Comparison of 2D and 3D uncertainty quantification in geostatistical modeling in stratigraphic formations

by

Amir Mahdi Latifi

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mining Engineering

Department of Department of Civil and Environmental Engineering
University of Alberta

Abstract

Under suitable geological conditions and depending on the modeling goals at the project's current phase, a geological modeler could justifiably choose to estimate natural resource concentrations by building geostatistical models on a 2D grid rather than a 3D grid. While the application of these 2D models is limited only to resource estimation, they take less time for a modeler to build and while generating similar overall results to 3D models, are simpler to build. Thus, in geological domains that have a flat and layer-like geometry, i.e., tabular mineral deposits, modelers might prefer these 2D models to 3D models.

While the implementation of 2D geostatistical modeling in these geological domains is well established, the quantification of uncertainty in these workflows is not, especially in comparison to 3D models in the same domains. Since the uncertainty in final resource estimations is a critical economic factor in deciding the feasibility and risk of investing in a mineral/hydrocarbon extraction campaign, an understanding of differences in uncertainty quantification between 2D and 3D models allows modelers to make informed choices on which type of modeling workflow is most appropriate for their specific project.

To make a comprehensive comparison of uncertainty quantification in 2D and 3D modeling, all aspects of uncertainty are considered and grouped into three categories: uncertainty in model parameters, uncertainty in the geometry, and residual uncertainty (which is the uncertainty in the models that include no parameter uncertainty). These categories are analyzed through two lenses: firstly, by building 2D and 3D models in the McMurray formation of Northern Alberta as a case study and

analyzing the results, and secondly, by looking at each category separately through analytical tests.

The results show that the smaller size of the 2D domain and data set leads to higher uncertainties in some categories, such as histogram uncertainty or residual uncertainty. However, this is not necessarily true in other categories, such as the uncertainty in the geometry, and uncertainty could be higher in the 2D or 3D workflow depending on modeler choices and domain characteristics. In the case study, the 3D workflow shows almost identical overall global uncertainty in the results compared to the 2D workflow. This emphasizes the importance of incorporating full uncertainty quantification in the results to have realistic resource estimations while indicating the loss of information in the 2D workflow that causes higher uncertainties in some aspects of the workflow. Additionally, the results from analytical tests provide some insight into contributing factors to differences between uncertainty quantification in 2D and 3D workflows, such as the vertical size of the domain, degree of vertical spatial continuity, and sample density that have an impact on the degree of differences between the two workflows. By considering these factors, the modeler could make a more informed decision on the type of workflow for their project with an understanding of probable differences between the two workflows.

To my girlfriend, Negin, and my parents, Maryam and Jamal, without whom I would not have made it here.

Acknowledgements

Firstly, I would like to express my gratitude to my academic supervisor, Dr. Jeff Boisvert, who has been a constant and unrelenting supporter during my master's studies and has provided me with professional, intellectual, and personal help whenever I needed it. Without his patience and mentorship, this work would not have been possible.

Secondly, I want to thank the Centre for Computational Geostatistics (CCG) for providing financial support during my master's studies and my colleagues at CCG for providing a warm and helpful environment of friendship and collaboration. Specifically, I would like to thank Ben Harding, whose help and guidance have been a major support in my graduate research experience during the COVID-19 pandemic. Lastly, I am eternally grateful to my girlfriend, Negin, and my parents for providing loving and kind support throughout the past two years.

Table of Contents

List of Figures

List of Symbols

Latin

CP	number of control points in each borehole.	
$LBH1$	length of the array of samples in borehole 1.	
$LBH2$	length of the array of samples in borehole 2.	
$OOIP$	Original Oil In Place.	m^3
$R1$	bounding limits of the elevation of control point in borehole 1.	m
$R2$	bounding limits of the elevation of control point in borehole 2.	m
So	Oil Saturation.	*ratio*
Tol	tolerance parameter in the pairing of samples.	*ratio*
$Vlimit$	maximum vertical index at a grid node.	

Greek

ϕ	Effective Porosity.	*ratio*

Abbreviations

BCR Bottom of the Continuous Reservoir.

OOIP Original Oil In Place.

SAGD Steam Assisted Gravity Drainage.

SGS Sequential Gaussian Simulation.

SIS Sequential Indicator Simulation.

TCR Top of the Continuous Reservoir.

Chapter 1

Introduction

1.1 Motivation and problem statement

Every natural resources extraction project, such as open-pit and underground mines and in-situ oil extraction projects, is developed based on an understanding of the spatial distribution of valuable minerals and hydrocarbons in the geological domain under consideration (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014; Darling 2011). This understanding usually takes the form of block models of the subsurface that contains estimated values of variables such as grades of metals and oil saturation (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014). These block models also provide a global estimation of the total tonnage of material and other variables which leads to some estimation of the contained resources, which in turn is a major factor in deciding the feasibility of investing in and extracting the resource under consideration (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014). One common way to construct these block models is through geostatistical modeling of these variables through relevant estimation and simulation workflows, which is the methodology followed in this research (Journel and Huijbregts 2003).

Most natural resource concentrations extend in three dimensions, which means to properly model and capture the spatial distribution of petrophysical properties and geological features, it is best to create 3D models of those resource concentrations (Journel and Huijbregts 2003). Such models are vital in many applications of geostatistical modeling, including flow simulations and mine schedules, and while they act as a basis for extraction campaigns of mining and petroleum resources, they are also the best way to apply and use all types of geospatial data (Journel and Huijbregts

2003; Pyrcz and Deutsch 2014). However, the process of creating 3D models in many cases can become too complex or computationally exhaustive (Ren *et al.* 2006). Many geological domains require fine-scale models, sometimes with grid cells in the order of centimeters, to capture the spatial variability of geological features. Creating models of these cells in domains that extend for multiple kilometers laterally and hundreds of meters vertically could be difficult. Plus, considering many important aspects of geostatistical modeling, like parameter uncertainty, could also be daunting in large 3D models (Ren *et al.* 2006).

On the other hand, in certain geological domains where the resource concentration has a flat and layer-like geometry that already demonstrates 2D characteristics and features, modelers could choose to model the domain on a 2D grid as an alternative to a 3D grid. An example of these geological domains is a stratigraphic formation where multiple strata are laid in parallel to each other over time and show a high degree of horizontal continuity along those layers (Mallet 2004). There are multiple examples of both mineral and hydrocarbon resource concentrations that are located within these types of geological structures (Figure 1.1). In the opening stages of assessing the economic viability of a mineral or hydrocarbon deposit where more advanced applications of geostatistical models such as flow simulation and mine scheduling are still far in the future and the modeling goal is mostly limited to gaining an approximate understanding of the economic potential and volume of resources in a domain, 2D models tend to be highly useful (Rossi and Deutsch 2014).

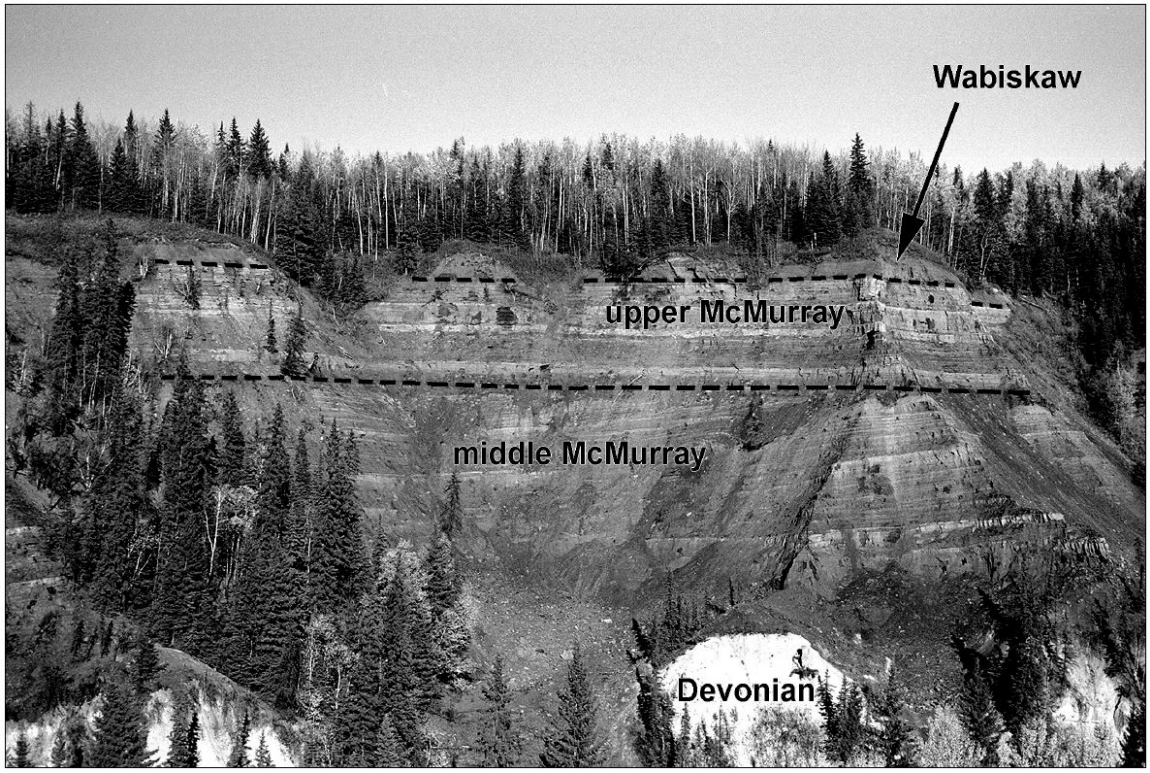

Figure 1.1: A photo taken from the Steepbank River in the McMurray formation of Northern Alberta, an example of a stratigraphic domain where multiple horizontal layers are laid in parallel to each other over time. The upper McMurray, middle McMurray, Devonian, and Wabiskaw layers can be seen here. This domain is the case study for this research and is a suitable candidate for 2D modeling (Langenberg *et al.* 2002).

The main advantage of modeling on a 2D grid rather than a 3D grid is that building 2D models take much less time for the modeler due to reduced complexity and smaller size of the domain and data sets (Ren *et al.* 2006). Additionally, this leads to less possible ground for errors in parameter selection and modeling decisions in the 2D workflow. However, transforming samples from a 3D to a 2D setting would inevitably lead to a loss of information in the process. Using samples initially collected in a 3D space on a 2D grid requires some form of grouping or averaging of the original samples until a single sample would be representing each well/borehole in the 2D data set. This grouping of samples from the 3D space to the 2D space leads to a reduction in the data set size that carries some inevitable loss of information with it. For instance, this would potentially remove many outlier samples that might indicate the presence of high-grade zones. Additionally, the use of these 2D models is limited

to resource estimation as most other applications of geostatistical models requires a full 3D model of the subsurface.

In most modern geostatistical analyses, rather than creating a single model using estimation methods such as Kriging or Inverse Distance Weighting, multiple models or realizations are constructed that are meant to honor the input statistical properties of the data, like reference histogram and variogram (Pyrcz and Deutsch 2014). Here, uncertainty in the estimations is quantified by these equiprobable realizations, where each realization would yield a different estimation of the resource and the full distribution of these estimations gives a range of estimations for that resource (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014; Khan and Deutsch 2015)(Figure 1.2). While in the past these realizations used constant input statistics and uncertainty was recognized only as the ergodic fluctuations from Monte-Carlo simulations, it is now understood that such an approach would lead to an underestimation of uncertainty (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014; Khan and Deutsch 2015). Modern geostatistical workflows are expected to also account for uncertainty in input statistics to perform a more comprehensive uncertainty analysis (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014; Khan and Deutsch 2015).

One of the reasons that 2D models are a suitable and frequently used option under the right geological conditions is the fact that 2D and 3D models tend to produce highly similar results when it comes to the averages of multiple resource estimations without parameter uncertainty (Mondal *et al.* 2015). However, the two workflows have been observed to generate significantly different results when it comes to the uncertainty in probabilistic resource estimations, usually with 2D models typically showing higher uncertainty in the results. Multiple undocumented studies have repeatedly demonstrated this disparity in different geological domains and with different modeling choices and frameworks (Figure 1.3).

Because the uncertainty of the distribution of resource estimations is a critical economic parameter that acts as the basis of assessing the risk associated with investing in a potential mineral or hydrocarbon resource, this disparity between uncertainty in the results of the two workflows makes the choice between them a highly consequential and critical decision for the modeler (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014; Khan and Deutsch 2015). If the 3D model is underestimating uncer-

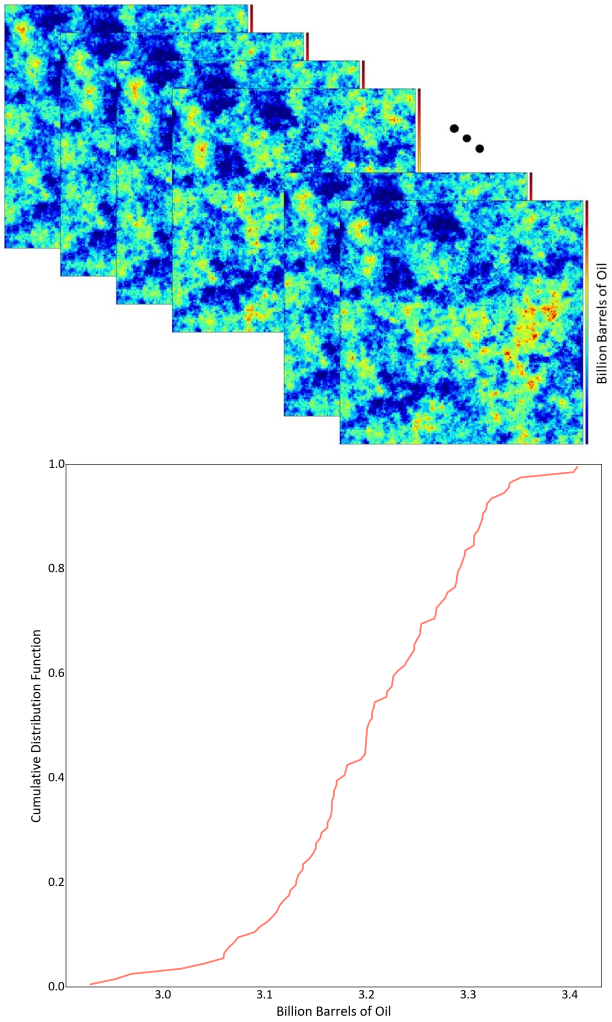

Figure 1.2: An illustration of probabilistic resource estimation, where instead of just one model of a variable of interest (Original Oil In Place in this figure), multiple models are created (the top plots), and their distribution is used to estimate a resource (the bottom plot).

tainty, it would be in turn underestimating the risk associated with the deposit or the reservoir, while if the 2D workflow is overestimating uncertainty, it would be giving an unrealistic range of positive or negative outcomes. Both cases are hazardous to any investment in a mineral or hydrocarbon resource and can greatly diminish its profitability.

The long-lasting negative impacts of an incorrect assessment of uncertainty motivate the need to reconcile and explain any disparity that exists between 2D and 3D models. Since there are certain advantages with both workflows, modelers should be able to make an informed decision when it comes to the uncertainty of the results

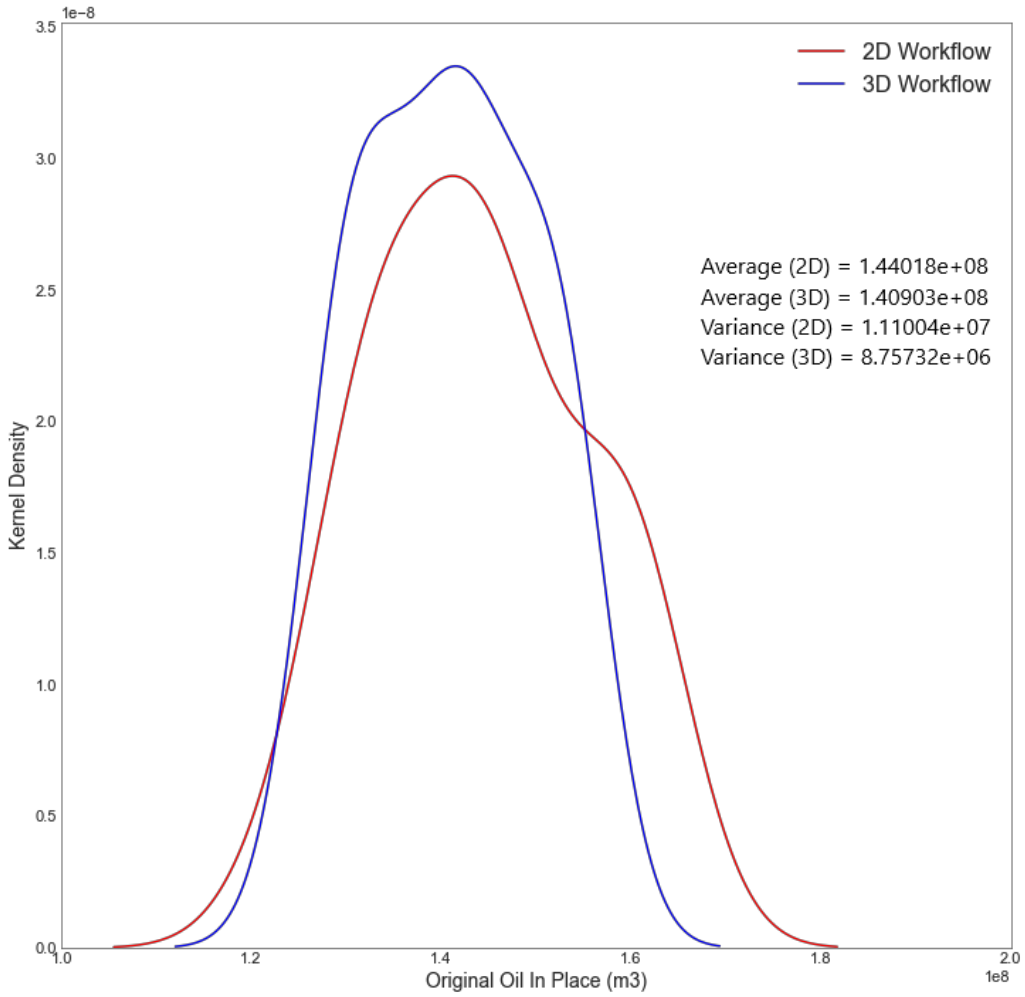

Figure 1.3: An example of estimations of Original Oil In Place (OOIP) (m^3)from 100 realizations in 2D (red) and 3D (blue) workflows (without parameter uncertainty), where the averages of the 100 estimations are almost equal but the variances show higher uncertainty in 2D results.

generated by these workflows. The purpose of this research is to explain disparities and differences in the quantification of uncertainty in 2D and 3D models. Due to the importance of uncertainty in model parameters and input statistics, the two models are compared with consideration of all relevant aspects of parameter uncertainty (Pyrcz and Deutsch 2014; Khan and Deutsch 2015).

1.2 Literature review

In this section, an introduction to the existing literature on probabilistic resource estimation uncertainty is added with previous attempts at comparing 2D and 3D modeling of geological domains using geostatistical tools to identify the research gap that this thesis aims at addressing. Additionally, the review of existing works on parameter uncertainty in geostatistical analysis helps identify significant aspects of parameter uncertainty that will be incorporated into this research and the right approach to quantifying their contribution.

1.2.1 Resource modeling

There are generally two approaches to natural resource modeling, geological and geostatistical modeling (Pyrcz and Deutsch 2014). The theory of geostatistical modeling was developed based on a need for numerical modeling of geological phenomena in pioneering works such as Krige (1951) and Journel and Huijbregts (2003), which have laid the foundation of geostatistical modeling. The basis of geostatistical modeling lies in spatial variability and dependency of variables, as laid out by Hohn (1991) and David (2012). This theory is itself based on the theory of the regionalized variables laid out by Matheron (1971) and is the basis for estimation methods such as Kriging (Krige 1951). In this context, samples of spatial variables are used to determine the covariance structure to provide estimated values of those variables in the entire domain (Matheron and Blondel 1962). This theory was developed further by Journel and Huijbregts (2003) into the foundations of resource modeling and estimations using geostatistics. As an alternative to estimation methods, Journel (1994a) and Journel (1996) laid out the methodology for resampling from stochastic simulations, which is the theoretical framework for simulation methods that exist today and allows for the incorporation of uncertainty into the geostatistical modeling workflow.

1.2.2 2D modeling

Multiple previous works have explored the advantages of 2D modeling and developed the workflows to implement it. Langenberg *et al.* (2002) detailed a 2D mapping approach to model large domains to generate resource estimations, specifically with

a focus on the oil sands of Northern Alberta, and how it could be implemented as an alternative to 3D modeling due to its simplicity. Bertoli *et al.* (2003) also reviewed the process of implementing 2D modeling on the 1A Shoot Nickel deposit in Western Australia and found significant advantages in 2D modeling, such as the lack of bias due to non-additivity. Mondal *et al.* (2015) reviewed the process of 2D geostatistical modeling in Western Onland Oil Field, India and compared the result of Kriging and E-Type mean from simulated realizations in 2D with the results from 3D modeling and found similarity in the averages, but no comparison of uncertainty quantification was provided. Multiple papers such as Harris and Perkins (1991) and Wang *et al.* (2022) have reviewed the process of upscaling 2D models and images to a 3D framework or building 3D models based on 2D images.

1.2.3 Parameter uncertainty

There has been extensive research in the areas of parameter uncertainty and uncertainty quantification in geostatistical modeling in general. Since early works in this area such as Journel (1996), Ballin *et al.* (1992), and Journel (1994b), geostatistics with uncertainty have grown to new heights and specific attention is being paid to parameter uncertainty (Chiles and Delfiner 2009). Wang and Wall (2003) show that if no parameter uncertainty is included in the workflow, the global uncertainty in the results is understated, and Babak and Deutsch (2009) further point out that a lack of parameter uncertainty leads to fluctuations above and below the average to cancel out and imply a small uncertainty in the final results. These works and others such as Khan and Deutsch (2015) and Wang *et al.* (2022) all point out the importance of including parameter uncertainty in a geostatistical modeling workflow.

One approach toward parameter uncertainty is to use maximum likelihood and Bayesian inference methods. Dowd and Pardo-Igúzquiza (2002) developed a method whereby inferring model parameters through maximum likelihood algorithms provides a way to estimate parameter uncertainty by including them in conditional simulation procedures. Diggle and Ribeiro Jr (2002) also developed a similar method where through Bayesian inference, model parameters and the uncertainty associated with their selection were captured and incorporated. Other than these Bayesian-based methods, another common method of accounting for parameter uncertainty

(specifically uncertainty in the reference distribution) is to resample the reference distribution, which leads to a measure of histogram uncertainty. Spatial bootstrap is the most common method in this category, which is developed on an earlier version of the bootstrap (Journel and Bitanov 2004; Efron and Tibshirani 1994). In this method, the reference distribution is resampled through several unconditional simulations and based on the spatial continuity of the variables. However, it has been demonstrated that spatial bootstrap leads to excessive uncertainty, and in response, other methods have been developed such as Babak and Deutsch (2009) that develops a conditional finite domain (CFD) method and Khan and Deutsch (2015) that focuses on the posterior uncertainty after the simulation workflow as a reference to the true uncertainty.

Another important aspect of parameter uncertainty is variogram uncertainty, which is in essence uncertainty in the covariance structure that is the input to many spatial unbiased predictors, such as Kriging (Wang and Wall 2003). Since the covariance structure, like the reference distribution, is derived from the sampled data set and is not inherently and exactly known, the experimental variogram and the variogram models that are based on this covariance structure are also uncertain (Wang and Wall 2003). While there have been many works in this area to incorporate and include variogram uncertainty in resource estimations, in many cases, especially in domains that are densely sampled, the true global uncertainty arising from variogram uncertainty is found to be often negligible (Putter and Young 2001).

1.3 Thesis outline

This research is focused on providing a detailed and comprehensive comparison of uncertainty quantification between 2D and 3D geostatistical modeling workflows, specifically as it relates to probabilistic resource estimation. This is achieved through both numerical and analytical comparison of uncertainty quantification in different aspects of parameter uncertainty between the two workflows and providing explanations into possible differences and contributing factors to these differences.

To explain differences and make comparisons, 2D and 3D geostatistical modeling is implemented with a full range of parameter uncertainty and aspects of uncertainty are

grouped into three main categories: uncertainty in model parameters, uncertainty in geometry, and residual uncertainty (the uncertainty in the results that are generated using no parameter uncertainty, which primarily comes from random fluctuations of Monte-Carlo simulations (Journel and Huijbregts 2003; Journel and Kyriakidis 2004a)). The results are compared between 2D and 3D workflows both using analytical tests and a numerical case study at the McMurray formation of Northern Alberta.

Chapter 2 focuses on the methodology of 2D and 3D geostatistical modeling and the differences between the two workflows in probabilistic resource estimation. Chapter 3 presents numerical results from the case study while Chapter 4 focuses on the results of some test cases that are set to compare histogram and residual uncertainty in 2D and 3D workflows, which provides some insight into differences between the two. Chapter 5 discusses a new methodology for pairing samples to provide a more optimal variogram, which was developed and used through this research as the need for better variogram calculation came up. Finally, Chapter 6 presents the overall results and conclusions and possible future works.

The case study for this research is at the McMurray formation of Northern Alberta. The Lower Cretaceous McMurray formation in north-eastern Alberta is home to one of the largest oil accumulations in the world, the Athabasca oil sands (Masliyah *et al.* 2011). Proven oil reserves in the entire province of Alberta are estimated at 165.4 billion barrels (bbl), while most of the profitable reserves lie in the Athabasca oil sands (Masliyah *et al.* 2011). The oil sands are currently under two main methods of production: surface mining and in-situ extraction methods (Masliyah *et al.* 2011). The McMurray formation sequence in the Athabasca oil sands are remnants of coastal plains of a large fluvial drainage system (Ren *et al.* 2006). This formation has an average thickness of 50 meters and mostly consists of unlithified quartz sand and shale (Ren *et al.* 2006). This formation is highly uncemented, with porosity going as high as 35 percent (Ren *et al.* 2006). There are three main facies in this formation which in ascending order are: Thick-bedded sand facies, Epsilon cross-stratified facies, and Argillaceous sand facies (Ren *et al.* 2006).

Chapter 2

Methodology

In this chapter, the methodology followed in this research is explained with each step in the geostatistical modeling process in both 2D and 3D settings described in detail. While the steps are relatively general to any natural resource concentration with the potential of modeling on a 2D grid, many steps and decisions are specific to the McMurray formation, which is the subject of the case study and focus of this research.

