water production in the regions being associated respectively with well group 1, 8, 9 and 10. The color scheme represents the most commonly visited cell blocks (blue – hardly visited, to red – highly visited), considering flow path from a specific well group, over the course of a whole simulation.

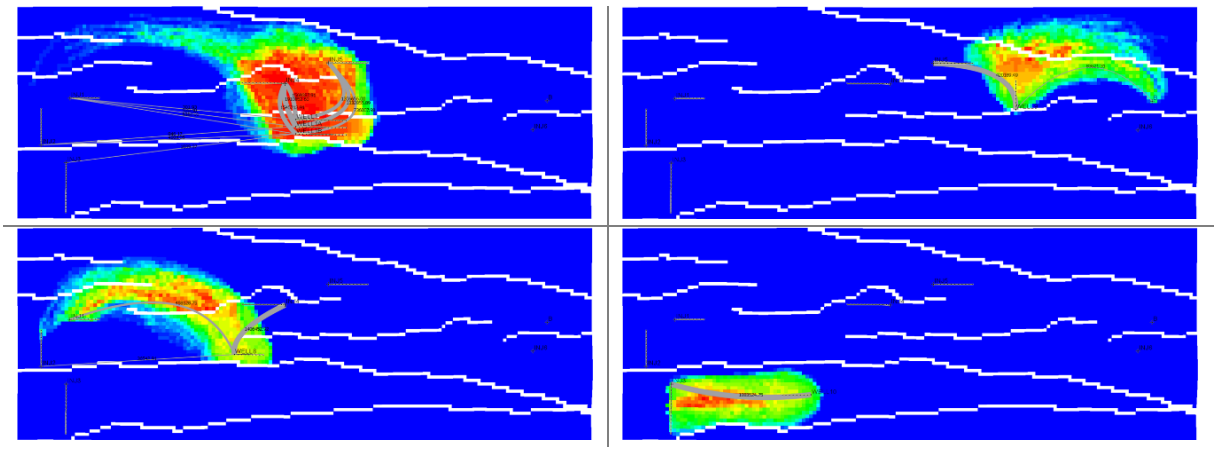


Figure 15 – Preferential flow path for water production, for well group 1 (top left), well group 9 (top right), well group 8 (bottom left) and well group 10 (bottom right).

In Figure 15, the grey lines on each figures, represent the total production of water from each individual well at the end of the streamline simulation. Each line connects a producing well to one or more injectors. The thicker the line, the higher the volume of produced water. Also note the existence of barriers to fluid flow, being introduced by the presence of some of the faults of the considered fault model (represented in white).

By aggregating the extracted information related to common flow paths for the dynamic model, presented above and taking into account a fault-based zonation criterion, the proposed regionalization method, composed of a hybridization of these two concepts, is presented in Figure 16.

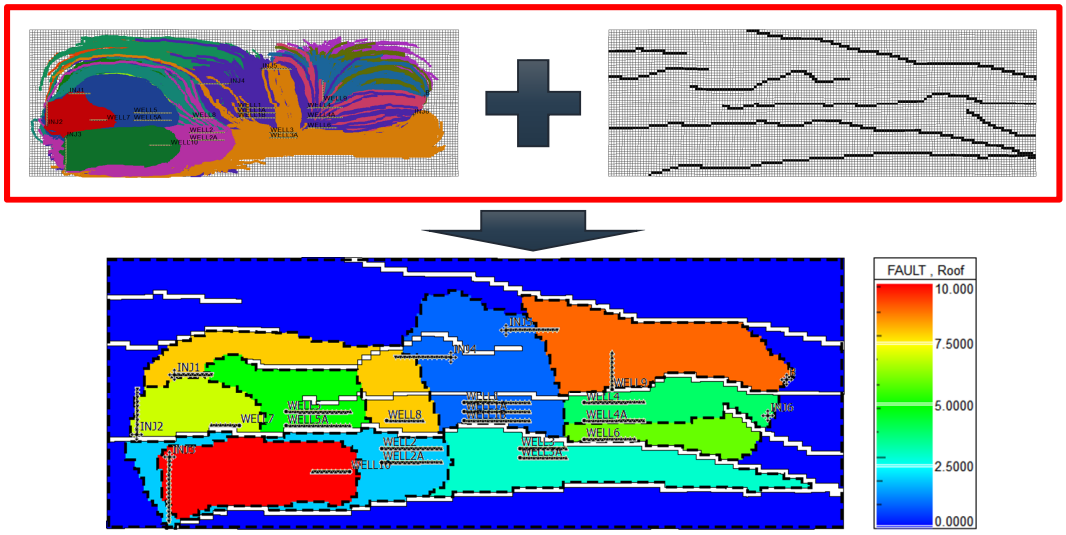


Figure 16 – Fault and Streamline zonation–based pattern.

Figure 17 is a 3D view of the adopted regionalization. It is possible to observe the pattern of overlapping and enveloping of flow regions, previously observed in the streamline study, being replicated.

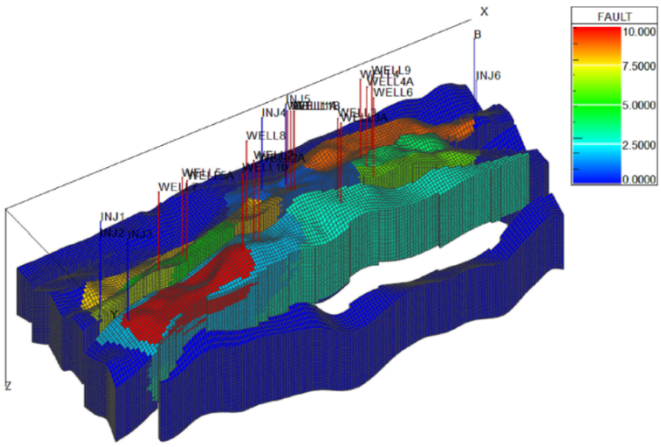


Figure 17 – 3D view of the Fault and Streamline zonation–based pattern.

3.3 Coupling with Adaptive Stochastic Sampling

One of the main objectives of this work is the development of a methodology able to couple GHM with Adaptive Stochastic Sampling. This allows the uncertainty quantifying on two different levels. GHM algorithm aims at quantifying uncertainty of the petrophysical properties assuming some *a priori* knowledge related with large scale geological parameters, e.g. the spatial continuity pattern. Adaptive Stochastic Sampling is able to quantify the uncertainty related with the geological parameters assumed by the geostatistical algorithms that are, for example, responsible for reproducing a given spatial continuity pattern. In addition, Adaptive Stochastic Sampling is able to assess uncertainty over engineering parameters that are set to be perturbed and that influence dynamic match.

The conventional GHM procedure used in the inner loop of the proposed approach (Chapter 3.2), assumes stationarity in the geostatistical parameters. However, there might be a huge lack of information about the correct interpretation and setting of these parameters. As a consequence, there is a lot of uncertainty on the resulting petrophysical models.

To quantify this uncertainty, the conventional history matching approach is coupled with an Adaptive Stochastic Sampling algorithm, to consider the uncertainty related with these parameters. In this specific case study, the considered uncertainty relates with the spatial continuity model of the data, represented by the variogram model parameters and the property histogram definition. Fault transmissibilities acquire a special relevance in terms of parameter uncertainty in GHM with a regionalization method based on fault presence. To tackle this issue, uncertainty quantification by perturbation of fault transmissibilities was also addressed.

PSO was used as the optimization algorithm for sampling the parameter space. A total of 17 parameters were perturbed in the coupling stage of the proposed multi-level methodology. As opposed to using a fixed value for the parameters mentioned above, uniform distributions (or discrete uniform, in the case of variogram *ranges of influence*) were assumed as a way of enveloping the *a priori* uncertainty for every specific parameter. The prior distributions are used by the stochastic sampling algorithm to sample the parameter space and converge to optimum parameter values.

3.3.1 Parameter uncertainty

3.3.1.1 Variogram perturbation

The variogram model is a key aspect to any geostatistical modeling procedure. Nevertheless, variogram modeling is a task that is highly prone to uncertainty due to the lack of experimental data for reliable calculation of experimental variograms. Since the true variogram of a given variable is unknown, a variogram model is fitted to an experimental variogram retrieved from the available experimental data. Unfortunately, variogram estimation from experimental data can be difficult due to a wide array of aspects, such as lack of data or inadequate data location.

The uncertainty parameters related with a variogram model, as for example variogram directions and *ranges of influence*, can be assessed by generating multiple variogram models that fit equally the experimental variogram. Added to that, when considering the specific variogram parametrization of a given petrophysical property, multiple stochastic simulations will result in a number of equiprobable spatial distribution models, allowing the assessment of the spatial uncertainty associated with a given property. It is of high importance to be able to quantify variogram uncertainty, in order to ensure that the variogram estimates are sufficiently accurate for the task at hand.

Within the scope of this work, uncertainty was assumed to exist in the horizontal and vertical *ranges of influence* for both porosity and permeability variograms. The considered semivariogram model was

exponential, with 0.1 *nugget effect*. The adopted parameterizations for porosity and permeability ranges are summarized in Table 2 and Table 3, respectively, and illustrated in Figure 18 and Figure 19 , respectively.

Table 2 – Uncertainty Quantification parameters for Porosity variogram ranges.

Property	Parameter	Variable Name	Distribution Type	Distribution Range
Porosity	Horizontal. Range	$pororange1	Discrete Uniform	[40, 80]
	Vertical. Range	$pororange2	Discrete Uniform	[10, 30]

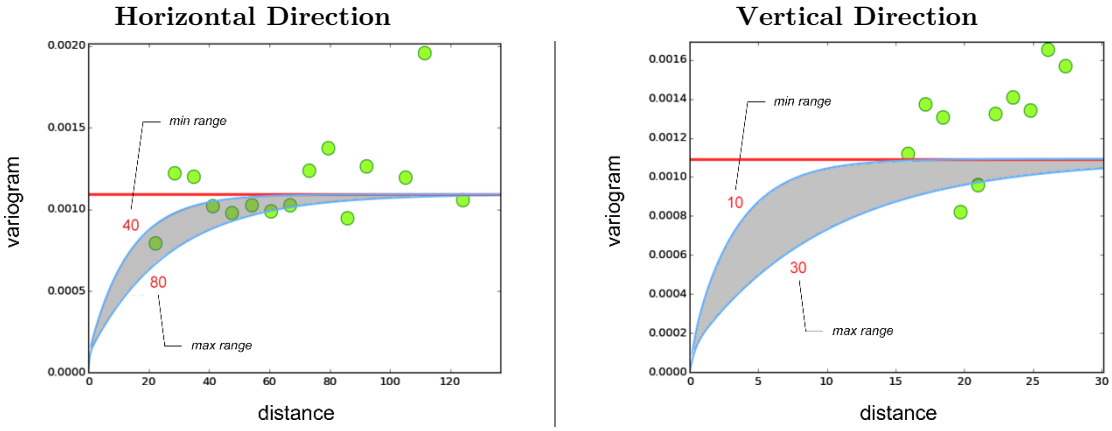


Figure 18 – Illustration of the uncertainty in Porosity ranges.

Table 3 – Uncertainty Quantification parameters for Permeability variogram ranges.

Property	Parameter Description	Variable Name	Distribution Type	Distribution Range
Permeability	Horizontal. Range	$permrange1	Discrete Uniform	[40, 80]
	Vertical. Range	$permrange2	Discrete Uniform	[10, 30]

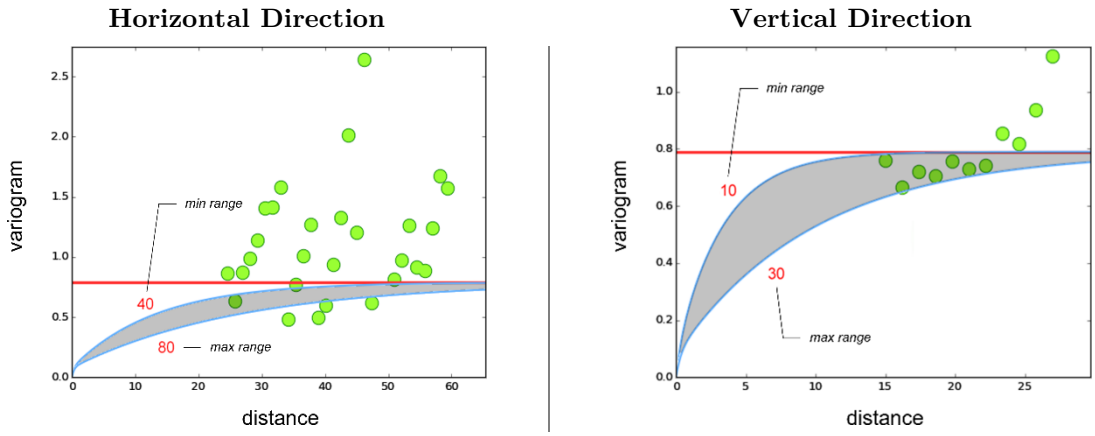


Figure 19 – Illustration of the uncertainty in Permeability ranges.

3.3.1.2 Histogram perturbation

Uncertainty is also part of existing experimental data, used as hard data for generating geostatistical realizations. In fact, available appraisal data might not reproduce the true histograms for porosity and permeability of a given field, due to the fact that appraisal wells are often aimed at target areas where petrophysical properties maximize the chance of hitting oil. The bias in terms of the location of appraisal wells affects the property histograms. To address this issue, perturbation on porosity and permeability histograms was considered and the distributions retrieved from the experimental data were approximated in terms of Gaussian mixture models. Each Gaussian family was optimized in terms of facies proportions and facies mean values. Table 4 summarizes the parameter distributions and thresholds assumed for the porosity histogram perturbation.

Table 4 – Uncertainty Quantification parameters for Porosity histogram perturbation.

Property	Parameter Description	Variable Name	Distribution Type	Distribution Range
Porosity	Facies 2 Proportion	$poro_fac_2	Uniform	[0.0, 0.3]
	Facies 1 Mean	$poro_mean_1	Uniform	[0.15, 0.19]
	Facies 2 Mean	$poro_mean_2	Uniform	[0.19, 0.23]

Figure 20 illustrates the parameter distributions and thresholds assumed for the porosity histogram perturbation.

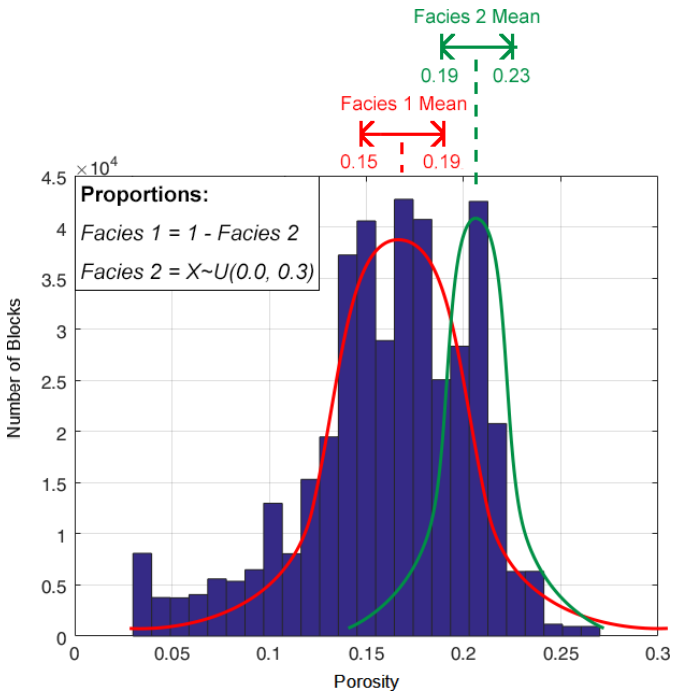


Figure 20 – Illustration of the uncertainty in the Porosity histogram.

Table 5 summarizes the parameter distributions and thresholds assumed for the permeability histogram perturbation:

Table 5 – Uncertainty Quantification parameters for Permeability histogram perturbation.

Property	Parameter Description	Variable Name	Distribution Type	Distribution Range
Permeability	Facies 2 Proportion	$perm_fac_2	Uniform	[0.7, 1.0]
	Facies 1 Mean	$perm_mean_1	Uniform	[–1.5, 1.5]
	Facies 2 Mean	$perm_mean_2	Uniform	[1.5, 3.0]

Figure 21 illustrates the parameter distributions and thresholds assumed for the permeability histogram perturbation.

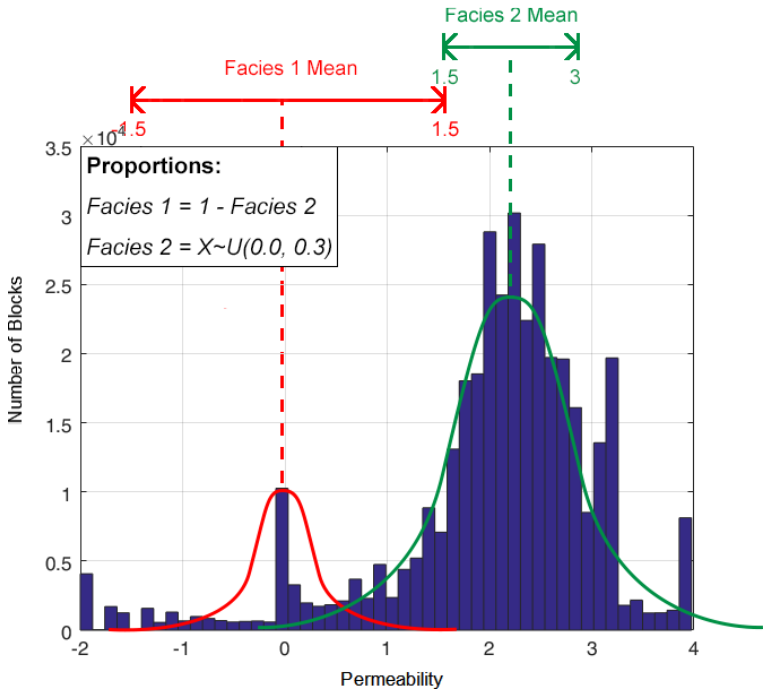


Figure 21 – Illustration of the uncertainty in the Permeability histogram.

3.3.1.3 Fault transmissibility perturbation

Fault properties are incorporated in dynamic simulation by using transmissibility multipliers. Transmissibility multipliers are a function of the static properties of the fault zone and of the grid–blocks in which they occur. By considering geological factors that influence the definition of fault zones, dynamic simulation benefits from the construction of high resolution, geologically driven, fault transmissibility models. However, fault interpretation is a task prone to uncertainty, depending on the

skills of the professional, regarding geological knowledge and sensitivity. As a result, factors like fault presence or location can be highly affected.

On the other hand, the mere identification of a fault is not enough, similar fault types and compositions may have different equivalent permeability. To tackle this issue, average values of fault permeability and thickness are predicted empirically from petrophysical and geometrical data of the reservoir model, which result in a single parameter describing fault transmissibility, assigned to each fault. To address this issue, uncertainty quantification regarding fault transmissibilities was carried out. Relevant faults were identified and marked out for parameter perturbation. Table 6 summarizes the assumed parameter distributions adopted for fault transmissibilities.

Table 6 – Uncertainty Quantification parameters for Fault transmissibility perturbation.

Property	Element	Variable Name	Distribution Type	Distribution Range
Fault Transmissibilities	Fault 2	$ftrans2	Uniform	[0.0, 1.0]
	Fault 3	$ftrans3	Uniform	[0.0, 1.0]
	Fault 5	$ftrans5	Uniform	[0.0, 1.0]
	Fault 6	$ftrans6	Uniform	[0.0, 1.0]
	Fault 7	$ftrans7	Uniform	[0.0, 1.0]
	Fault 8	$ftrans8	Uniform	[0.0, 1.0]
	Fault 9	$ftrans9	Uniform	[0.0, 1.0]

Figure 22 illustrates the assumed parameter distributions adopted for fault transmissibilities:

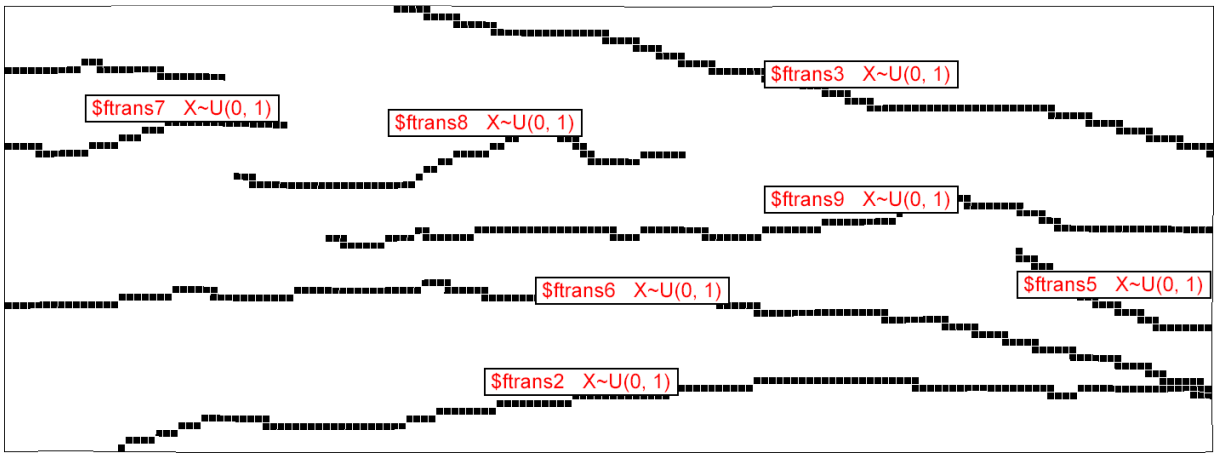


Figure 22 –Illustration of the uncertainty in fault transmissibilities.

4 Case Study – "The Watt Field"

The *Watt Field* case study (also known as IPE-1 field) was developed by the *Uncertainty Quantification Group* of the *Heriot–Watt University*. It is a semi–synthetic case study, based on the incorporation of synthetic and real data from a North Sea oil field. According to the authors, The *Watt Field* case is primarily used to describe a realistic field example seen through appraisal into the early development life stage (Arnold D. , *et al.*, 2013).

Traditional, Oil and Gas case studies (synthetic or real) are developed in order to test reservoir behavior by exploring the uncertainty on a particular set of features. Numerous case studies are used for addressing issues regarding property upscaling and distribution, uncertainty prediction, production optimization and history matching, among others. Nevertheless, few of the available reservoir benchmarks are designed primarily as a platform for the study of these issues in a multi–level and hierarchical approach.

In addition to the ever present direct and indirect measurement errors, lack of data and poor judgements based on a lack of domain knowledge or any of the factors previously pointed out, uncertainty is also generated and propagated throughout all of the reservoir modelling workflow. Interpretational choices and the choice of techniques used to model a given aspect of the reservoir, have an effect on the final outcome and quality of the reservoir description.

The goal behind the development of the *Watt Field* case study was primarily to address interpretational uncertainty, integrated throughout the reservoir modelling workflow. The interpretational choices considered were the following (Arnold D. , 2012):

- Top structure seismic interpretation – seismic interpretation of the top structure defines the enclosure size and can be ambiguous from the data and the depth conversion;
- Fault network definition – fault location, dimensions and the connectivity of the network defines the partitioning of the reservoir;
- Grid resolution – pay off between number of iterations and resolution of the model to capture the reservoir features adequately;
- Facies modelling approach – choice of algorithm to populate model with properties;
- Facies interpretations – facies estimation from well logs through cut–offs;
- Petrophysical property prediction from the available well data – how to predict permeability and porosity across the reservoir;
- Relative permeability data.

By including the possibility to consider different interpretational choices concerning reservoir geology, and the possibility of perturbing features and properties that are not included in traditional history matching purposed benchmarks, there is an opportunity to explore a multitude of geological scenarios and to subsequently address uncertainty quantification over the combination of these parameters on the production of the reservoir. Figure 23 shows the different possible interpretational choices for the *Watt Field* case study.

Model property	Description	File name	
Grid	100 m by 100 m by 5 m	G-1	Total of 81 different combinations of these properties
	100 m by 100 m by 10 m	G-2	
	200 m by 200 m by 5 m	G-3	
Top Structure	1	TS-1	
	2	TS-2	
	3	TS-3	
Fault Model	1	FM-1	
	2	FM-2	
	3	FM-3	
Facies Model (Cutoffs)	0.6	CO-1	
	0.7	CO-2	
	0.8	CO-3	
Modelling approach	Data is provided for different possible field depositional models based on different outcrop analogues. Data from these sources could be used in the construction of the geological model		
Relative permeability data	Coarse sand relperms	RP_0_1	
		RP_0_2	
		RP_0_3	
	Fine sand relperms	RP_1_1	
		RP_1_2	
		RP_1_3	
Simulation model	100 m by 100 m by 5 m	100 100 5	
	100 m by 100 m by 10 m	100 100 10	
	200 m by 200 m by 5 m	200 200 5	

Figure 23 – Possible interpretational choices for the *Watt Field* case study (Arnold D. , 2012).

4.1 Watt field description

As previously mentioned, the *Watt Field* case study was developed with the primary intent of addressing interpretational uncertainty integrated throughout the reservoir modelling workflow. For this effect, the benchmark provides a number of different model realizations, which are intended to represent typical uncertainties and interpretational choices encountered in a reservoir geomodelling workflow.

The reservoir has a $12.5 \times 2.5 km^2$ surface area (East–West direction), with a thickness of approximately 190 m, much of which is below the OWC.

Initial depth of the reservoir is located around $1555m$ subsurface with an initial reservoir pressure around the order of $2500psi$ as measured from repeat formation tests and well test data. The OWC was identified from wireline and RFT data at a constant $1635m$ subsurface.