Figure 2.1 demonstrates the geostatistical modeling workflow and resource estimation in a 2D setting. In this figure, the green box refers to the process of generating simulated realizations of the continuous variables on a 2D grid, which is laid out in more detail in Figure 2.3. Continuous variables are the spatial variables that are used in resource estimation and are sampled in the domain. The yellow box refers to the process of modeling thickness (m) (the vertical length of the reservoir), which is simulated as a continuous variable in a workflow like the one depicted in Figure 2.3 on a 2D grid. Realizations of thickness are the representation of the domain's volumetric geometry on a 2D grid, and these realizations are used along with those of the continuous variables to generate resource estimations.

Figure 2.2 presents the same overview of geostatistical modeling and resource estimation on a 3D grid. The green box in this figure lays out the process of simulating continuous variables (Figure 2.3), and the yellow box refers to the process of modeling the geometry of the domain in a 3D grid, which is a combination of modeling bounding surfaces and thickness. In the 3D workflow, simulations of continuous variables are generated separately for each rock type or facies and are clipped to realizations of

the rock type, in addition to being limited to simulated realizations of the geometry of the reservoir.

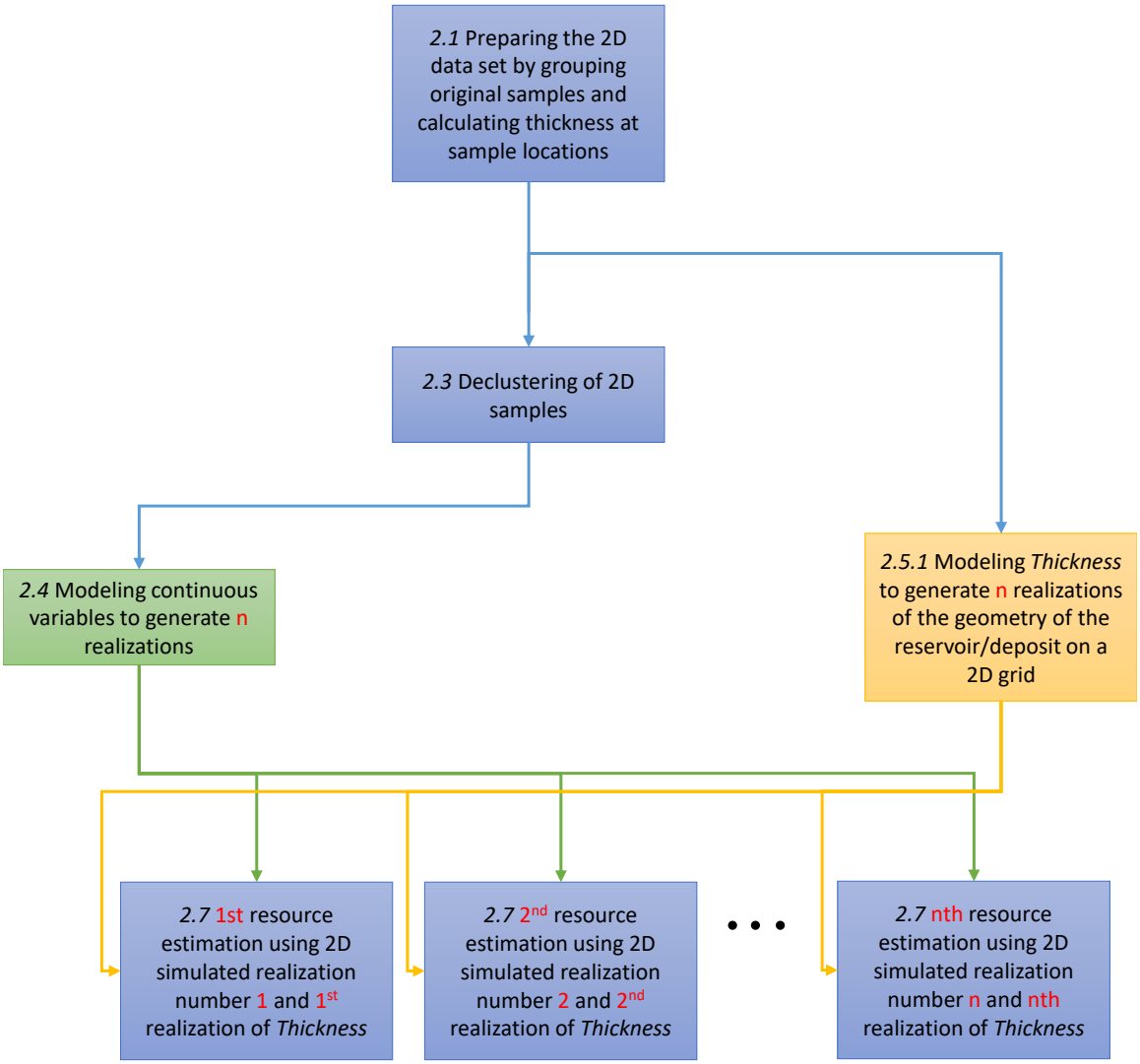

Figure 2.1: A general depiction of geostatistical modeling on a 2D grid. The steps are numbered with the corresponding following sections.

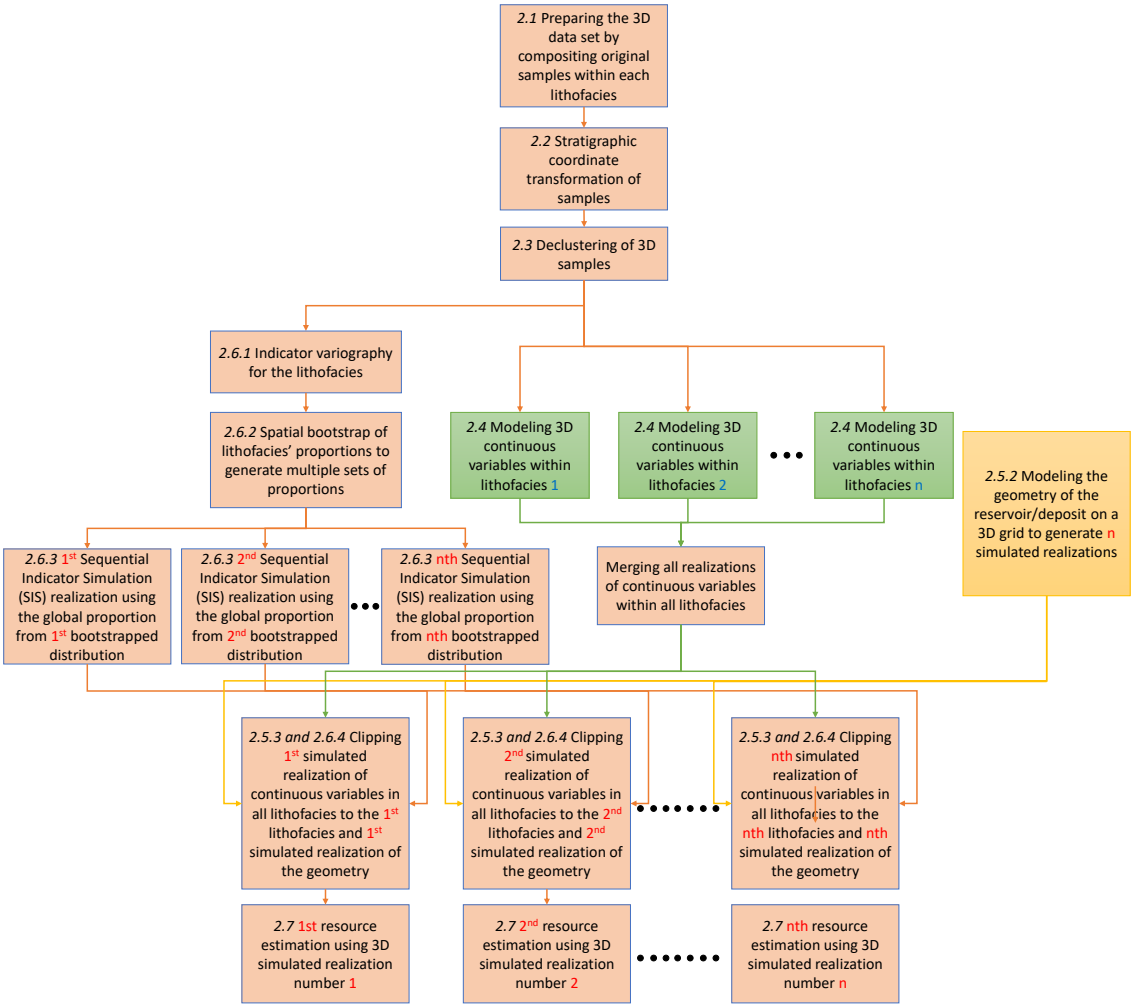

Figure 2.2: A general depiction of steps in geostatistical modeling on a 3D grid. The steps are numbered with the corresponding following sections.

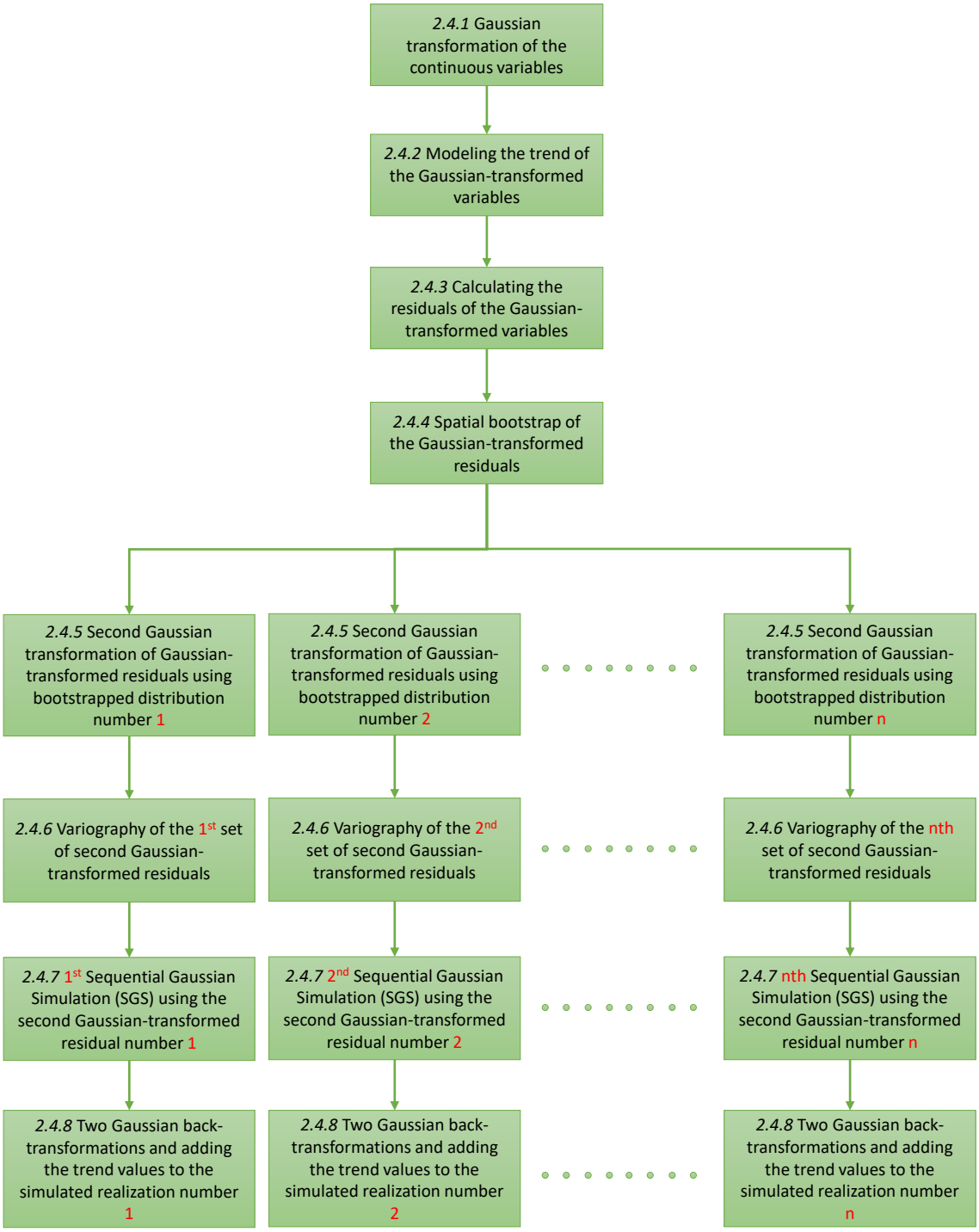

Figure 2.3: A general depiction of steps in the simulation of continuous variables. The steps are numbered with the corresponding following sections.

The case study is on an 8*8-kilometer domain in the McMurray formation of North-

ern Alberta. There are 393 vertical wells and 268549 samples within this domain of two continuous variables, effective porosity (ϕ) (ratio) and oil saturation (So) (ratio), as well as five lithofacies as the categorical variable (clean sand (1), sandy HIS (2), sand with 30-70 percent mud (3), mud (4), and breccia (5)). Elevations of the top of the continuous reservoir (TCR) (m) and the bottom of the continuous reservoir (BCR) (m) are also available for each well, which in turn leads to values of thickness for each vertical well, which is the third continuous variable in the 2D workflow. The workflows are implemented in 100 realizations using aerial grid dimensions of 50*50 meters in both 2D and 3D workflows and 1 meter in the vertical direction for the 3D workflow.

2.1 Preparing data sets

The first step is to transform the input samples collected from the domain into 2D and 3D settings. Figure 2.5 demonstrates a plan view of the vertical wells in the domain. The 3D data set is prepared by compositing hard data samples taken from wells. This includes both categorical, such as facies or rock types, and continuous spatial variables, such as effective porosity and oil saturation. This step also includes other auxiliary tasks such as data cleaning and outlier management. In the case study, samples are composited to 1-meter intervals and with a 40 percent minimum acceptance rate.

The 2D data set is generated by combining samples along the length of the continuous bitumen. In the case study, samples within the continuous reservoir of the McMurray formation are averaged to form the 2D data set (Figure 2.4), with one sample representing each well.

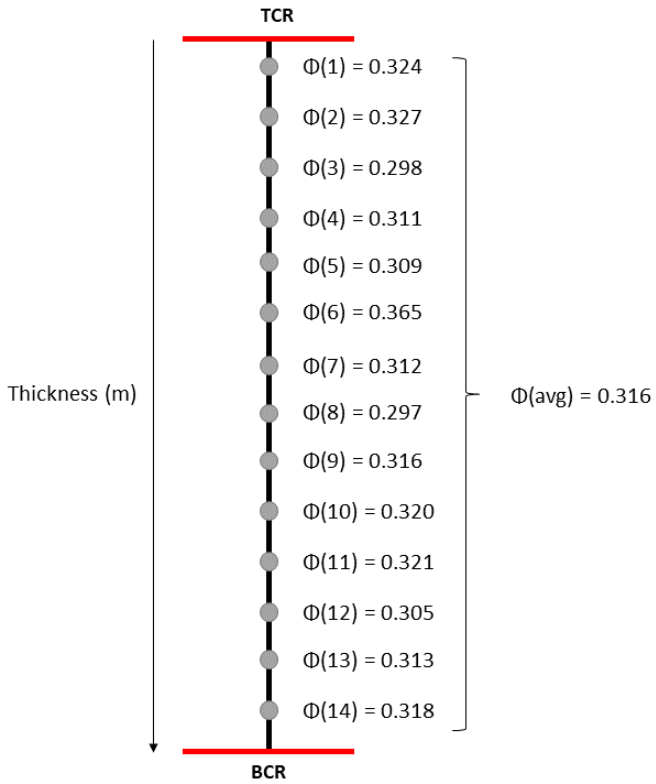

Figure 2.4: An example of a well with 14 samples of effective porosity between the *TCR* and *BCR* surfaces. These samples are averaged to produce a single sample in the 2D data set. The difference between the elevation of the *TCR* and *BCR* surfaces is the thickness of the reservoir in that well.

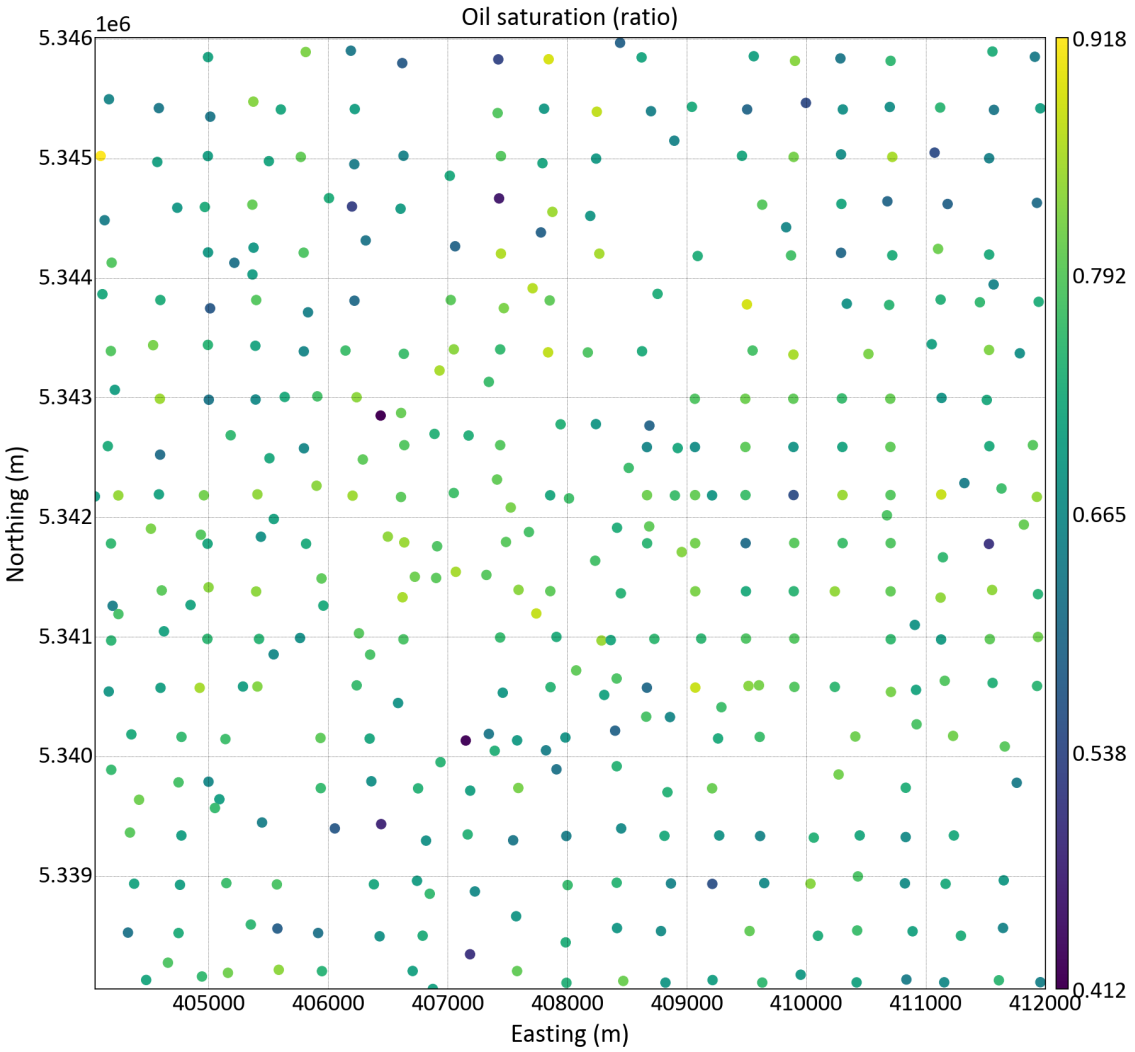

Figure 2.5: A map of the spread of sample wells in the domain, with averaged values of oil saturation in the 2D data set.

When a spatial variable that is sampled in a 3D space is transformed into a 2D setting by averaging those initial samples, the distribution of the values of that spatial variable would end up with lower variability. In this process, many of the extreme values that define the upper and lower ends of the 3D distribution are merged with other samples. Figure 2.6 illustrates an example of this, where the 2D and 3D distribution of effective porosity in the case study are plotted together.

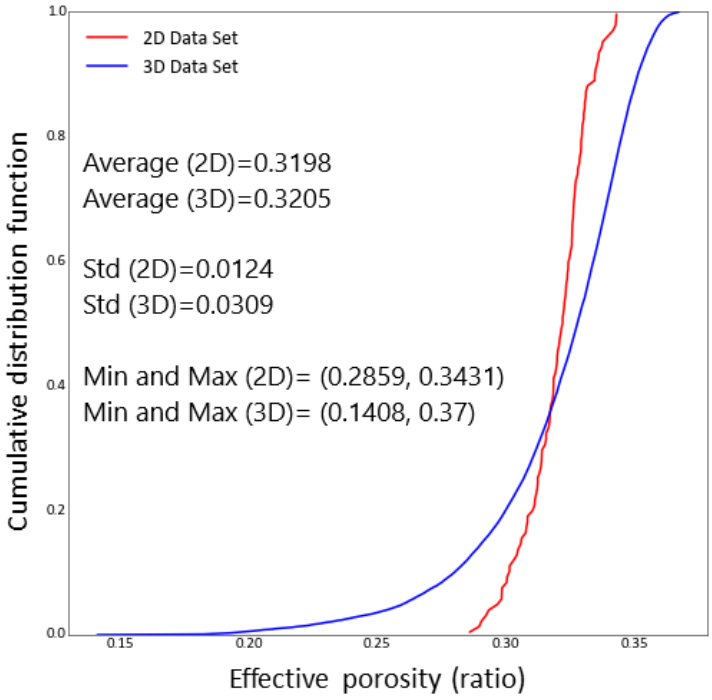

Figure 2.6: The distribution of effective porosity in the 2D and 3D data sets. The average is the same in both settings but the variance is smaller in 2D, which is due to averaging.

Apart from effective porosity and oil saturation, another variable that is used in the 2D workflow is thickness. If the angle of the intersection of the well in the ground is not perfectly vertical, the detected length might have to be converted to the true thickness (Bertoli *et al.* 2003). In the 2D workflow, thickness represents the geometry of the target resource purely from a volumetric perspective and is modeled along with the other variables as a continuous variable (Figure 2.7).

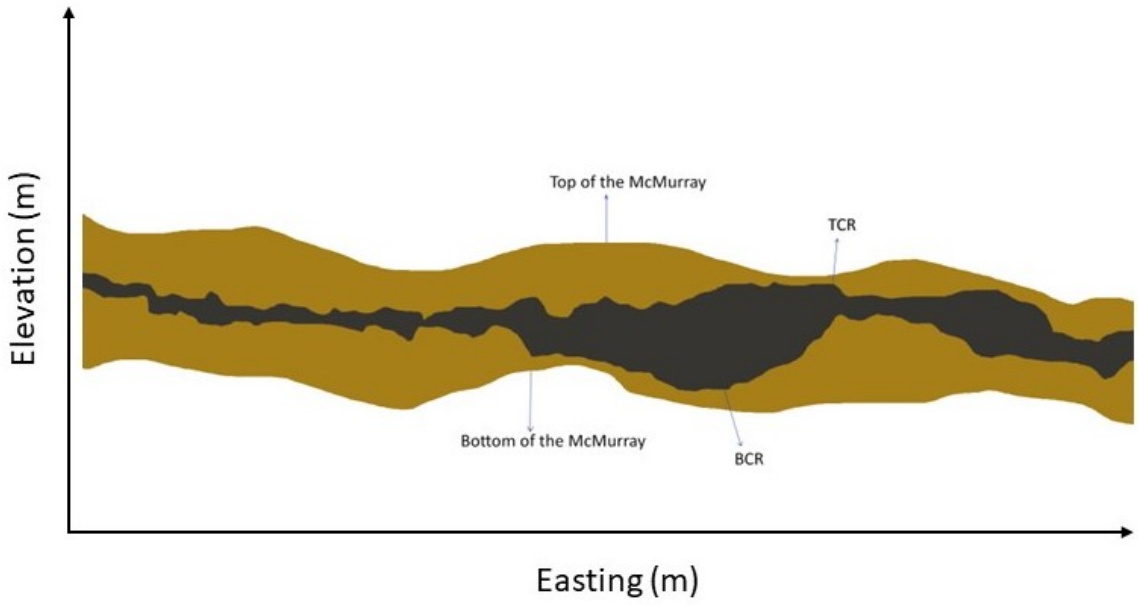

Figure 2.7: A conceptual example of the cross-section view of the McMurray formation. The black area is the continuous reservoir with the TCR and BCR surfaces pointed out in the picture. The outer brown area is the entire McMurray layer.

2.2 Stratigraphic coordinate transformation

In stratigraphic formations, such as that of the McMurray formation of Northern Alberta, one necessary step in geostatistical modeling is to transform the coordinates of samples to reflect the geological genesis of the formation. This process is called stratigraphic coordinate transformation and involves the alteration of the elevation of samples in relation to one or two bounding surface(s) (Latifi and Boisvert 2022). In the case study, samples are shifted in relation to the top bounding surface of the McMurray formation in the 3D data set. This step is exclusive to the 3D workflow. The improvement in spatial continuity of effective porosity and oil saturation from the data set can be seen in Figure 2.8.

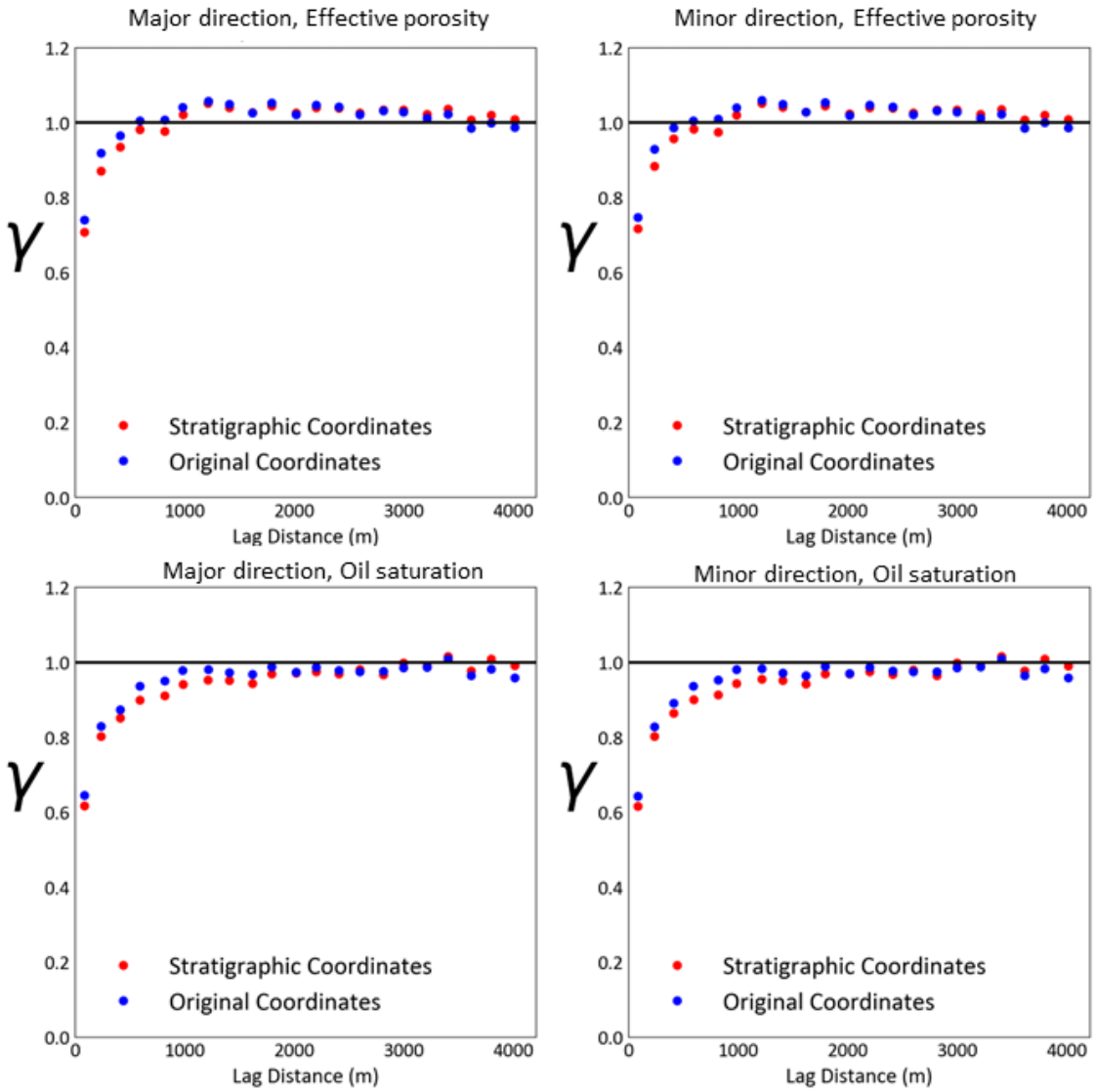

Figure 2.8: The experimental semi-variogram of the variables effective porosity and oil saturation in the major and minor directions (azimuth of 0 and 90 degrees, respectively) with original (blue) and stratigraphic (red) coordinates.

2.3 Declustering

After creating the 2D and 3D data sets, declustering must be implemented to have a representative distribution of variables that is a key input in the Gaussian transformation of variables. While there are different approaches to declustering, cell declustering was used in this research and the case study (Deutsch and Deutsch 2015).

Cell declustering gives higher weights to samples in more sparsely sampled areas and lower weights to samples in more densely sampled regions. This is accomplished by dividing the sampled area into equal-length squares and giving equal weight to all samples within each square which leads to densely sampled squares sharing the same weight between more samples (Deutsch and Deutsch 2015).

While there are differences between 2D and 3D cell declustering, since wells are vertical in the 3D data set, 2D cell declustering is implemented for both workflows with cells of 500*500 meters based on well spacing. The differences between distributions of the variable oil saturation with and without the use of declustering weights can be seen in Figure 2.9.

In some cases, the variables must also go through despiking to facilitate Gaussian transformation in the simulation process but this is not needed in the case study of this research.

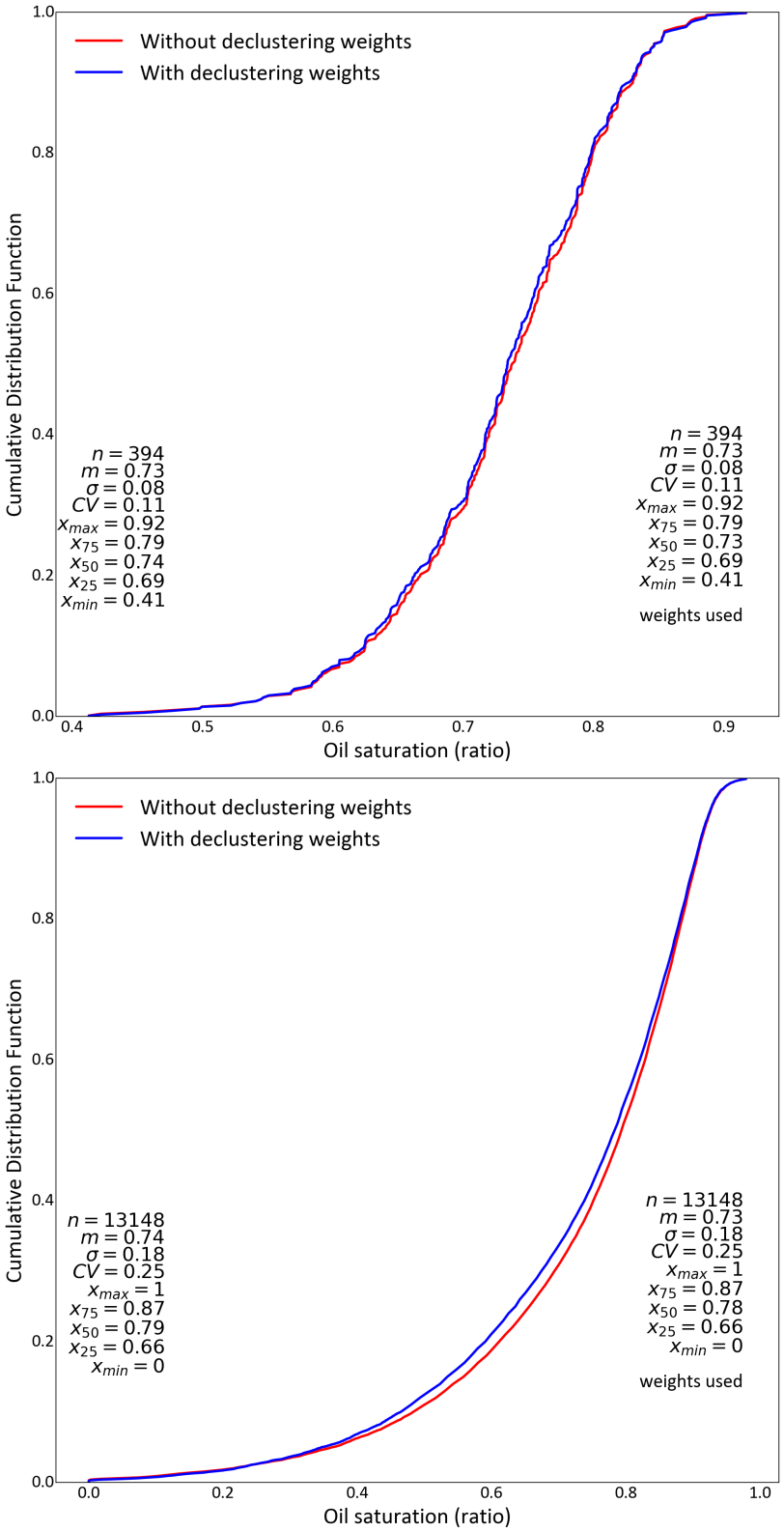

Figure 2.9: Distribution of oil saturation in 2D (top plot) and 3D (bottom plot) data sets with (blue distribution and statistics on the right) and without (red distribution and statistics on the left) the use of declustering weights.

2.4 Modeling continuous variables

After creating the data sets and other necessary pre-modeling steps such as declustering and stratigraphic coordinate transformation, the continuous variables must be modeled in both 2D and 3D workflows. The steps outlined in this section are general to every continuous variable in both 2D and 3D settings, and whenever any continuous variable is modeled in this research, including effective porosity, oil saturation, thickness, or surface elevations, the following workflow is followed regardless of the nature of the variable.

2.4.1 Gaussian transformation

Declustering weights provide a representative distribution of continuous variables. These variables must then be transformed into a Gaussian variable to model them in the next subsequent steps. This transformation is necessary for Sequential Gaussian Simulation (SGS) that leads to simulated realizations of the variables across the domain (Pyrcz and Deutsch 2014).

2.4.2 Trend modeling

To account for the assumption of stationarity in the subsequent modeling steps, the trends of the continuous variables must be modeled across the domain (Journel and Huijbregts 2003; Pyrcz and Deutsch 2014; Rossi and Deutsch 2014). There are multiple approaches to trend modeling and a moving window averaging method is used in this research (Manchuk and Deutsch 2011). It must be pointed out that while the 2D workflow only has an aerial trend model, the 3D workflow requires a 3D trend as well. For the case study, a moving window of 4*4 kilometers with 70 percent of all samples in 2D and 3D workflows was chosen. The results can be seen in Figures 2.10 and 2.12. Figures 2.11 and 2.13 illustrate the swath plots of these trend models in relation to the reference data in the 2D and 3D models, respectively.

In theory, since the trend model uses as input several parameters and is thus subjective to some degree, parameter uncertainty could also extend into trend model uncertainty. In that case, each simulated realization is modeled using a different trend model. In this research, trend model uncertainty was included and since its

contribution to overall uncertainty was negligible, it was not incorporated in the final results.

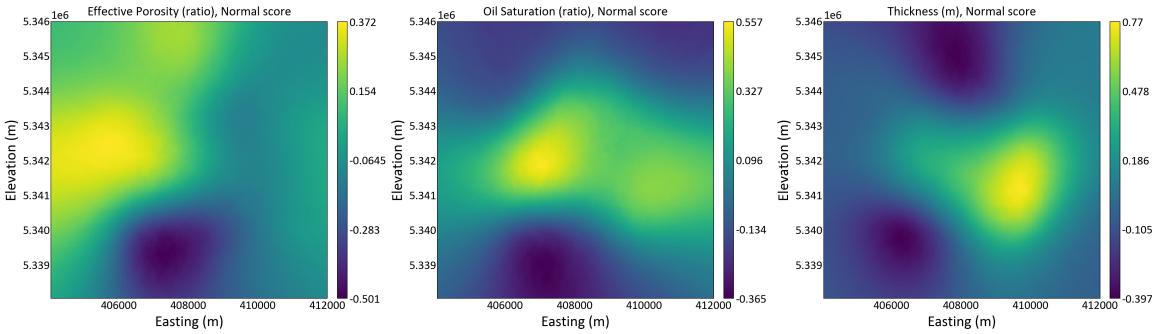

Figure 2.10: The areal trend model of the three variables in the 2D workflow.

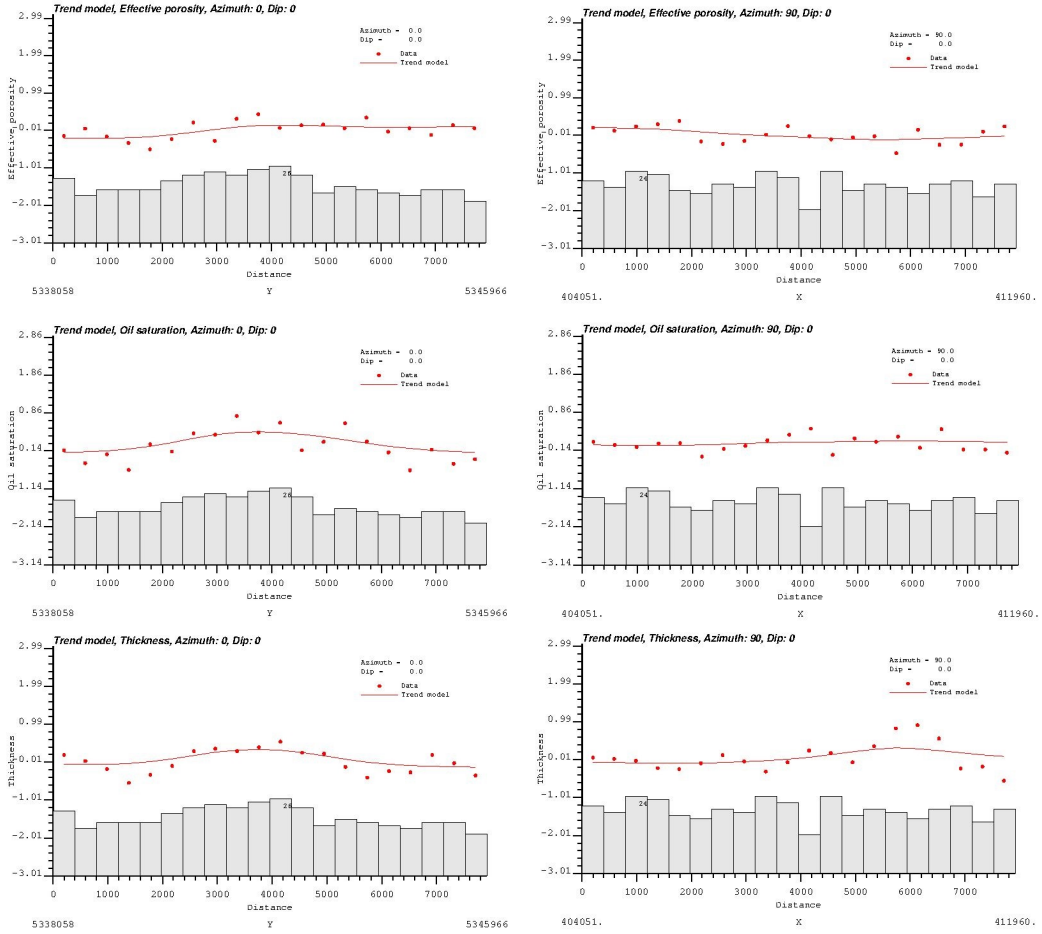

Figure 2.11: The swath plots of the trend models in the 2D workflow in relation to the reference data in the major (0 azimuth) and minor (90 azimuth) directions for the three variables in the 2D workflow.

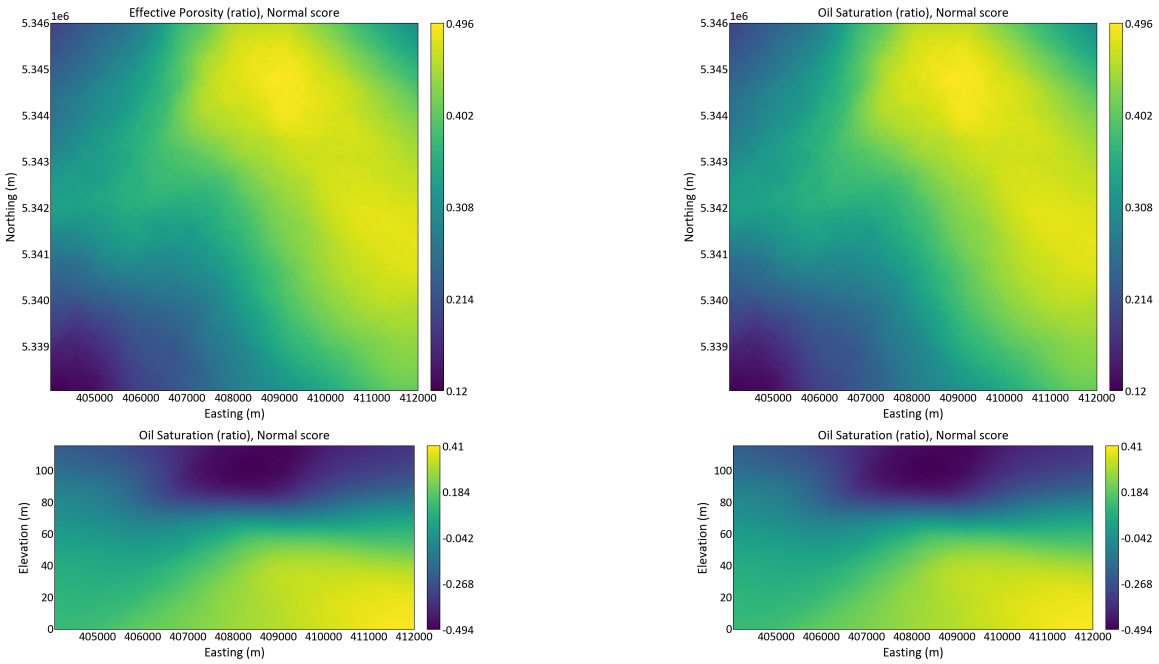

Figure 2.12: The 3D trend model of the two continuous variables in the 3D workflow, with plan views at the top and cross-sections at the bottom.

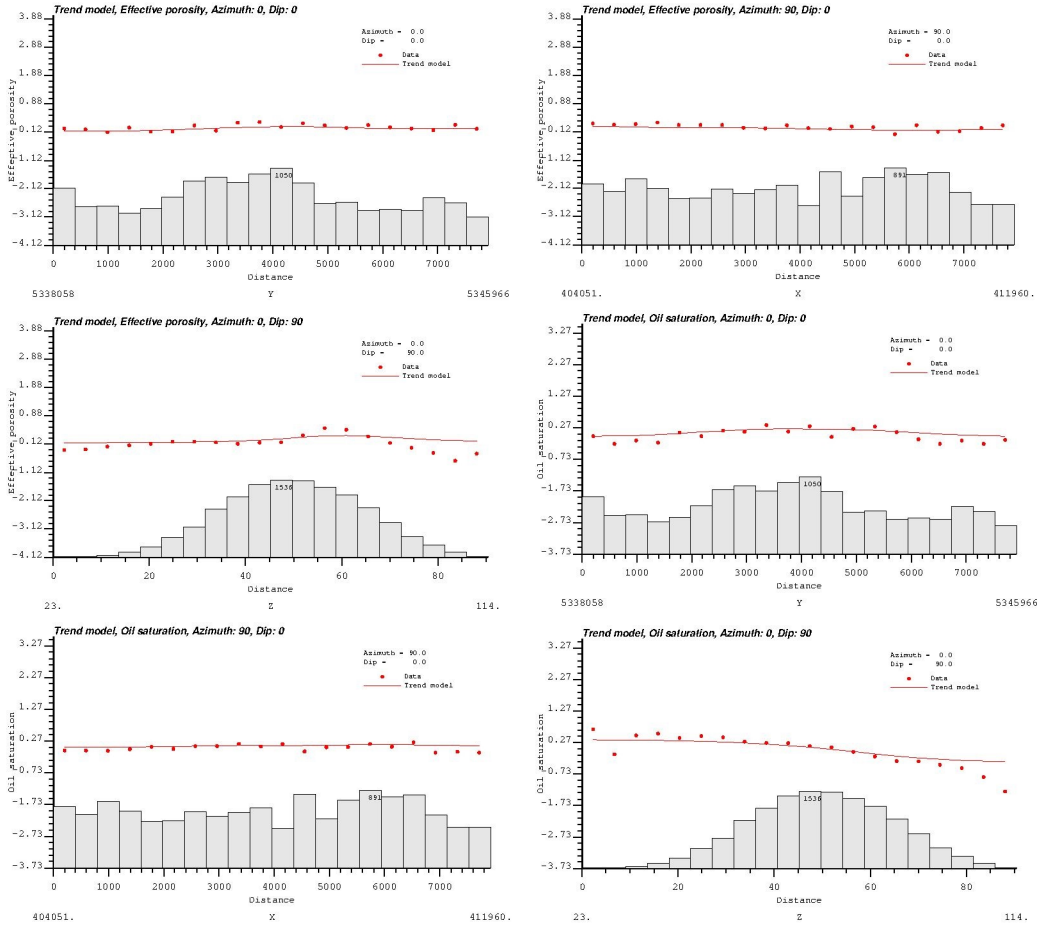

Figure 2.13: The swath plots of the trend models in the 3D workflow in relation to the reference data in the major (0 azimuth), minor (90 azimuth), and vertical directions for the two variables in the 3D workflow.

2.4.3 Residual calculation

After the trend is modeled, the residual of the trend must be derived. In this research, residuals are calculated by subtracting collocated trend values from sample values at sample locations (McLennan 2007; Qu and Deutsch 2018; Wackernagel 2003; Harding and Deutsch 2021).

2.4.4 Spatial bootstrap

As discussed in the introduction, to account for histogram uncertainty, the weighted distribution of the residuals is used along with the variogram of the residuals to generate multiple bootstrapped distributions (depending on the number of intended

realizations) (Journel and Bitanov 2004; Vincent and Deutsch 2019). Figures 2.14 and 2.15 demonstrate these bootstrapped distributions along with the original residual distribution. In order to quantify histogram uncertainty in the final resource estimations, the entire modeling workflow is implemented once with the inclusion of histogram uncertainty (using multiple bootstrapped distributions), and once without histogram uncertainty (using a fixed reference distribution) and the reduction in overall uncertainty is measured to quantify histogram uncertainty.

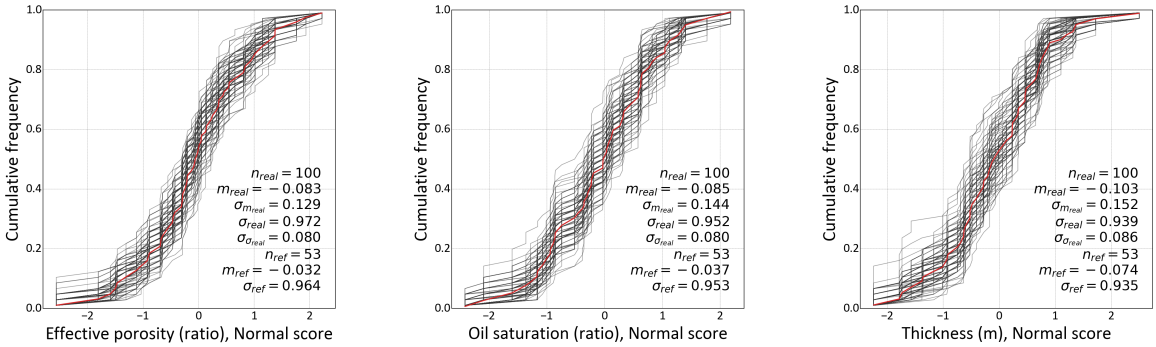

Figure 2.14: 100 bootstrapped distributions (grey distributions) of the residual of the continuous variables in the 2D workflow and the original reference distribution of the residual (red distribution).

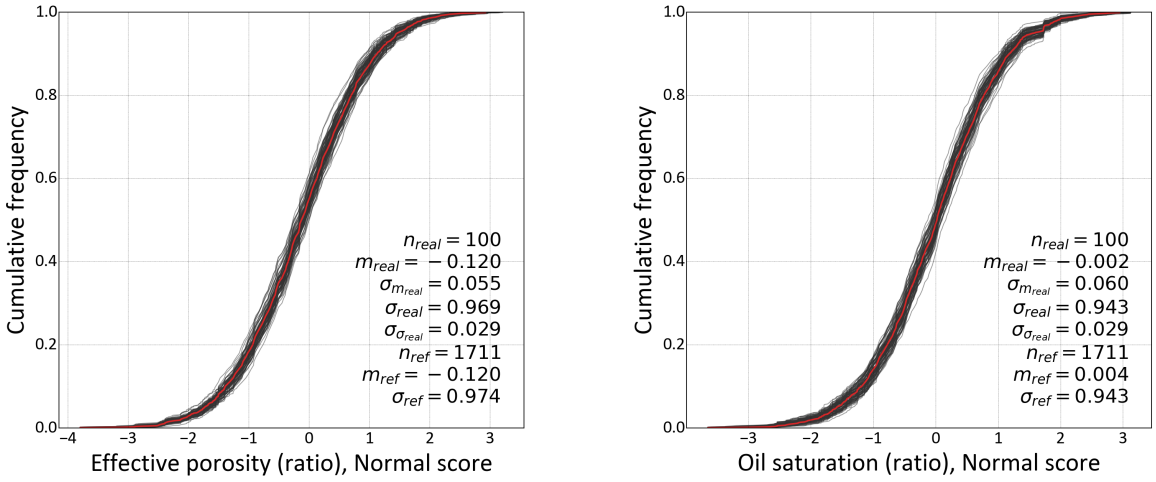

Figure 2.15: 100 bootstrapped distributions (grey distributions) of the residual of the continuous variables in the 3D workflow and the original reference distribution of the residual (red distribution).

2.4.5 Second Gaussian transformation

Using the bootstrapped distributions generated in the previous step, the residuals of the continuous variables are transformed into a Gaussian variable again. This time, however, this transformation takes place multiple times, depending on the number of intended realizations of the domain. Each transformation uses a separate bootstrapped distribution from the previous step as a reference distribution that leads to a separate Gaussian transformed variable.

2.4.6 Variography of the Gaussian residual

As was pointed out in the previous step since multiple bootstrapped distributions of the original residual variable are used to generate multiple sets of Gaussian residual variables, each variable would lead to a different experimental variogram and variogram model. This is utilized to incorporate variogram uncertainty, where each Gaussian residual variable is used with a different variogram model in the SGS workflow to capture variogram uncertainty. The results can be seen in Figures 2.16 and 2.17. As is evident in these figures, the degree of variability in the variograms in the 2D workflow is much higher compared to the variograms in the 3D workflow. This is a direct cause of the lower variability of bootstrapped distributions in the 3D workflow compared to the 2D workflow, as illustrated in Figures 2.14 and 2.15. The larger variability in bootstrapped distributions in the 2D workflow leads to more variability in the Gaussian residual variables generated in section 2.4.5 and thus more variability in the experimental variogram calculated from them.

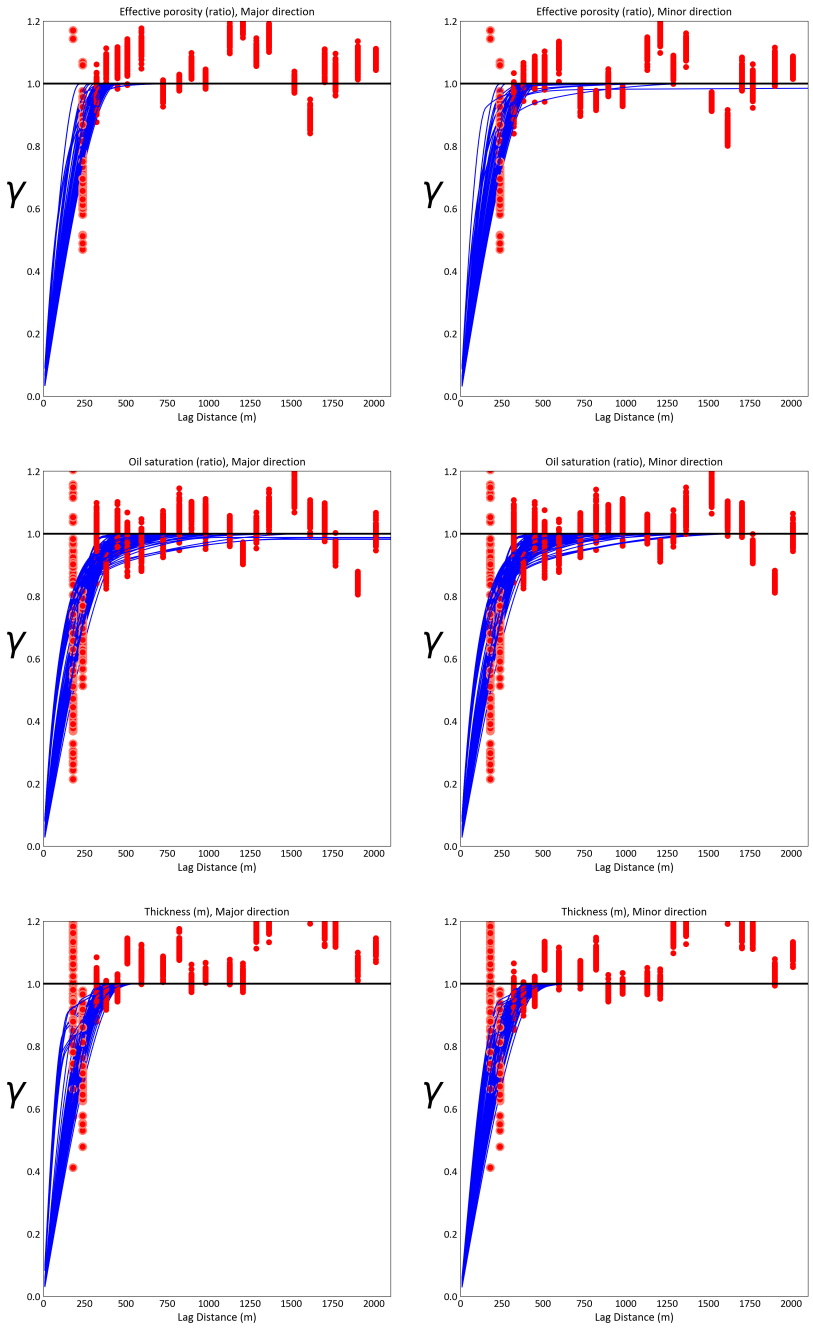

Figure 2.16: The results of Gaussian residual variography in the 2D workflow. The red points indicate the experimental variogram for each Gaussian residual variable and the blue lines refer to the variogram model fit to the variables. Since there are 100 realizations in the case study workflow, there are 100 experimental variograms and 100 variogram models generated in this step here.