Figure 24 shows the OWC level in the *Watt Field* case study. Different interpretational choices regarding Top Structure geometry, from seismic, are also illustrated (Yellow – TS1, Red – TS2, Blue – TS3).

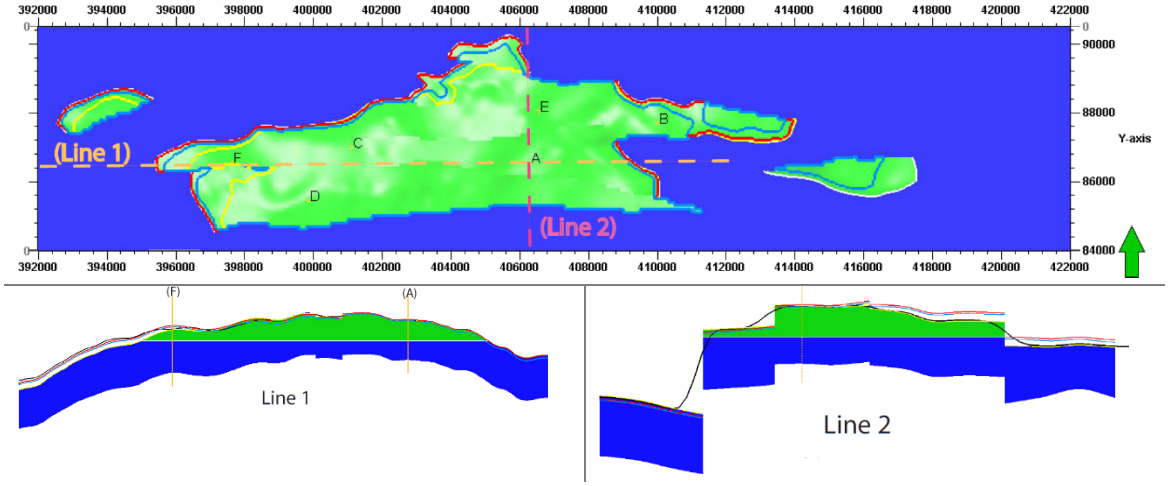


Figure 24 – Top-view of the OWC level and Top Structures for the *Watt Field* (Arnold, 2012).

The depositional environment for the field is considered to be a braided river system, with common facies types including fluvial channel sands, overbank fine sands, and background shales (Figure 25).

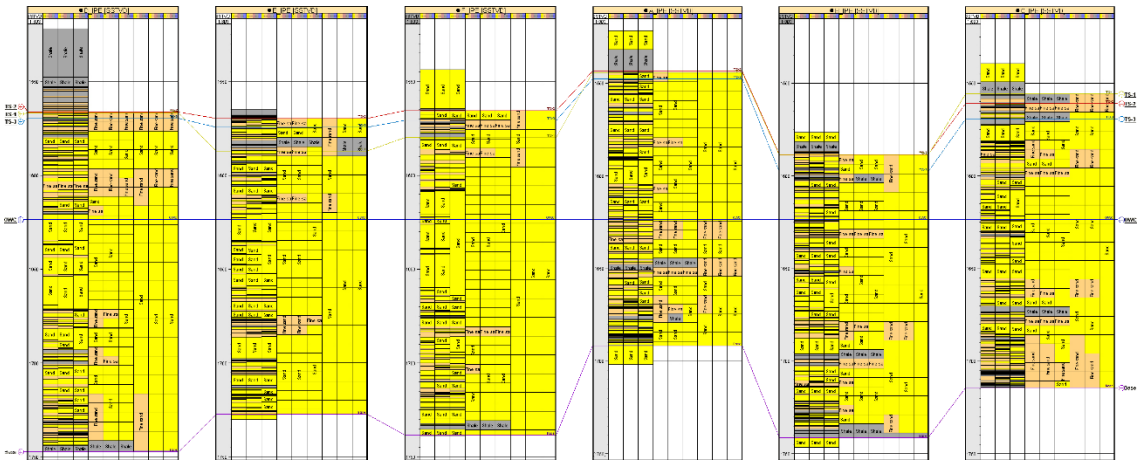


Figure 25 – Facies interpretations for the *Watt Field* (Arnold D. , 2012).

Initial appraisal of the field was made through a set of 6 wells (Wells A–F). Porosity and permeability data is available for Wells A and C, while core and neutron density porosity data is provided for all appraisal wells (Figure 26). Coarse and fine sand facies have each a set of three relative permeabilities being provided, with the intent of reproducing uncertainty in the choice of the true representative relative permeability behavior.

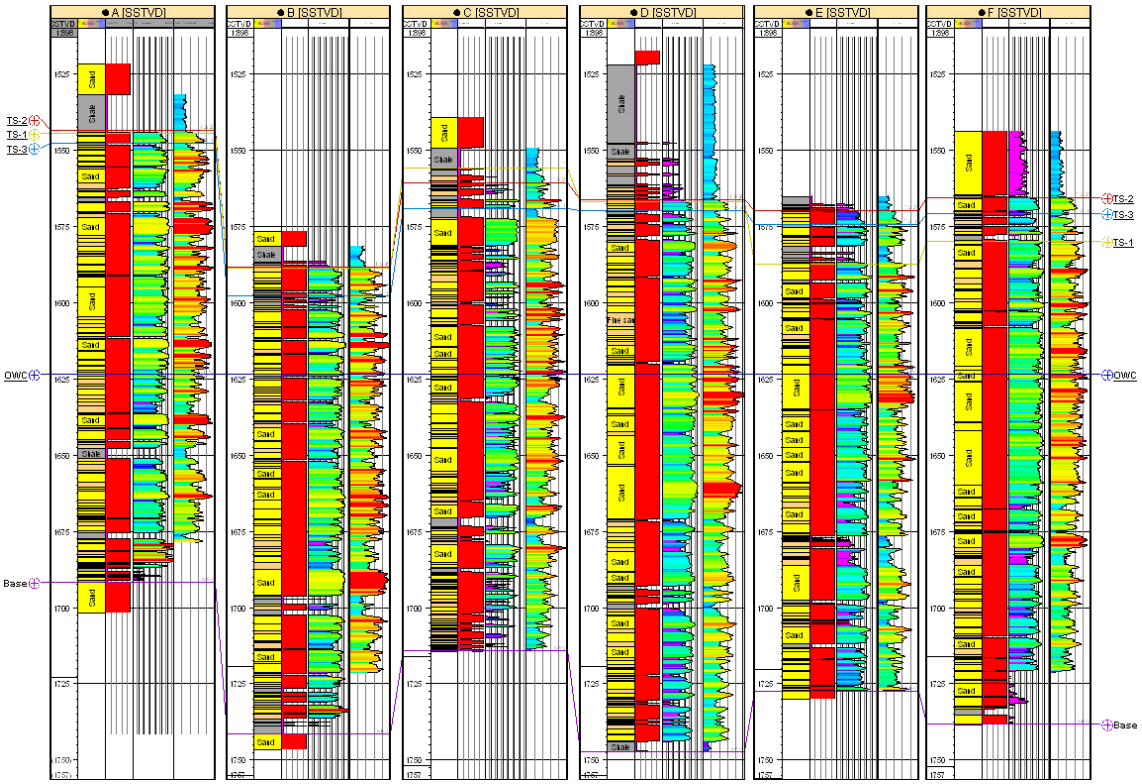
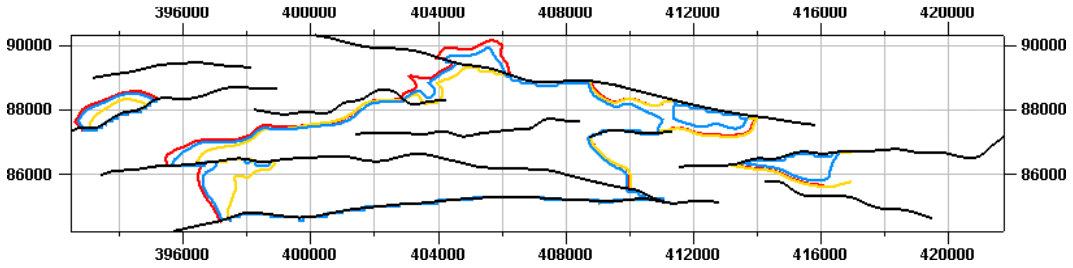


Figure 26 – Well correlation panel of Wells A–F, showing the top structure well pics for TS–1, TS–2 and TS–3 (Arnold D. , 2012).

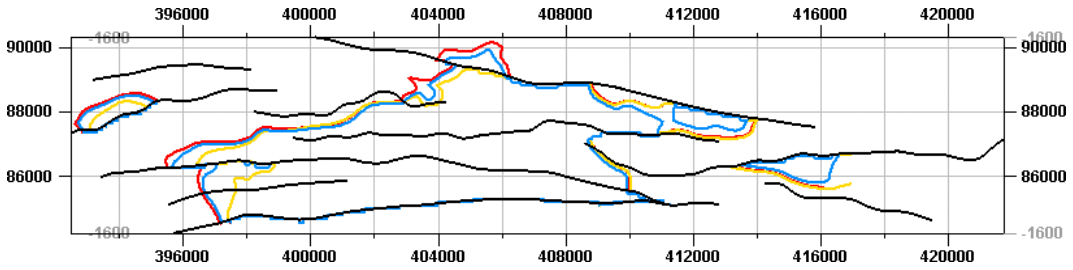
An interpretation of the top structure and major faults was picked from the seismic in the time domain and then depth converted. The consideration of uncertainty in this process resulted in the arriving of three possible top structure geometries, with three possible structural models, resulting in a combination of nine resulting top surface compositions (Figure 27).

In addition, three different model grid resolutions were provided (coarse, intermediate and fine), along with three different facies interpretations. This results in a total of 81 possible simulation models to choose from.

Fault Model 1



Fault Model 2



Fault Model 3

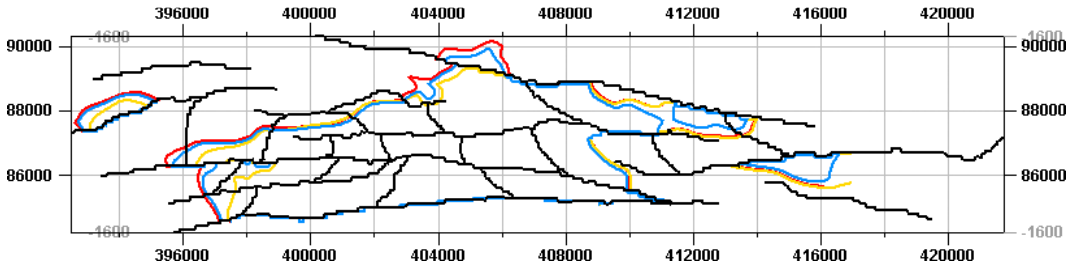


Figure 27 – Fault network definition for the Watt Field, adapted from Arnold (2012).

The reservoir is produced by means of 16 horizontal production wells located across the central parts of the reservoir. An additional 5 horizontal and 2 vertical injectors are developed around the edges of the reservoir to assist in production by providing pressure support. Horizontal multi-lateral wells were adopted to address the low relief of target zones and to maximize distance from the OWC, in order to maximize Oil production rates, while minimizing Water cuts. Table 7 and Table 8 provide a list of the operation parameters for all producer and injector wells:

Table 7 – Operation parameters for all producer wells of the Watt Field.

Well Name	Producer Type	Start of Production	Control Mode	Flow Rate (*stb/day*)	BHP target (*psia*)
WELL1	OIL, GAS	01/01/2011	Liquid Rate	8000	1000
WELL1A	OIL, GAS	01/01/2011	Liquid Rate	8000	1000
WELL1B	OIL, GAS	01/01/2011	Liquid Rate	8000	1000
WELL2	OIL, GAS	07/11/2011	Liquid Rate	6500	1000
WELL2A	OIL, GAS	07/11/2011	Liquid Rate	6500	1000
WELL3	OIL, GAS	05/06/2011	Liquid Rate	8000	1000
WELL3A	OIL, GAS	05/06/2011	Liquid Rate	8000	1000
WELL4	OIL, GAS	01/01/2011	Liquid Rate	8000	1000
WELL4A	OIL, GAS	01/01/2011	Liquid Rate	8000	1000
WELL5	OIL, GAS	10/03/2012	Liquid Rate	8000	1000
WELL5A	OIL, GAS	10/03/2012	Liquid Rate	8000	1000
WELL6	OIL, GAS	01/01/2011	Liquid Rate	8000	1000
WELL7	OIL, GAS	11/06/2012	Liquid Rate	8000	1000
WELL8	OIL, GAS	05/06/2011	Liquid Rate	8000	1000
WELL9	OIL, GAS	01/01/2011	Liquid Rate	6500	1000
WELL10	OIL, GAS	12/09/2012	Liquid Rate	7000	1000

Table 8 – Operation parameters for all injector wells of the Watt Field.

Well Name	Injector Type	Control Mode	Flow Rate *(stb/day)*	BHP target (psia)	Flow Rate (*stb/day*)	BHP target (*psia*)
INJ1	WATER	Flow Rate	12500	3500	19000	3500
INJ2	WATER	Flow Rate	15000	3500	19000	3500
INJ3	WATER	Flow Rate	15000	3500	19000	3500
INJ4	WATER	Flow Rate	15000	3500	19000	3500
INJ5	WATER	Flow Rate	15000	3500	19000	3500
INJ6	WATER	Flow Rate	15000	3500	19000	3500
B	WATER	Flow Rate	10000	3500	11000	3500

This configuration of well profiles and production schedules results in a simulation model that is representative of the production complexities seen in a real oil field development planning.

Production data for the reservoir was generated, based on a truth case scenario, for a period of 2903 days (8 years) and for all wells.

Oil, Water and Gas rates and bottom-hole pressure (BHP) data are provided and a Gaussian noise of 15% of the value was added to the data. Figure 28 exemplifies the production history for FIELD vectors (FOPR, FWPR and FGPR) and for WBHP for Well1.

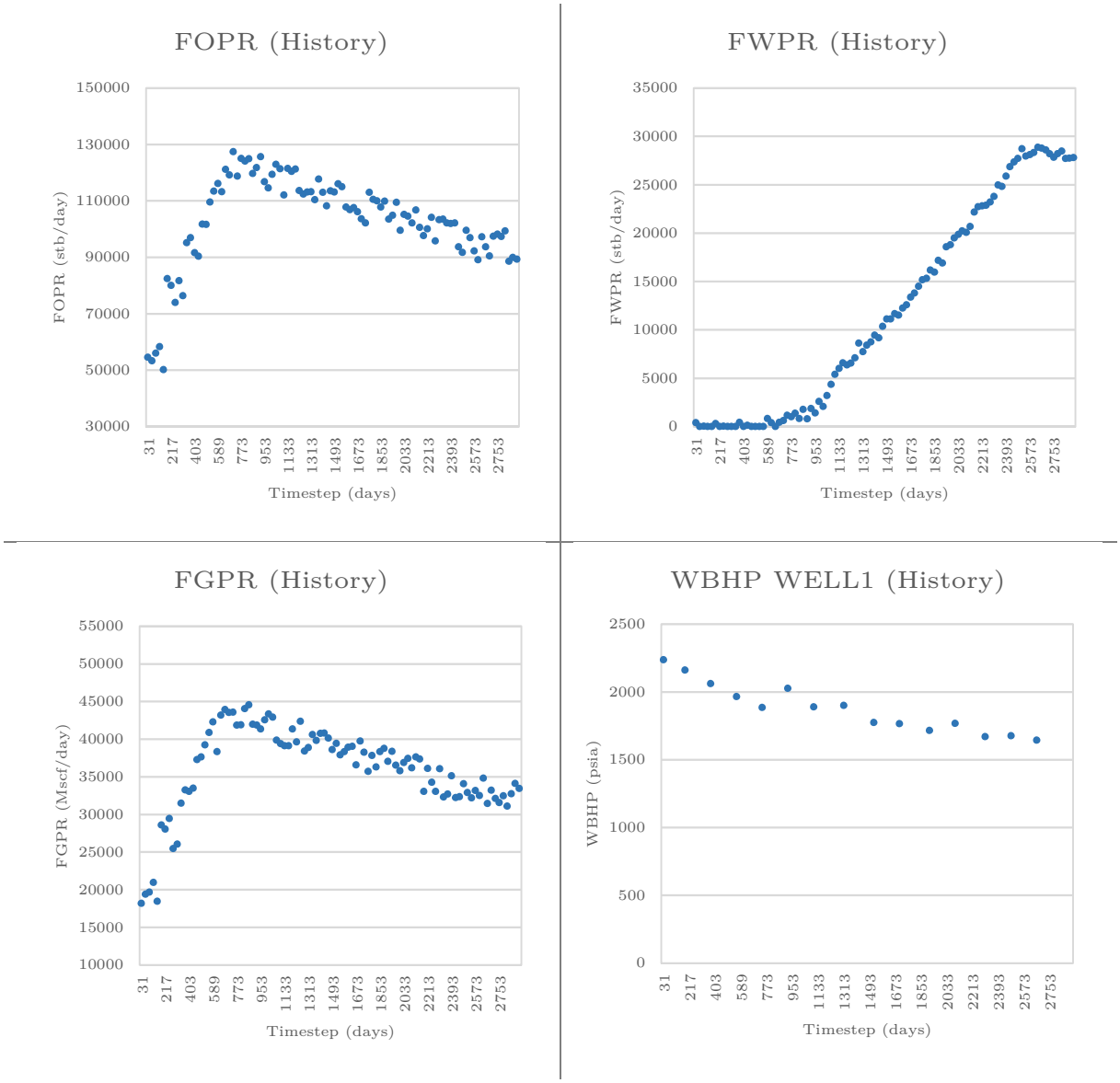


Figure 28 – Production history for FOPR, FWPR, FGPR and WBHP Well1.

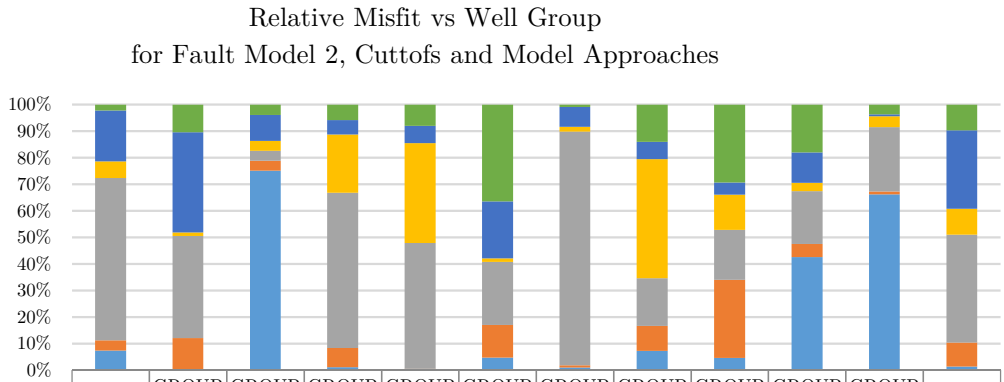
4.2 Base scenario selection

As previously described, the *Watt Field* benchmark offers a choice of 81 possible scenarios with different interpretational choices, ready for dynamic simulation. For the purpose of this work, there was a need to choose a single base scenario in order to test the proposed history matching algorithm. Due to time and computational constraints, the selection was made according to the purpose of the work and regarding several aspects of its concept. Integration of geologic consistency in the history matching workflow was one of the purposed objectives of the project. This geologic consistency, in the form of

geologic consistent regionalization for geostatistical conditional assimilation of properties, is addressed primarily by considering a form of integrating fault presence in the adopted regionalization patterns. Perturbation of fault transmissibilities is also a concern, so, parsimonious fault quantity in the selected fault model, should be an issue. As such, by looking at the available fault models for the *Watt Field*, as shown in the previous chapter, Fault Model 2 was selected, due to it offering a good balance between number of available faults and number of fault transmissibility parameters to be perturbed.

Over the course of the work, the overall observation regarding the multitude of base scenario responses allowed to conclude that relative permeabilities for coarse sand type (CO1, CO2 and CO3) often replicate observed data better than fine sand relative permeabilities. The same was observed regarding the top structure selection, with Top Structure 1 being the type of interpretation more prone to better observed data reproduction.

As such a set of test runs was made for the remaining interpretational choices for the benchmark in order to decide the property cut–offs and their respective modeling approach. Figure 29 summarizes the results obtained for the set of runs. Quality of the runs was represented by discretizing the relative misfits by FIELD and well groups, over the remaining combination of interpretation uncertainties.



	FIELD	GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5	GROUP 6	GROUP 7	GROUP 8	GROUP 9	GROUP 10	TOTAL
Misfit CO3 PIXEL	123	325136	1469	1057	77399	1925	384	391	1062	1829	1020	411795
Misfit CO3 OBJECT	1027	1167547	3558	977	61409	1137	3017	179	167	1168	202	1240387
Misfit CO2 PIXEL	334	42297	1378	3960	358666	70	751	1241	480	313	1104	410594
Misfit CO2 OBJECT	3270	1194397	1360	10526	451641	1256	35593	498	681	2018	6595	1707836
Misfit CO1 PIXEL	207	369431	1359	1299	2328	650	361	259	1061	503	289	377748
Misfit CO1 OBJECT	388	827	27503	188	1740	242	362	198	165	4311	18024	53948

Figure 29 – Result comparison of Relative Misfit vs Well group, over selected scenarios.

From the obtained results, the choice for the base scenario fell on the CO1 OBJECT model. Although other scenarios offer lower misfit values regarding field productions, the total of misfits given by the CO1 OBJECT is much lower than the other available choices. The criteria behind this selection was to select a model with good assumptions regarding local matches (well group matches), but with a high enough global mismatch for better observing the effect of the proposed history matching algorithm.

Figure 30 shows a 3D view of the porosity field for the selected base scenario, along with its respective histogram:

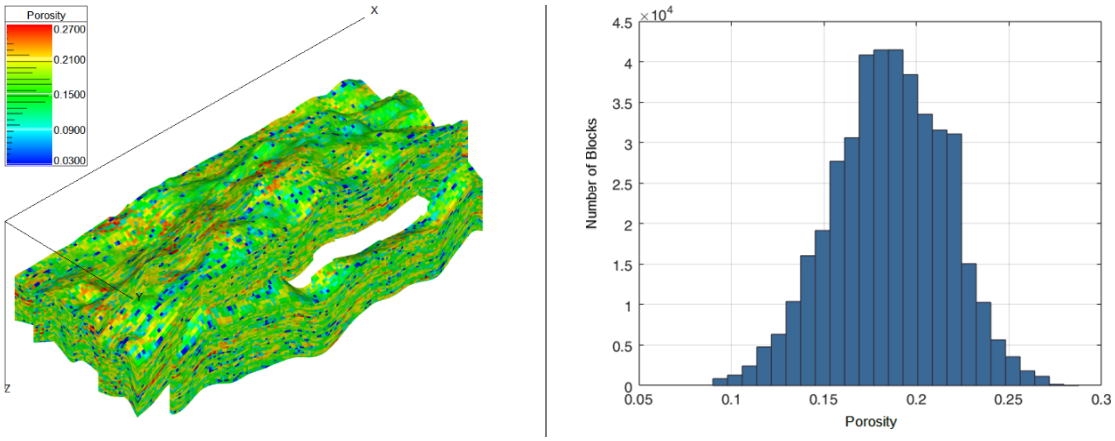


Figure 30 – Porosity field for the selected base scenario (left), Histogram for porosity (right).

Table 9 summarizes the property statistics of the model, for the variable of porosity:

Table 9 – Model statistics for porosity.

Property	Porosity
Min:	0
Max:	0.27
Mean:	0.1619
Std. dev.	0.0464
Variance:	0.0022

Figure 31 shows a 3D view of the permeability field for the selected base scenario, along with its respective histogram for $log_{10}(k)$:

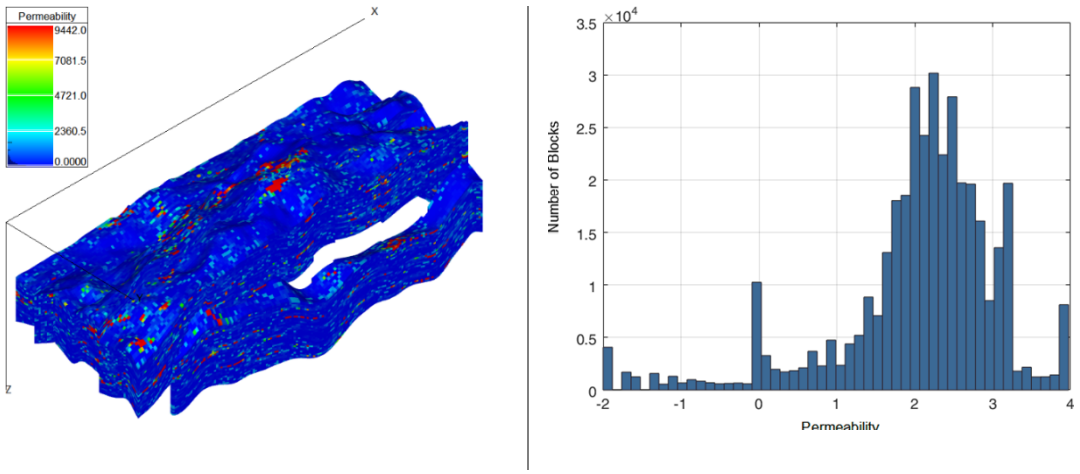


Figure 31 – Permeability field for the selected base scenario (left), Histogram for permeability of $\log_{10}(k)$ (right).

Table 10 summarizes the property statistics of the model, for the variable of porosity:

Table 10 – Model statistics for permeability.