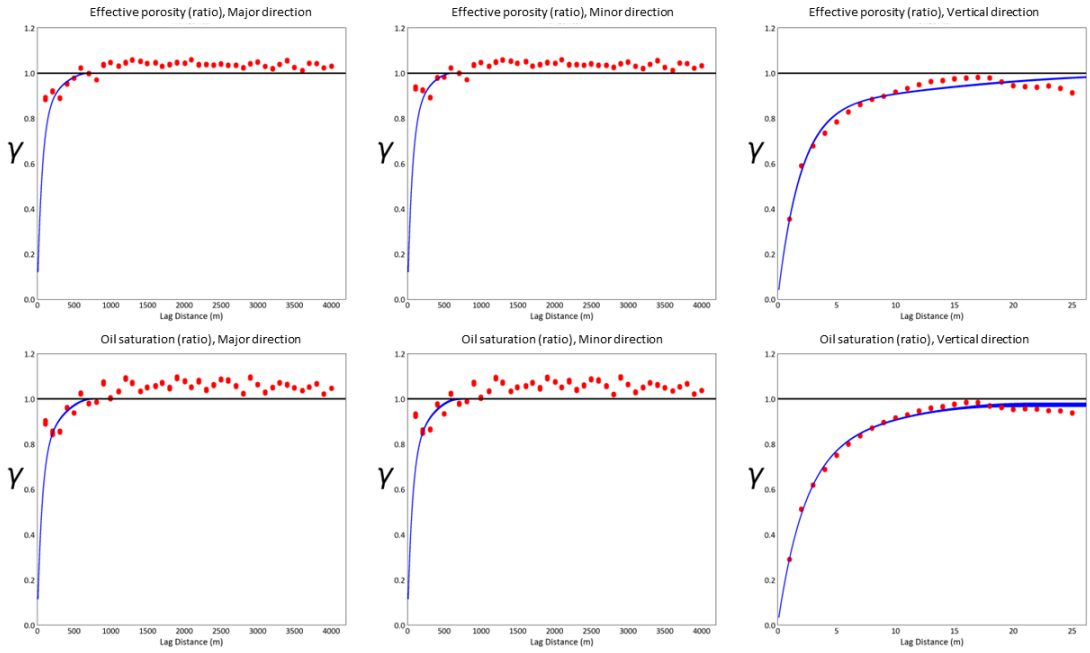

Figure 2.17: The results of Gaussian residual variography in the 3D workflow. The red points indicate the experimental variogram for each Gaussian residual variable and the blue lines refer to the variogram model fit to the variables. Since there are 100 realizations in the case study workflow, there are 100 experimental variograms and 100 variogram models generated in this step here.

2.4.7 SGS

Using each one of the Gaussian residual variables and its corresponding variogram model, SGS is implemented and multiple simulated realizations of the continuous variables are generated (Deutsch, Journel, *et al.* 1992; Gómez-Hernández and Journel 1993; Journel and Kyriakidis 2004b).

2.4.8 Back transformation

The simulated realizations generated in the previous step are in Gaussian residual units and must be back-transformed into original units. To do this, the second Gaussian transformation must first be back-transformed according to the corresponding transformation table generated in section 2.4.5. Subsequently, the back-transformed values which are in original residual units must be added back with the trend values and reverse the process in step 2.4.3. Lastly, the final Gaussian back-transformation is implemented using the transformation table generated at step 2.4.1. The results

of these back-transformations are simulated realizations in the original units of the continuous variable.

2.5 Modeling geological boundaries

Modeling the geometry and boundaries of a geological formation is a critical part of geostatistical modeling and resource estimation and is also very diverse in nature and approach (Pyrcz and Deutsch 2014; Rossi and Deutsch 2014). In general, since there is uncertainty also in the modeling of the boundaries of a geological domain, it must also include uncertainty to reflect a more realistic estimation of resources. In both 2D and 3D workflows in this research, the geometry and boundaries of the domain are modeled in a geostatistical workflow with the inclusion of all aspects of uncertainty.

2.5.1 2D geological boundary modeling

In this research, modeling geological boundaries in the 2D workflow is limited to modeling the thickness of the continuous reservoir. Simulated realizations of thickness, as generated in section 2.4, lead to a 2D understanding of the volume and geometry of the reservoir and are directly used in resource estimation (Figure 2.18). While it might be necessary to model the aerial boundaries of the domain as well, the McMurray formation of the case study does not have a specific aerial lease limit and thus this step is not included in the workflows.

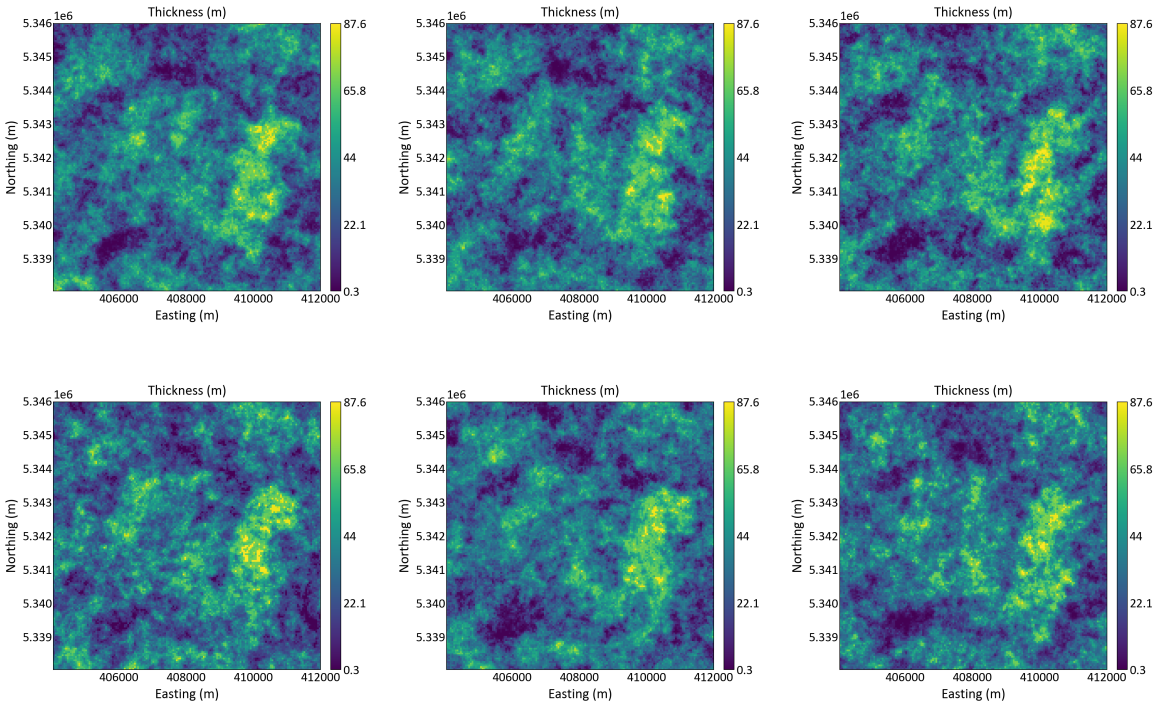

Figure 2.18: Six different realizations of thickness within the same domain. Each realization of thickness leads to a different realization of the geometry of the domain.

2.5.2 3D geological boundary modeling

In this case study and based on the geological information available, realizations of the elevation of the upper bounding surface are simulated on a 2D grid and following the workflow laid out in section 2.4 and are added with the simulated value of thickness at each grid node to infer elevations of the lower bounding surface. This process leads to 3D simulated realizations of the geometry of the reservoir (Figure 2.19). Like the 2D workflow, no aerial limit is conceived for the reservoir.

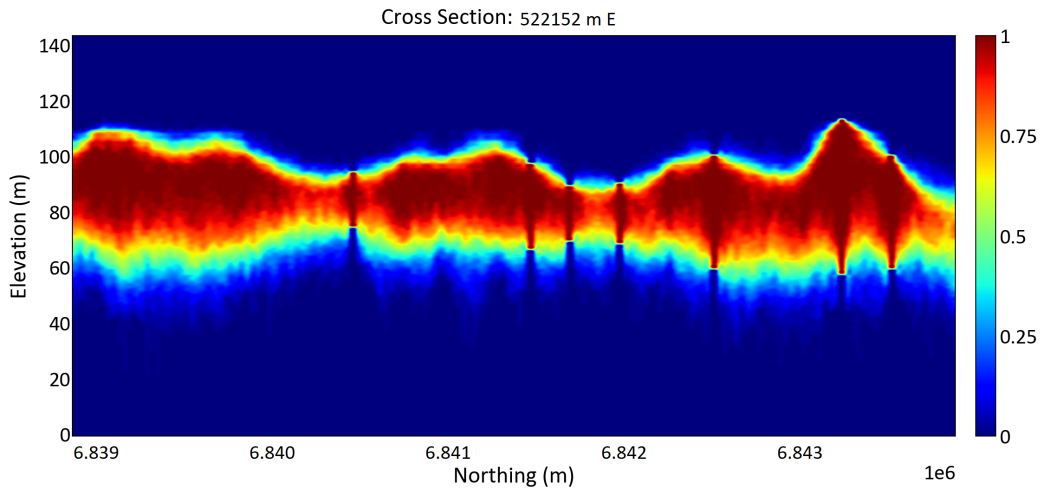

Figure 2.19: A cross-section of the average of 100 realizations of the reservoir in the 3D model. Each point that falls within the boundaries of the reservoir gets an indicator value of 1, and 0 if it falls outside. The averaging of these indicator realizations shows the extent of the variation of the boundary surfaces within these realizations. The dark red vertical lines represent exploration wells. Notice how the wells control the realizations of the reservoir.

2.5.3 Clipping continuous variables to the boundaries of the domain

Only in the 3D workflow, simulated realizations of variables are limited to the area between upper and lower bounding surfaces, as generated in the previous section.

2.6 Modeling categorical variables

In most resource modeling projects, categorical variables, such as lithofacies or rock types, carry significant geological information and are modeled in a geostatistical workflow (Journel and Huijbregts 2003; Pyrcz and Deutsch 2014; Rossi and Deutsch 2014). The 2D data set is the product of averaging variables into a single sample per well and the product of averaging facies in a well does not carry significant geological information and cannot be properly modeled on a 2D grid. Thus, in this research, categorical variable modeling is limited to the 3D workflow. In this section, a general description of categorical variable modeling is explained. In the case study, five main lithofacies form the primary categorical variable of interest: clean sand, sandy IHS, sand with 30-70 percent mud, mud, and breccia.

2.6.1 Indicator variography

In the context of categorical variable modeling, indicator variables are binary variables that indicate the presence of each instance (in this context lithofacies) of the categorical variable at sample locations or simulated grid nodes. Figures 2.20 to 2.24 show the variograms of these indicator variables for the five lithofacies used in the case study.

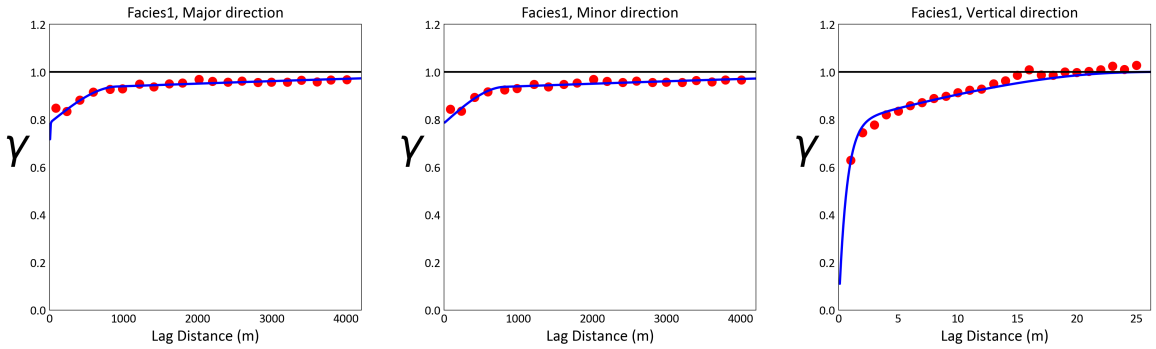

Figure 2.20: The experimental variogram of the indicator variable of the first lithofacies used in the case study and the fitted variogram model.

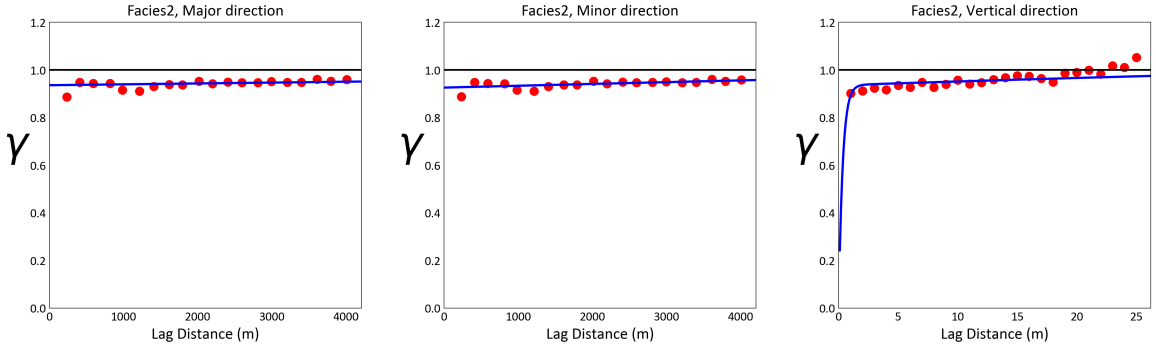

Figure 2.21: The experimental variogram of the indicator variable of the second lithofacies used in the case study and the fitted variogram model.

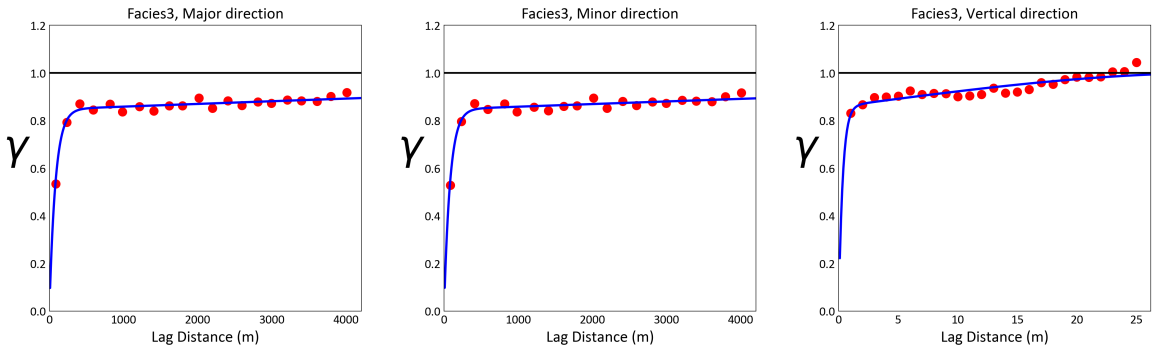

Figure 2.22: The experimental variogram of the indicator variable of the third lithofacies used in the case study and the fitted variogram model.

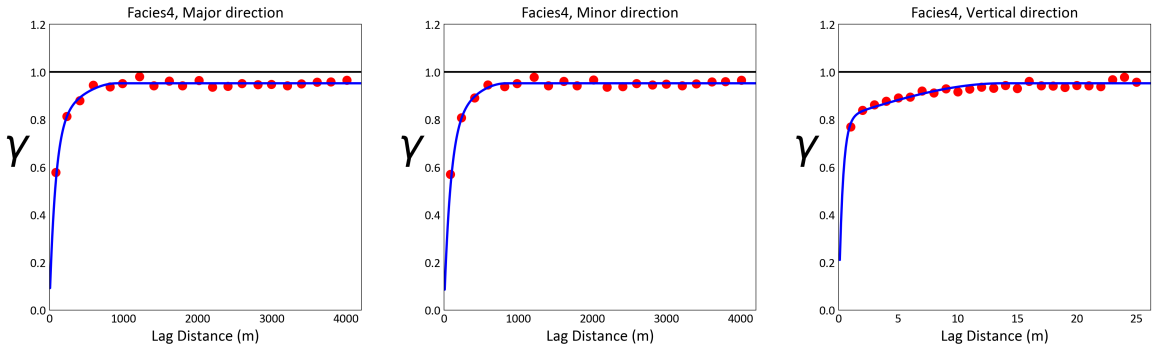

Figure 2.23: The experimental variogram of the indicator variable of the fourth lithofacies used in the case study and the fitted variogram model.

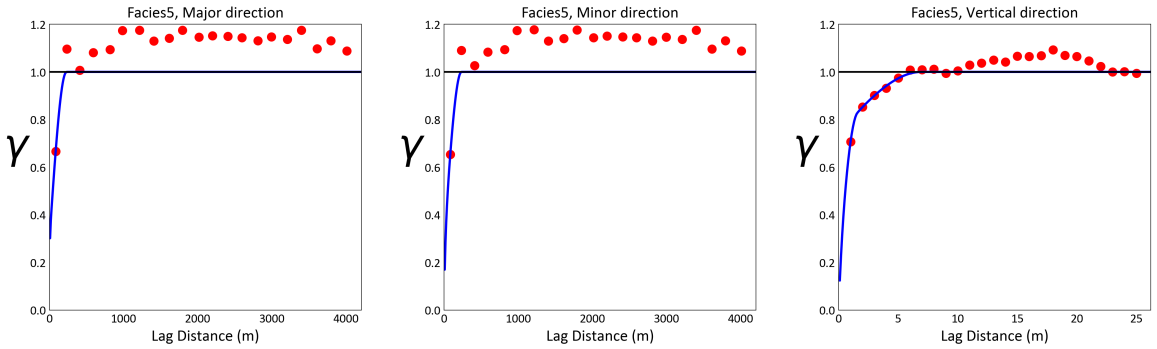

Figure 2.24: The experimental variogram of the indicator variable of the fifth lithofacies used in the case study and the fitted variogram model.

2.6.2 Spatial bootstrap

Similar to the continuous variable modeling workflow in section 2.4.4, spatial bootstrap is also implemented on the categorical variables to generate several different

representative distributions that in turn lead to a number of different global proportions of the categorial variables. These proportions are used as input in the Sequential Indicator Simulation (SIS) workflow to incorporate histogram uncertainty in the categorical variable modeling workflow (Deutsch 2006).

2.6.3 SIS

Using the indicator variogram models, global proportions of the lithofacies generated using the spatial bootstrap workflow, and the original samples collected in the 3D space and composited, simulated realizations of the categorical variables are generated with a trend using the SIS workflow (Figure 2.25 show these trends used in the SIS workflow and Figures 2.26, 2.27, 2.28, 2.29, and 2.30 show the swath plots for these categorical trend models) (Deutsch 2006). In the continuous variable modeling workflow and in section 2.4.6, each realization of spatial bootstrap leads to a separate distribution of reference samples and a separate experimental semi-variogram. The spatial bootstrap workflow in the categorical modeling workflow is implemented differently, with each realization of spatial bootstrap leading to a different set of global proportions for the lithofacies. These proportions are used directly in the SIS workflow. Thus, unlike the continuous variable modeling workflow where there are multiple experimental variograms that could capture variogram uncertainty, a single indicator variogram is used here. In order to incorporate variogram uncertainty, parameters of indicator variogram models are selected randomly within a realistic range of values that fit the experimental variograms but allow for variogram uncertainty to be included.

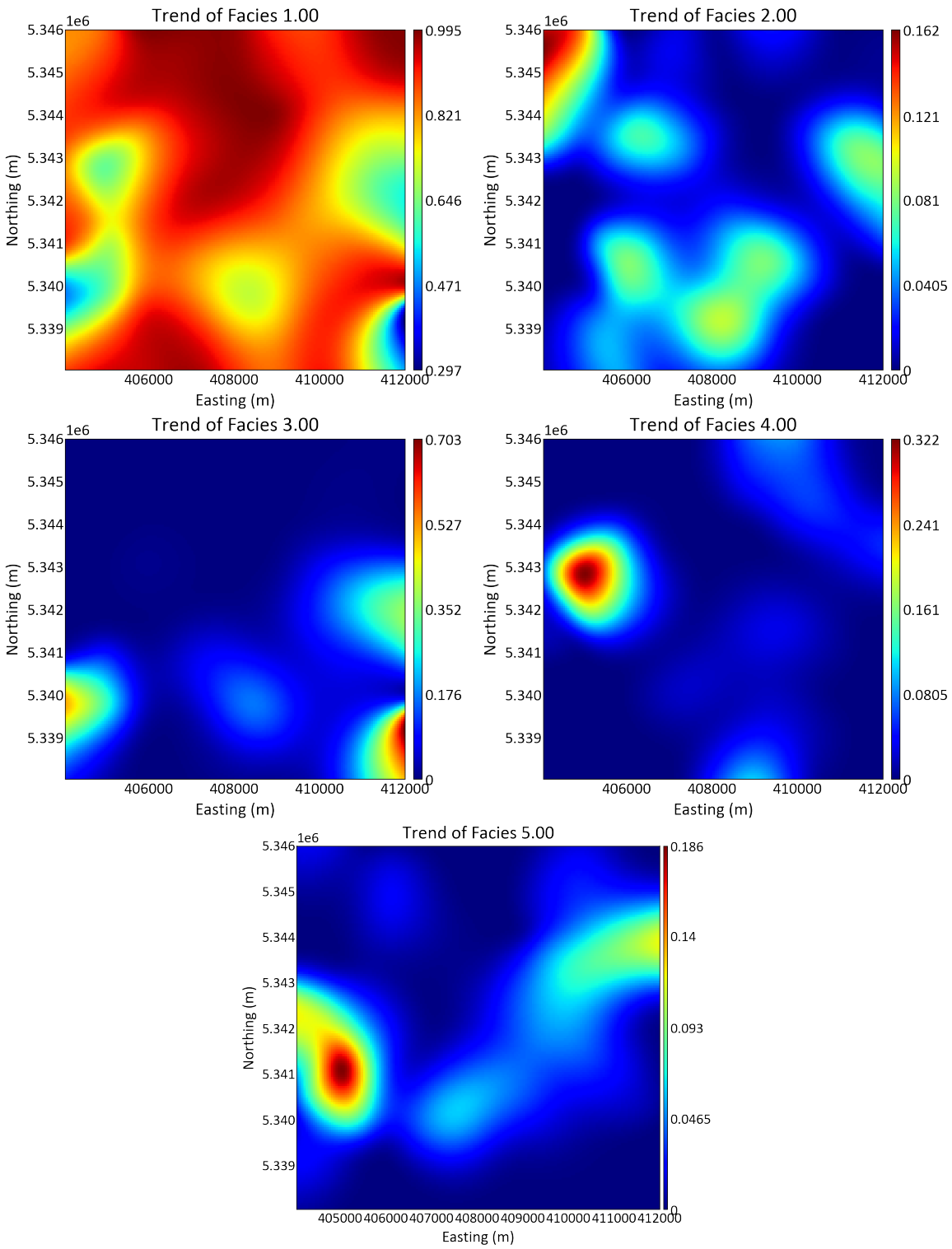

Figure 2.25: A plan view of the trends for the five lithofacies used in the SIS workflow at an elevation of 55 meters.

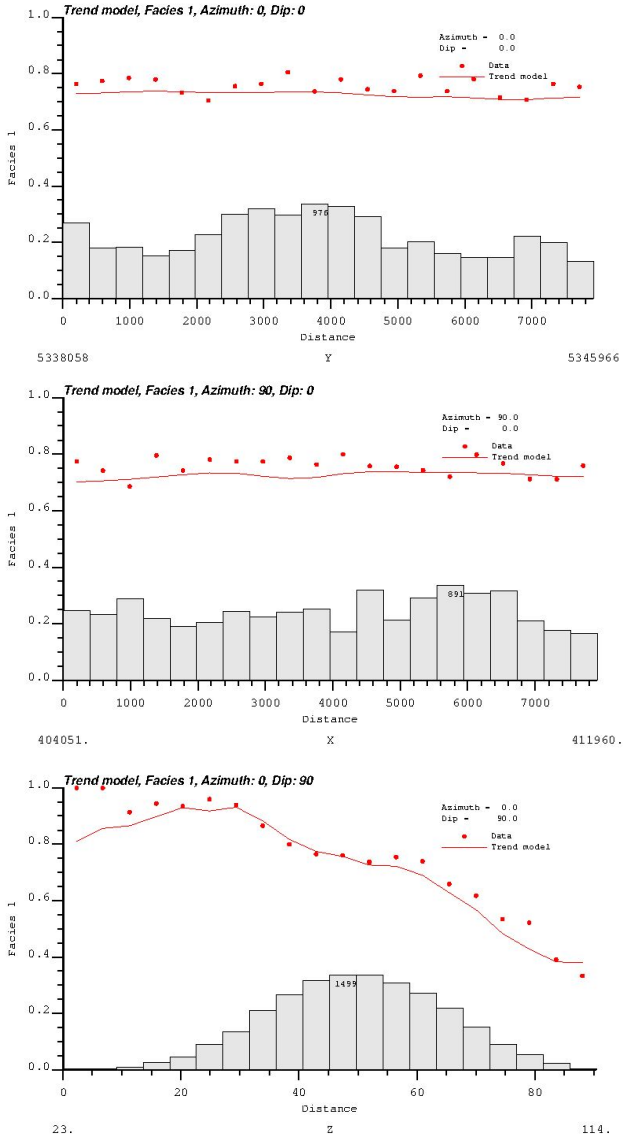

Figure 2.26: The swath plots of the trend models for the first facies in the major (0 azimuth), minor (90 azimuth), and vertical directions in relation to the reference data.

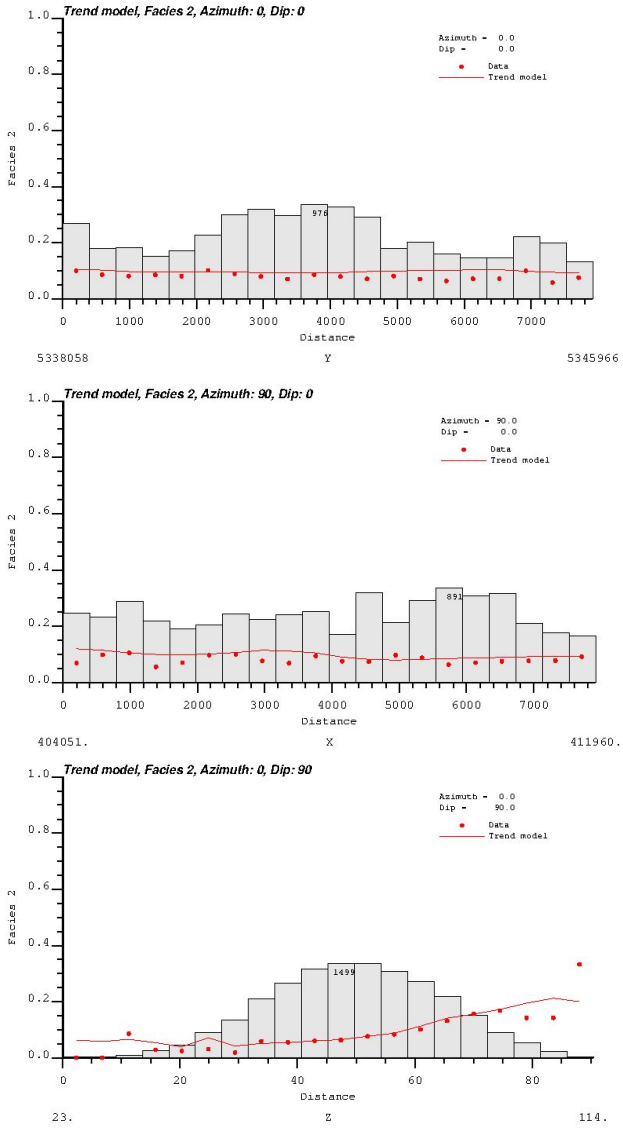

Figure 2.27: The swath plots of the trend models for the second facies in the major (0 azimuth), minor (90 azimuth), and vertical directions in relation to the reference data.

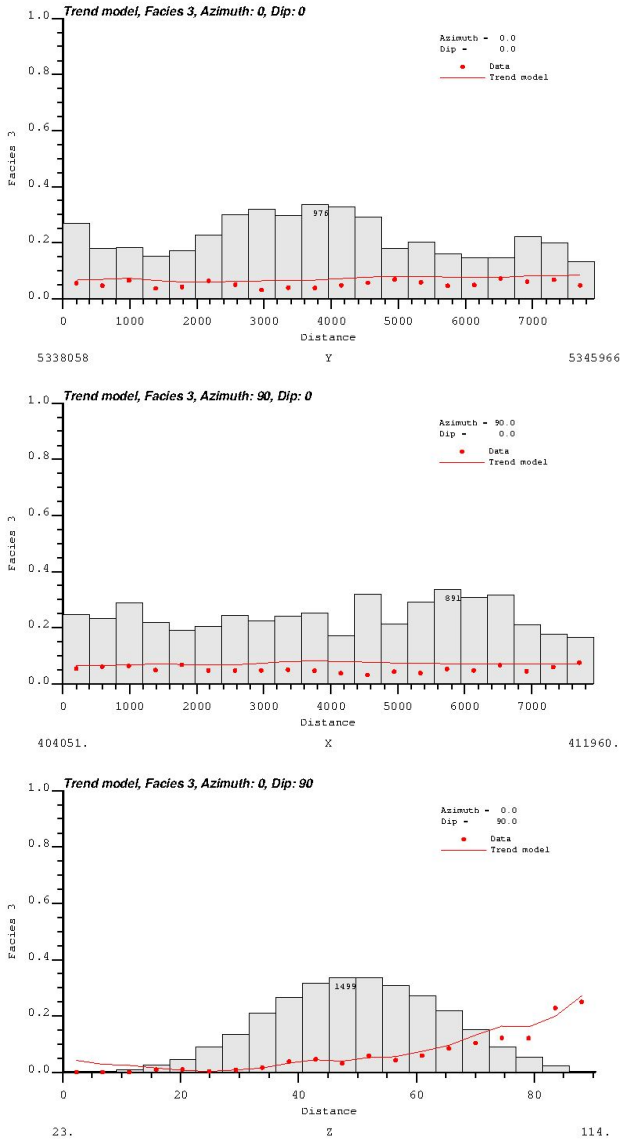

Figure 2.28: The swath plots of the trend models for the third facies in the major (0 azimuth), minor (90 azimuth), and vertical directions in relation to the reference data.

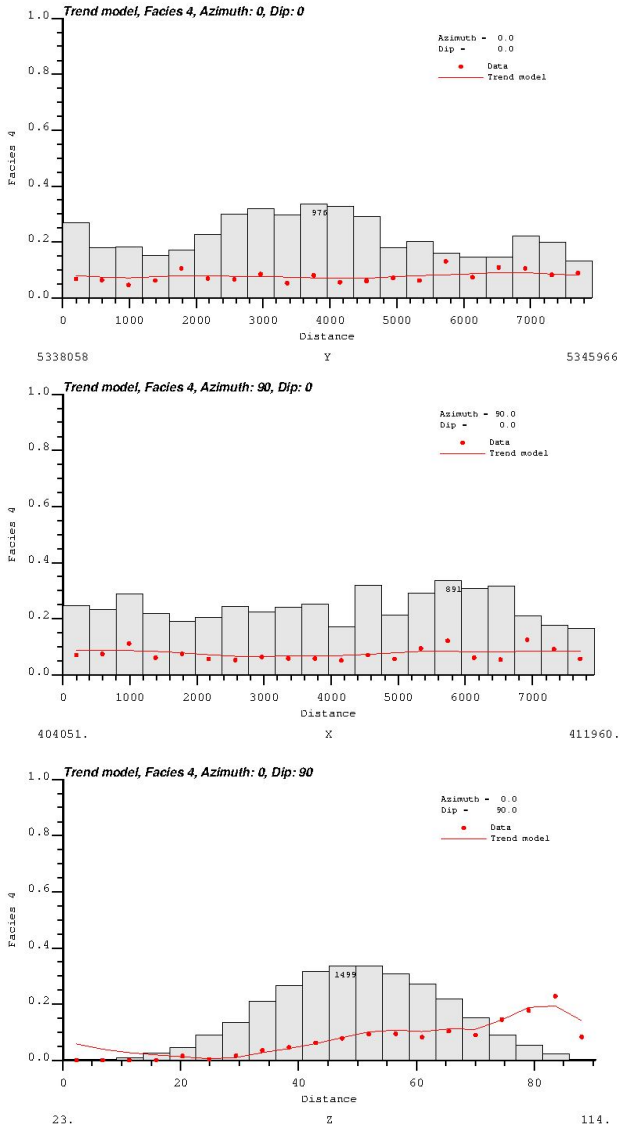

Figure 2.29: The swath plots of the trend models for the fourth facies in the major (0 azimuth), minor (90 azimuth), and vertical directions in relation to the reference data.

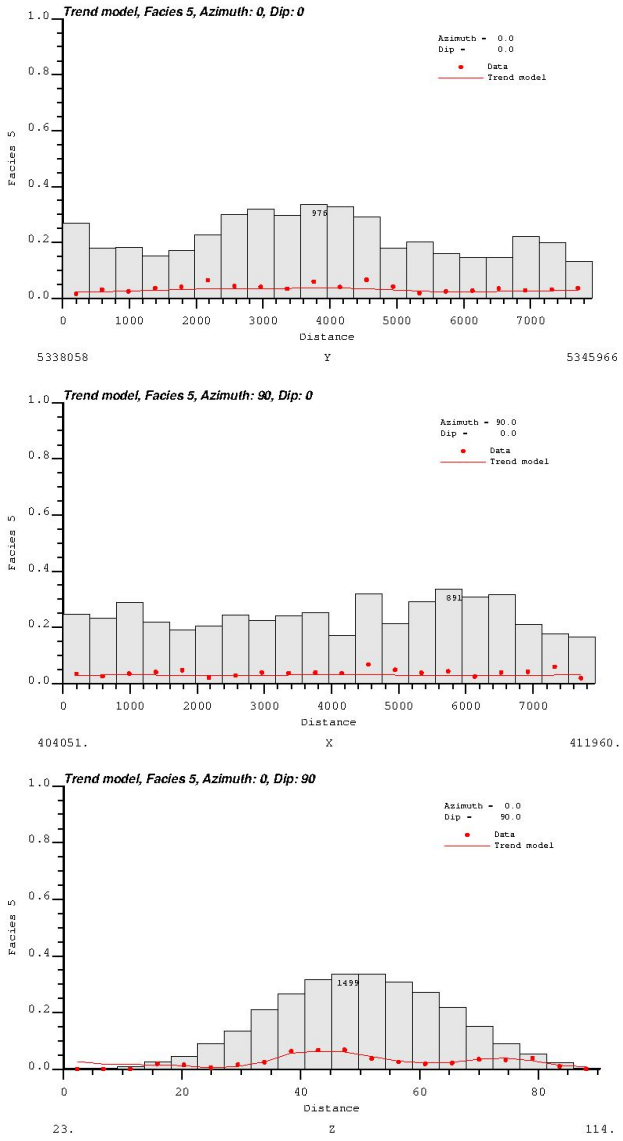

Figure 2.30: The swath plots of the trend models for the fifth facies in the major (0 azimuth), minor (90 azimuth), and vertical directions in relation to the reference data.

2.6.4 Clipping continuous variables to categorical models

The workflow detailed in section 2.4 must be repeated for samples of continuous variables within each facies. Then, to each grid node, the simulated value of the continuous variables is assigned from the workflow of the facies that the simulated result of the SIS workflow shows. For instance, if the simulated value of the categorical modeling workflow shows the value of a grid node to be facies 2, the simulated values of So and ϕ are assigned from the workflow of facies 2 to that grid node. Through this

process, continuous variables are modeled within each facies separately as a different domain.

2.7 Probabilistic resource estimation

After realizations of the continuous variables are generated and clipped by simulated realizations of the boundaries of the reservoir and simulated realizations of lithofacies, they can be used for resource estimation. In the case study, realizations of effective porosity and oil saturation are multiplied at each 2D grid node with the realization of thickness at that location to calculate the *Original Oil In Place (OOIP)* (m^3). This calculation is made using expression 2.1, where $\phi(u)$is the effective porosity at location u, $So(u)$ is the oil saturation at location u, $Thickness(u)$ is the thickness of the reservoir at location u, and $S(u)$ is the surface of the grid cell (m^2) at location u, for all locations of u at the domain. Expression 2.2 shows this calculation in a 3D setting, where thickness is removed and $V(u)$ is the vertical size of the grid cell (m) at location u.

$$OOIP(m^3) = \sum(\phi(u) * So(u) * Thickness(u) * S(u)) \forall u \in Domain \tag{2.1}$$

$$OOIP(m^3) = \sum(\phi(u) * So(u) * S(u) * V(u)) \forall u \in Domain \tag{2.2}$$

Chapter 3

Application of uncertainty quantification

In this chapter, the analyses and workflows discussed previously to measure uncertainty quantification in probabilistic resource modeling are implemented on a data set of sampled wells in the McMurray formation as a case study. The main purpose of the case study is to quantify the contribution of the three categories of uncertainty to overall global uncertainty in the probabilistic resource estimation of OOIP in the 2D and 3D workflows.

To do this, probabilistic resource estimation is implemented in three cases: 2D workflow, 3D workflow with no lithofacies, and 3D workflow with lithofacies (Figure 3.1). Since lithofacies are not included in the 2D modeling workflow, for better comparison between 2D and 3D settings the results are generated with and without the inclusion of lithofacies in the 3D workflow. This means that the steps outlined in Chapter 2 related to modeling categorical variables are skipped and continuous variables are simulated across the domain and within the geological boundaries. This would also help determine the impact of uncertainty in the modeling of lithofacies in the 3D workflow.

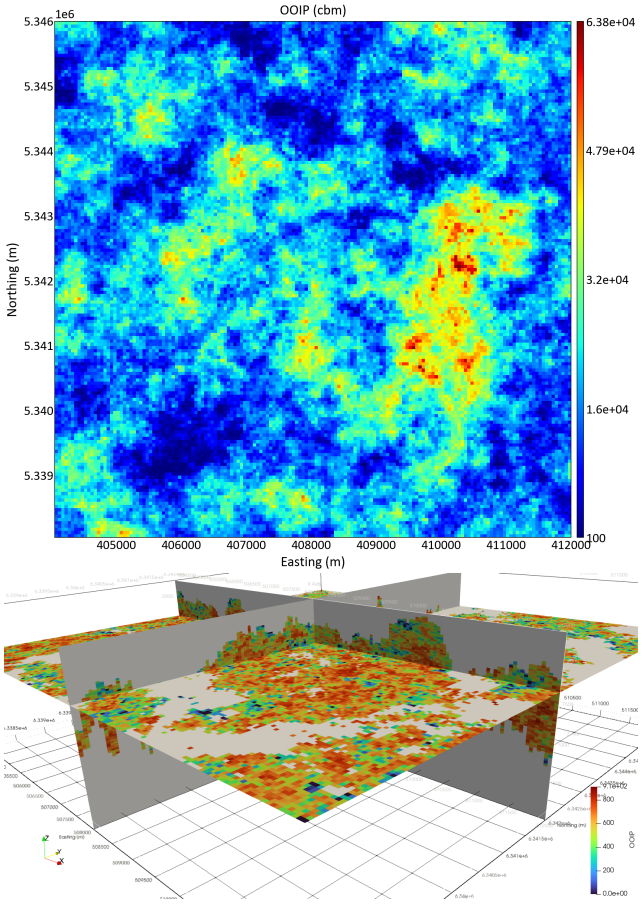

Figure 3.1: A single realization of OOIP in 2D (top plot) and 3D (bottom plot). The color bars indicate OOIP in cubic meters.

As discussed previously, the McMurray formation has a flat and layer-like geometry and is considered a stratigraphic formation, which makes it an ideal candidate for 2D modeling. Furthermore, modeling the McMurray formation on a 2D grid is already common practice within the industry, and analysis of 2D and 3D uncertainty quantification in this setting could be extra beneficial (Ren *et al.* 2006) (Figure 1.1).

3.1 Model checks and validation

The final simulated realizations generated in the methodology detailed in Chapter 2 need to be checked and validated before their results are interpreted. In this section, some important checks and validations from the results are reviewed. These checks consist of the reproduction of reference histograms, reproduction of variograms, reproduction of sample values, swath plots to inspect the reflection of trends in the

results, and vertical proportion curves.

Starting with some visual inspections, Figure 3.2 is the average and variance of simulated realizations of the three continuous variables in the 2D workflow. Figure 3.3 shows a single realization of the lithofacies in the 3D workflow. In this figure, grid nodes that fall outside the geological boundaries are colored as white. The visual checks show a general fidelity to sample value, but abrupt and very short-range changes to simulated values, which is in line with the geological context of the models and the variograms reviewed in Chapter 2, especially for porosity and oil saturation.

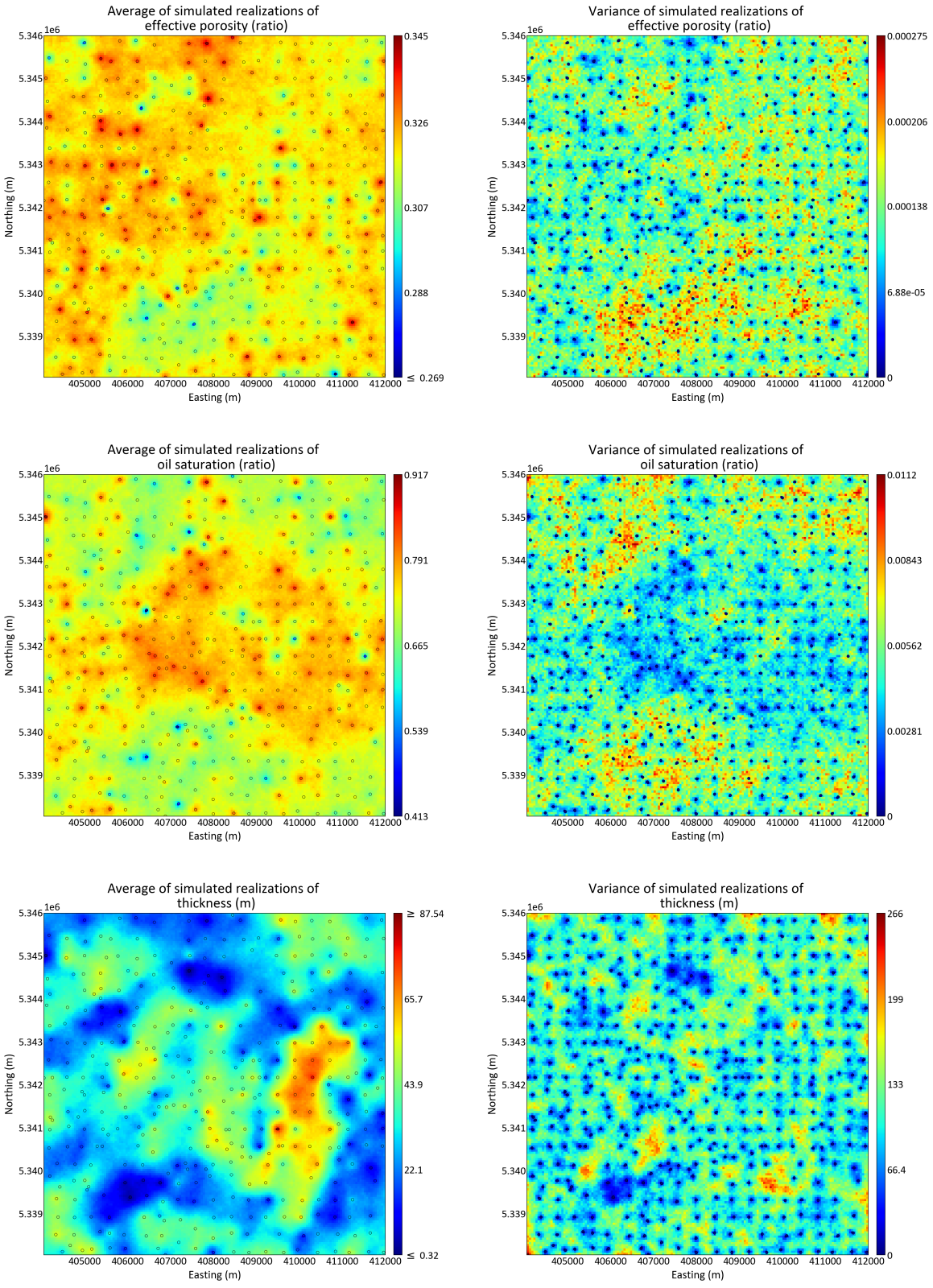

Figure 3.2: The averages and variances of 100 simulated realizations of continuous variables in the 2D workflow. The black points represent sample locations.

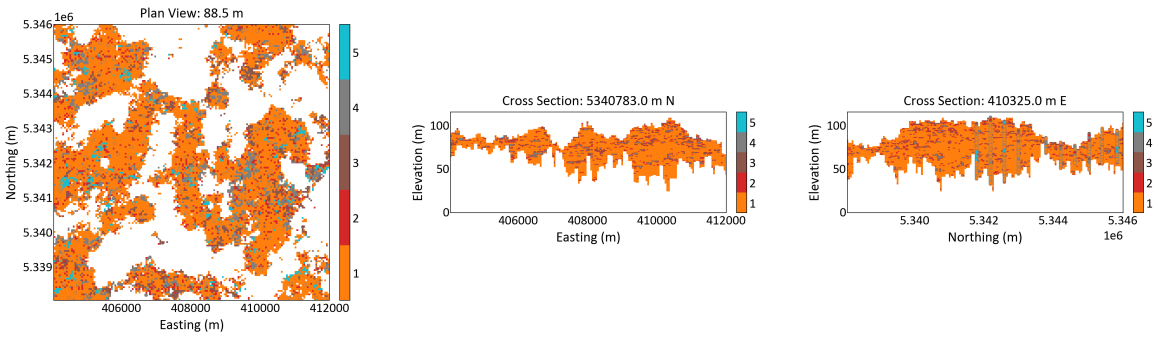

Figure 3.3: An example of the lithofacies simulations at specific plan views and cross sections. The white cells indicate parts of the domain that are outside the boundaries of the geological domain. The numbers indicate the five lithofacies: clean sand (1), sandy HIS (2), sand with 30-70 percent mud (3), mud (4), and breccia (5)

Figures 3.4 and 3.5 demonstrate the histogram reproduction of the continuous variables in the 2D and 3D without facies workflows, respectively. These histogram reproduction plots indicate high fidelity to reference distributions, in terms of averages, standard deviations, and the distribution itself, with a range of uncertainty that appears to be higher in the 2D workflow.

Figures 3.6 and 3.7 demonstrate the sample reproduction of the continuous variables in the 2D and 3D without facies workflows, respectively. The sample reproduction shows complete adherence to sample values. The points in these plots that appear to show different realization values from the truth values are due to the fact that these checks show several simulated values all in one plot and the simulation software used in this research produces a simulated value for grid nodes where there are more than one samples inside the node, hence different values for the same truth value in those nodes.

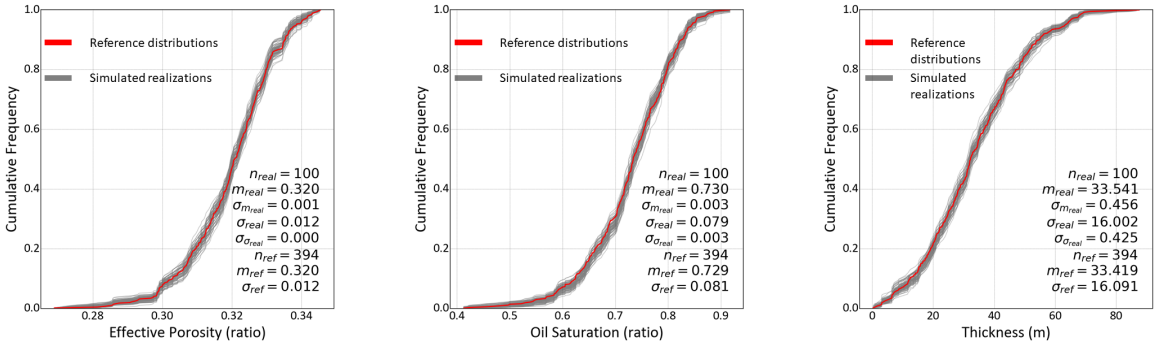

Figure 3.4: The histogram reproduction of the three continuous variables in the 2D workflow.

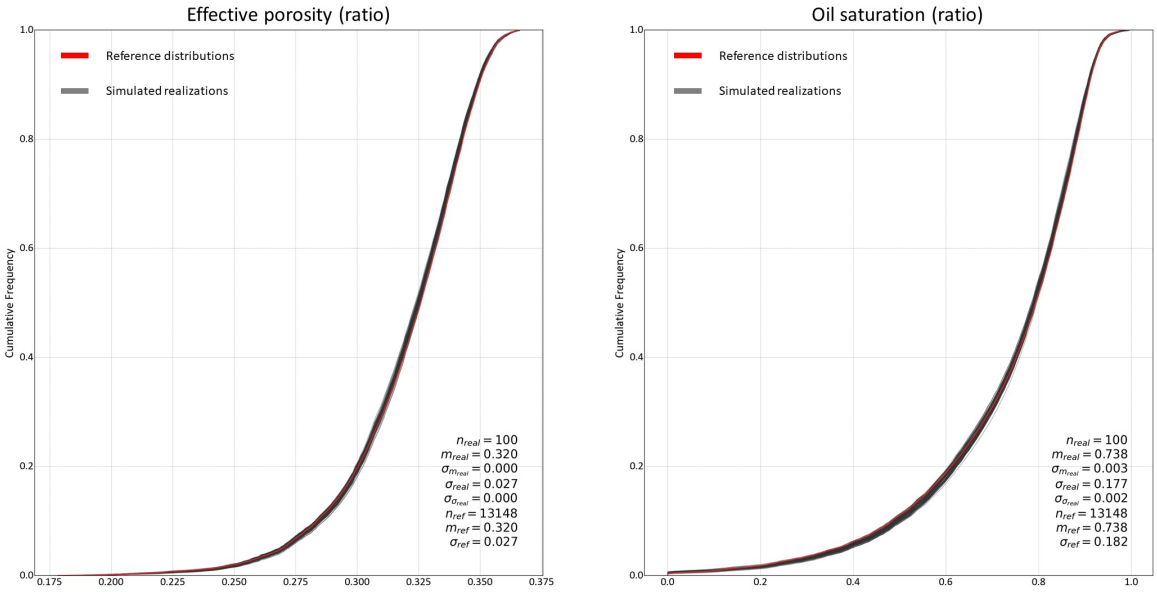

Figure 3.5: The histogram reproduction of the two continuous variables in the 3D workflow without facies.

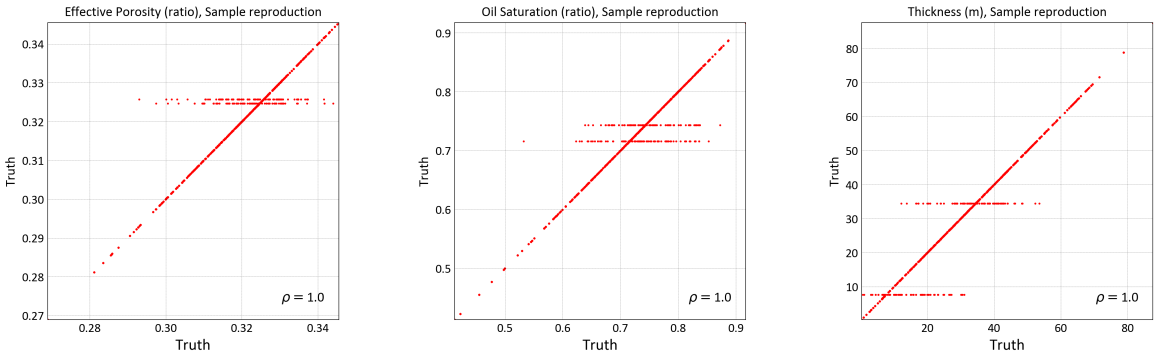

Figure 3.6: The sample reproduction of the three continuous variables in the 2D workflow.