Property	Permeability
Min:	0
Max:	9442
Mean:	537.5942
Std. dev.	1417.934
Variance:	2010535
Variance:	2010535

Figure 32 shows the saturation regions concerning the respective facies modeling approach for the selected base scenario:

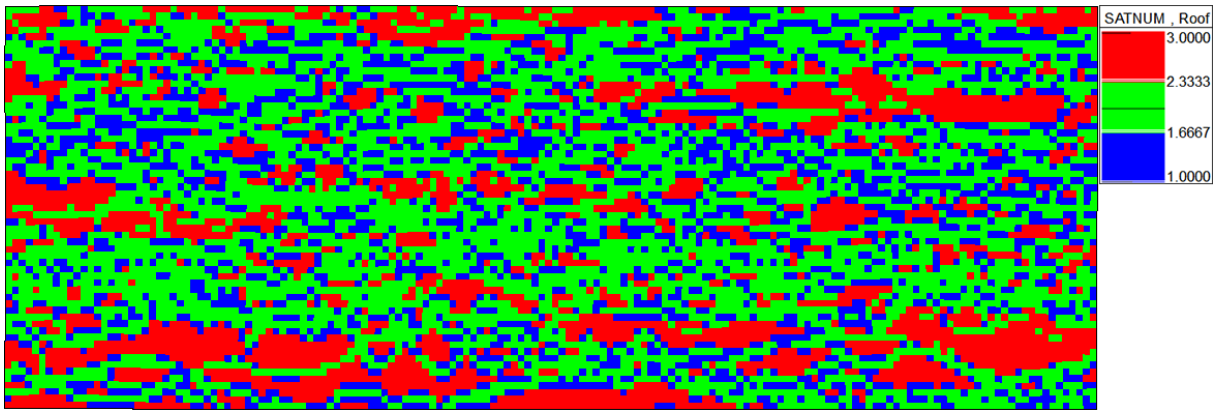


Figure 32 – Saturation regions for the selected base scenario.

5 Results and Discussion

This section introduces the results obtained by applying the methodology described in Chapter 3 and applied to the case study introduced in Chapter 4.2.

5.1 Geostatistical history matching

The first stage concerns only the conventional GHM with a regionalization model. This section compares the misfit values obtained for the different proposed regionalization methods. This part of the work consists on the inner loop of the second stage of the work.

For this stage, a total of 10 iterations, with 10 simulations each (resulting in a total of 100 realizations) is performed for every zonation method. The best–fit porosity and permeability are shown here. Plots for minimum misfit per iteration, standard deviation of misfit value per iteration, best minimum misfit along all iterations and average misfit per iteration are presented as a form of illustration of match quality and convergence obtained.

5.1.1.1 Global zonation

Figure 33 and Figure 34 illustrate the best–fit models of porosity and permeability obtained from the application of the GHM algorithm with Global zonation.

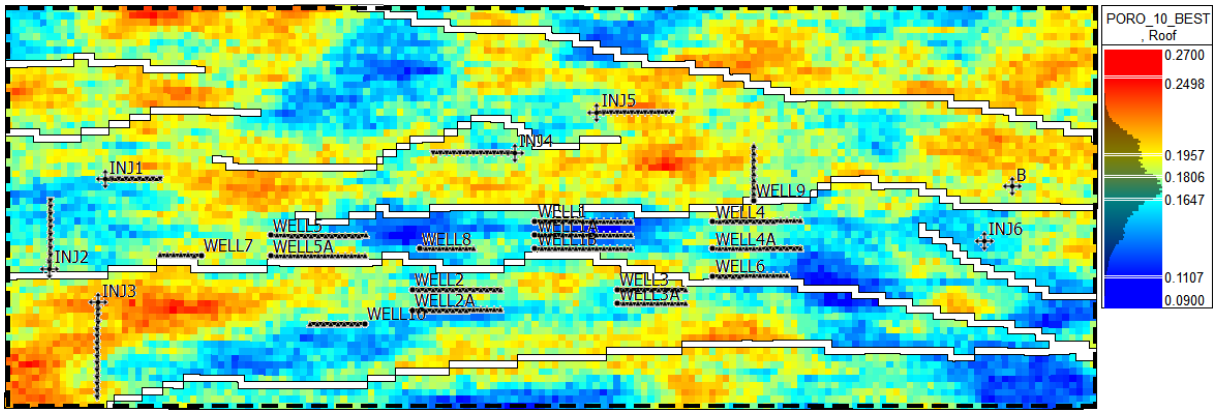


Figure 33 – Best–fit Porosity for Global zonation method.

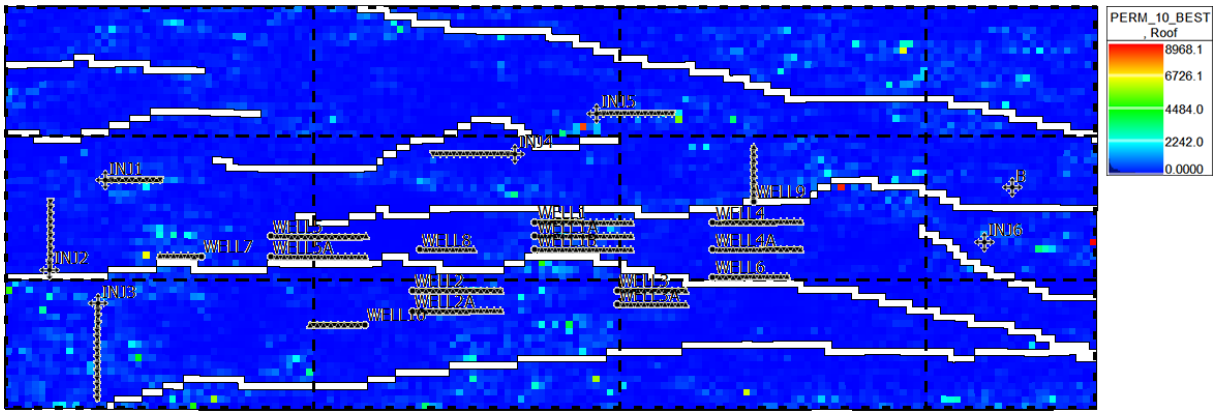


Figure 34 – Best–fit Permeability for Global zonation method.

Results, in terms of field values, for the Global zonation GHM run are show in Figure 35.

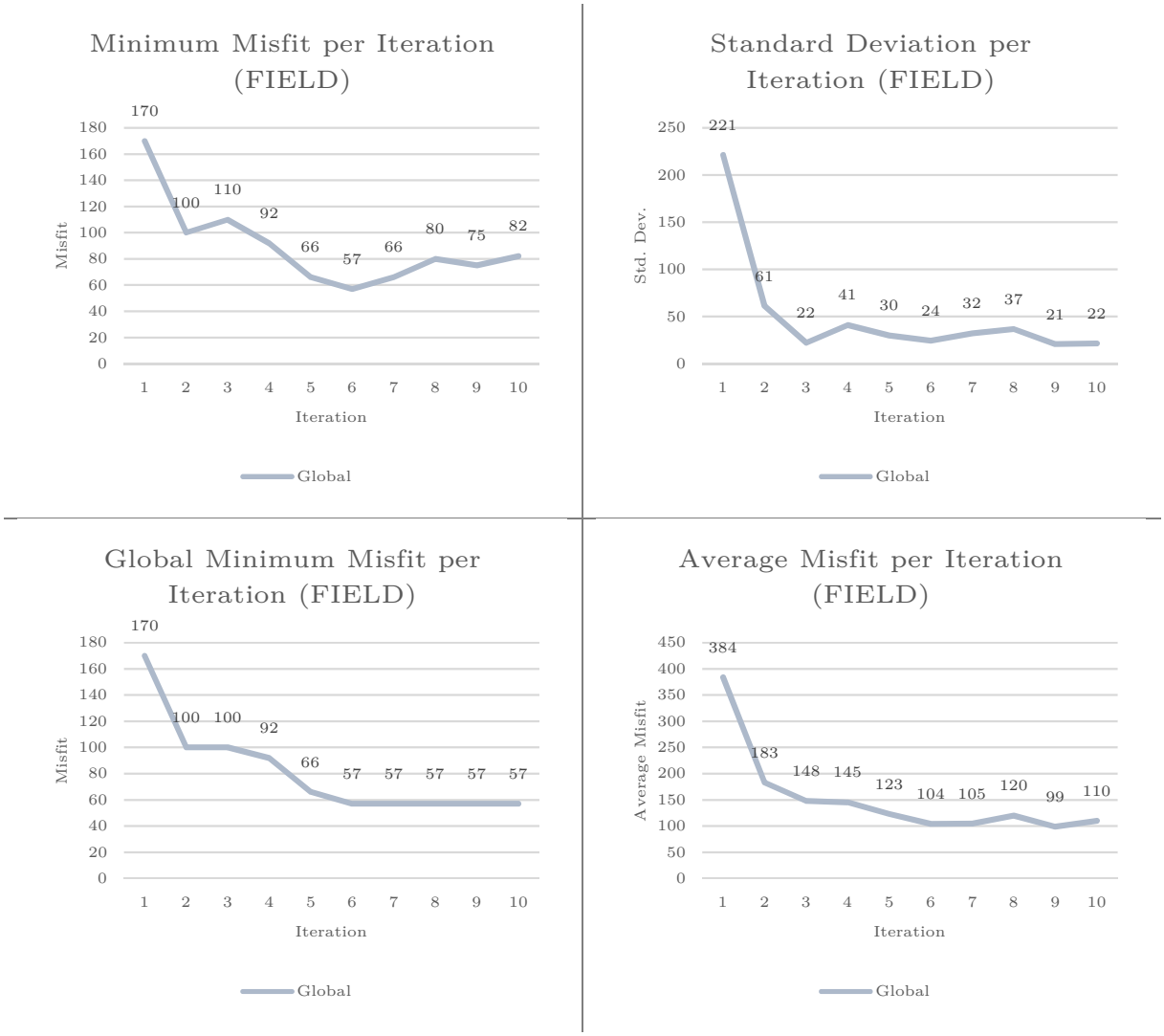


Figure 35 – Global Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right).

The general pattern for the run can be characterized by a considerable drop in the results from iteration 1 to iteration 2. After this drop a slow and steady progress of the values is observed until the end of the run. This can be explain simply by looking at the standard deviation per iteration plot, which illustrates how the runs start showing little variation after only iteration 3. Such a quick drop in standard deviation values can be explained by the type of correlation coefficient used for the global zonation method, which is a global correlation coefficient, being calculated by the FIELD vectors of FOPR and FWPR. Convergence of the run is obtained at iteration 6 with a global minimum misfit of 57 for the run.

5.1.1.2 Square zonation

Figure 36 and Figure 37, illustrate the best–fit models of porosity and permeability for the GHM algorithm with Square zonation.

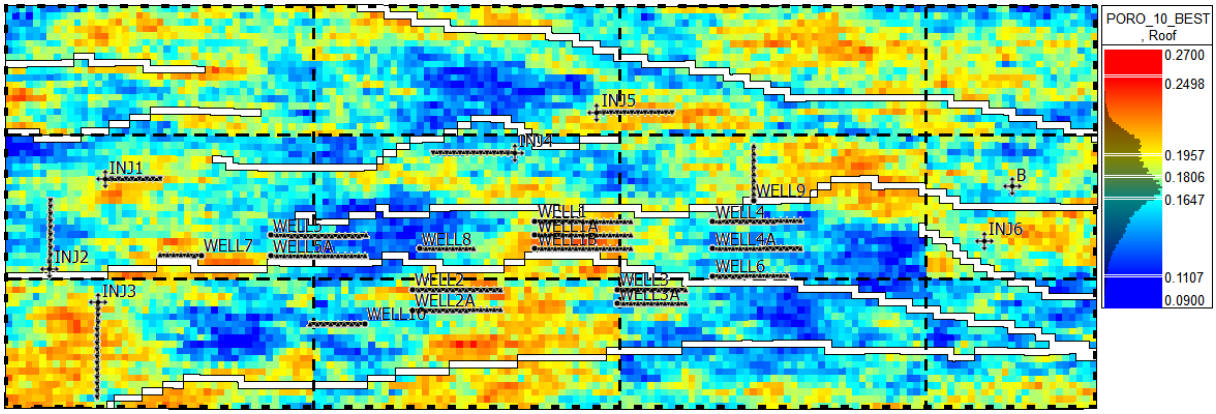


Figure 36 – Best–fit Porosity for Square zonation method.

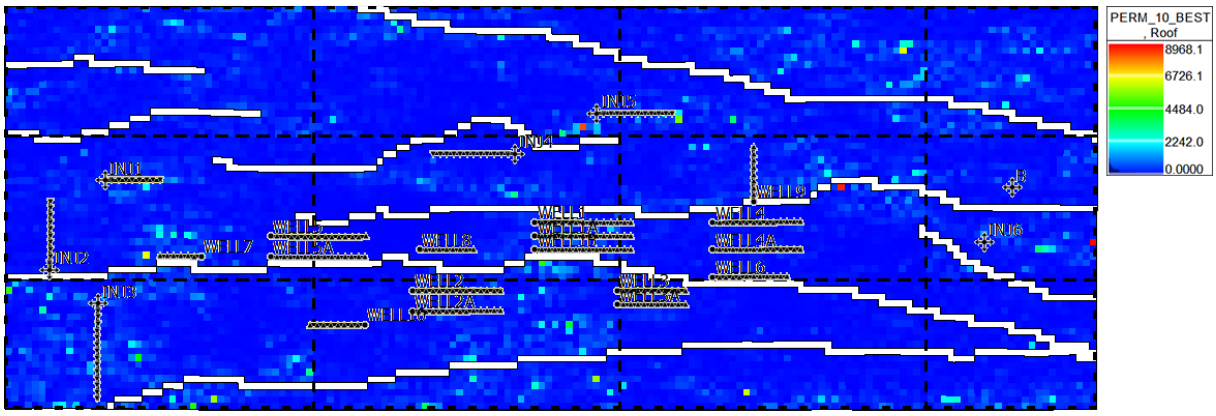


Figure 37 – Best–fit Permeability for Square zonation method.

The results, in terms of field values, for the squared zonation GHM run are show in Figure 38. Squared zonation shows an overall poor match. Although it is possible to observe a convergence occurring in terms of global misfit value, the minimum misfit value is very high (120), which given the 96 timesteps of the fluid flow simulation, means that in average, the deviation between observed and simulated data

is outside the admissible error. Also, there is significant variability in terms of minimum misfit per iteration, standard deviation and average misfit, which means that the history matching algorithm cannot stabilize and progress towards a minimum value. This shows the inability of the adopted regionalization pattern in replicating the characteristics of geological consistency required for reliable reservoir modeling and characterization.

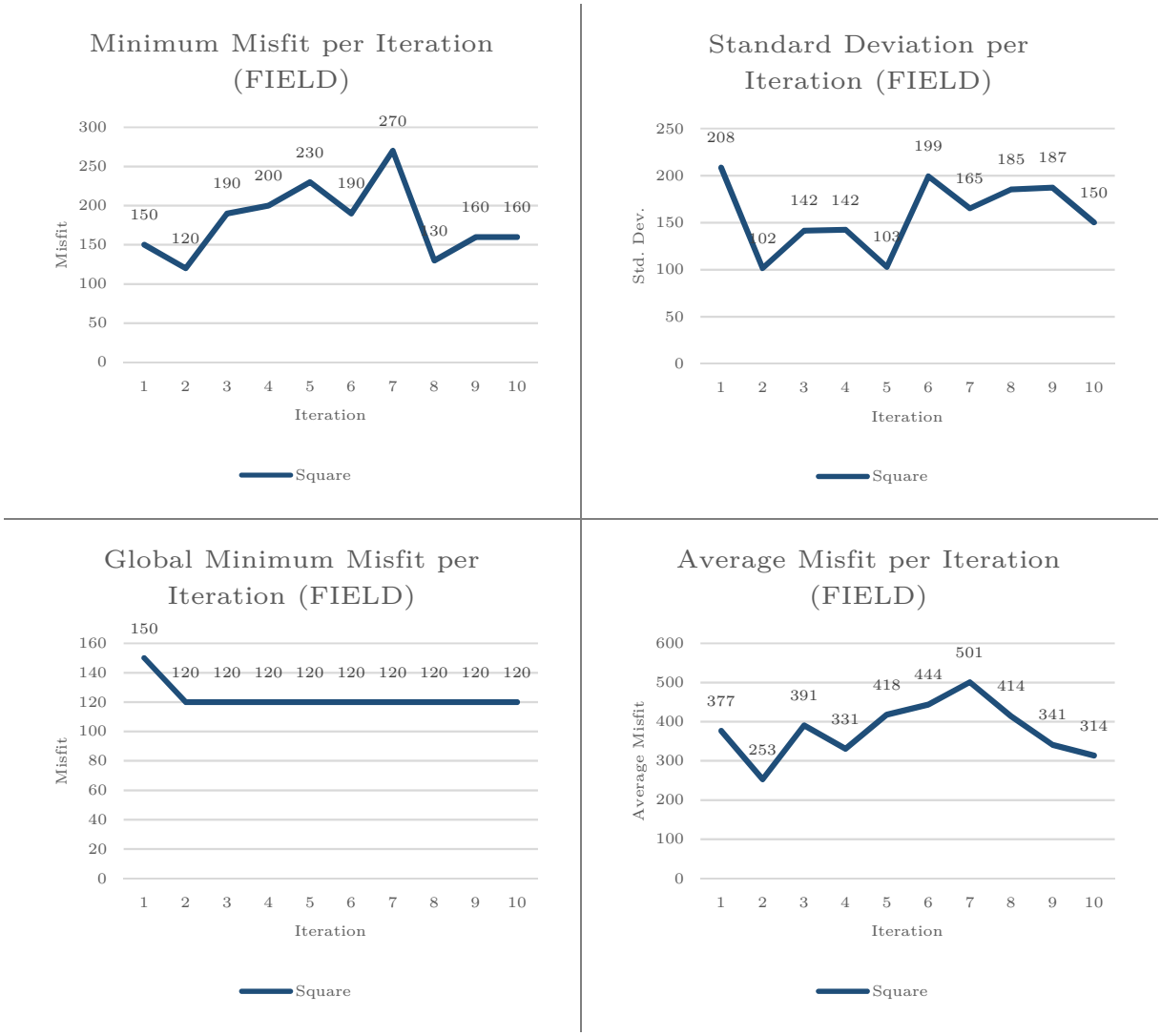


Figure 38 – Square Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right).

5.1.1.3 Circular zonation

Figure 39 and Figure 40 illustrate the best–fit models of porosity and permeability obtained from the application of the GHM algorithm with circular zonation.

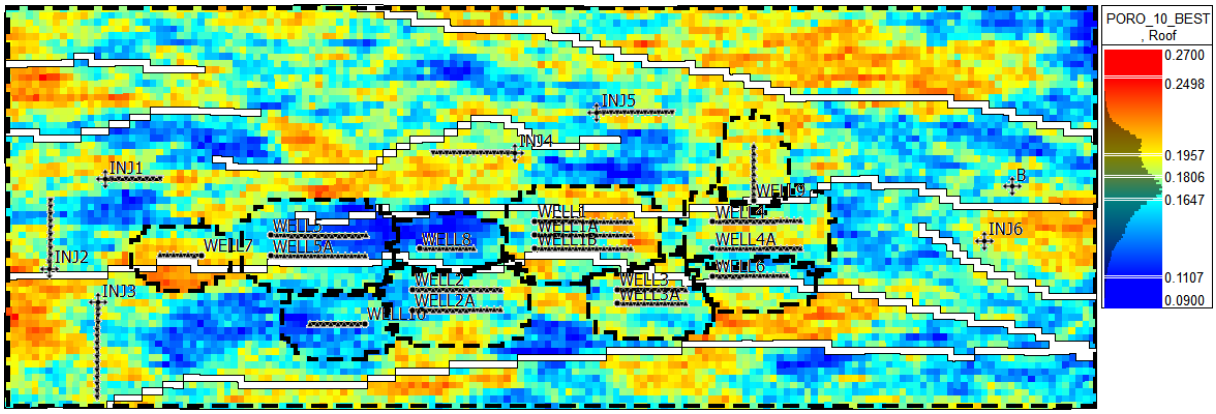


Figure 39 – Best–fit Porosity for Circular zonation method.

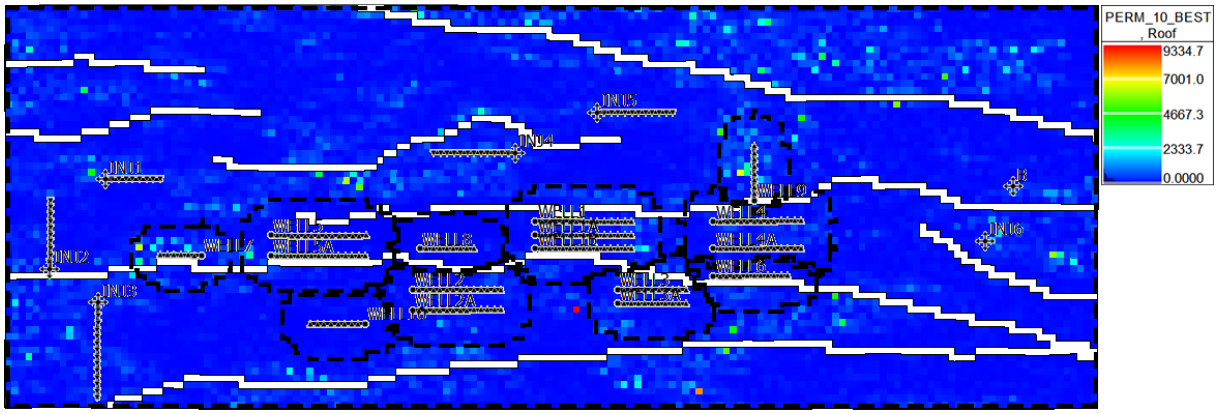


Figure 40 – Best–fit Permeability for Circular zonation method.

The results, in terms of field values, for the circular zonation GHM run are show in Figure 41.

With this regionalization model the results obtained show a good match achieving a global minimum at iteration 3. However, this result occurs very early in the run, with the next iterations having higher variability in terms of minimum misfit value. This could probably mean that the misfit value of 61, obtained at iteration 3, could be a product of chance. Also, the fact that after iteration 3 consequent runs rarely return to misfit values close to 61 could mean that the conditionally assimilated geostatistical changes that occurred on the petrophysical models, were not significant enough to be reproduced as better matches in the dynamic simulation. Standard deviation per iteration shows a strong decrease along the course of the run, with the same pattern being observed on the average misfit per iteration.

These results show that well radius of influence could be an important aspect to be considered in a zonation–based methodology. However, the high variability observed in the minimum misfits per iteration could mean that this simple regionalization model is not enough to fully characterize flow patterns and aggregate different zones of the reservoir.

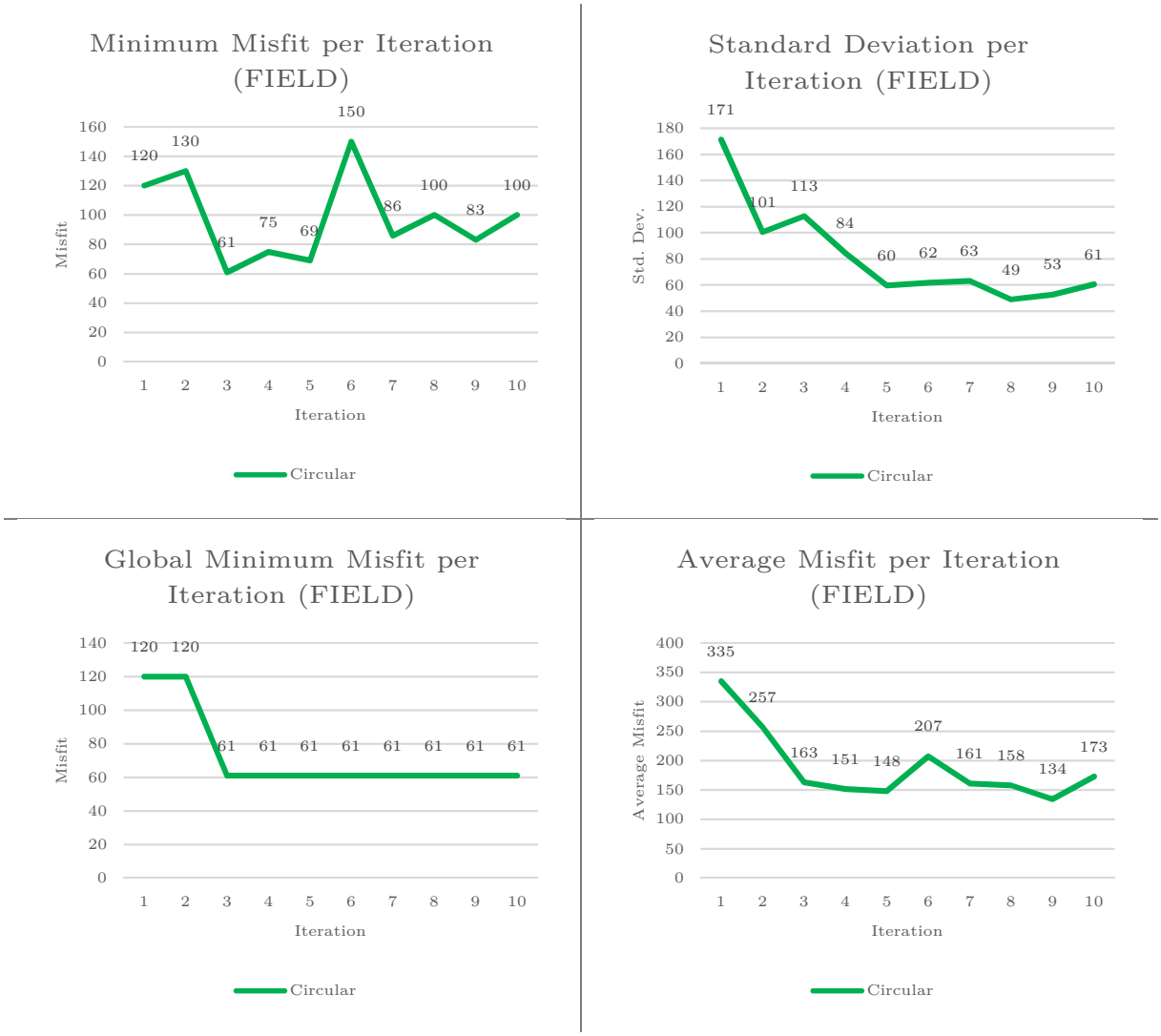


Figure 41 – Circular Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right).