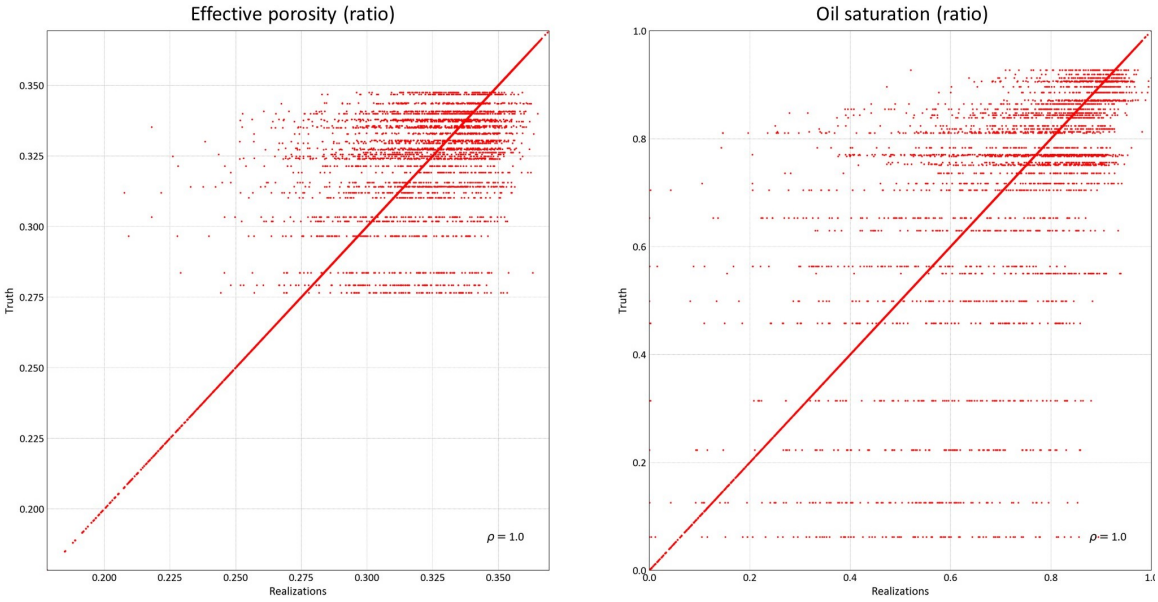

Figure 3.7: The sample reproduction of the two continuous variables in the 3D workflow without facies.

Figures 3.8 and 3.9 show the variogram reproduction in the 2D and 3D without facies workflows, respectively. In these plots, the variogram of the simulated realizations is compared with both the reference experimental variogram and the variogram model. As discussed in Chapter 2, a different variogram is used for each realization, which is why there are several experimental variograms and variogram models in these plots. In these plots, the reproduced variograms follow both experimental and model variograms closely, with deviations happening in the short ranges. This is due to sparse sampling and extreme changes in the very short ranges. Additionally, it also appears that 2D simulations show higher variability in 2D than 3D.

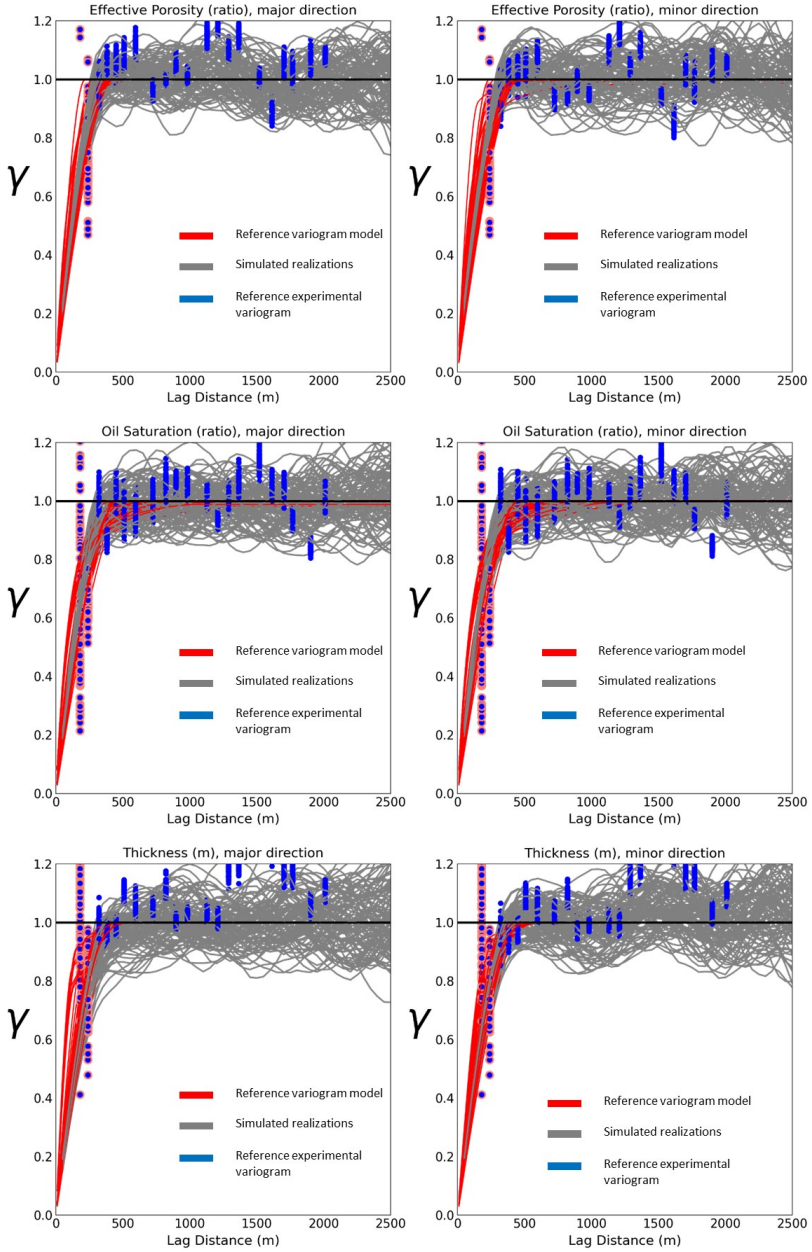

Figure 3.8: The variogram reproduction of the three continuous variables in the 2D workflow in the major (0 azimuth) and minor (90 azimuth) directions.

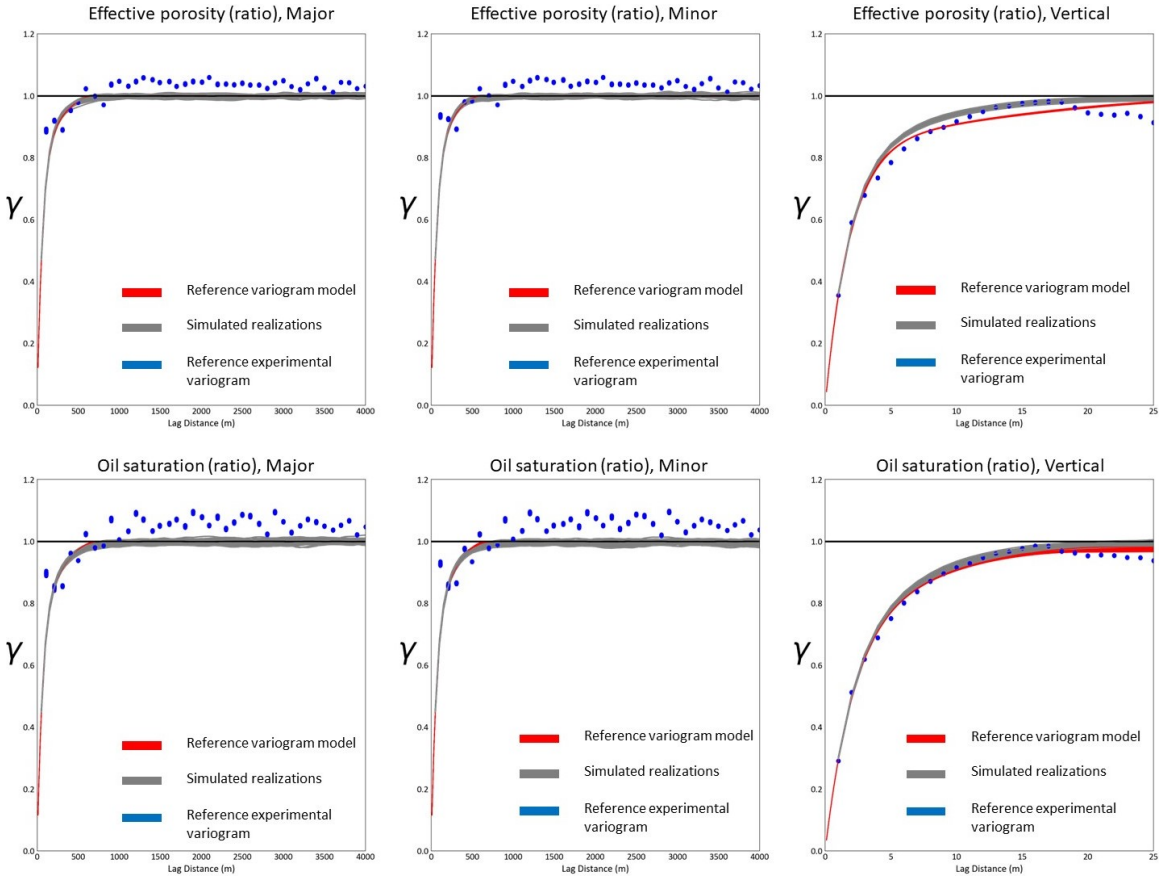

Figure 3.9: The variogram reproduction of the two continuous variables in the major (0 azimuth), minor (90 azimuth), and vertical directions in the 3D workflow without facies.

Figure 3.10 shows the proportion reproduction check for the lithofacies in the simulated realizations, while Figure 3.11 shows the reproduction of the indicator variograms of the five lithofacies.

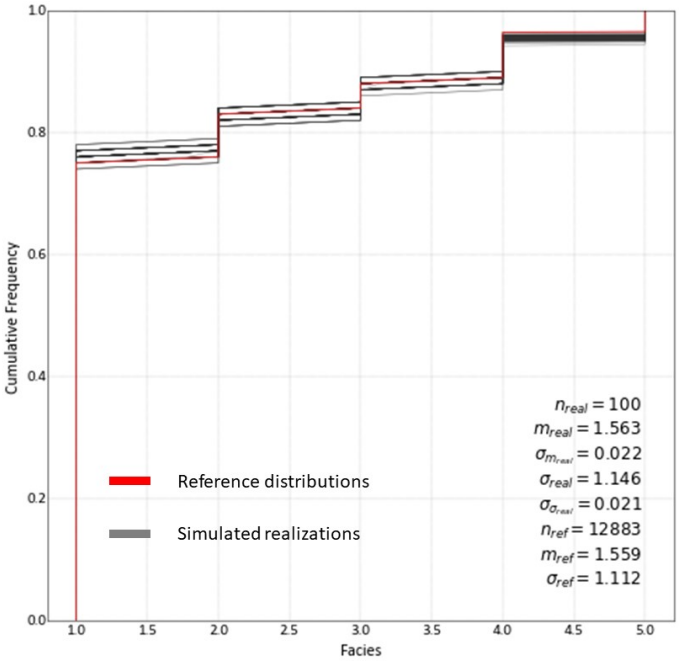

Figure 3.10: The reproduction of the reference proportions of the five lithofacies in the simulations.

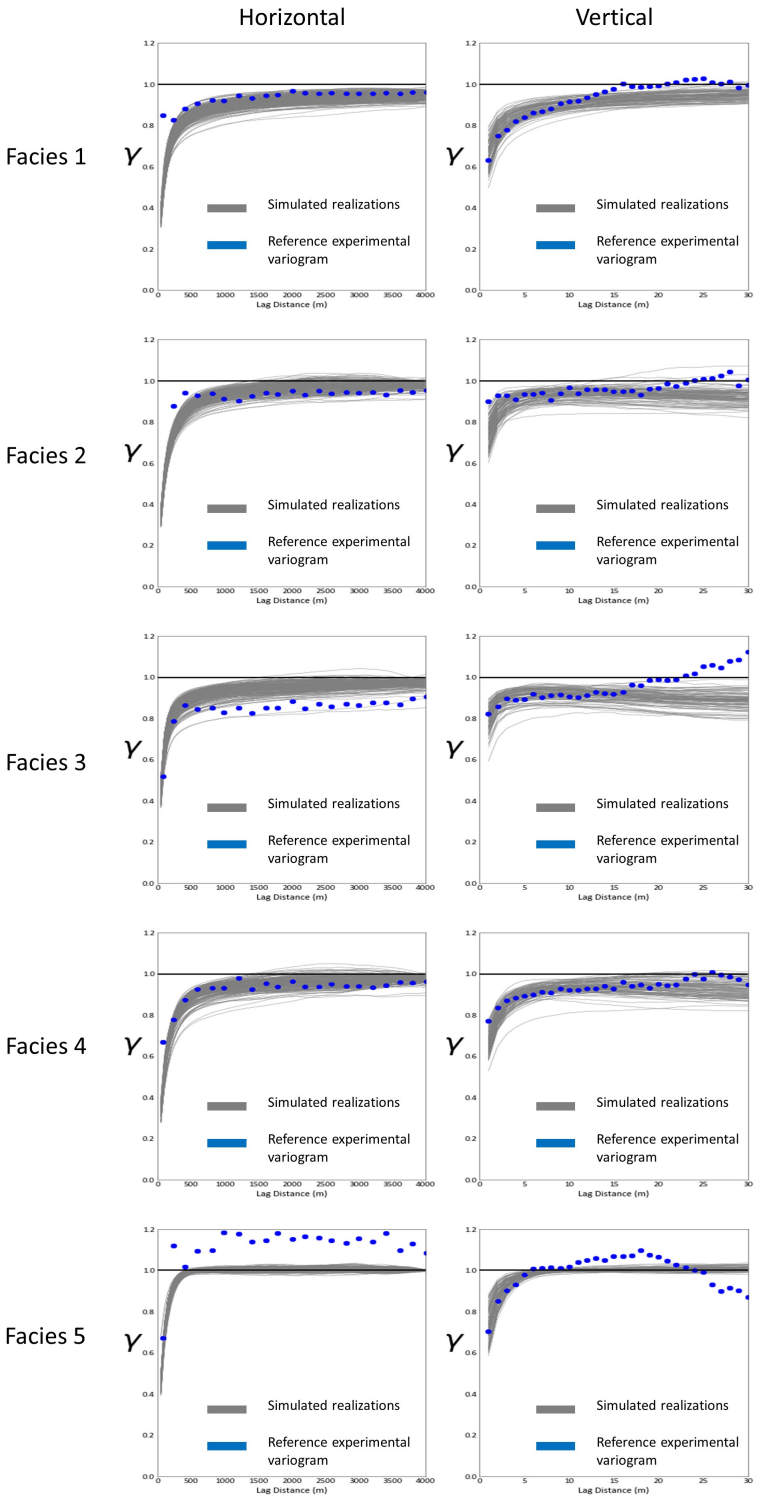

Figure 3.11: The reproduction of the indicator variograms of the five lithofacies in the horizontal (omni-directional variogram) and vertical directions.

Swath plots of the average of 100 simulated realizations of the continuous variables

are plotted along with the reference data. Figure 3.12 shows these swath plot checks in the major and minor directions for the 2D workflow and Figure 3.13 is the same checks but for the 3D workflow without facies in the major, minor, and vertical directions.

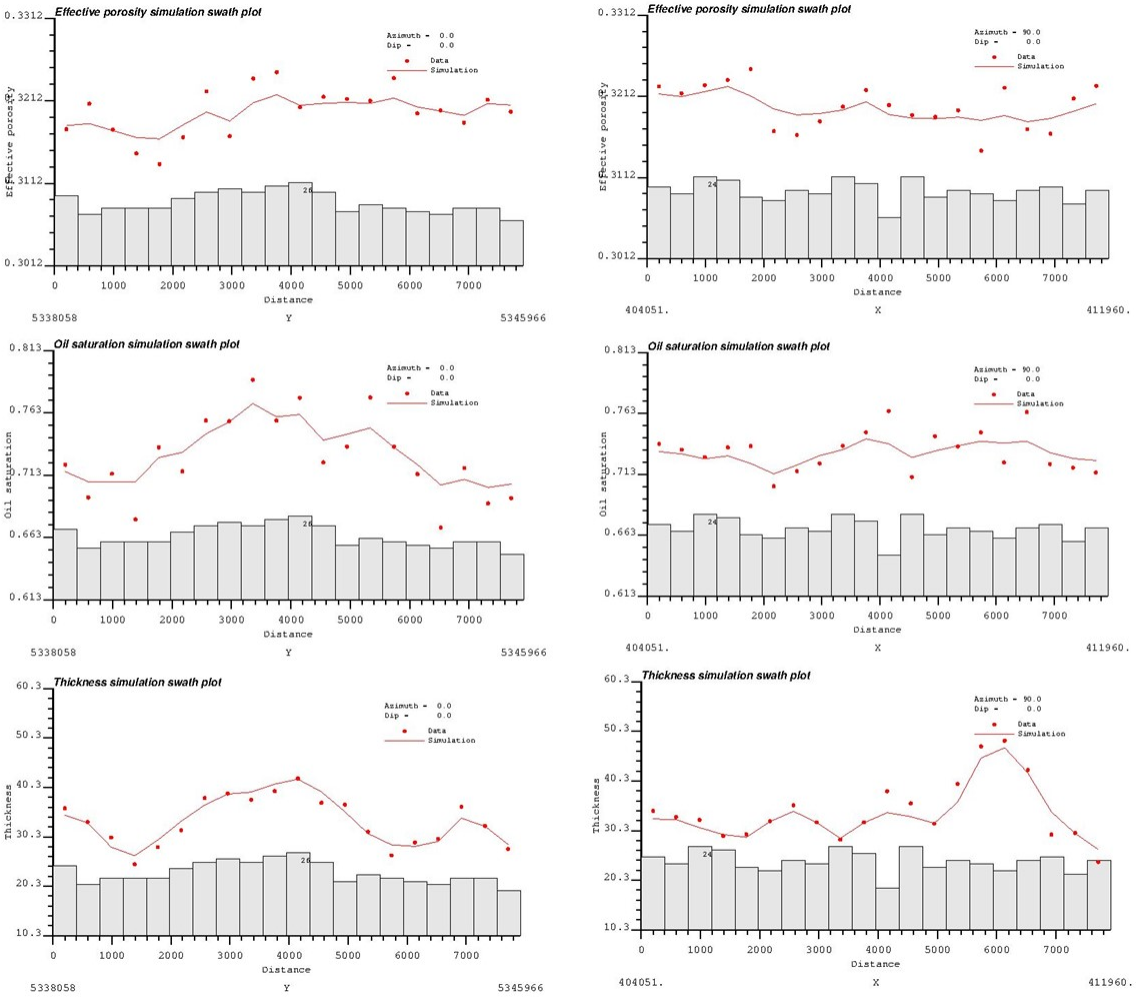

Figure 3.12: The swath plot of the average of 100 simulated realizations of the continuous variables in the 2D workflow in the major (0 azimuth) and minor (90 azimuth) directions plotted against the data. The red dots represent the average of sample values at given ranges, the red line shows the average of simulated realizations at different ranges, and the histogram bars at the bottom of the plots show the frequency of samples being located at the given ranges.

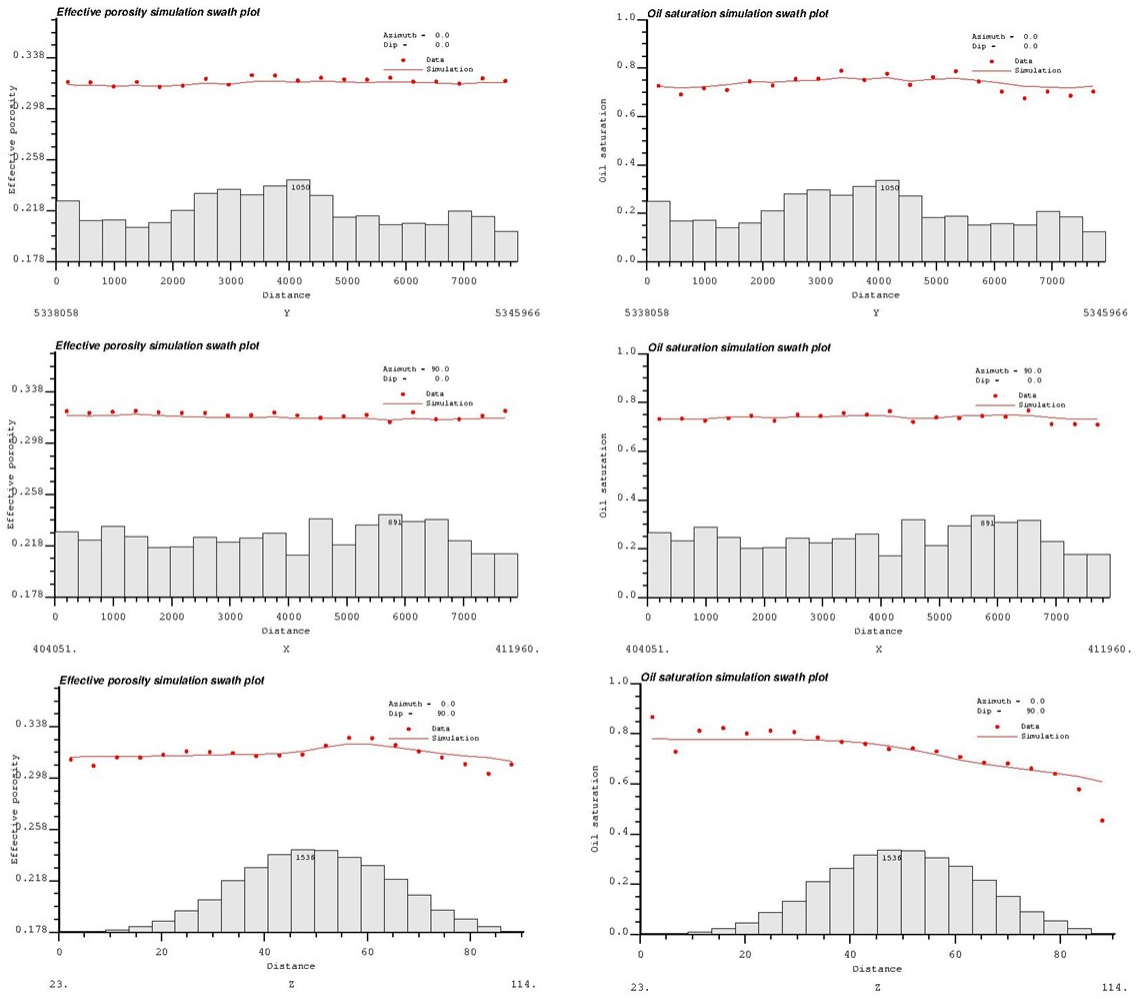

Figure 3.13: The swath plot of the average of 100 simulated realizations of the continuous variables in the 3D workflow without facies in the major (0 azimuth), minor (90 azimuth), and vertical directions plotted against the data. The red dots represent the average of sample values at given ranges, the red line shows the average of simulated realizations at different ranges, and the histogram bars at the bottom of the plots show the frequency of samples being located at the given ranges.

All the swath plots show an accurate reflection of trends in the models. Overall, the model checks indicate reliability in the models in terms of fidelity to the reference data sets and how realistic they are considering the geological context of the models.

3.2 Uncertainty quantification in 2D and 3D workflows

As discussed in Chapter 2, the contribution of each aspect of uncertainty is quantified by implementing the entire modeling workflow with and without the inclusion of each aspect, and the reduction in global uncertainty of the results is determined as the contribution of that aspect. This process can be demonstrated using multiple distributions of OOIP estimations in Figures 3.14, 3.15, 3.16 where the exclusion of one aspect of uncertainty leads to a distribution of estimations with lower variability in the results.

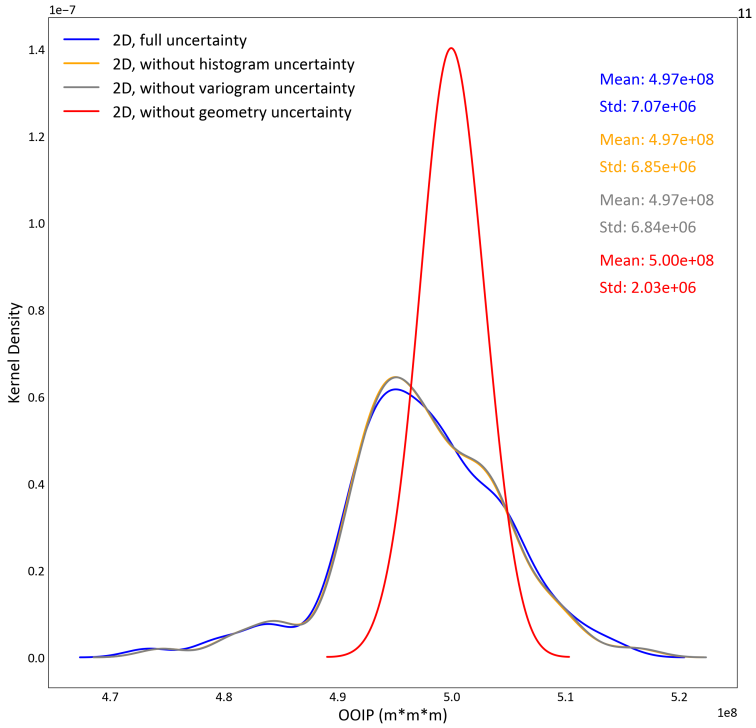

Figure 3.14: The distribution of 100 realizations of OOIP in the 2D workflow with the exclusion of one aspect at each step.

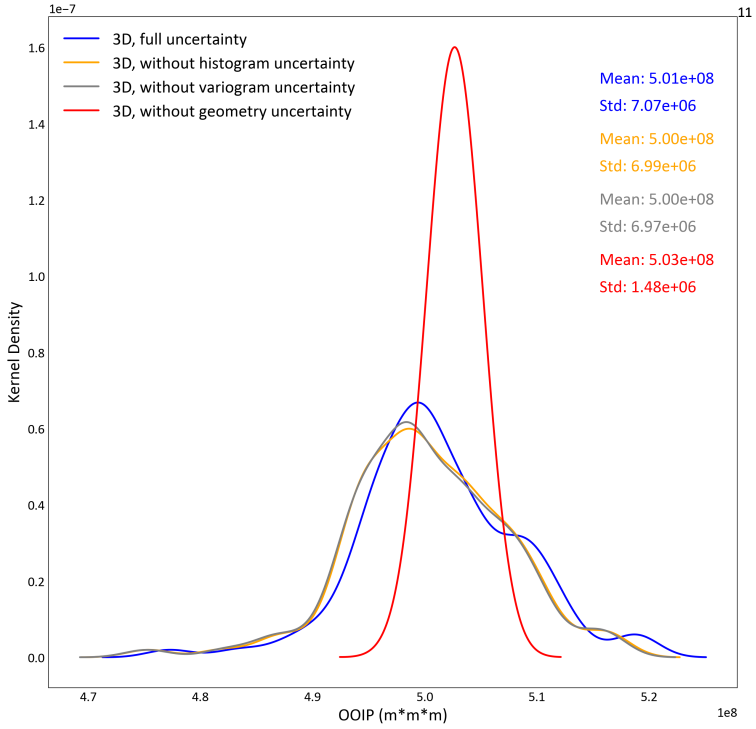

Figure 3.15: The distribution of 100 realizations of OOIP in the 3D workflow without lithofacies with the exclusion of one aspect at each step.

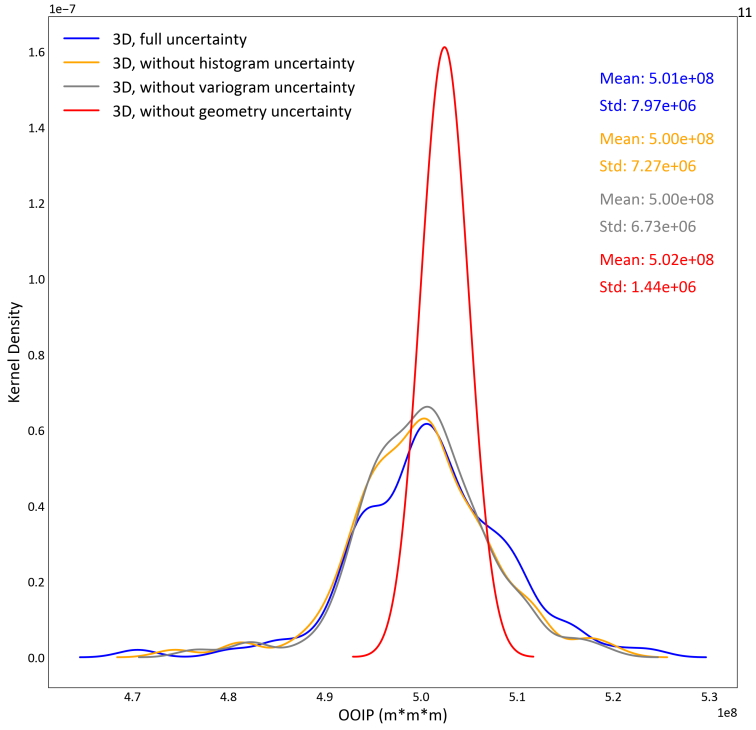

Figure 3.16: The distribution of 100 realizations of OOIP in the 3D workflow with lithofacies with the exclusion of one aspect at each step.

Figure 3.17 shows the breakdown of uncertainty into four main aspects: histogram uncertainty, variogram uncertainty, uncertainty in the geometry, and residual uncertainty. The first two aspects would form uncertainty in model parameters when combined. The charts in Figure 3.17 show the standard deviation in estimations of OOIP in the three workflows.

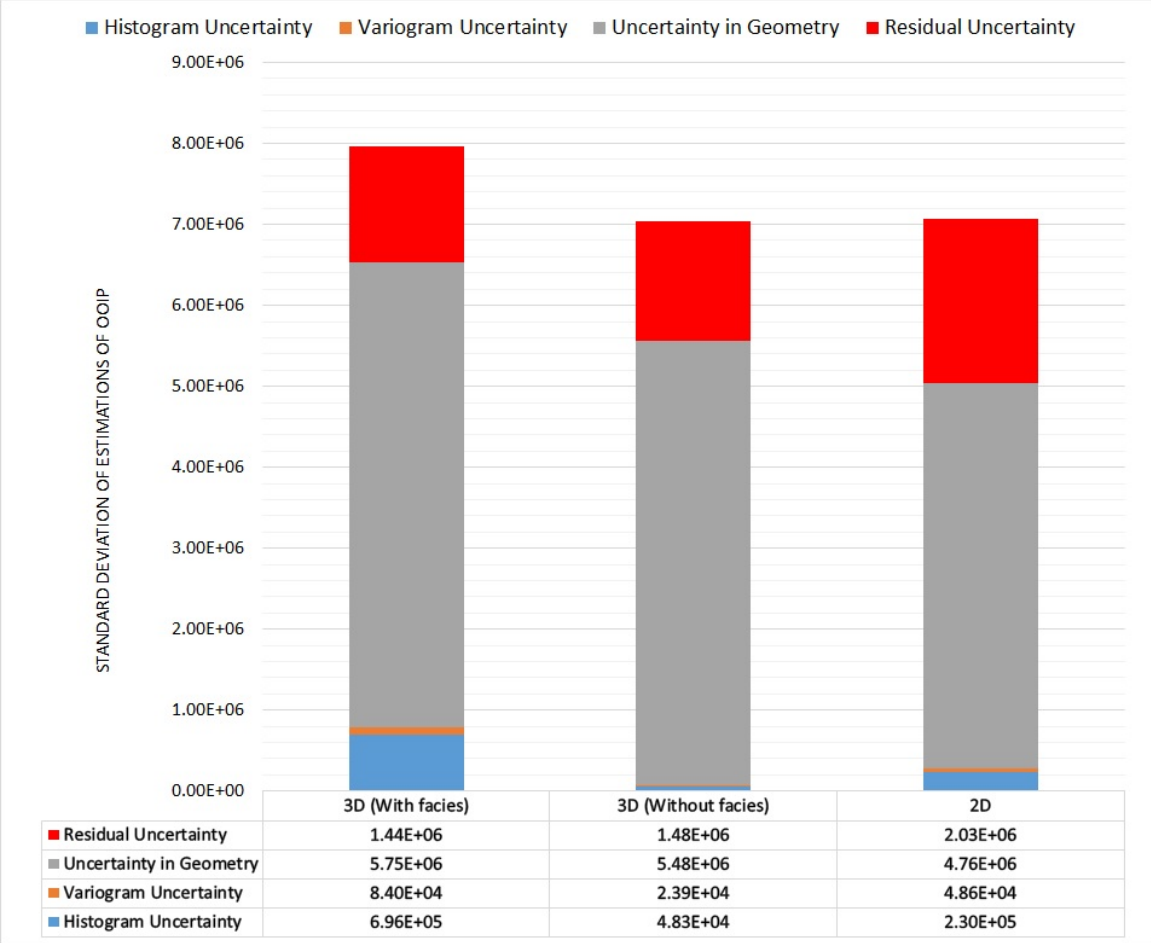

Figure 3.17: Uncertainty in the estimations of OOIP in the three workflows along with the contribution of each aspect to overall uncertainty.

The results in Figure 3.17 demonstrate that while the 2D workflow and the 3D workflow without facies have similar overall global uncertainties in estimations of OOIP, the 3D workflow with facies shows a significantly higher uncertainty. While the 3D workflow without facies includes histogram and variogram uncertainty for the two continuous variables, the 3D workflow with facies also incorporated the uncertainty in the global proportions of the categorical variables, which is a function of

categorical reference distribution uncertainty, as well as the uncertainty in the indicator variogram models used in SIS. This additional uncertainty in the 3D workflow with facies which is absent in the 3D workflow without facies is the primary reason the former shows significantly higher histogram and variogram uncertainties.

In terms of the contribution of each aspect to overall uncertainty, uncertainty in the geometry of the reservoir is the dominant aspect in all three workflows, followed by residual uncertainty, histogram uncertainty, and variogram uncertainty. Uncertainty in the geometry shows higher contributions in the 3D workflows compared to 2D, whereas residual uncertainty is higher in 2D compared to the 3D workflows. While histogram uncertainty is higher in the 2D workflow compared to the 3D workflow without facies, it is higher than all of them in the 3D workflow with facies, primarily due to the additional uncertainty from categorical proportions.

While the comparison of global uncertainty in estimations of OOIP reveals important information about the contribution of each aspect to overall uncertainty, it is also important to compare the local distribution of uncertainty in the two workflows. However, the 2D and 3D models are generated on different grids, which makes direct comparison of results at each grid node impossible. To make such comparisons, the 3D estimations are transformed into a 2D grid by summing values of OOIP at all grid nodes in the same aerial location, represented as a vertical column of cells at each aerial location. Equation 3.1 lays out this transformation in more detail. In this equation, $OOIP_{2D}(i,j)$ is the estimated value of OOIP at the aerial grid nodes located at indices i and j, and $OOIP_{3D}(i,j,k)$ is the estimated value of OOIP at the grid node located at aerial indices i and j and the vertical index of k. $Vlimit$ is the maximum vertical index at aerial grid nodes of i and j. Figure 3.18 demonstrates the local distribution of averages and variances of OOIP estimations in the 2D and 3D workflow, generated using Equation 3.1 for the latter.

$$OOIP_{2D}(i,j) = \sum_{k=1}^{Vlimit(i,j))} OOIP_{3D}(i,j,k) \tag{3.1}$$

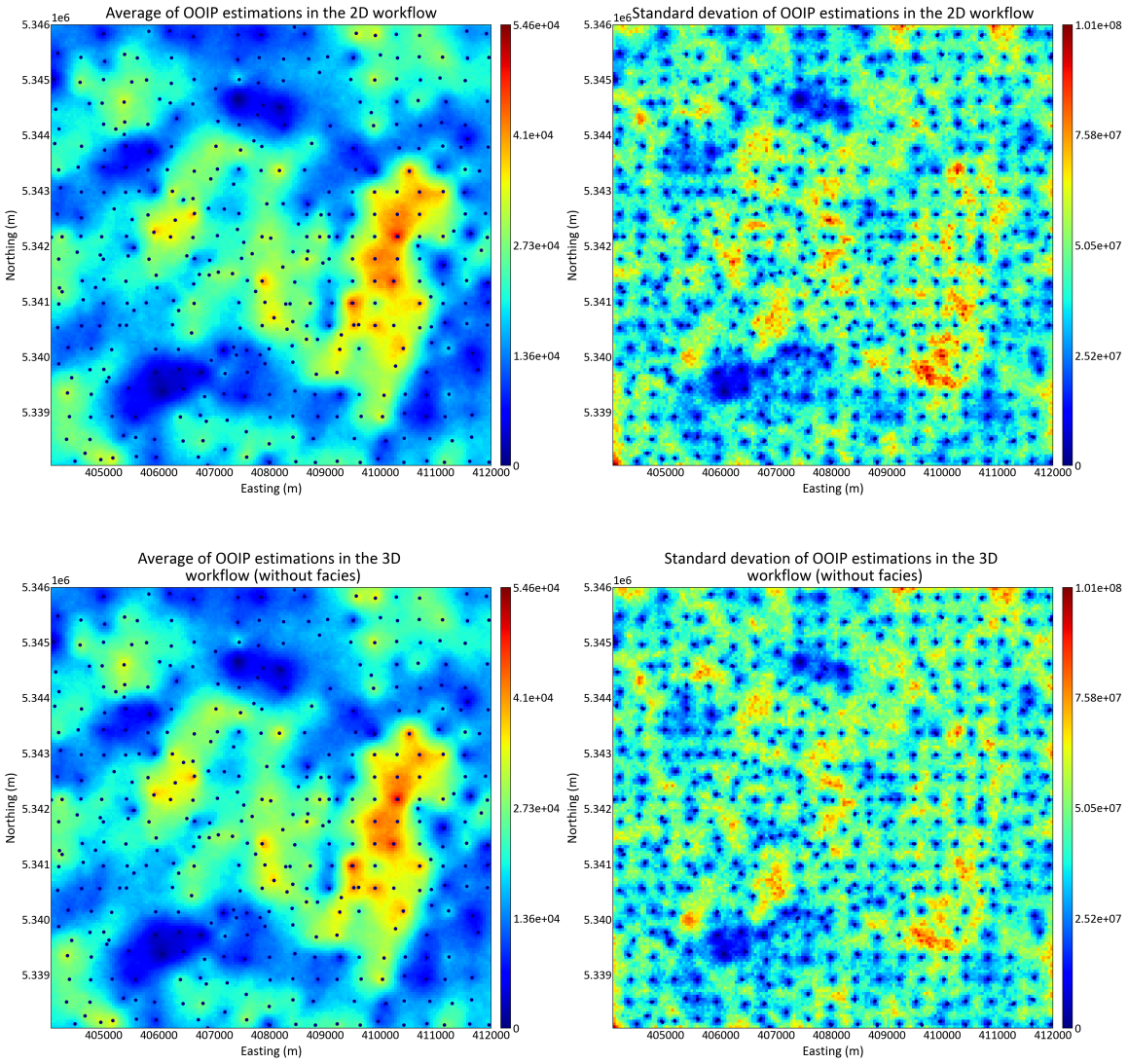

Figure 3.18: The map of averages and variances of OOIP estimations in the 2D and 3D (without facies) workflows.

3.3 Pad-size analysis

In this section, some of the applications of probabilistic resource models in the context of the case study are further illustrated and discussed. Probabilistic resource modeling gives an assessment of uncertainty in the estimated resources in a geological domain, which can then be used in economic decisions related to the feasibility of that resource and the risk associated with extracting it. Since the uncertainty in estimations is a critical parameter in these decisions and it has been demonstrated that the choice of 2D or 3D modeling impacts the quantification of uncertainty, it is useful to show how

this choice between 2D and 3D settings could change these economic decisions.

One of the most common methods of in-situ hydrocarbon extraction in the McMurray formation is steam-assisted gravity drainage (SAGD) (Butler 1994). In this method, two wells are drilled into a designated area, which is referred to as a SAGD pad, with one well injecting hot steam into the reservoir and one well extracting the fluids. The first well, called the injector well, is tasked with lowering the viscosity of the fluids, which would, in turn, cause them to be displaced. The second well called the producing well then extracts the displaced fluids from the pad to the surface (Butler 1994).

Figure 3.19 shows two plots, with the left plot showing the difference in the standard deviation of OOIP estimations in the 2D and 3D without facies workflows and the right plot showing the same difference, but in a grid of 200*200 meters, which is an upscaled model with dimensions approaching that of a SAGD pad. The right plot shows more clearly how quantification of uncertainty impacts different pads located in different areas of the domain, with red cells indicating potential SAGD pads where 2D quantification of uncertainty is higher than 3D, and blue areas showing the reverse. This indicates that in addition to decisions that are impacted by uncertainty in estimations of a resource in an entire domain, more locally focused decisions are also influenced by the type of modeling workflow that is selected in that domain.

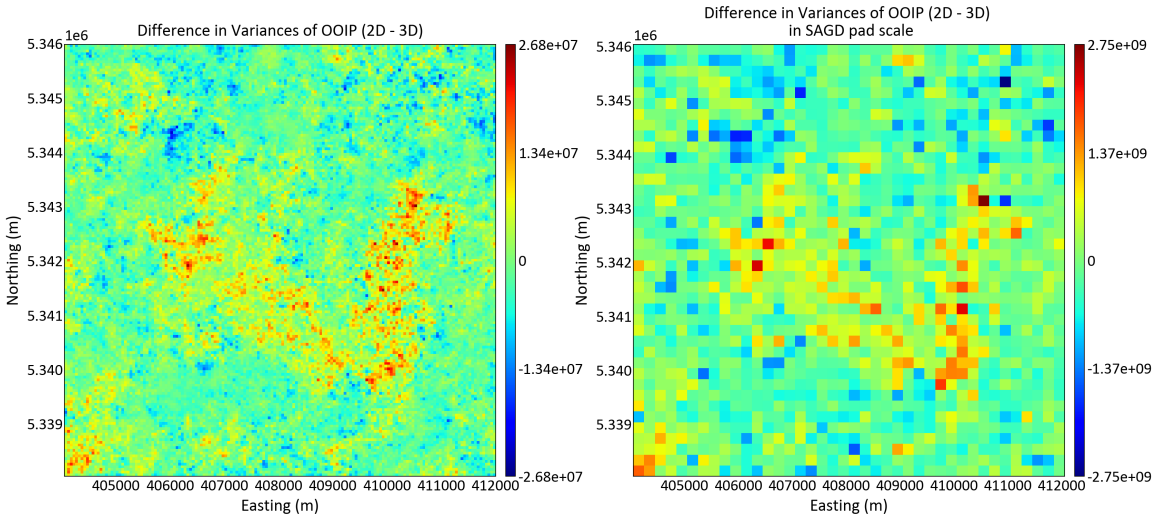

Figure 3.19: Left: the map of differences in the standard deviation of OOIP estimations between 2D and 3D (without facies) workflows. Right: the same map but in 200*200-meter grid blocks.

Figure 3.20 further demonstrates the same concept with three simulated pads being singled out, wherein for each pad the results are different between the two workflows (2D workflow and 3D workflow without facies). Deciding which of the two results is appropriate for each pad is a major decision.

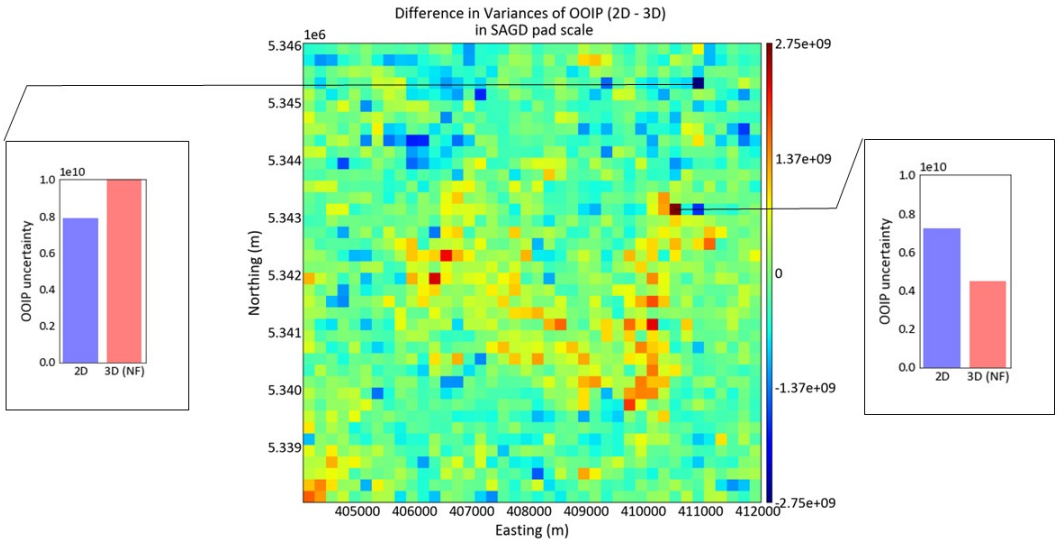

Figure 3.20: The quantification of uncertainty is shown as an example in two pad-size blocks. In the right block, 2D shows a significantly higher uncertainty, whereas in the left block, the 3D workflow shows higher variance.

3.3.1 Sample density analysis

The results and analyses demonstrated previously are generated using samples from 393 vertical wells in both workflows. The same uncertainty analyses are implemented on 2D and 3D workflows with decreasing number of wells, where in each step a given number of wells are removed from the data set randomly, and the simulation workflow is repeated for the new data set (Figure 3.21).

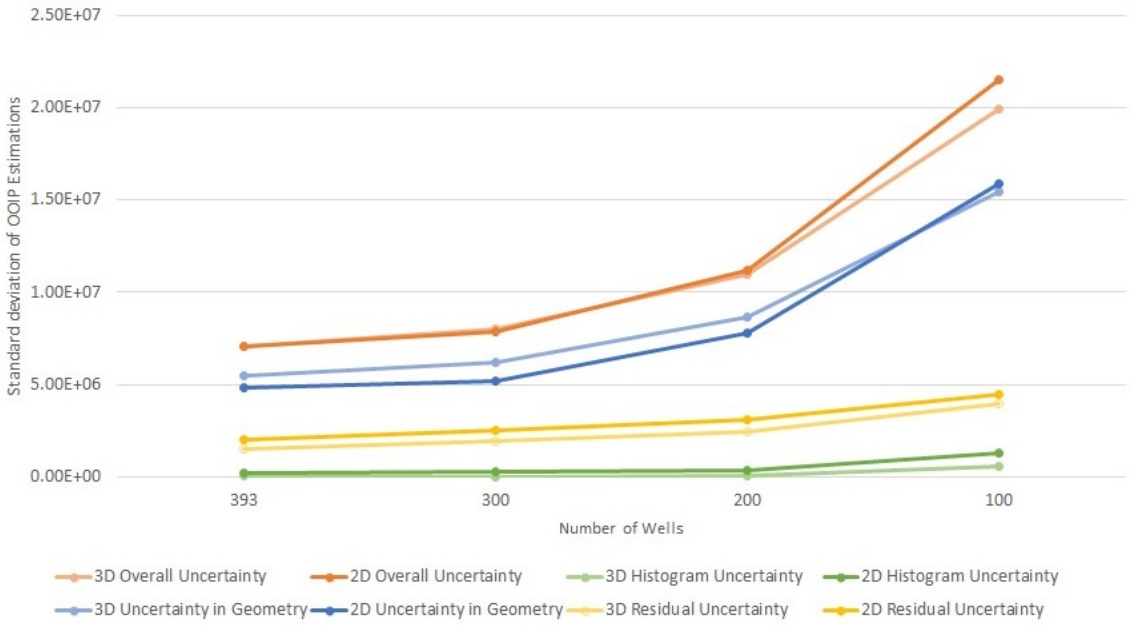

Figure 3.21: Global uncertainty in 100 estimations of OOIP in 2D and 3D (without facies) workflows with decreasing number of input sampled wells, with overall uncertainty in estimations as well as the contribution of three important aspects: histogram uncertainty, uncertainty in geometry, and residual uncertainty. Although the results are similar in cases with higher sample density, 2D uncertainty starts to grow faster as the number of wells drops.

Here, although the two workflows show similar global uncertainty in OOIP estimations in the higher density of samples, a growing disparity in favor of 2D appears between them as the number of wells decreases. In terms of each aspect, both workflows show a very similar trend, except for uncertainty in the geometry, which appears to increase for the 2D workflow at the lowest number of wells.

3.4 Summary

In this chapter, the quantification of the contribution of each aspect of uncertainty is presented in the 2D and the 3D workflow (with and without the inclusion of lithofacies in the latter). The results in the case study show similar overall uncertainty in resource estimations in the 2D and 3D workflow without facies but higher uncertainty in the 3D workflow with facies, which is due to the inclusion of parameter uncertainty in the categorical variable modeling workflow. In terms of global uncertainty, uncertainty in geometry has the largest contribution in all workflows, followed

by residual uncertainty and model parameter uncertainty. The results are also analyzed on a local scale, which shows significant differences in the local distribution of uncertainty between the two workflows, and the impact of this difference is exemplified in a SAGD-pad analysis. Finally, the analysis of results in multiple different workflows with changing sample densities reveals a similar composition of aspects of uncertainty in different sample densities.

Chapter 4

Analysis of residual and histogram uncertainty quantification

Analysis of results from the implementation of probabilistic resource modeling in Chapter 3 reveals important information about differences in the results of 2D and 3D modeling workflows regarding the contribution of each aspect of uncertainty and the overall uncertainty within the results. However, these results did not point to the reasons for these differences.

In this chapter, two categories of uncertainty are analyzed for the 2D and 3D workflows: residual and histogram uncertainty. The contribution of these two categories of uncertainty is compared given different modeling parameter choices and characteristics within synthetic data sets. These test models are designed to identify the contribution of different modeling parameters and data set characterizations to differences between 2D and 3D models and the sensitivity of these differences to each of these factors.

4.1 Test model description

The test models used to analyze residual and histogram uncertainty are based on a synthetic data set of drill holes in a 100*100-meter domain with block sizes of 2*2 meters. The synthetic drill holes consist of samples of a single random Gaussian variable that are created based on variable input parameters, such as vertical and horizontal spatial continuity of the variable and the drill hole spacing, and the contribution of these parameters is evaluated. The samples within drill holes are uniformly dis-

tributed in the vertical direction across all drill holes, assuming a simple stratigraphic formation that is fixed for all the 500 realizations.

As in Chapter 2, the synthetic data set is initially constructed as a 3D collection of samples and is averaged into a 2D data set. Subsequently, probabilistic resource modeling is implemented within 2D and 3D spaces for the random Gaussian variable within 500 realizations. Finally, histogram and residual uncertainty are measured and compared.

4.2 Residual uncertainty

Probabilistic resource modeling is implemented with one input variable parameter for each parameter while fixing other parameters to measure the impact of each factor on residual uncertainty in 2D and 3D workflows. The standard deviation of the mean of 500 realizations is compared between the 2D and 3D workflows as a measure of residual uncertainty. For residual uncertainty tests, unconditional simulation is implemented.

The base case for all the following tests has 2*2-meter blocks in a 100*100-meter domain with 30-meter domain thickness, a nugget effect of 0.0, and a two-structure spherical variogram with a total major and minor range of 50 meters and a vertical range of 15 meters. These ranges are divided equally between the two structures. A single realization of the test model in the horizontal and vertical directions is depicted in Figure 4.1. The base case variogram model can also be seen in Figure 4.2.