5.1.1.4 Voronoi zonation

Figure 42 and Figure 43 illustrate the best–fit models of porosity and permeability obtained from the GHM algorithm with Voronoi zonation.

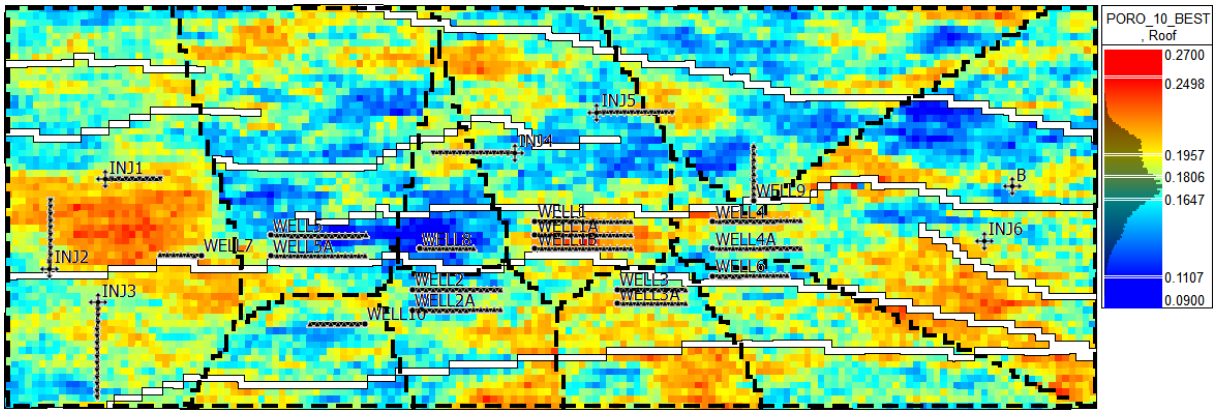


Figure 42 – Best–fit Porosity for Voronoi zonation method.

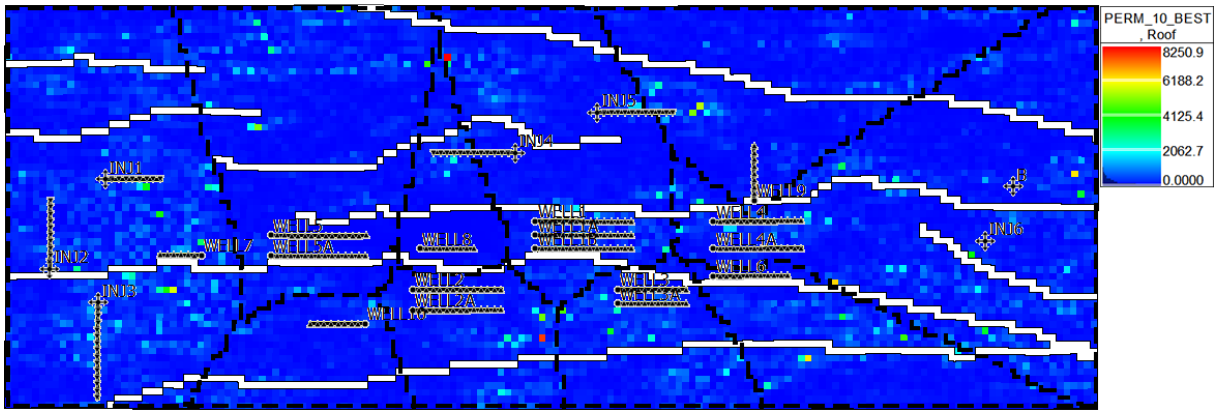


Figure 43 – Best–fit Permeability for Voronoi zonation method.

The results, in terms of field values, for the Voronoi zonation GHM run are show in Figure 44.

The Voronoi zonation method, shows a good match profile, with minimum misfits per iteration being observed with low variability for the course of the run. In terms of minimum misfits per iteration and average misfit per iteration, an abrupt decrease is observed after iteration 1, with the run progressing steadily for the following iterations in a smooth trend towards smaller misfit values. The minimum misfit value occurred at iteration 6, with consequent results in terms of all presented metrics, showing a stabilization of the results.

These results show that the Voronoi regionalization approach is capable of replicating some of the characteristics necessary to discretize the reservoir model in terms of petrophysical variables and fluid flow patterns. However, high variability of the misfit standard deviation in the first 6 iterations means

that changes occurring from iteration to iteration reproduce a wide range of results, implying that additional aspects in terms of a zonation–based methodology should be taken into consideration

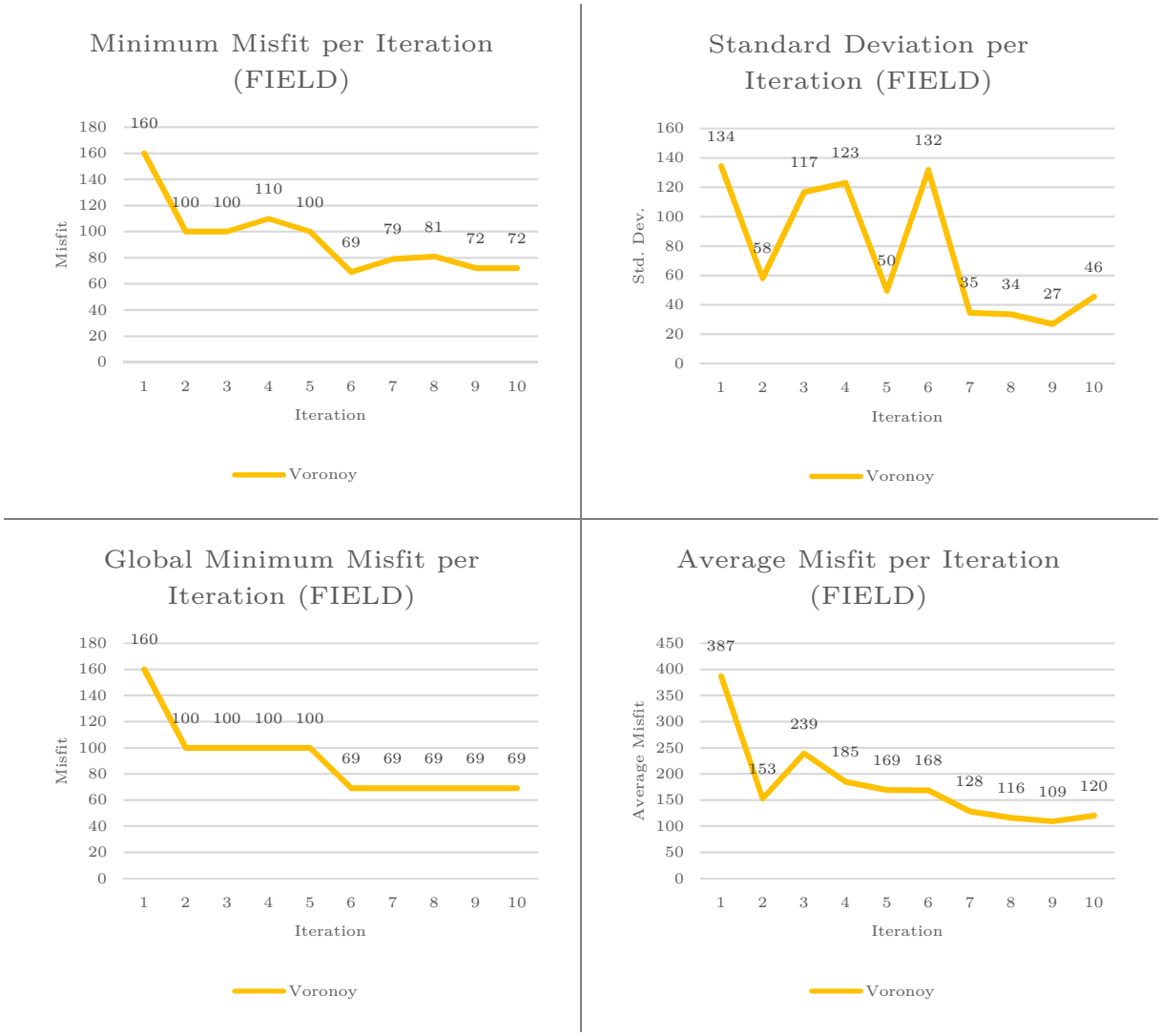


Figure 44 – Voronoi Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right).

5.1.1.5 Fault and Streamline zonation

Figure 45 and Figure 46 illustrate the best–fit models of porosity and permeability obtained from the GHM algorithm with Fault and Streamline zonation.

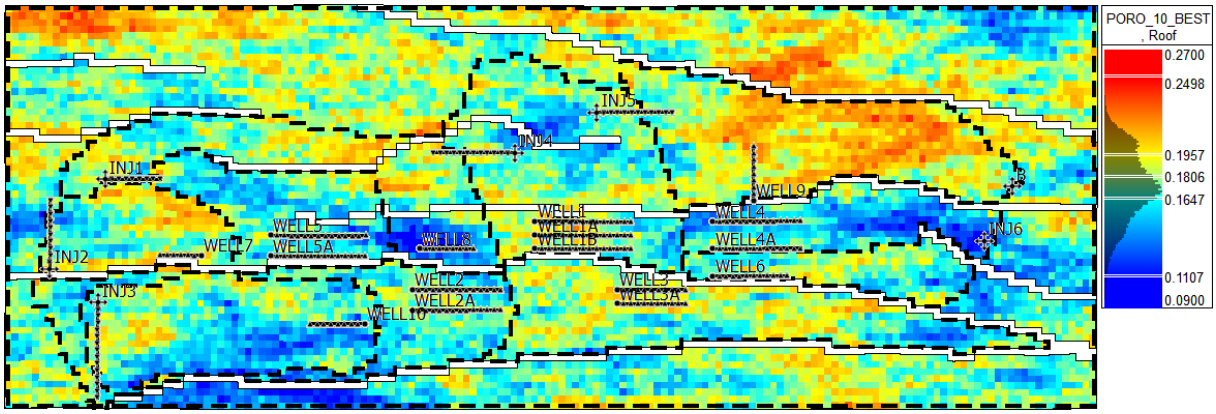


Figure 45 – Best–fit Porosity for Fault and Streamline zonation method.

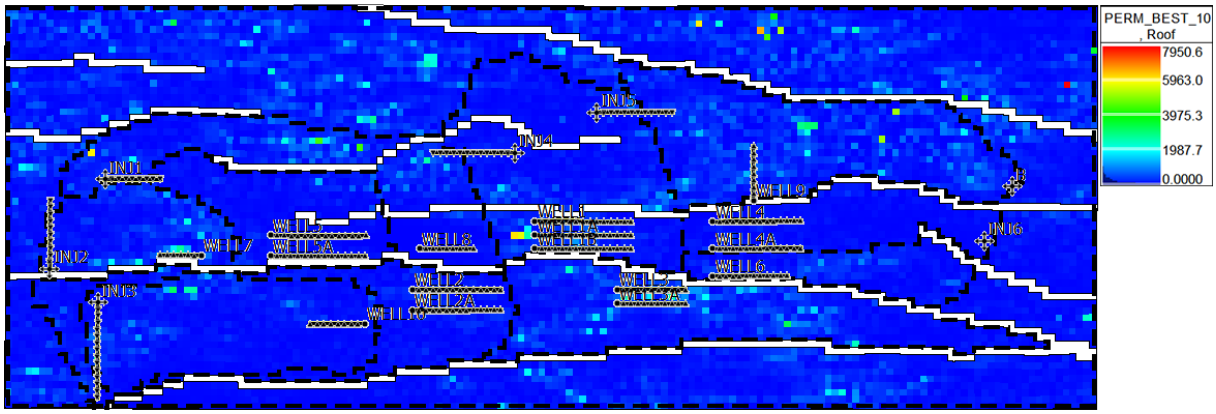


Figure 46 –Best–fit Permeability for Fault and Streamline zonation method.

The final zonation–based methodology being tested is the Fault and Streamline zonation methodology. The assumptions and details behind this approach are fully explored in Chapter 3 and can be summarized as an attempt to fully integrate various concepts related to geological consistency: well area of influence, fault zonation, production streamline volumetric and overlapping and enveloping of production areas along depth. This is an attempt to arrive at a more efficient reservoir discretization to be used by geostatistical history match through soft conditioning, which will serve as the inner core to the Adaptive Stochastic Sampling algorithm to be used.

The results, in terms of field values, for the Fault and Streamline zonation GHM run are show in Figure 47.

Fault and Streamline zonation resulted in a good match of the observed reservoir production data. The minimum misfit value was achieved at iteration 6, with a value of 56. The minimum misfit per iteration and misfit standard deviation show a global downward trend through the course of the run, meaning that the geostatistical assimilation of properties through soft–data conditioning results in progressively lower misfit values being obtained over the course of the run.

These results enhance the idea behind the assumptions made for this zonation approach, which integrates different aspects of both petrophysical and dynamic nature. An aspect of geologic consistency is kept considering fault presence and applying a fault based zonation, while attempting to characterize fluid flow consistency by considering production streamlines. The results showed an overall better performance of the fault and streamline zonation methodology that can be obtained.

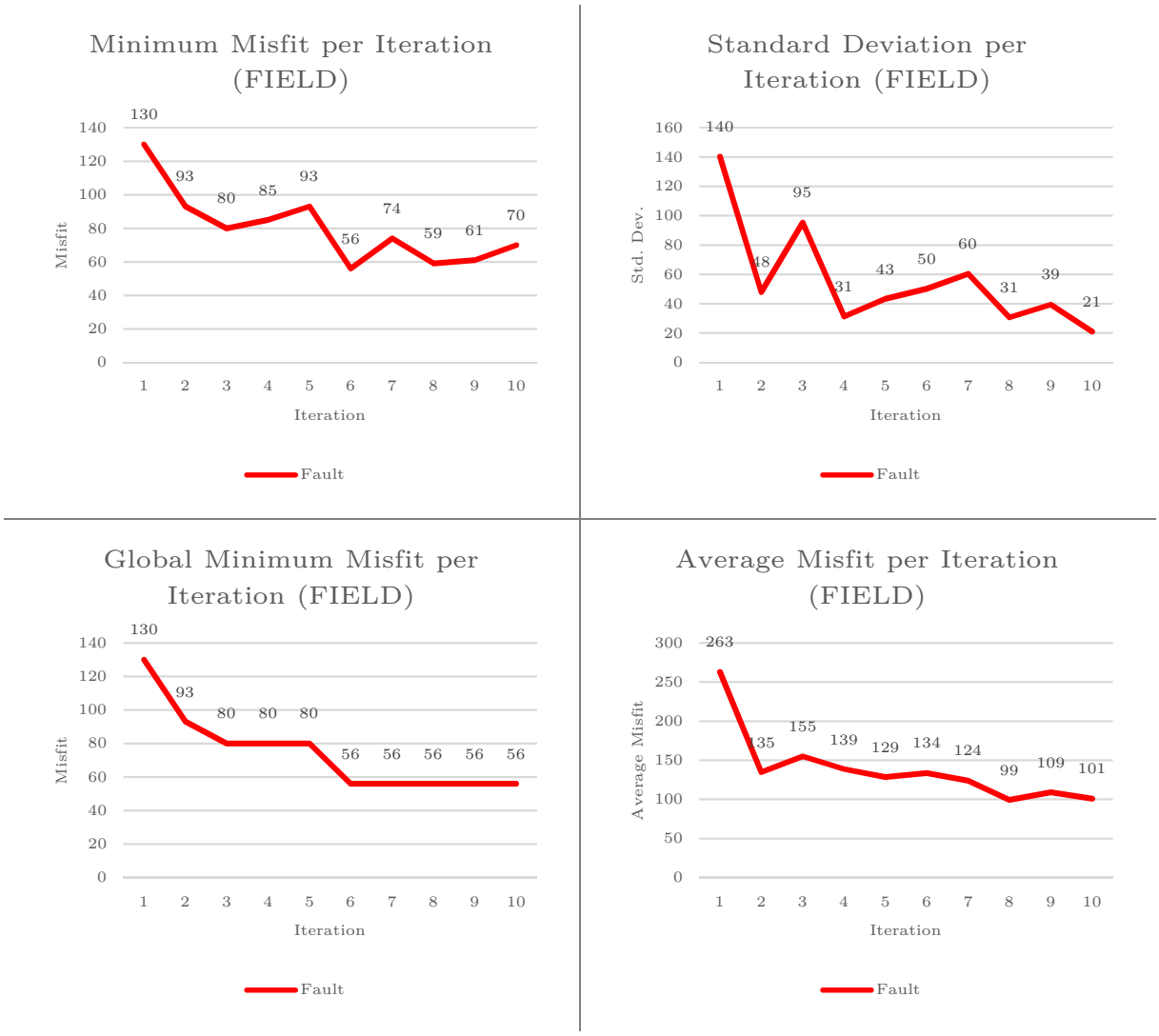


Figure 47 – Fault and Streamline Zonation – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right).

Figure 48 shows the comparison of results between the different adopted zonation methods.

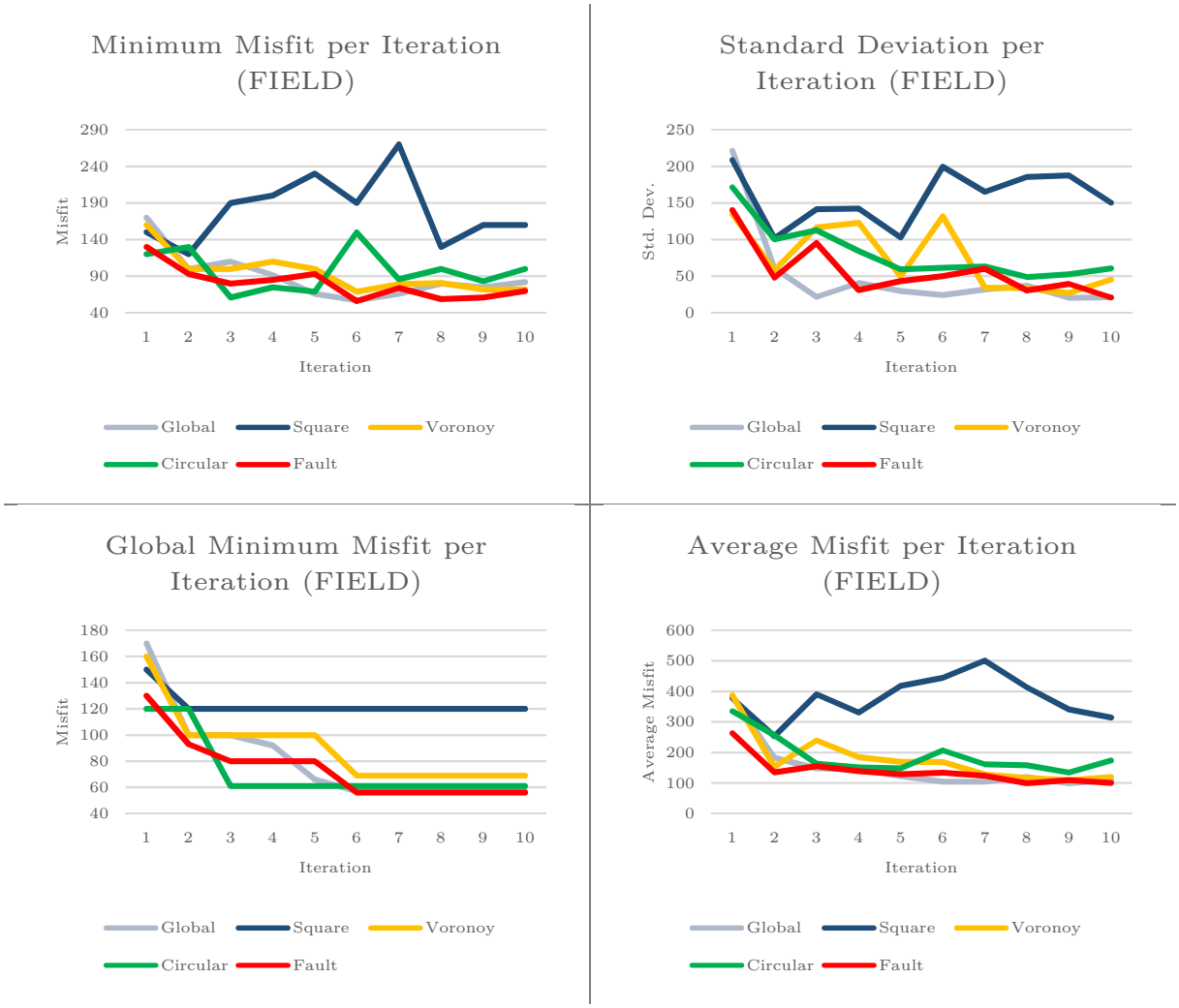


Figure 48 – Comparison between different zonation methods – Minimum Misfit per Iteration (top left), Standard Deviation per Iteration (top right), Global Minimum Misfit per Iteration (bottom left) and Average Misfit per Iteration (bottom right).

A a brief summary table with the key comparison indicators is presented in Table 11.

Table 11 – Summary table of comparison between zonations.

	Misfit			**Std. Dev**		**Avg. Misfit**	
	Minimum	Convergence (Iteration)	Behaviour	Stabilization (Iteration)	Behaviour	Stabilization (Iteration)	Behaviour
Global	57	6	Good	3	Good	3	Good
Square	120	2	Bad	No convergence	Bad	No convergence	Bad
Circular	61	3	Average	5	Good	3	Good
Voronoi	69	6	Average	7	Bad	7	Good
Fault and Streamline	56	6	Good	4	Good	2	Good

In order to provide a wider analisys over the behaviour of the different zonation methods, Figure 49 shows the misfit evolution per iteration. It is possible to observe the full misfit convergence of the different methods and the overal better performance of the Fault and Streamline zonation methodology.

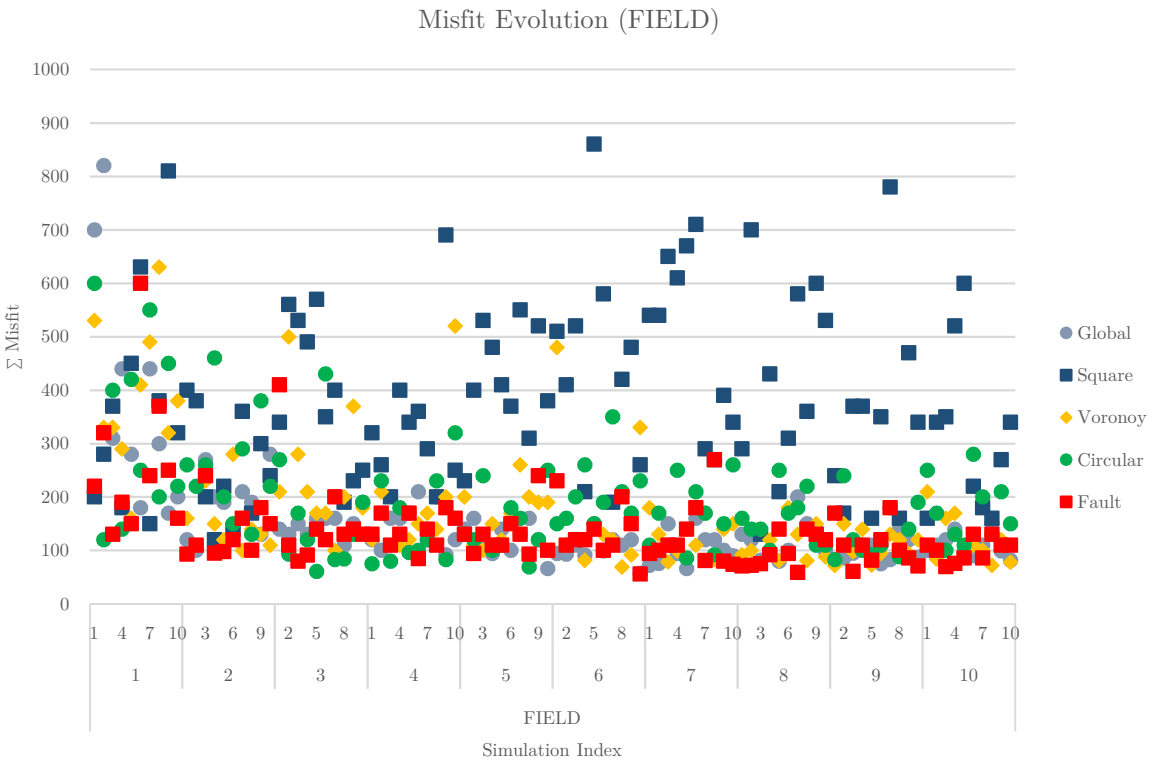


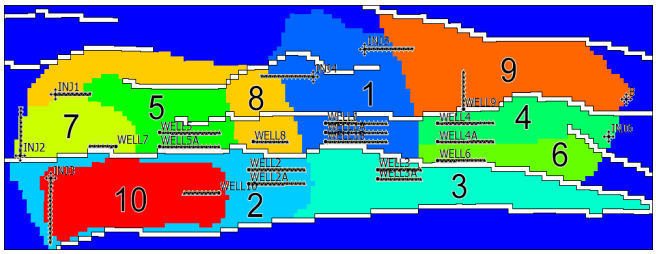
Figure 49 – Misfit evolution comparison between all zonation methods.

For exemplification of the local matches, fault and streamline zonation was compared to the best geometrical zonation method tested (circular) in terms of FIELD misfit. For this, maps of average porosity over the whole depth of the reservoir are displayed for the best matches of both regionalization methods. For better reading, the zonation scheme adopted for the fault and streamline method is also displayed on top. In either zonation cases, the zone numeration is attributed to be the same as the numeration of the well group (or single well) that each specific zone includes. This way, misfits per well group (or zone) can be compared with the help of the misfits per zone values to the right of Figure 50.

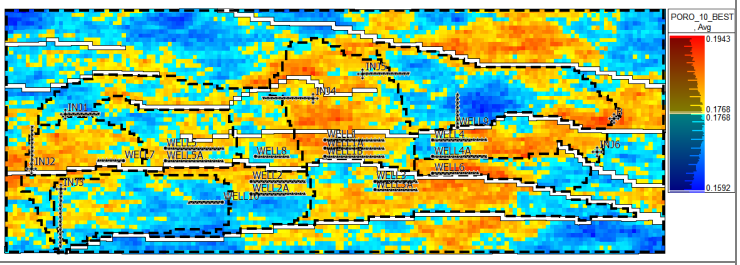
By looking at Figure 50, it is possible to observe that the fault and streamline zonation method provided better total and FIELD misfits. The method also provided better results for zones 2, 3, 4, 7 and 10, with zones 1, 5, 6, 8 and 9 producing worse misfits, although by a much smaller difference. Zone 2 produced the biggest misfit difference of the two methods. By looking at Figure 50 it is possible to observe that the circular area of influence is not able to fully characterize the much wider low porosity zone that the fault and streamline method used, in order to achieve a lower mismatch. In the case of zone 9, it is possible to observe that the misfit difference is very small between both methods, which is also depicted on the respective zone, with both methods achieving a low porosity average assigned to the vicinity of

WELL9. Zone 7 presented the same misfit value for both methods, although porosity averages are not equally distributed between both methods. This can be explained by the proximity of WELL7 to the two injectors adjacent to this zone, and the delimiting of this zone by a sealing fault, to the south. The pressure support given by the injectors and the confining of the well by the fault presence, could be outweighing the importance the perturbation of properties in this area, thus resulting in equally good matches.

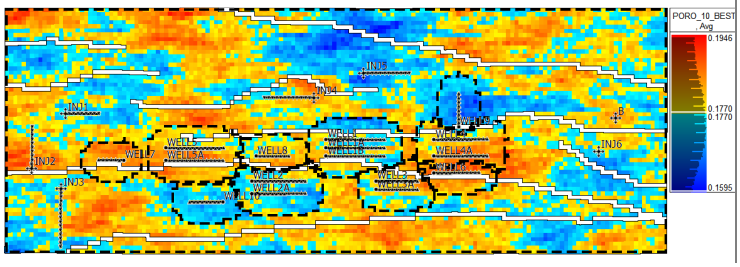
Fault and Streamline zones



Best average Porosity (Fault and Streamline)



Best average Porosity (Circular)



	Fault and Streamline	Circular
Field	56	61
Group 1	850	793
Group 2	32346	175793
Group 3	571	2282
Group 4	1883	1896
Group 5	283	46
Well 6	1988	769
Well 7	61	61
Well 8	662	381
Well 9	284	271
Well 10	297	344
Total	**39280**	**182696**

Figure 50 – Comparison of Best average Porosity between fault and streamline and circular zonation. (left) Total misfit values per well group and field (right).

5.2 Coupling with Adaptive Stochastic Sampling

5.2.1 Production match analysis

As the outcome of the comparison between different zonation methods made on Chapter 5.1, fault and streamline based zonation was selected to be integrated on the GHM loop, for coupling with Adaptive Stochastic Sampling. The inner zonation–based history matching loop was set to run 5 simulations at every iteration. For the Adaptive Stochastic Sampling outer loop, PSO was selected to optimize the

perturbed parameters over a total of 223 iterations. A total of 1115 fluid flow simulations were, therefore, ran.

A total of 17 parameters were perturbed, with 4 variogram parameters for porosity and permeability (horizontal and vertical ranges of both properties), 6 histogram parameters (facies means and proportions of both properties) and 7 fault transmissibilities (for faults 2, 3, 5, 6, 7, 8 and 9). Table 12 shows a summary of all the selected prior distributions used for the selected parameters to match (see Chapter 3.3.1 for more details).

Table 12 – Selection of prior distribution ranges.

Property	Parameter	Variable Name	Distribution Type	Prior Range
Fault Transm.	All Fault Transm.	*$ftrans (2,3,5,6,7,8,9)*	Uniform	[0.0, 1.0]
Permeability	Horizontal. Range	*$permrange1*	Discrete Uniform	[40, 80]
	Vertical. Range	*$permrange2*	Discrete Uniform	[10, 30]
	Facies 2 Proportion	*$perm_fac_2*	Uniform	[0.7, 1.0]
	Facies 1 Mean	*$perm_mean_1*	Uniform	[–1.5, 1.5]
	Facies 2 Mean	*$perm_mean_2*	Uniform	[1.5, 3.0]
Porosity	Horizontal. Range	*$pororange1*	Discrete Uniform	[40, 80]
	Vertical. Range	*$pororange2*	Discrete Uniform	[10, 30]
	Facies 2 Proportion	*$poro_fac_2*	Uniform	[0.0, 0.3]
	Facies 1 Mean	*$poro_mean_1*	Uniform	[0.15, 0.19]
	Facies 2 Mean	*$poro_mean_2*	Uniform	[0.19, 0.23]

Figure 51 shows the overall misfit progress of the run, highlighting the 5 best iterations of the run, along with their misfits. As a way of visualizing the match improvement obtained from the implementation of the algorithm, the misfit calculated for the base scenario dynamic evaluation is also shown.

From the observation of the misfit evolutions for FIELD data (Figure 51), it is possible to interpret that convergence occurred at around iteration 100, when 4 of the 5 lowest misfits were obtained. The lowest misfit belongs to iteration 91, with a value of 32.01. Notice the significant improvement not only in comparison with the base scenario misfit value (388), but also in comparison with the lowest misfit observed (56), for the testing of the fault–streamline methodology made on Chapter 5.1.1.5

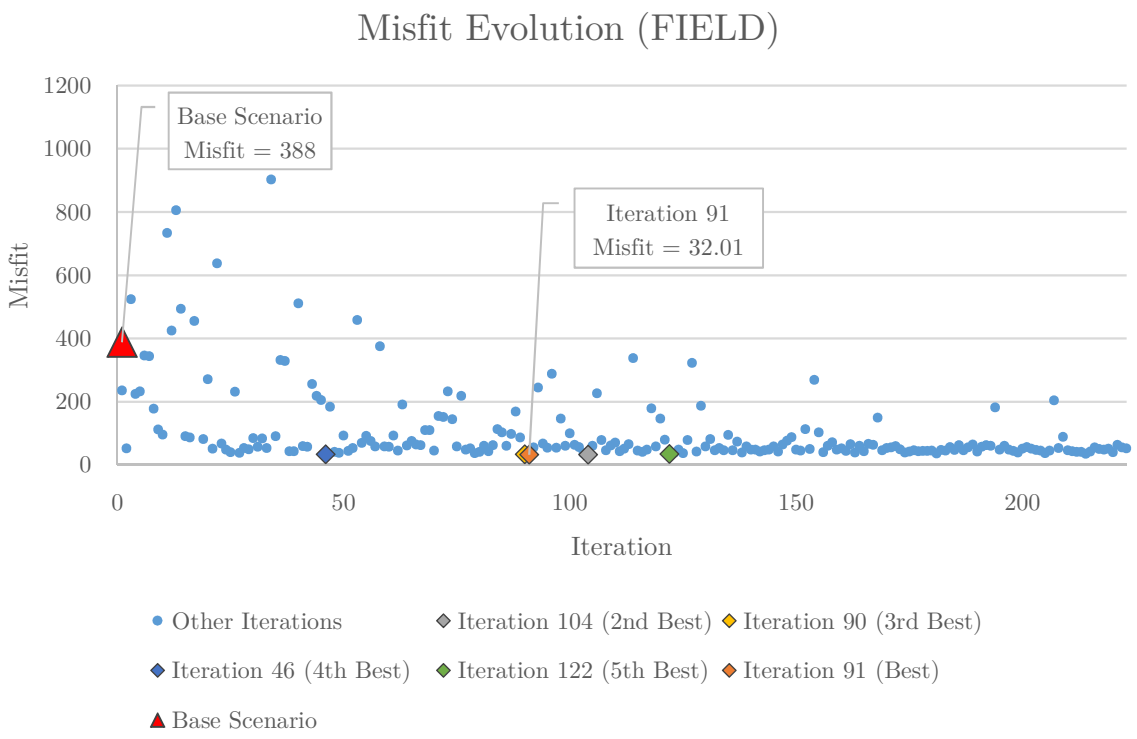


Figure 51 – Misfit Evolution of the zonation–based GHM algorithm, coupled with Adaptive Stochastic Sampling.

Figure 52 shows the fluid flow response for the best 5 iterations of the run in terms of FIELD values.

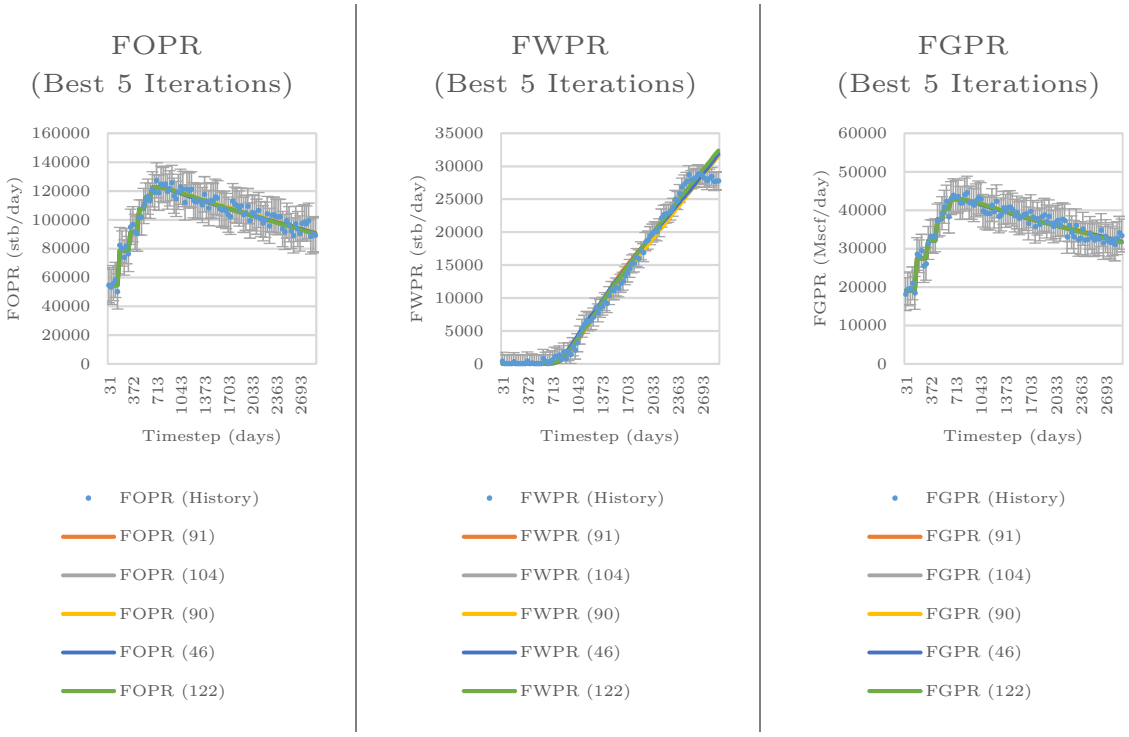


Figure 52 – Fluid flow response for FOPR (left), FWPR (middle) and FGPR (right) for GHM coupled with Adaptive Stochastic Sampling.

Figure 53 is an overview of the misfit evolutions obtained for every zone of the reservoir.

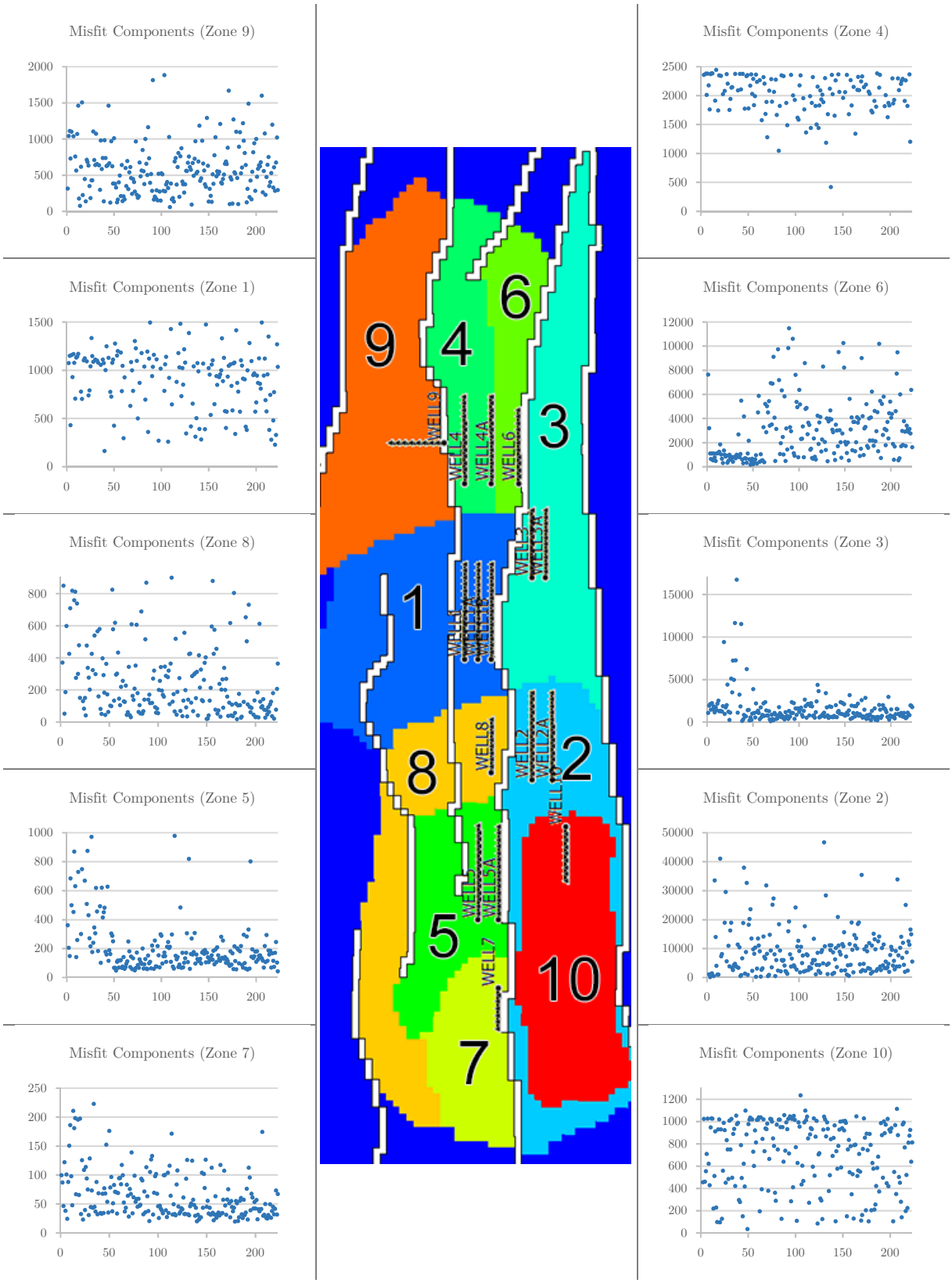


Figure 53 – Misfit evolution over different zones (x axis – Iteration, y-axis – Misfit).

By looking at Figure 53, an overview of the misfit evolutions obtained for every zone of the reservoir is obtained. Its interpretation shows that zones 5, 7 and 8 have a noticeable convergence pattern along the iterations. On the other hand, some zones show erratic or even divergent behavior. This can be explained by taking into account the fact that some zones have more than one well associated and also the fact that there is excessive water production occurring in the reservoir. Not only can water production shift from one well to another within the same well group, but it also can shift from one well group, or zone, to another, depending on changes being made to the model by the history matching algorithm.

In fact, by looking at the misfit values for Zone 2, 5 and 7 (Figure 53), it is possible to observe that misfit values start dropping at around iteration 30. Such an abrupt change can only be attributed to petrophysical changes occurring in the model due to the GHM inner loop. The consequence was the shifting of water production to a different zone. In fact, the spikes interpreted for Zone 3 between iterations 30 to 50, correspond to the shift in water production to this zone. When misfit values for Zone 3 decrease at around iteration 50, a sudden increase in misfit is observed in Zone 6. In short, the described pattern of misfit evolution illustrates the shifting of water production, occurring in early iterations, from Zone 2, 5 and 7 to Zone 3 and then to Zone 6, with the final pattern of having zone 6 produce the excessive water in the reservoir being considered as preferential in terms of achieving lower global misfit values throughout the remainder of the run.

Figure 53 also allows interpreting that some of the zones do not show a pattern of convergence, meaning that conditional assimilation of properties and the adopted parameter perturbation from Adaptive Stochastic Sampling is missing something to allow the match on these zones. Zones like zone 1, 4 and 10 show no convergence, regardless of the perturbation. This can be explained by taking into account the fact that zone 1 is composed of a group of 3 wells (WELL1,1A and 1B) and zone 4 is composed of 2 wells (WELL 4 and 4A) with two additional wells to the north and south (WELL9 and WELL6). The observed non convergent patterns for these two zones are explained by the compensation of the mentioned wells composing the groups, in terms of production. For zone 10, the reason for the non-convergence is due to the fact that this zone is entrapped by the production area for zone 2. Zone 2 is characterized by excessive production of water, having the highest misfits of the whole reservoir. Because of this, Zone 10 is affected in its mismatch convergence.

Nevertheless, regardless of the poor misfit values and convergence for some of the zones, global misfit values for FIELD data show that these zones were not representative enough in terms of production when compared to the global production of the reservoir. In fact, by looking at the five best production responses (Figure 52), it is possible to observe that, more than matching FIELD production for the reservoir, they are practically identical not only in shape but also in terms of misfit value. By considering the scope of 223 iterations, and considering that the 5 best matches found occurred approximately in the first half of the run (5^{th} best was at iteration 122), and that they show identical response behavior, it is safe to assume that the model was exhausted in terms of further improving matches with the type of perturbation applied.

Figure 54 shows an example of the matches obtained for the 5 best iterations of the run, in terms of Oil and Water Production rates for WELL5 and WELL7. This figure shows the quality of the matches with production curves for the 5 best iterations, presenting a very close profile, with flow rates being matched throughout the total production time. See Appendix A for other production curves.

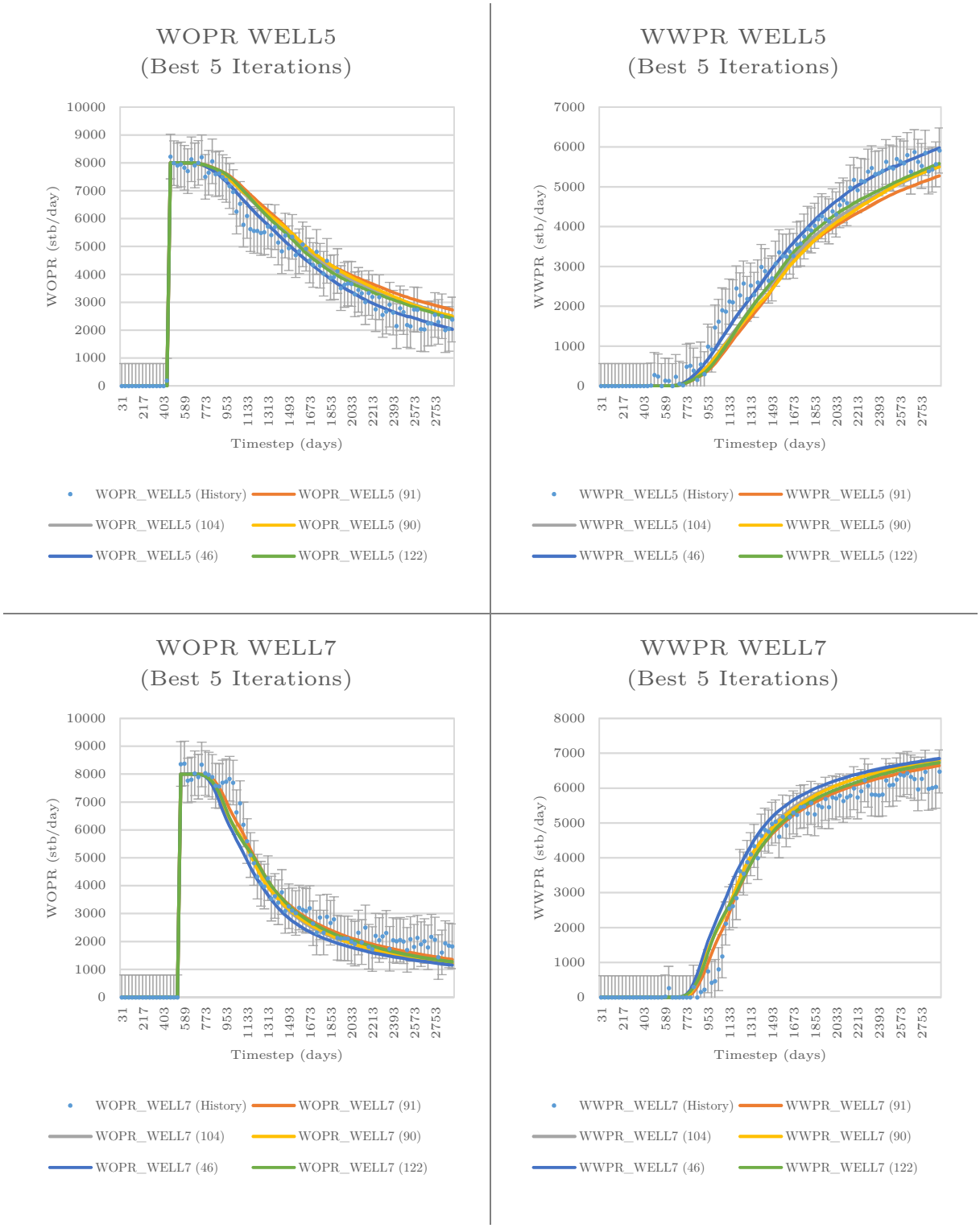


Figure 54 – Oil and Water production for WELL5 (top left, top right) and WELL7 (bottom left, bottom right) for the 5 best iterations.

For match quality illustration, a comparison is made between misfits obtained from the base scenario selected and the best iteration from the coupling algorithm on Figure 55. As previously observed, by looking at Figure 51, for FIELD values (FOPR, FWPR and FGPR), the base scenario returned a cumulative misfit of 388, while the best coupling iteration (iteration 91, simulation 5) achieved a misfit of 32.

For individual well matches, Figure 55 displays the sum of misfits for WOPR, WWPR and WBHP of each individual well. The picture shows the variation of misfits along the different areas of the reservoir, by plotting the summation of all well misfits in the same zone. This way, there is the possibility to see which zones matched better, compared to others and were the problematic zones are located.

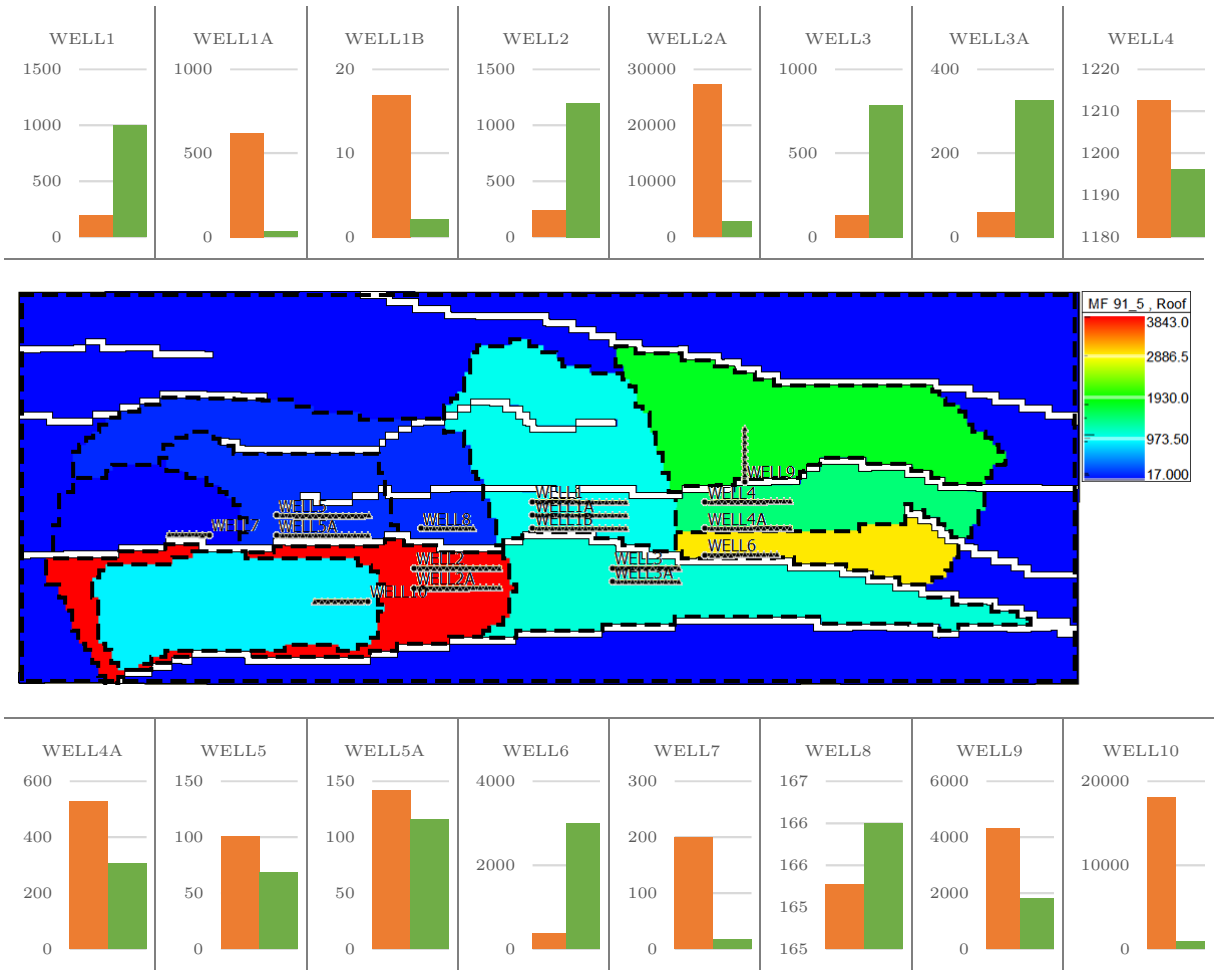


Figure 55 – Comparison of match quality by well between Base Scenario (Orange) and GHM coupling (Green), (y axis – Sum of misfits for WOPR, WWPR AND WBHP).

Figure 55 makes it possible to visualize the improvements made on the production matches of specific wells. Improvements were observed on 10 of the 16 wells (namely, WELL1A, 1B, 2A, 4, 4A, 5, 5A, 7, 9 and 10), while the remaining wells showed worse matches. Nevertheless, the cumulative improvement of several wells was reproduced by a tenfold reduction of FIELD matches (388 misfit value from base scenario to 32 misfit value from GHM coupled with Adaptive Stochastic Sampling).

Well groups 1 and 2, are cases where one well in a group achieves a higher misfit, but is compensated by the other wells in the same group, in terms of production mismatch. Well group 3 and 6 show an increase in misfit due to excessive water production (see analysis made on misfit evolution over different zones: Figure 53). Nevertheless, higher misfits for WELL3 and WELL6 weren't reproduced on higher global misfits.

5.2.2 Parameter perturbation analysis

For the particular case of fault transmissibilities, Figure 56 presents the parameter value vs iteration pattern being observed during the course of the run.

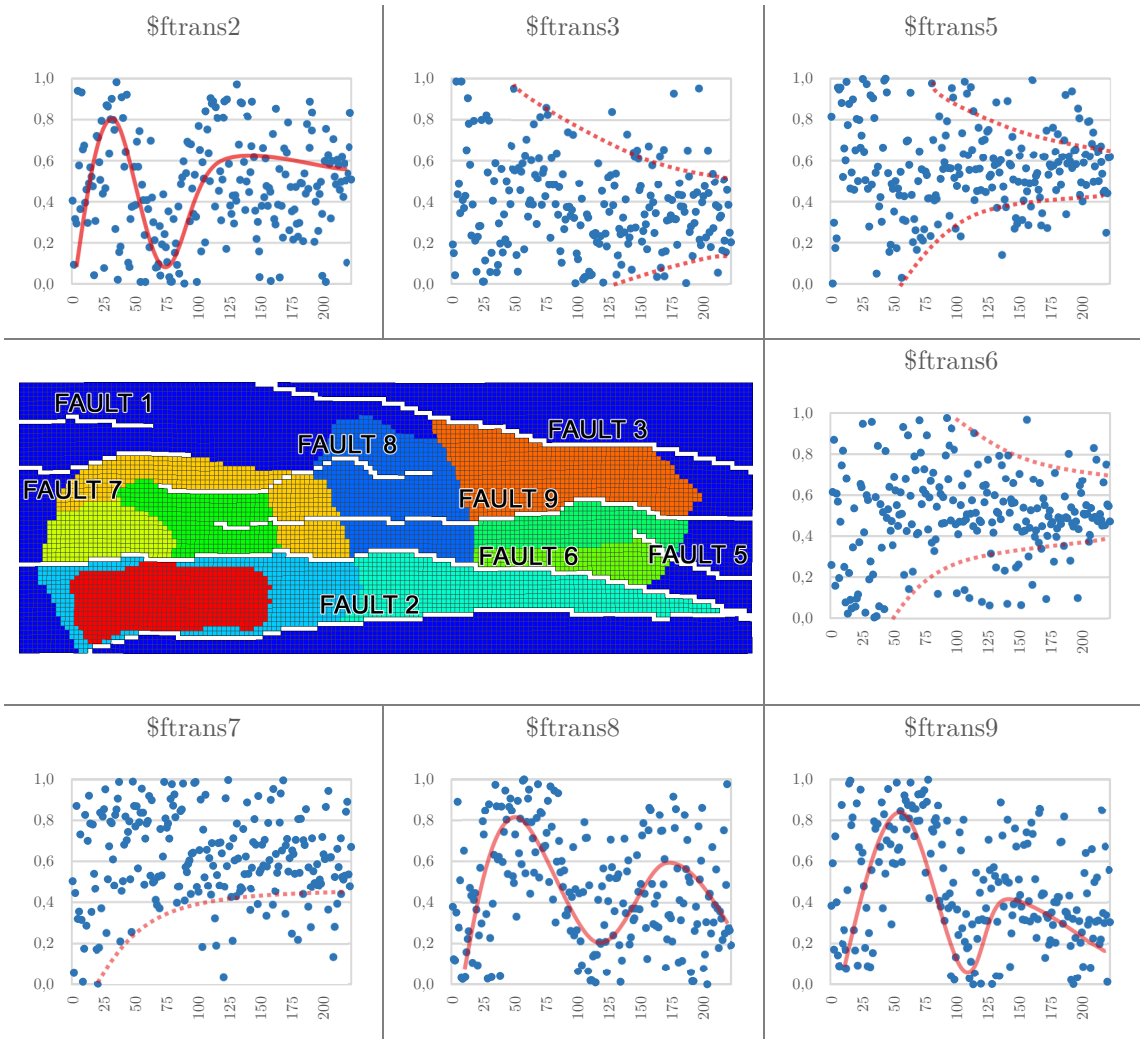


Figure 56 – Fault transmissibility parameter value (y axis) vs Iteration (x axis).

Figure 56 shows a lot of variability in the sampling of the parameter space. On some of the parameters, it is even possible to observe a *wavy–like* pattern of parameter space search (e.g. $ftrans2, $ftrans8 and $ftrans9). This can be explained by changes occurring in the static model whenever a best misfit value is being obtained along the course of the run. These changes correspond to a significant deviation from

previous iterations, in terms of dynamic response. This may sometimes force the PSO sampling algorithm to search the parameter space for a combination of parameter values that will go into a different direction from what was previously being observed. This could also mean that for some of the faults, their transmissibilities are not significant enough in terms of their influence regarding dynamic response of the model.

Figure 57 shows the Parameter vs Misfit plots for the perturbation of fault transmissibilities. It shows a big dispersion of values along the lower values of misfit. As previously explained, this is a result from the type of search of the parameter space that occurred during the run. However, it is possible to interpret some of the concentrations of most frequent parameter values for lower misfit iterations, namely on faults 3, 5, 6, with remaining fault transmissibility values being too sparse to assign a specific interval with a degree of confidence.

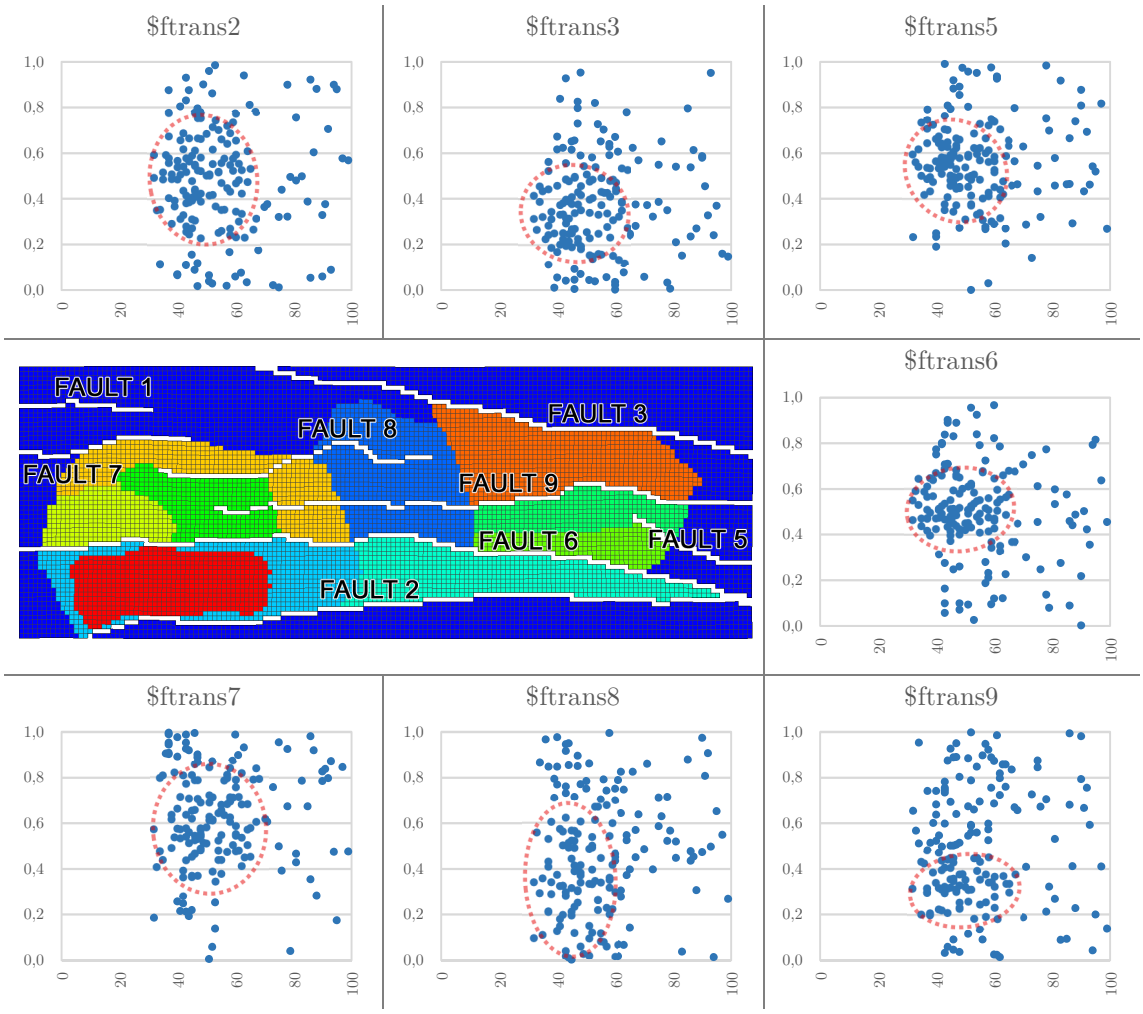


Figure 57 – Fault transmissibility parameter value (y axis) vs Misfit (x axis).

In what variogram parameter perturbation is concerned Figure 58 shows the values obtained for horizontal and vertical Permeability ranges vs Iteration.

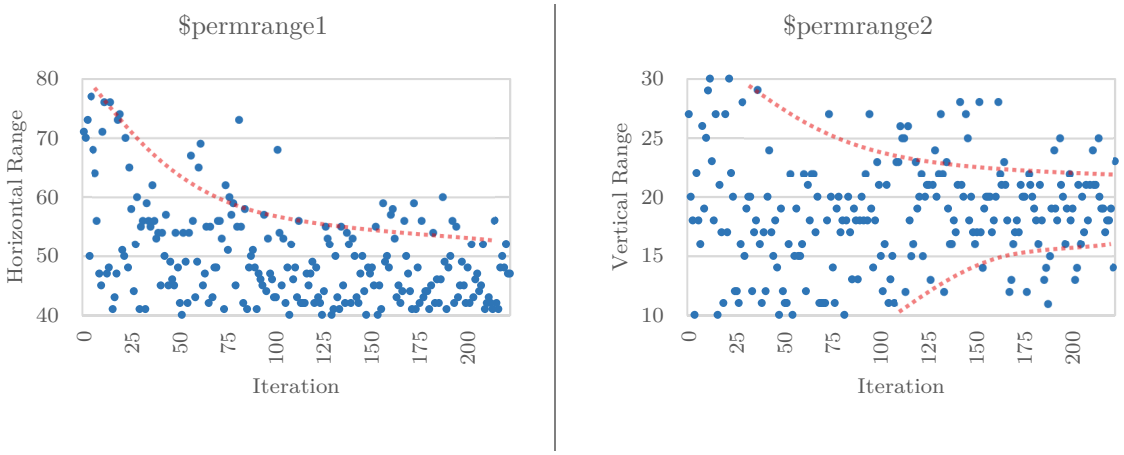


Figure 58 – Permeability range vs Iteration (left – Horizontal range, right – Vertical range).

By looking at Figure 58 , it is possible to observe a tendency for lower values of horizontal ranges ($permrange1) being adopted. In terms of vertical ranges ($permrange2), towards the final half of the iterations, a tendency for central values regarding the adopted distribution is being more recurrent.

Figure 59 shows the concentration of parameter values for horizontal and vertical permeability ranges over the course of the iterations. A concentration of lower values for horizontal permeability ranges ($permrange1) is observed, while for vertical ranges, the spread of parameter values corresponding to lower misfits is much higher. This could mean that vertical ranges for the property of permeability are not important enough in terms of dynamic response for the sampling algorithm to detect. In other words, regardless of the adopted range for vertical permeability within the assumed parameter value uncertainty, dynamic simulation is not sufficiently affected. Another explanations could lie in the definition of the distribution intervals, which could eventually favor better matched responses and thus, higher concentration of parameter values, for higher than assumed ranges of vertical permeability.

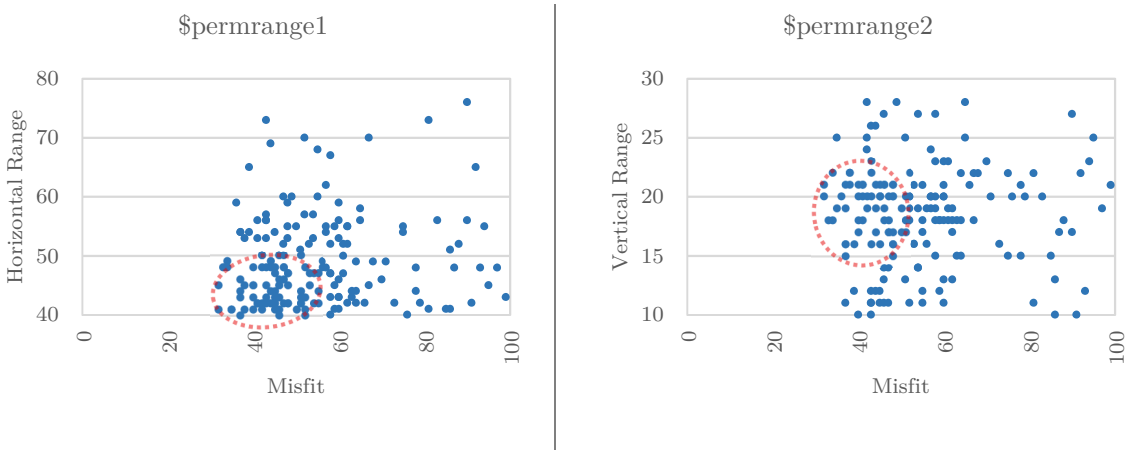


Figure 59 – Permeability range vs Misfit (left – Horizontal range, right – Vertical range).

Figure 60 shows a top view of the Permeability field for the best simulation (Iteration 91, simulation 5).

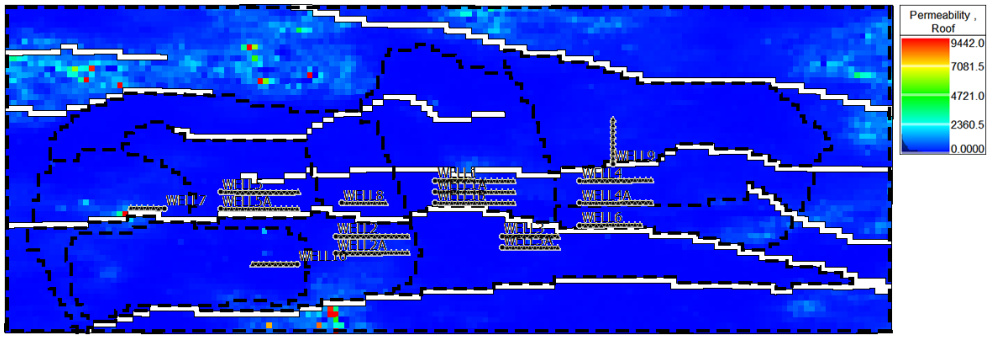


Figure 60 – Top view of the best-fit Permeability realization (Iteration 91, Simulation 5).

Regarding histogram perturbation for permeability, Figure 61 shows the results obtained for histogram perturbation vs Iteration. A tendency for sampling the higher values for facies 2 proportions ($perm_fac_2), can be observed, which is the facies associated with the highest permeability values (Chapter 3.3.1.2). Regarding means for both facies, it is possible to observe that the tendency is not clear for facies 1, with the same *wavy–like* sampling pattern being observed as previously shown regarding fault transmissibilities. For facies 2, convergence is observed towards a range of values located at the mid to top range of the perturbation distribution interval.

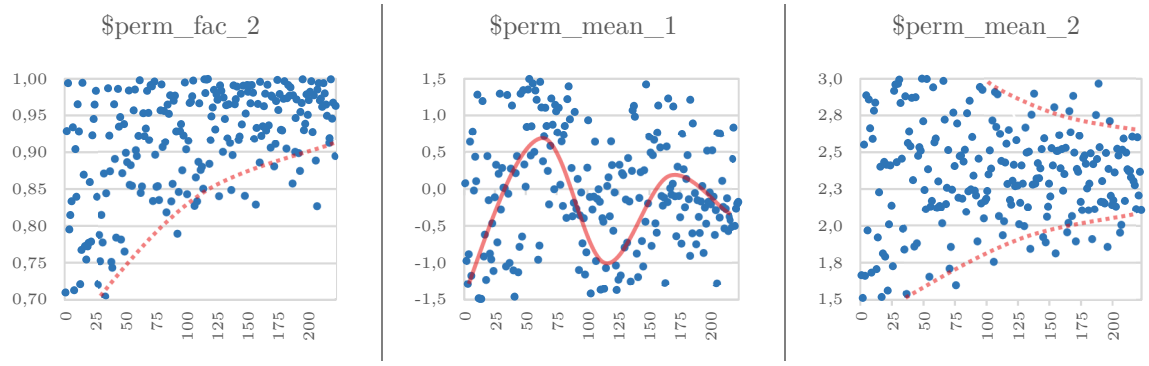


Figure 61 – Permeability Histogram perturbation (y axis – parameter value) vs Iteration (x axis), (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean).

Figure 62 shows the concentration of parameter values obtained for histogram perturbation, over the course of the iterations. There is a concentration of values for the proportion of Permeability Facies 2 occurring at higher values. Permeability means for Facies 1 does not display an identifiable pattern, meaning that probably the perturbation for this parameter does not have an effect that is relevant enough on dynamic response, for the sampling algorithm to detect. As for permeability means for facies 2, the same a concentration of parameter values, at the top half part of the parameter distribution interval, is observed.

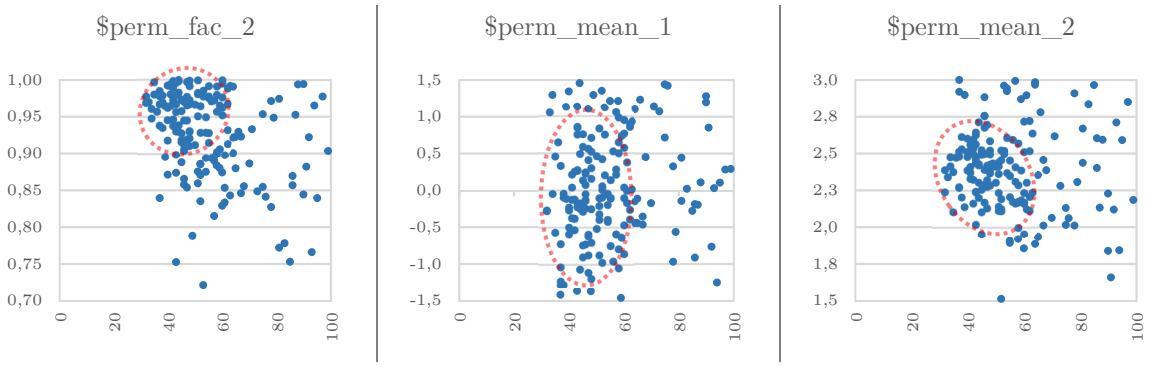


Figure 62 – Permeability Histogram perturbation (y axis – parameter value) vs Misfit (x axis), (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean).

Figure 63 is a comparison between the base scenario $\log_{10}(k)$ histogram and the histograms resulting from the best realization (iteration 91, simulation 5).

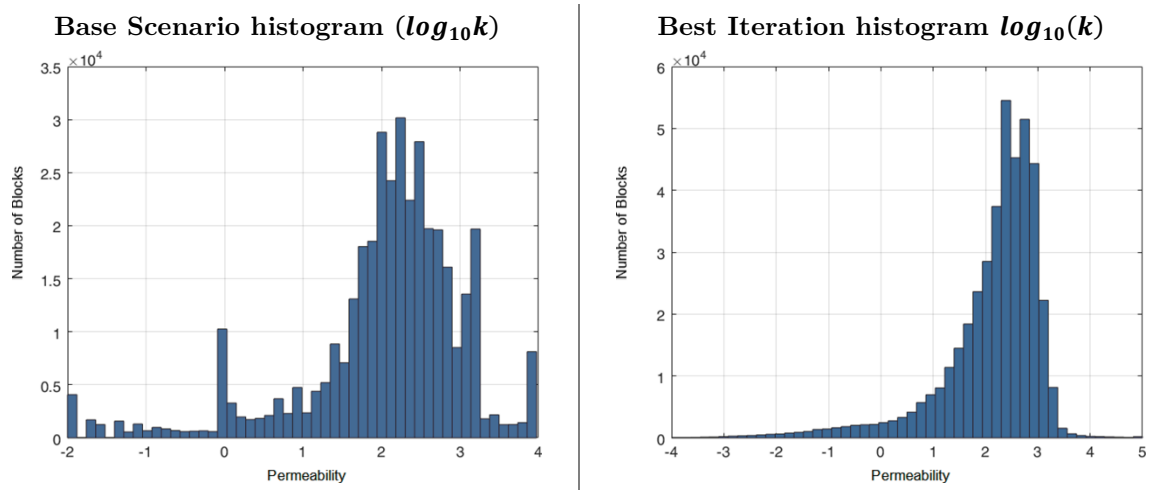


Figure 63 – Comparison between Permeability histogram for Base Scenario and Best Iteration.

For variogram parameter perturbation of porosity, Figure 64 shows the values obtained for horizontal and vertical Porosity ranges vs Iteration. A trend for lower values of horizontal ranges ($pororange1) can be observed. In terms of vertical ranges ($pororange2), it is possible to observe a trend for higher values of vertical ranges.

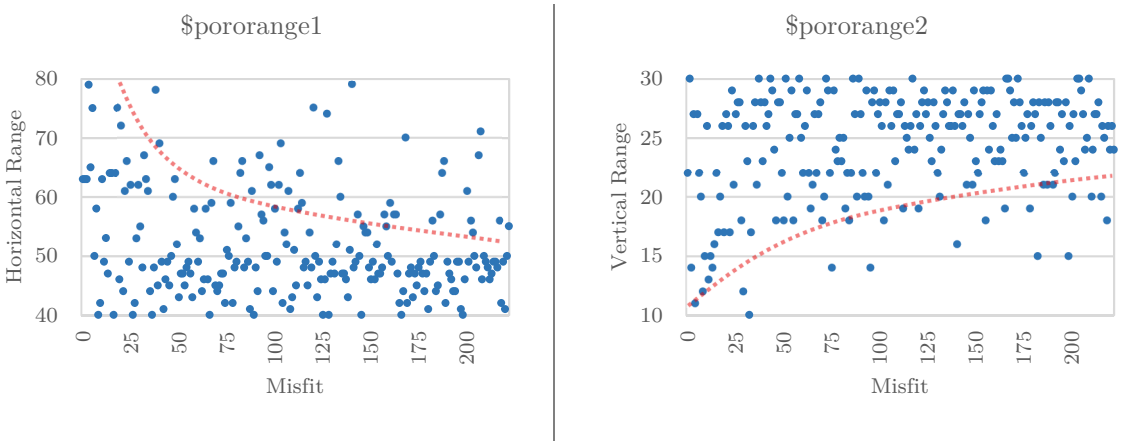


Figure 64 – Porosity range vs Iteration (left – Horizontal range, right – Vertical range).

Figure 65 shows the concentration of parameter values for horizontal and vertical porosity ranges over the course of the iterations. A concentration of lower values for horizontal porosity ranges ($pororange1) can be observed, while for vertical ranges, the concentration tends towards higher values of the adopted perturbation distribution.

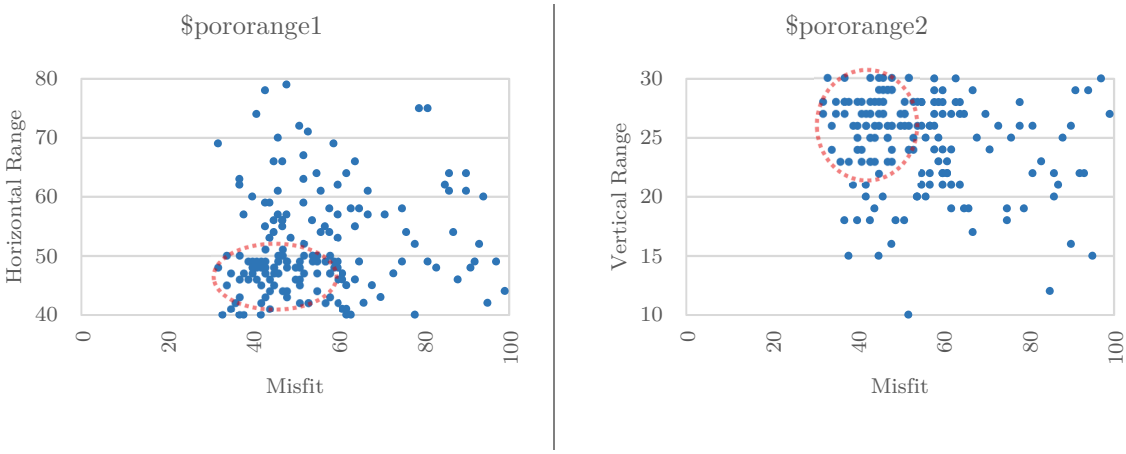


Figure 65 – Porosity range vs Misfit (left – Horizontal range, right – Vertical range).

Figure 66 shows a top view of the Permeability field for the best simulation (iteration 91, simulation 5).

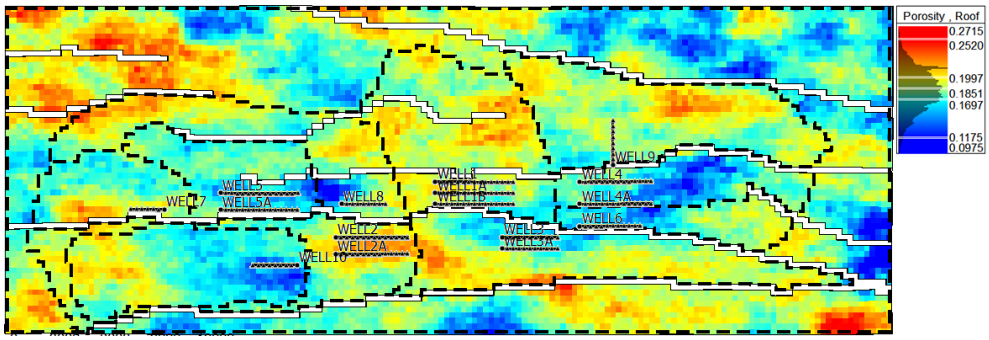


Figure 66 – Top view of the best-fit Porosity realization (Iteration 91, Simulation 5).

Regarding histogram perturbation for porosity, Figure 67 shows the results obtained for histogram perturbation vs Iteration. Values for facies 2 proportions ($poro_fac_2) tend to be sampled in the mid region of the adopted distribution interval, towards later iterations. As for facies distribution means, it is possible to observe the same *wavy–like* observed in other cases of parameter perturbation pattern of searching in the parameter space, with an apex at around iteration 50. Nevertheless, for both facies, sampling seems to be preferring higher values for perturbation, although for Facies 2, the mean seems to converge with more accuracy towards a value of 0.22±0.5.

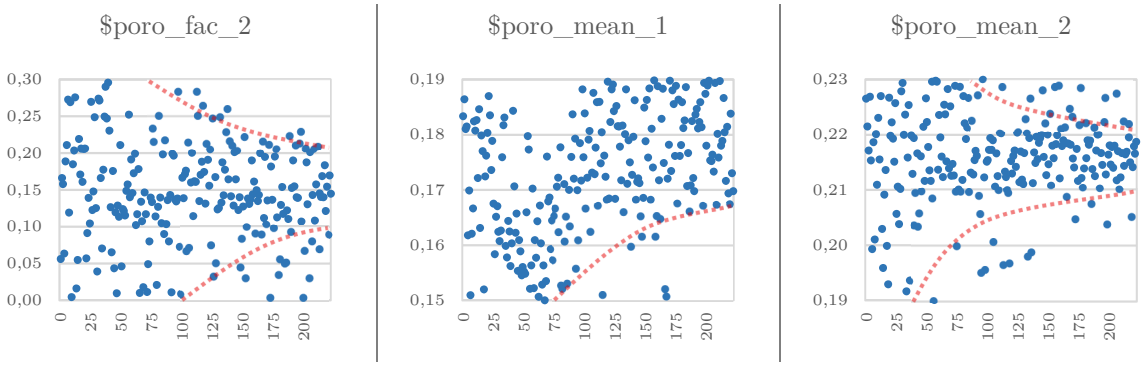


Figure 67 – Porosity Histogram perturbation (y axis – parameter value) vs Iteration (x axis), (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean).

Figure 68 shows the concentration of parameter values obtained for histogram perturbation, over the course of the iterations. There is a concentration of values for Porosity Facies 2 proportions occurring at the mid–region of the perturbation distribution interval. Porosity means for Facies 1 does not display an identifiable pattern, while Porosity means for facies 2, shows a concentration of parameter values at the top half part of the interval, for values between 0.21 and 0.22.

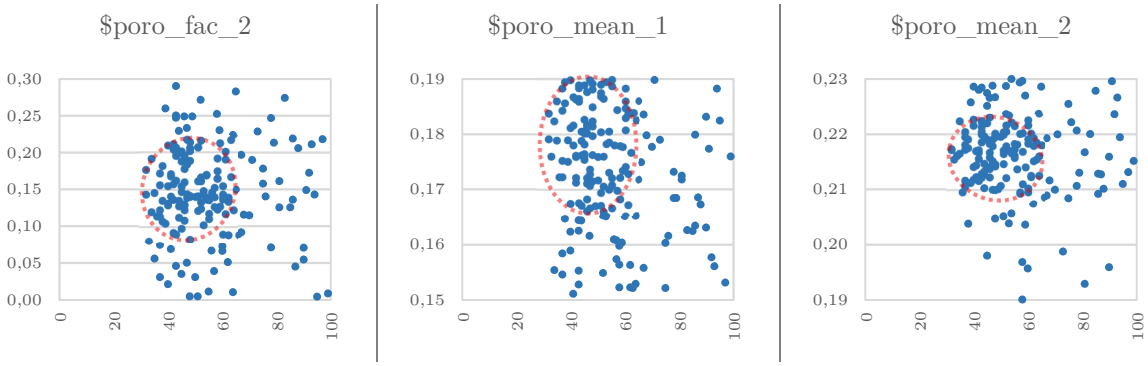


Figure 68 – Porosity Histogram perturbation (y axis – parameter value) vs Misfit (x axis) (left – Facies 2 Proportion, middle – Facies 1 Mean, right – Facies 2 Mean).

Figure 69 is a comparison between the base scenario porosity histogram and the histograms resulting from the best realization (iteration 91, simulation 5).

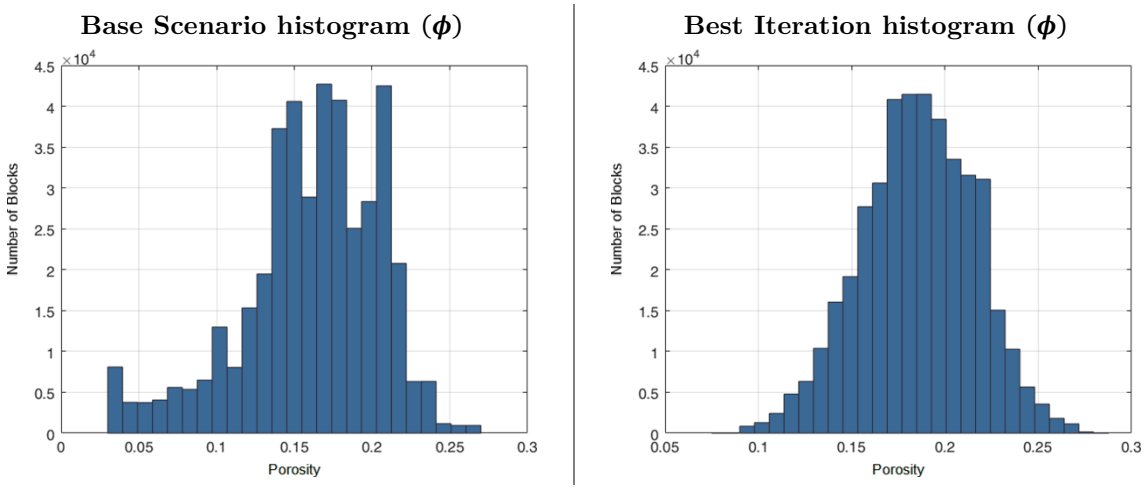


Figure 69 – Comparison between Porosity histogram for Base Scenario and Best Iteration.

6 Conclusions and Future Work

The motivation behind the work presented in this thesis was to contribute with an integrated workflow regarding geostatistical history matching. The purpose of this work was the development of an algorithm that was able to couple GHM with Adaptive Stochastic Sampling, addressing the local match of production data, under a zonation-based methodology. Another contribution of this work was to provide a comparison between several methods of zonation for carrying out GHM.

A workflow was presented, presenting very positive and promising results. The application of this method can be extended to other case studies, parametrization or regionalization methodologies.

The advantages of the proposed methodologies are:

- Addressing perturbation in a geologically consistent manner;
- Respecting fluid production flow pattern, by usage of production streamlines;
- Reduced parametrization of static properties with the discretization of the reservoir;
- Ability to perturb static and engineering parameters and quantifying their uncertainty while integrating them on the history matching process;
- Ability to provide better results than other standard regionalization methods, under less time;
- Ability to generate multiple matched models;

The application of the proposed methodology showed promising results, with the achieving of multiple history matched models with considerably low misfits. The creation of an ensemble of multiple history matched models by solving the ill–posed calibration problem, allows the prediction of reservoir behavior with a degree of uncertainty.

Nevertheless, a limitation of the proposed methodology is its dependency on the right choice of prior ranges, which is essential for good optimization results.

In the case of application to this reservoir case study, water production matching could probably be improved with a deeper investigation on the perturbation of other relevant parameters. Other methodologies or other choice of parameter perturbation and parametrization, could improve match results further (e.g. OWC or relative permeabilities). Other modeling related techniques, such as regridding, combination of different scenario grid properties, like top structures, could also assist in the finding of better matches. Finally, further work on the presented methodology can be done by exploring the benefits of using a Multi-Objective History Matching methodology for improvements on individual well matches.

The successful application of the proposed methodology to a challenging semi-synthetic reservoir case study delivers good perspectives for its application to real cases. Further research on this area could be focused towards other types of parameter perturbation or towards its integration with seismic inversion.

References

Abdollahzadeh, A., Christie, M., Corne, D., Davies, B., & Williams, B. (2012). Bayesian Optimization Algorithm applied to Unceratinty Quantification. *SPE Journal*, 865-873.

Agterberg, F. P. (1974). *Geomathematics, Mathematical Background and Geoscience Applications.* Amsterdam: Elsevier Scientific Publishing Company.

Arnold, D. (2008). *Geological parameterisation of petroleum reservoir models for improved uncertainty quantification.* PhD Thesis Institute of Petroleum Engineering, Heriot-Watt University, U.K., Edinburgh.

Arnold, D. (2012). *Watt Field Case Study.* Heriot-Watt University, Uncertainty Quantification Group, Edinburgh, UK.

Arnold, D., & Demyanov, V. (2015). Modelling and Management. *Modelling and Management Course Material*. Edinburgh, UK: Heriot-Watt University.

Arnold, D., Demyanov, V., Tatum, D., Christie, M., Rojas, T., Geiger, S., & Corbett, P. (2013). Hierarchical benchmark case study for history matching, uncertainty quantification and reservoir characterisation. *Computers and Geosciences, 50*, 4–15.

Bennett, F., & Graf, T. (2002). Use of Geostatistical Modeling and Automatic History Matching to Estimate Production Forecast Uncertainty-A Case Study. *SPE International Petroleum Conference and Exhibition in Mexico.* Society of Petroleum Enginneers.

Caers, J. (2002). Geostatistical history matching under training-image based geological model constraints. *SPE annual technical conference and exhibition.* Society of Petroleum Engineers.

Caers, J. (2003). Efficient gradual deformation using a streamline-based proxy method. *Journal of Petroleum Science and Engineering, 39*(1), 57-83.

Caers, J., & Hoffman, T. (2006). The probability perturbation method: A new look at Bayesian inverse modeling. *Mathematical geology, 38*(1), 81-100.

Caragea, P., & Smith, R. L. (2006). Approximate Likelihoods for Spatial Processes. *Technical Report, Department of statistics, Iowa State University.*

Carrera, J., Alcolea, A., Medina, A., & Hidalgo, J. (2005). Inverse problem in hydrogeology. *Hydrogeology journal, 13*(1), 206-222.

Christie, M., Demyanov, V., & Erbas, D. (2006). Uncertainty quantification for porous media flows. *Journal of Computational Physics, 217*(1), 143-158.

Correia, P. (2013, March). *Student Garden Geostatistics course.* Retrieved from ResearchGate: https://www.researchgate.net/publication/268811925_Student_Garden_Geostatistics_course

Erbas, D., & Christie, M. A. (2007). Effect of sampling strategies on prediction uncertainty estimation. *In SPE Reservoir Simulation Symposium.* Society of Petroleum Engineers.

Evensen, G., Hove, J., Meisingset, H., & Reiso, E. (2007). Using the EnKF for assisted history matching of a North Sea reservoir model. *SPE Reservoir Simulation Symposium.* Society of Petroleum Engineers.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation.* New York: Oxford Press.

Gross, H., Thiele, M. R., Alexa, M. J., & Caers, J. (2004). Streamline-based history matching using geostatistical constraints: Application to a giant, mature carbonate reservoir. *SPE Annual Technical Conference and Exhibition.* Society of Petroleum Engineers.

Hajizadeh, Y., Christie, M. A., & Demyanov, V. (2009). Application of differential evolution as a new method for automatic history matching. *Kuwait International Petroleum Conference and Exhibition.* Society of Petroleum Engineers.

Hoffman, B. T., & Caers, J. (2005). Regional probability perturbations for history matching. *Journal of Petroleum Science and Engineering, 46*(1), 53-71.

Hoffman, B. T., & Caers, J. (2007). History matching by jointly perturbing local facies proportions and their spatial distribution: Application to a North Sea reservoir. *Journal of Petroleum Science and Engineering, 57*(3), 257-272.

Hohn, M. E. (1999). *Geostatistics and Petroleum Geology.* Kluwer Adademic Publishers.

Hu, L. Y., Blanc, G., & Noetinger, B. (2001). Gradual deformation and iterative calibration of sequential stochastic simulations. *Mathematical Geology, 33*(4), 475-489.

Journel, A. G. (1994). *Modelling uncertainty: some conceptual thoughts, Geostatistics for the next century.* Kluwer, Dordrecht, Holland: (eds R. Dimitrakopoulos).

Kashib, T., & Srinivasan, S. (2006). A probabilistic approach to integrating dynamic data in reservoir models. *Journal of Petroleum Science and Engineering, 50*(3), 241-257.

Kavetski, D., Kuczera, G., & Franks, S. W. (2006). Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. *42*(3).

Kennedy, J., & Eberhart , R. C. (1995). *Particle swarm optimization. Proceedings of IEEE International Conference on Neural Networks.* Piscataway, NJ.

Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy, 52*(6), 119-139.

Le Ravalec-Dupin , M., & Da Veiga, S. (2011). Cosimulation as a perturbation method for calibrating porosity and permeability fields to dynamic data. *Computers & geosciences, 37*(9), 1400-1412.

Le Ravalec-Dupin, M., & Fenwick, D. H. (2002). A Combined Geostatistical and Streamline-Based History Matching Procedure. *SPE Annual Technical Conference and Exhibition.* San Antonio, Texas: Society of Petroleum Engineers.

Maschio, C., Vidal, A. C., & Schiozer, D. J. (2008). A framework to integrate history matching and geostatistical modeling using genetic algorithm and direct search methods. *Journal of Petroleum Science and Engineering, 63*(1), 34–42.

Mata-Lima, H. (2008). Reservoir characterization with iterative direct sequential co-simulation: integrating fluid dynamic data into stochastic model. *62*(3), 59-72.

Matheron, G. (1963). Principles of geostatistics. *Economic geology, 58*(8), 1246-1266.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in applied probability,*, 439-468.

Mohamed, L. (2011). *Novel sampling techniques for reservoir history matching optimisation and uncertainty quantification in flow prediction.* PhD Thesis Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, UK.

Mohamed, L., Christie, M. A., & Demyanov, V. (2011). History matching and uncertainty quantification: multiobjective particle swarm optimisation approach. *SPE EUROPEC/EAGE annual conference and exhibition.* Society of Petroleum Engineers.

Nunes, R., Soares, A., Azevedo, L., & Pereira, P. (2016). Geostatistical Seismic Inversion with Direct Sequential Simulation and Co-simulation with Multi-local Distribution Functions. *Mathematical Geosciences*, 1-19.

Robert, C. P., & Casella, G. (2004). *Monte Carlo statistical methods, volume 319.* Citeseer.

Rojas, R., Feyen, L., & Dassargues, A. (2009). Sensitivity analysis of prior model probabilities and the value of prior knowledge in the assessment of conceptual model uncertainty in groundwater modelling. *Hydrological Processes, 23*(8), 1131-1146.

Sambridge, M. (1999). Geophysical inversion with a neighbourhood algorithm—II. Appraising the ensemble. *Geophysical Journal International, 138*(3), 727-746.

Sichel, H. S. (1952). New methods in the statistical evaluation of mine sampling data. *Bull. Inst. Min. Metall.*, 261-288.

Silva, A. C. (1997). *Simulação de Reservatórios Petrolíferos, Curso de Mestrado em Geo-Recursos.* Lisboa: Departamento de Engenharia de Minas, Secção de Mineralogia e Planeamento Mineiro, Instituto Superior Técnico.

Soares, A. (2001). Direct sequential simulation and cosimulation. *Mathematical Geology, 33* (8), 911–926.

Soares, A. (2006). *Geostatística para as Ciências da Terra e do Ambiente, 2ª Edição.* Lisboa: IST Press.

Soares, A., Azevedo, L., Focaccia, S., & Carneiro, J. (2014). Geostatistical history matching conditioned to seismic data. *Mathematics of Planet Earth*, 65-68.

Streamsim. (2016). Retrieved from http://www.streamsim.com: http://www.streamsim.com/sites/default/files/genmodelingworkflow.png

Suzuki, S., & Caers, J. (2006). History matching with an uncertain geological scenario. *SPE Annual Technical Conference and Exhibition.* Society of Petroleum Engineers.

Tureyen, O. I., & Caers, J. (2003). A two-level optimization method for geostatistical integration of production data on non-uniform grids. *SPE Annual Technical Conference and Exhibition.* Society of Petroleum Engineers.

Unisim. (2016). Retrieved from http://www.unisim.cepetro.unicamp.br/: http://www.unisim.cepetro.unicamp.br/br/pesquisas/tecnicas-de-calibracao-de-modelos-com-dados-dinamicos/ajuste-integrado-com-caracterizacao-de-reservatorios

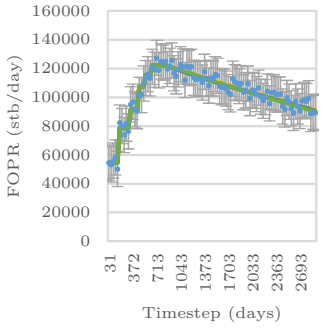
Williams, A. M., Keating, J. F., & Barhouty, M. F. (1998). The Stratigraphic Method: A Structured Approach to History Matching Complex Simulation Models. *SPE Reservoir Engineering, 1*(2), 169-176.

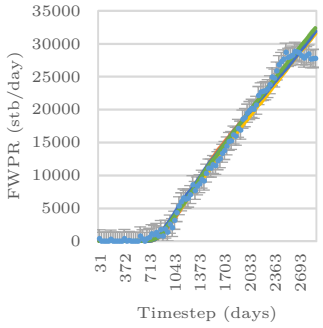
Yaseen, M., Hamm, N. A., Tolpekin, V., & Stein, A. (2013). Anisotropic kriging to derive missing coseismic displacement values obtained from synthetic aperture radar images. *Journal of applied remote sensing, 7*(1), 073580-073580.

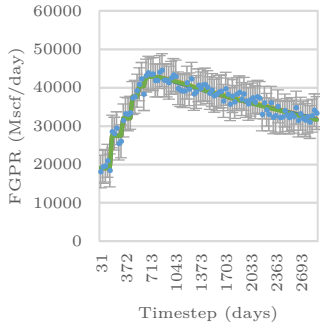
Yustres, Á., Asensio, L., Alonso, J., & Navarro, V. (2012). A review of Markov Chain Monte Carlo and information theory tools for inverse problems in subsurface flow. *Computer Geosciences, 16*(4), 1-20.

Appendix A – Best 5 Realizations (GHM coupling)

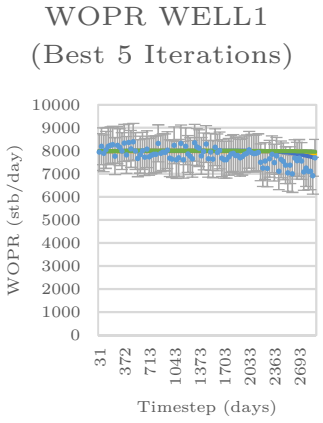
Field

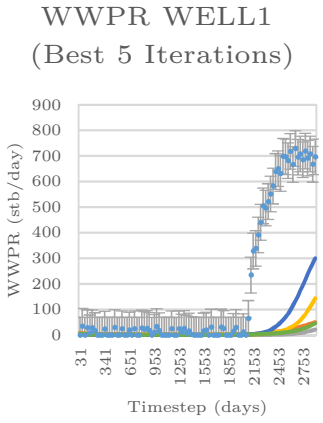






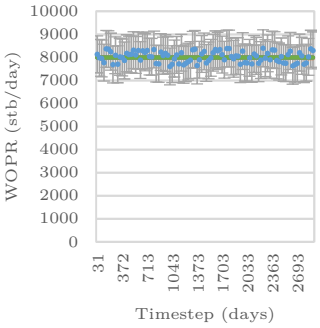
WELL1





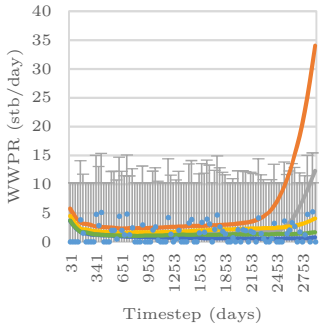
WELL1A

WOPR WELL1A
(Best 5 Iterations)



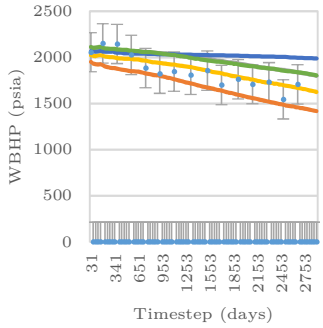
- WOPR_WELL1A (History)
- WOPR_WELL1A (91)
- WOPR_WELL1A (104)
- WOPR_WELL1A (90)
- WOPR_WELL1A (46)
- WOPR_WELL1A (122)

WWPR WELL1A
(Best 5 Iterations)



- WWPR_WELL1A (History)
- WWPR_WELL1A (91)
- WWPR_WELL1A (104)
- WWPR_WELL1A (90)
- WWPR_WELL1A (46)
- WWPR_WELL1A (122)

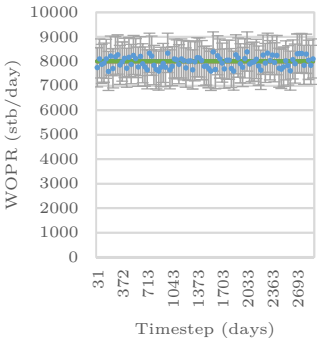
WBHP WELL1A
(Best 5 Iterations)



- WBHP_WELL1A (History)
- WBHP_WELL1A (91)
- WBHP_WELL1A (104)
- WBHP_WELL1A (90)
- WBHP_WELL1A (46)
- WBHP_WELL1A (122)

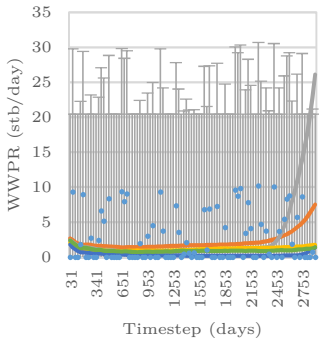
WELL1B

WOPR WELL1B
(Best 5 Iterations)



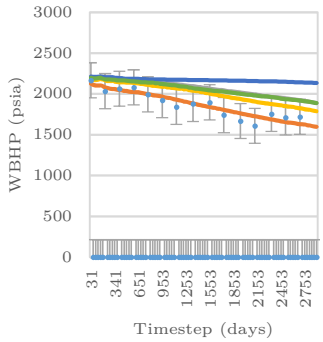
- WOPR_WELL1B (History)
- WOPR_WELL1B (91)
- WOPR_WELL1B (104)
- WOPR_WELL1B (90)
- WOPR_WELL1B (46)
- WOPR_WELL1B (122)

WWPR WELL1B
(Best 5 Iterations)



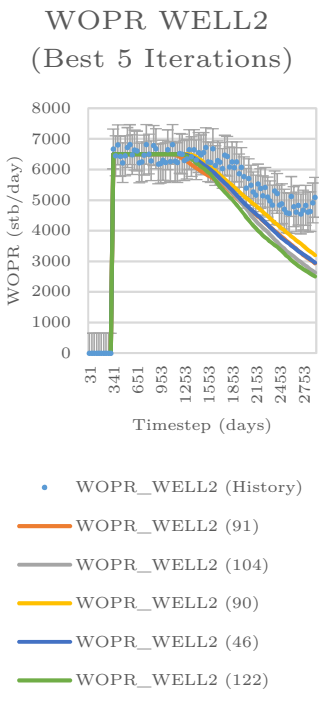
- WWPR_WELL1B (History)
- WWPR_WELL1B (91)
- WWPR_WELL1B (104)
- WWPR_WELL1B (90)
- WWPR_WELL1B (46)
- WWPR_WELL1B (122)

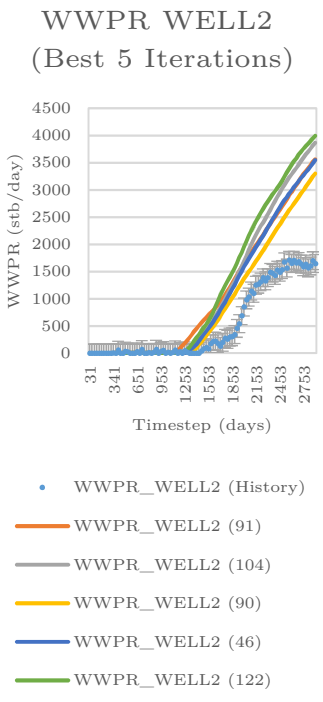
WBHP WELL1B
(Best 5 Iterations)

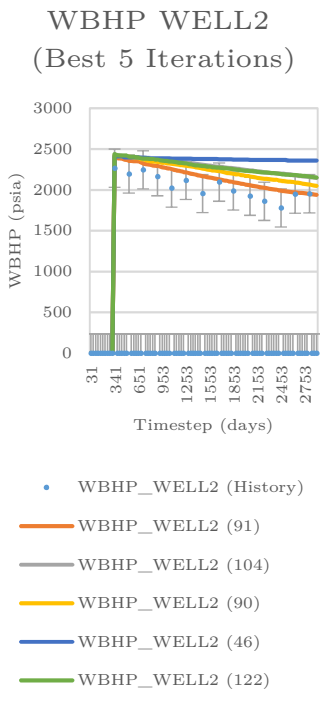


- WBHP_WELL1B (History)
- WBHP_WELL1B (91)
- WBHP_WELL1B (104)
- WBHP_WELL1B (90)
- WBHP_WELL1B (46)
- WBHP_WELL1B (122)

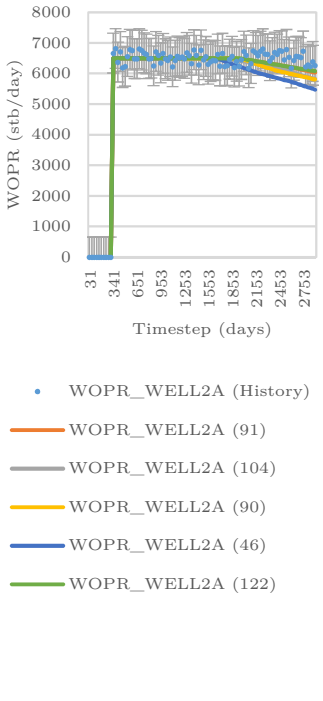
WELL2

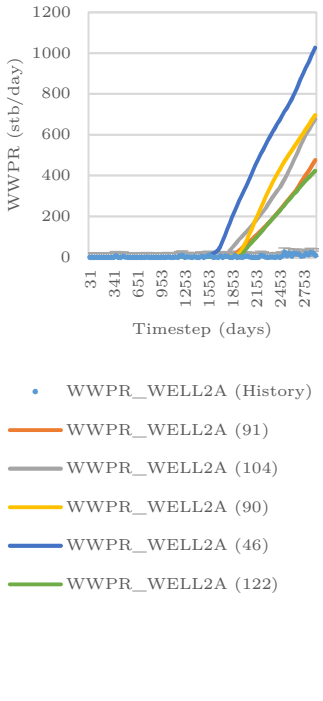


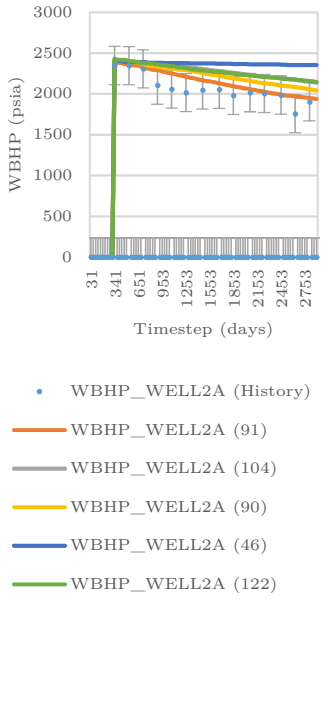




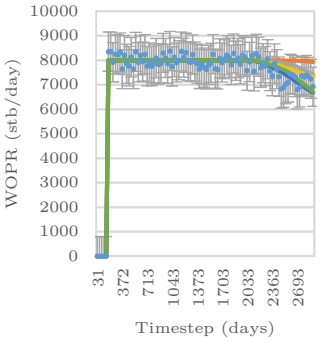
WELL2A

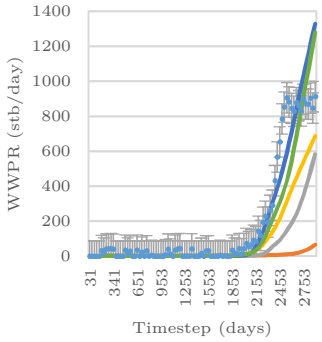


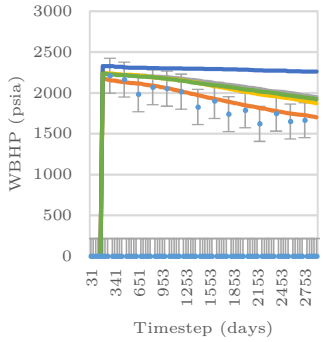




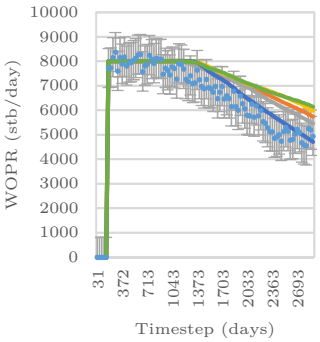
WELL3

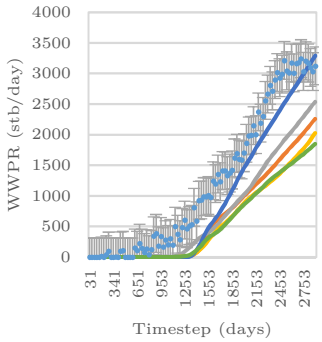


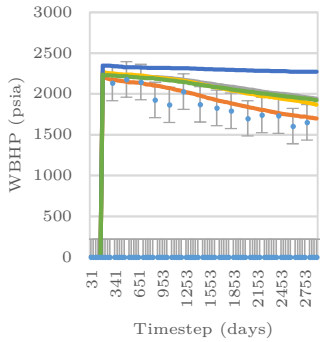




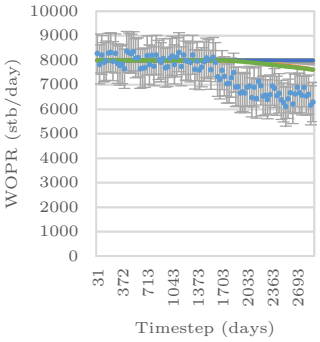
WELL3A

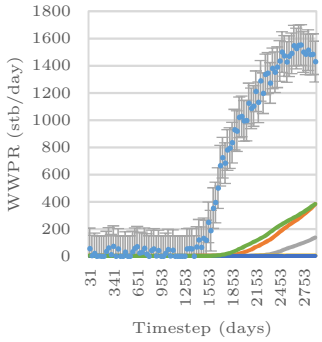


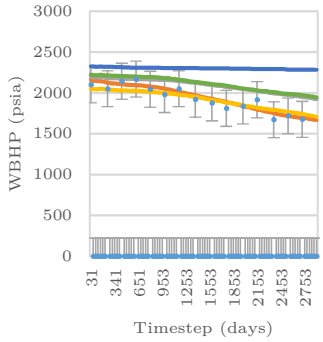




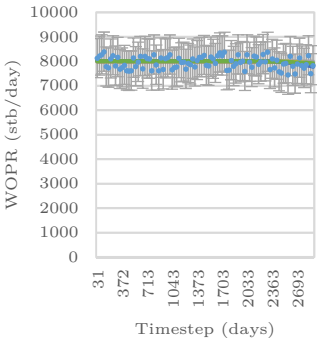
WELL4

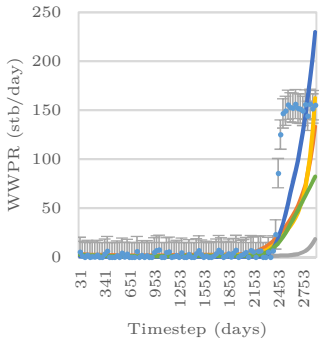


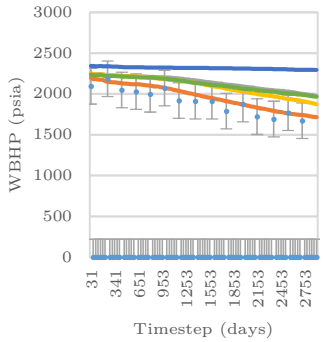




WELL4A

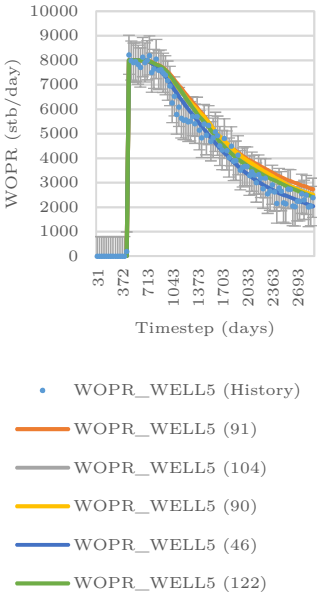




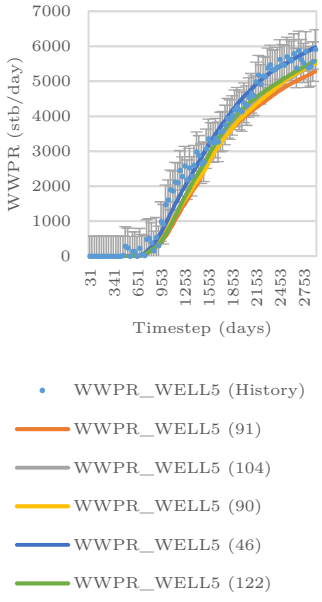


WELL5

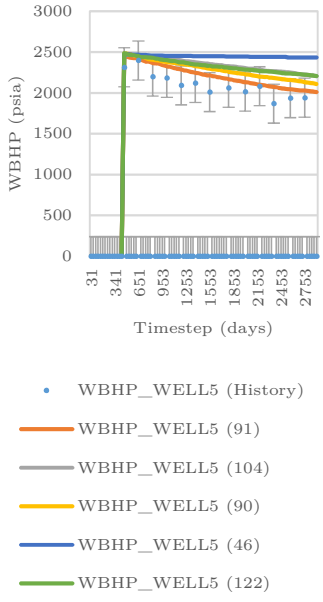
WOPR WELL5 (Best 5 Iterations)



WWPR WELL5 (Best 5 Iterations)

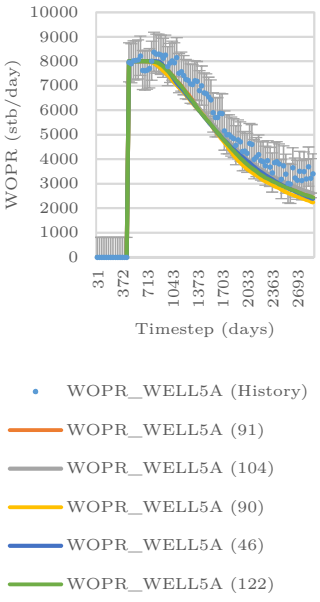


WBHP WELL5 (Best 5 Iterations)

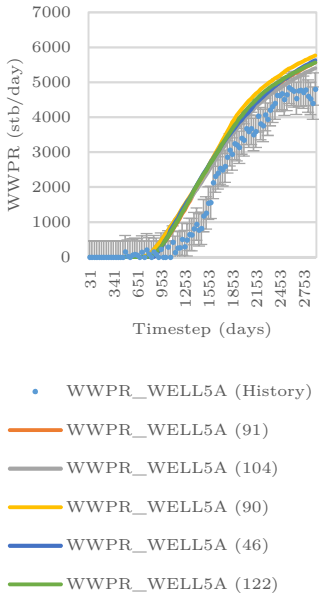


WELL5A

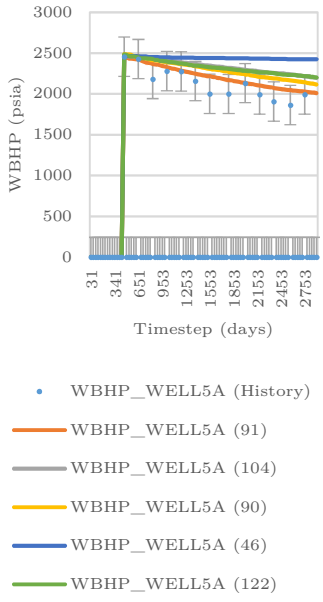
WOPR WELL5A (Best 5 Iterations)



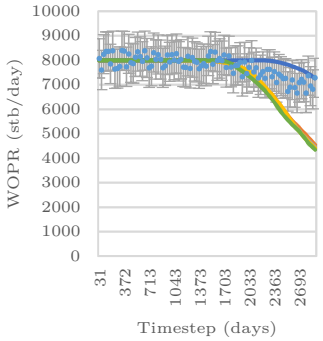
WWPR WELL5A (Best 5 Iterations)

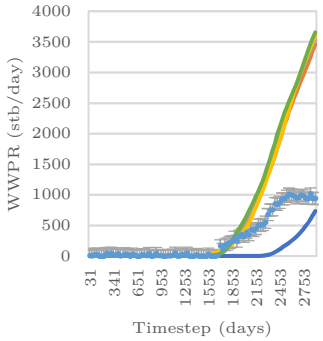


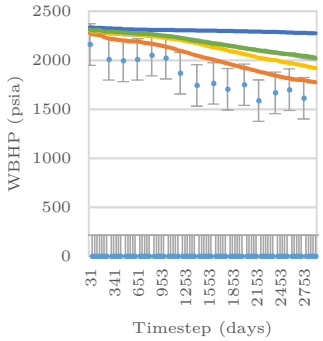
WBHP WELL5A (Best 5 Iterations)



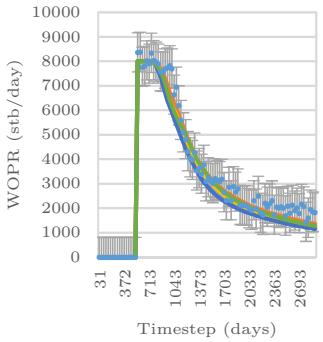
WELL6

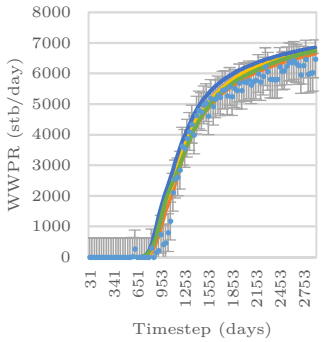

WOPR WELL6
(Best 5 Iterations)
WOPR (stb/day)
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
31
372
713
1043
1373
1703
2033
2363
2693
Timestep (days)
WOPR_WELL6 (History)
WOPR_WELL6 (91)
WOPR_WELL6 (104)
WOPR_WELL6 (90)
WOPR_WELL6 (46)
WOPR_WELL6 (122)

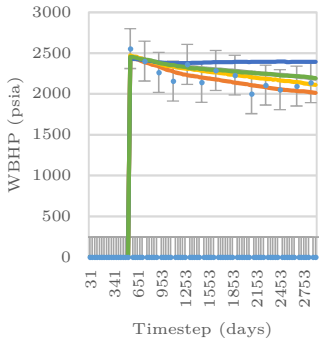

WWPR WELL6
(Best 5 Iterations)
WWPR (stb/day)
4000
3500
3000
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WWPR_WELL6 (History)
WWPR_WELL6 (91)
WWPR_WELL6 (104)
WWPR_WELL6 (90)
WWPR_WELL6 (46)
WWPR_WELL6 (122)


WBHP WELL6
(Best 5 Iterations)
WBHP (psia)
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WBHP_WELL6 (History)
WBHP_WELL6 (91)
WBHP_WELL6 (104)
WBHP_WELL6 (90)
WBHP_WELL6 (46)
WBHP_WELL6 (122)

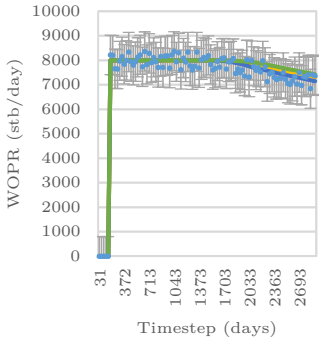
WELL7

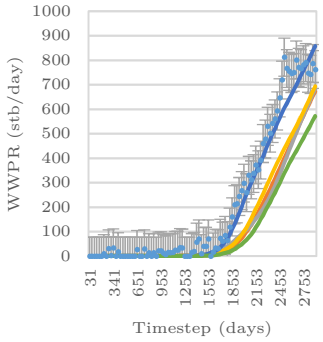

WOPR WELL7
(Best 5 Iterations)
WOPR (stb/day)
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
31
372
713
1043
1373
1703
2033
2363
2693
Timestep (days)
WOPR_WELL7 (History)
WOPR_WELL7 (91)
WOPR_WELL7 (104)
WOPR_WELL7 (90)
WOPR_WELL7 (46)
WOPR_WELL7 (122)

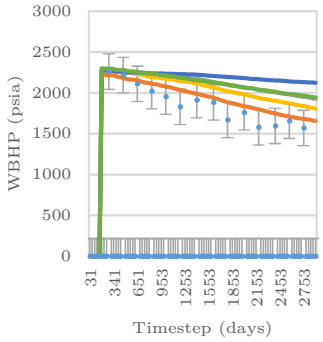

WWPR WELL7
(Best 5 Iterations)
WWPR (stb/day)
8000
7000
6000
5000
4000
3000
2000
1000
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WWPR_WELL7 (History)
WWPR_WELL7 (91)
WWPR_WELL7 (104)
WWPR_WELL7 (90)
WWPR_WELL7 (46)
WWPR_WELL7 (122)


WBHP WELL7
(Best 5 Iterations)
WBHP (psia)
3000
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WBHP_WELL7 (History)
WBHP_WELL7 (91)
WBHP_WELL7 (104)
WBHP_WELL7 (90)
WBHP_WELL7 (46)
WBHP_WELL7 (122)

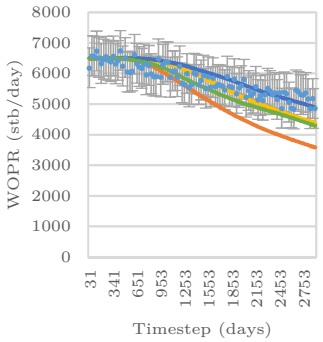
WELL8

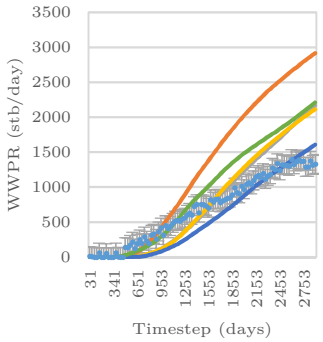

WOPR WELL8
(Best 5 Iterations)
WOPR (stb/day)
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
31
372
713
1043
1373
1703
2033
2363
2693
Timestep (days)
WOPR_WELL8 (History)
WOPR_WELL8 (91)
WOPR_WELL8 (104)
WOPR_WELL8 (90)
WOPR_WELL8 (46)
WOPR_WELL8 (122)

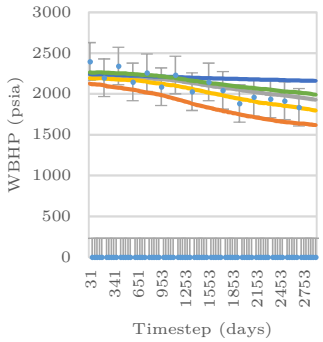

WWPR WELL8
(Best 5 Iterations)
WWPR (stb/day)
1000
900
800
700
600
500
400
300
200
100
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WWPR_WELL8 (History)
WWPR_WELL8 (91)
WWPR_WELL8 (104)
WWPR_WELL8 (90)
WWPR_WELL8 (46)
WWPR_WELL8 (122)


WBHP WELL8
(Best 5 Iterations)
WBHP (psia)
3000
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WBHP_WELL8 (History)
WBHP_WELL8 (91)
WBHP_WELL8 (104)
WBHP_WELL8 (90)
WBHP_WELL8 (46)
WBHP_WELL8 (122)

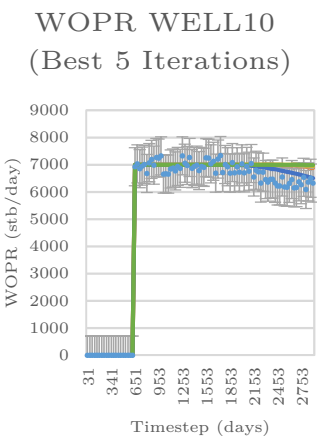
WELL9

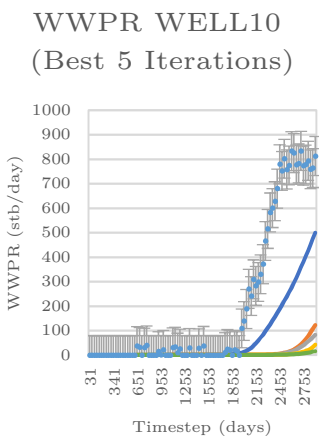

WOPR WELL9
(Best 5 Iterations)
WOPR (stb/day)
8000
7000
6000
5000
4000
3000
2000
1000
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WOPR_WELL9 (History)
WOPR_WELL9 (91)
WOPR_WELL9 (104)
WOPR_WELL9 (90)
WOPR_WELL9 (46)
WOPR_WELL9 (122)

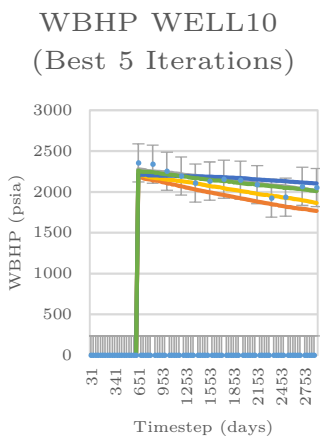

WWPR WELL9
(Best 5 Iterations)
WWPR (stb/day)
3500
3000
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WWPR_WELL9 (History)
WWPR_WELL9 (91)
WWPR_WELL9 (104)
WWPR_WELL9 (90)
WWPR_WELL9 (46)
WWPR_WELL9 (122)


WBHP WELL9
(Best 5 Iterations)
WBHP (psia)
3000
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WBHP_WELL9 (History)
WBHP_WELL9 (91)
WBHP_WELL9 (104)
WBHP_WELL9 (90)
WBHP_WELL9 (46)
WBHP_WELL9 (122)

WELL10


WOPR WELL10
(Best 5 Iterations)
WOPR (stb/day)
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WOPR_WELL10 (History)
WOPR_WELL10 (91)
WOPR_WELL10 (104)
WOPR_WELL10 (90)
WOPR_WELL10 (46)
WOPR_WELL10 (122)


WWPR WELL10
(Best 5 Iterations)
WWPR (stb/day)
1000
900
800
700
600
500
400
300
200
100
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WWPR_WELL10 (History)
WWPR_WELL10 (91)
WWPR_WELL10 (104)
WWPR_WELL10 (90)
WWPR_WELL10 (46)
WWPR_WELL10 (122)


WBHP WELL10
(Best 5 Iterations)
WBHP (psia)
3000
2500
2000
1500
1000
500
0
31
341
651
953
1253
1553
1853
2153
2453
2753
Timestep (days)
WBHP_WELL10 (History)
WBHP_WELL10 (91)
WBHP_WELL10 (104)
WBHP_WELL10 (90)
WBHP_WELL10 (46)
WBHP_WELL10 (122)