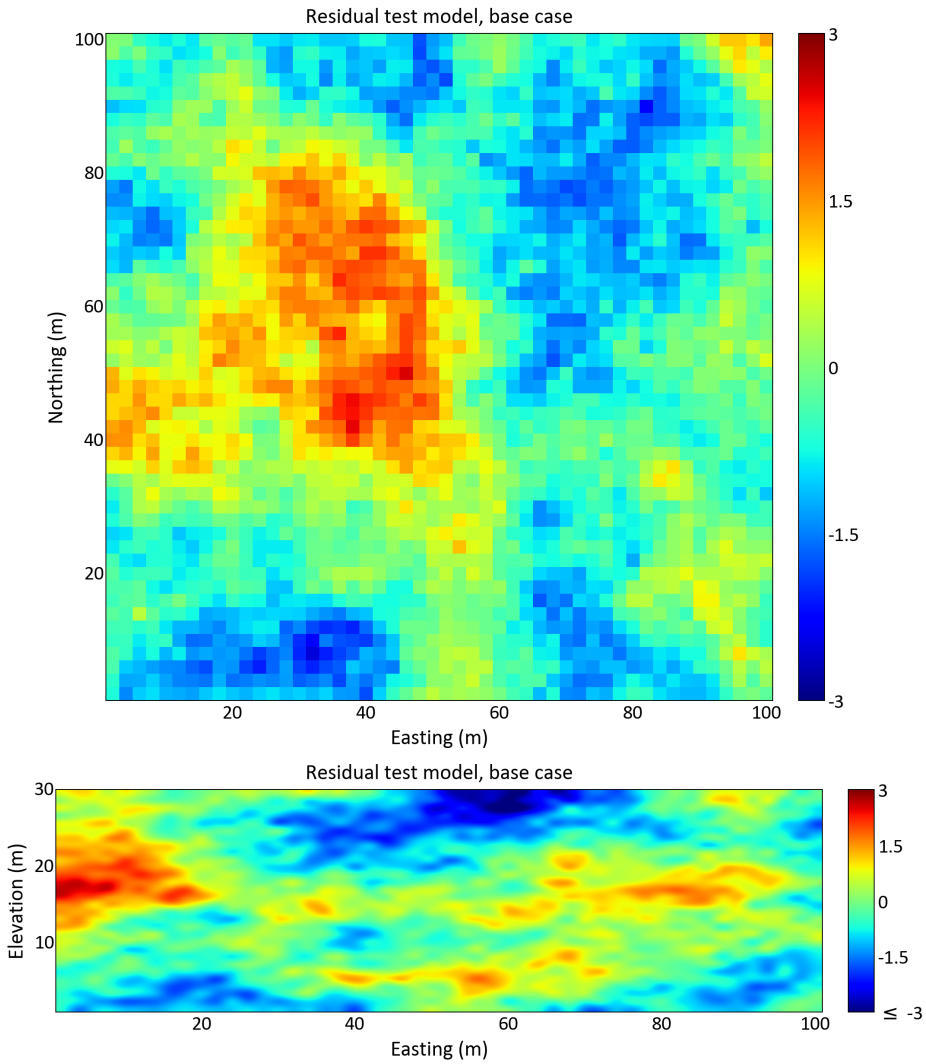

Figure 4.1: The top plot demonstrates a single simulated realization of the base case test model from a plan view and the bottom plot shows the same realization from a cross-section view.

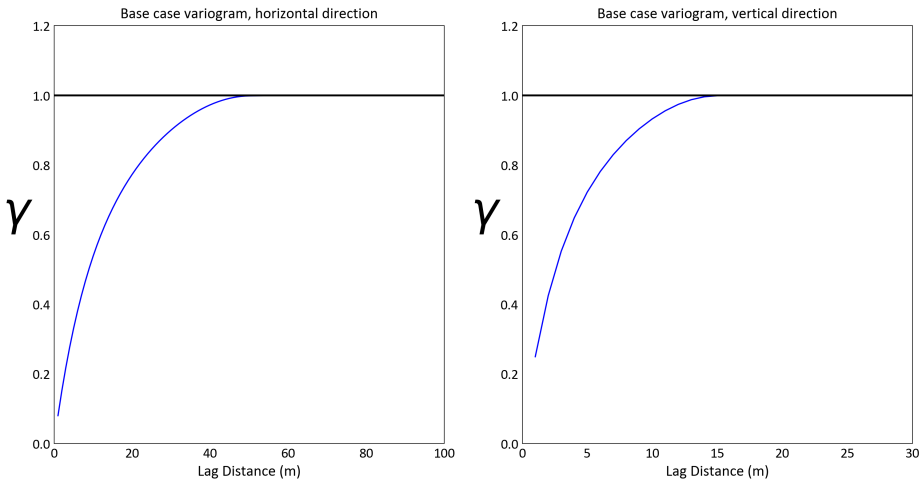

Figure 4.2: The base case variogram model used in the residual tests in the horizontal and vertical directions.

4.2.1 Thickness

The thickness of the stratigraphic formation which is equal to the vertical size of the drill holes is the first factor that is analyzed for its impact on residual uncertainty. Fixing other input parameters as described in the base case and changing the vertical dimension of drill holes in both 2D and 3D workflows results in Figure 4.3.

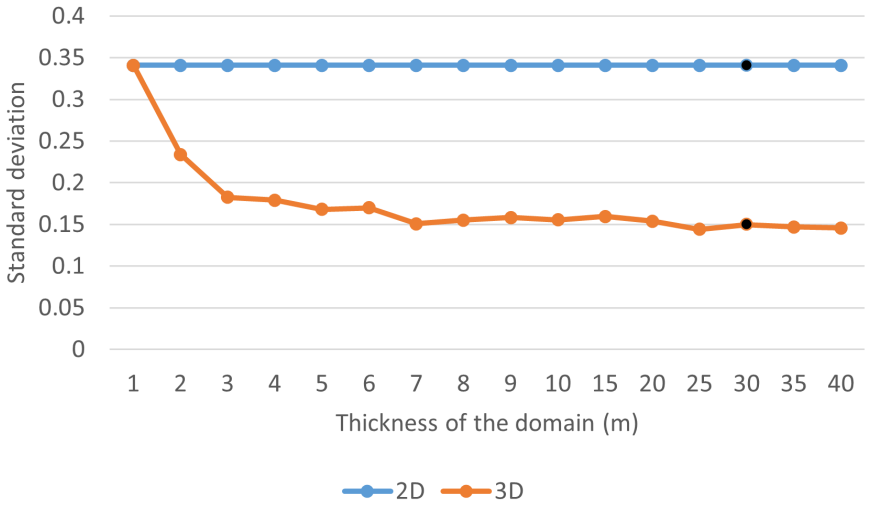

Figure 4.3: Standard deviation of the mean of 500 simulated realizations of the random variable with varying values for the thickness of the domain in 2D and 3D workflows. The black dots indicate the base case.

The lines in Figure 4.3 show that with increasing values of thickness residual un-

certainty in 2D remains the same, which is expected since no change is made to parameters in the 2D workflow and unconditional simulation leads to the same exact simulated values for all realizations. However, for the 3D tests, the higher the thickness of the domain, the lower the residual uncertainty. This could potentially be due to more averaging in the results with thicker domains and a constant variogram. This trend starts to flatten out at around 4-5 meters and sees lower changes after that.

The results in Figure 4.3 indicate that the gap between residual uncertainty in 2D and 3D workflows increases with larger vertical size of the domain. This also is in line with the observations made in Chapter 3 and Figure 3.17 regarding residual uncertainty in the 2D and 3D workflows in the case study, which showed higher residual uncertainty in the 2D workflow.

The results from the numerical case studies and the analytical tests here show that the 2D workflow tends to produce higher residual uncertainty. Whether the interpretation that higher thickness always leads to a higher difference between the two workflows is true depends on more comparative analysis and more conditional simulations to verify the observation made here.

4.2.2 Horizontal spatial continuity

The base case included a two-structure spherical variogram with a range of 50 meters. Changing this range to different values in the unconditional simulations for both 2D and 3D workflows leads to Figure 4.4. This indicates that differences in horizontal spatial continuity affect the two workflows similarly.

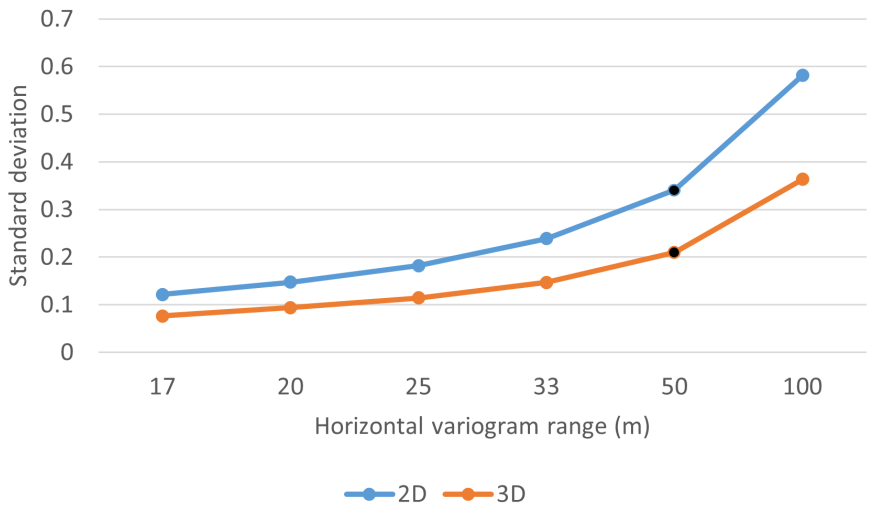

Figure 4.4: Standard deviation of the mean of 500 simulated realizations of the random variable with varying values for the horizontal range of the variogram in 2D and 3D workflows. The black dots indicate the base case.

4.2.3 Vertical spatial continuity

Similar to the horizontal variogram ranges, sensitivity to the vertical variogram ranges are also considered and the results can be seen in Figure 4.5. The residual uncertainty for all ranges in the 2D workflow is the same because there is no vertical spatial continuity in the 2D models, which is due to the exclusion of any vertical spatial continuity in that setting. However, the 3D residual uncertainty increases with higher vertical variogram ranges, indicating a shrinking difference between the two workflows with higher spatial continuity. This can be interpreted as more 2D-like behavior by the 3D models in vertically continuous domains.

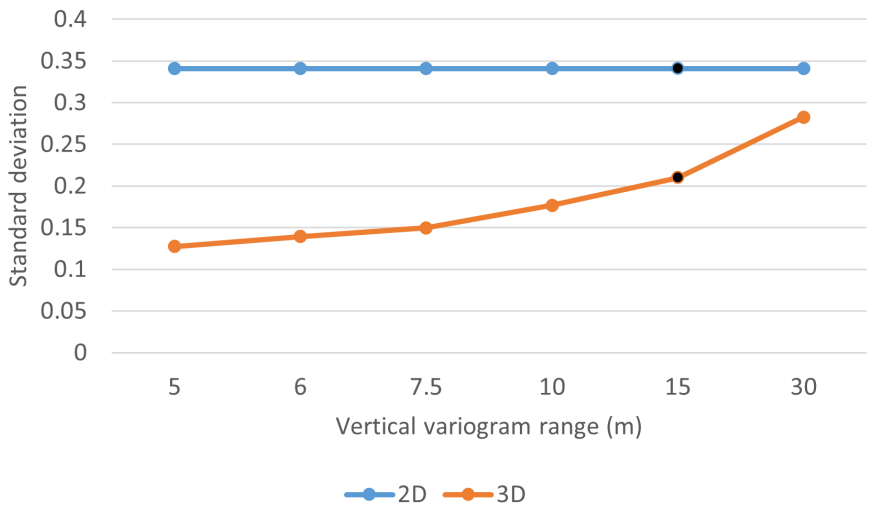

Figure 4.5: Standard deviation of the mean of 500 simulated realizations of the random variable with varying values for the vertical range of the variogram in 2D and 3D workflows. The black dots indicate the base case.

4.2.4 Nugget effect

Changing the nugget effect in the variogram used for unconditional simulations leads to Figure 4.6, which shows different behaviors in the 2D and 3D workflows. The 2D results show a consistent drop in residual uncertainty with a higher nugget effect, whereas the 3D results show an almost parabolic shape. It is hard to interpret much information from the 3D results, but the 2D results show a lower uncertainty in domains with a high nugget effect.

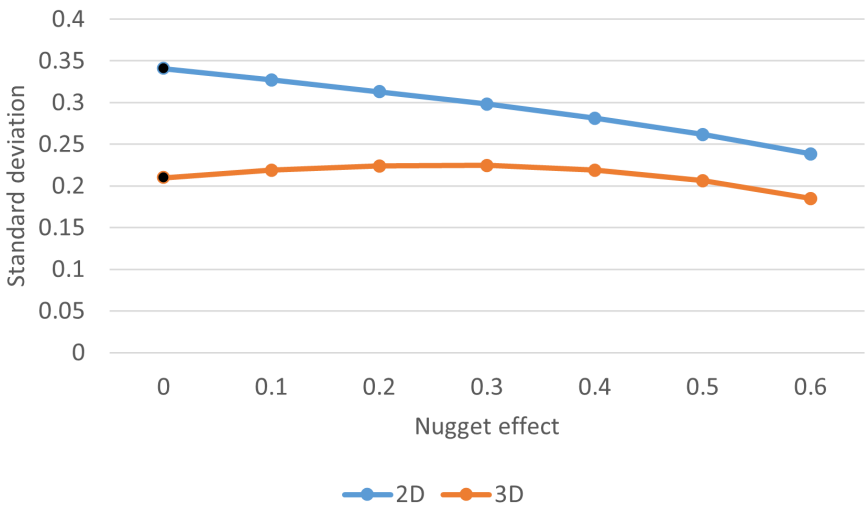

Figure 4.6: Standard deviation of the mean of 500 simulated realizations of the random variable with varying values for the nugget effect of the variogram in 2D and 3D workflows. The black dots indicate the base case.

4.3 Histogram uncertainty

As demonstrated in Chapter 3, histogram uncertainty is the dominant aspect of uncertainty in model parameters and contributes the most to global uncertainty. To analyze uncertainty in model parameters, tests are designed to measure histogram uncertainty in 2D and 3D probabilistic modeling workflows.

The test model described earlier is set up with conditional simulation with a number of synthetic drill hole samples with 20 meters spacing and a normal distribution. One important difference between these tests and the residual tests described earlier is that variogram ranges are only an input to the generation of the synthetic drill hole samples and the variograms used in the simulation are automatically fitted to the experimental variograms of the samples in a process similar to the one described in Chapter 2. Thus, variable variogram ranges only lead to a change in the level of spatial continuity of the input samples.

The workflows described in Chapter 2 are implemented with and without the presence of histogram uncertainty for 500 realizations. Histogram uncertainty is calculated as the difference between the standard deviation of simulated means with and without histogram uncertainty. Similar to residual uncertainty, the impact of multiple

factors on 2D and 3D histogram uncertainty is assessed while fixing other parameters according to the base case described above. Figure 4.7 demonstrates an example of the layout of drill hole samples from a plan and cross-section view in the domain.

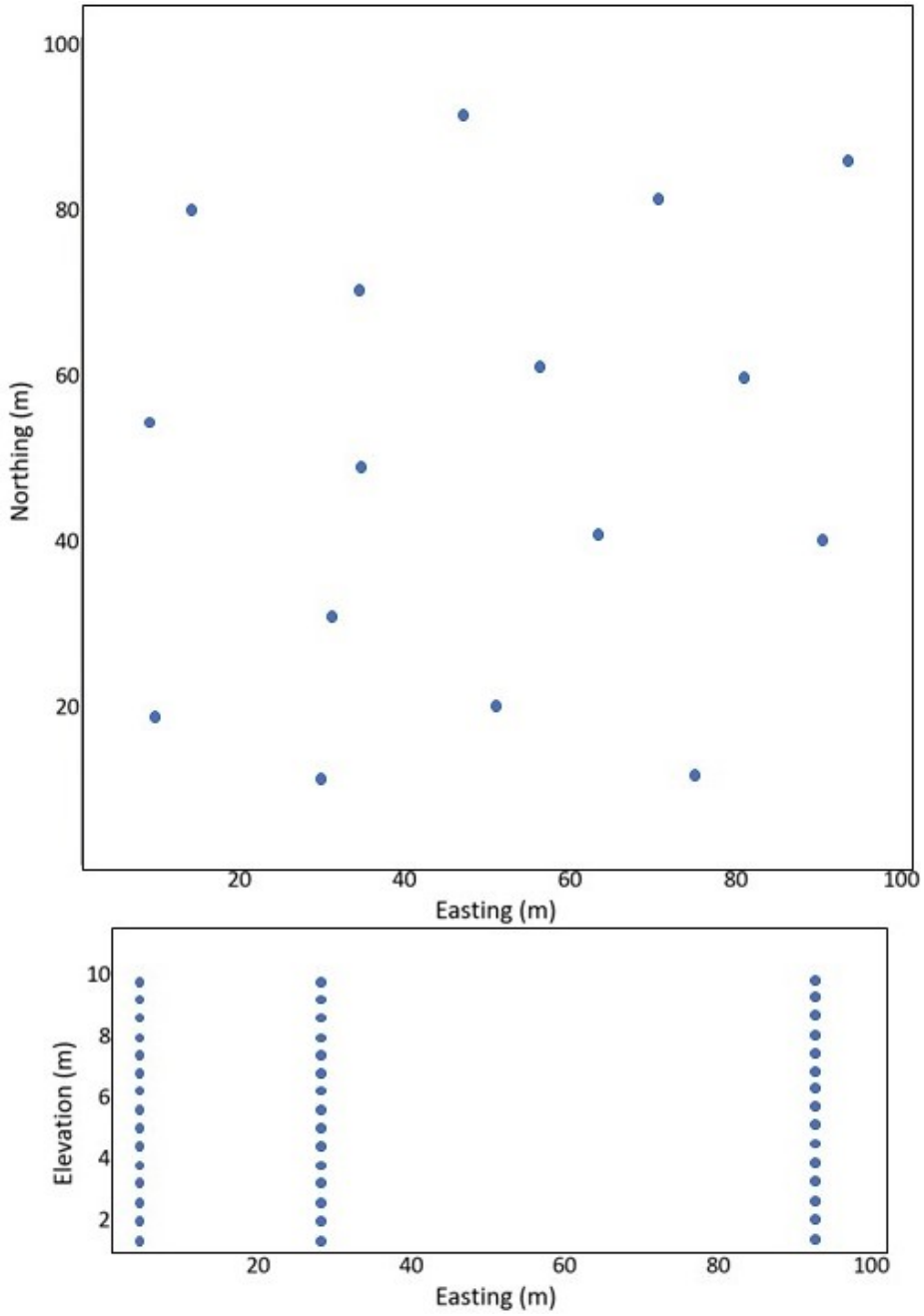

Figure 4.7: A conceptual demonstration of the distribution of samples in the synthetic data set used in histogram uncertainty tests.

4.3.1 Drill hole spacing

Histogram uncertainty in 2D and 3D with variable drill hole spacings is summarized in Figure 4.8. While the histogram uncertainty in 3D seems to remain flat for most spacings, the 2D results show a consistent increase with higher drill hole spacing. This is expected, as higher drill hole spacing leads to less sample density and more uncertainty in the results. However, the presence of more than one sample per drill hole seem to lessen the impact of decreasing drill holes on the 3D results.

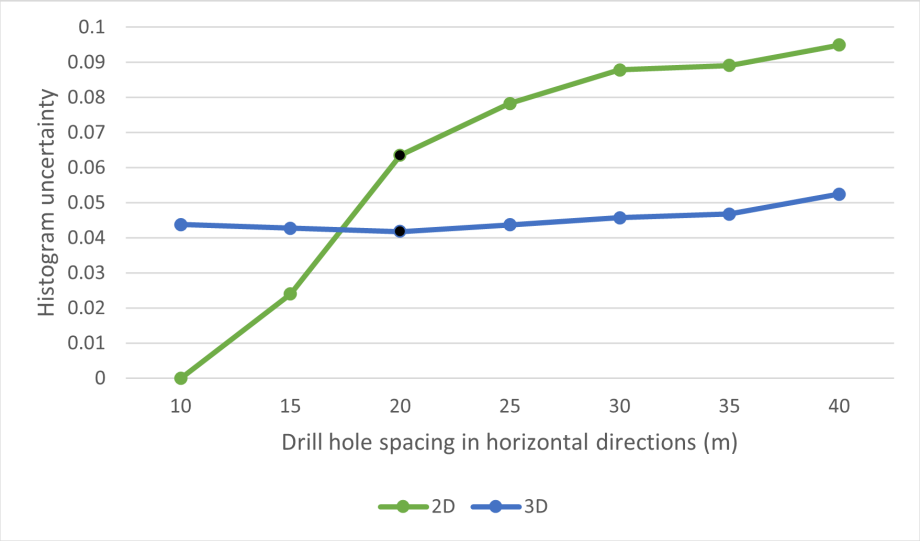

Figure 4.8: Histogram uncertainty with varying values for the horizontal spacing of drill holes in 2D and 3D workflows. The black dots indicate the base case.

4.3.2 Horizontal spatial continuity

Figure 4.9 demonstrates the change in histogram uncertainty given different horizontal ranges of the variogram in both workflows.

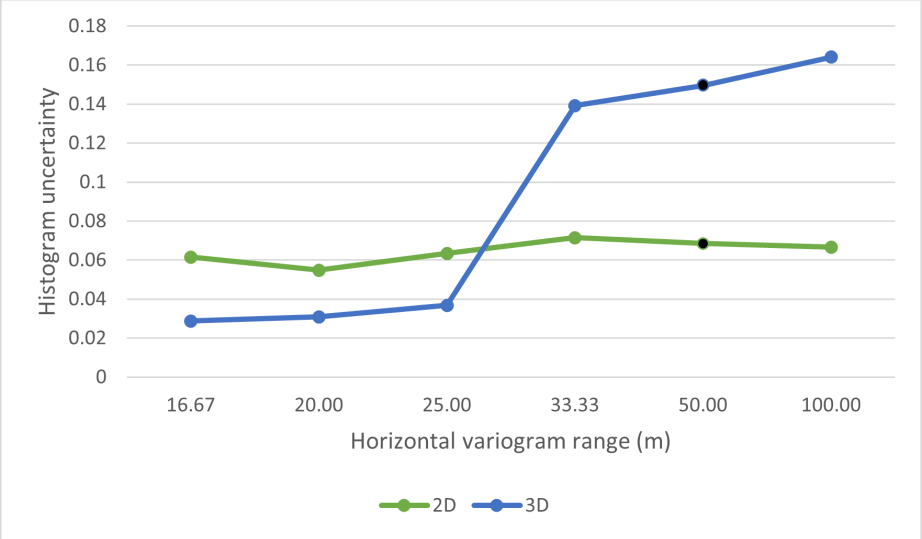

Figure 4.9: Histogram uncertainty with varying values for the horizontal range of the variogram in 2D and 3D workflows. The black dots indicate the base case.

The results in Figure 4.9 show higher histogram uncertainty in 2D for lower variogram ranges and higher values in 3D for higher ranges. 2D results show little changes over different ranges, but 3D results have a significant increase in histogram uncertainty towards higher ranges. It could be interpreted that changes to the spatial structure of the input samples appear to impact the 3D workflow more.

4.3.3 Vertical spatial continuity

Setting variable values for the vertical range of the variogram in the test model lead to results summarized in Figure 4.10. In this case, both 2D and 3D results show an overall increase in histogram uncertainty with higher vertical ranges. More spatial structure in the input samples translates to higher histogram uncertainty. The 2D results show a consistent increase with higher vertical ranges.

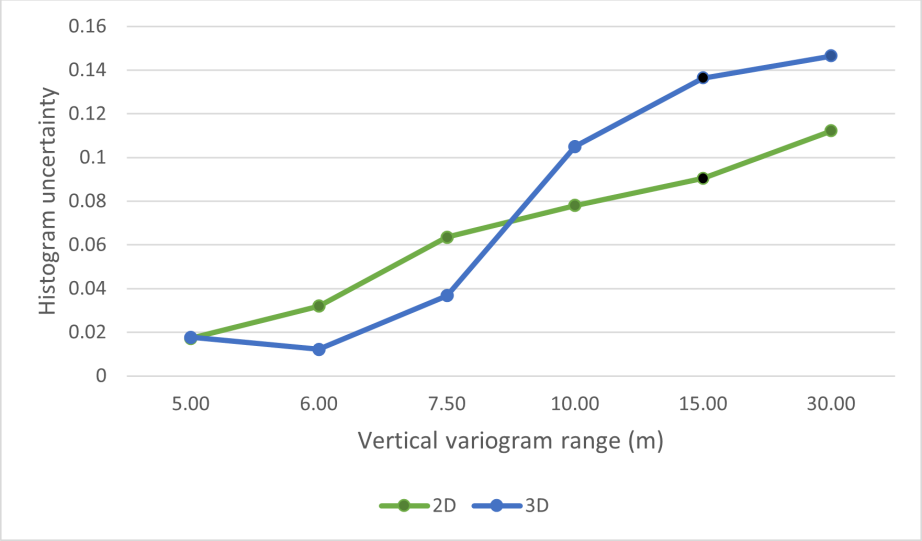

Figure 4.10: Histogram uncertainty with varying values for the vertical range of the variogram in 2D and 3D workflows. The black dots indicate the base case.

4.3.4 Nugget effect

Changing the value for the nugget effect used in creating the input samples leads to Figure 4.11. 2D and 3D results show opposite trends with increasing degrees of nugget effect, with 2D results consistently dropping, as opposed to 3D results which increase with higher nugget effects.

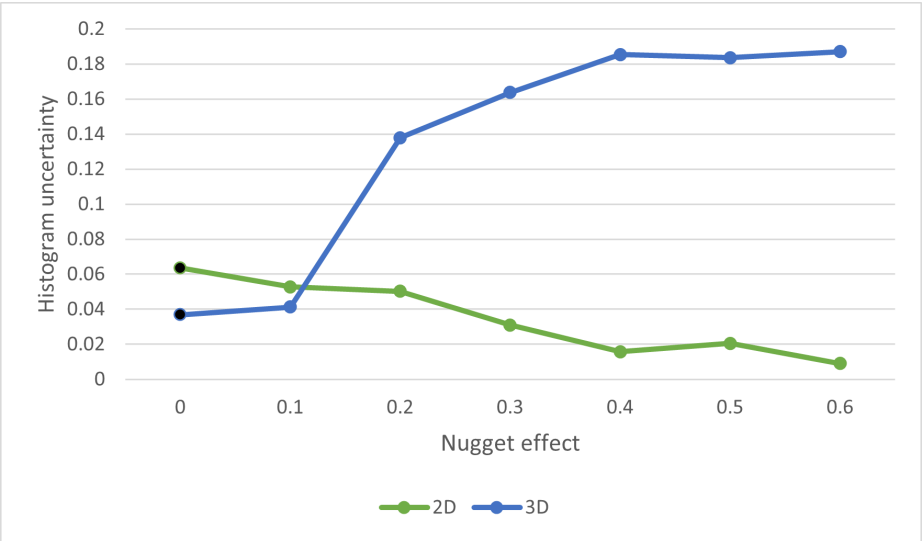

Figure 4.11: Histogram uncertainty with varying values for the nugget effect of the variogram in 2D and 3D workflows. The black dots indicate the base case.

4.3.5 Summary

The test models implemented for analyzing residual uncertainty reveal important patterns and information about the differences between uncertainty quantification in 2D and 3D workflows. In line with the results in Chapter 3, residual uncertainty is observed to be higher in the 2D workflow in all test cases and with different input parameters. This is interpreted to be due to the vertical dimension in the modeling grid and vertical spatial continuity, as is observed in Figure 4.3 and 4.5.

Higher spatial continuity in the vertical direction leads to less pronounced differences between the two workflows and the 3D results become more similar to the 2D. Vertical spatial continuity indicates the presence of information in input samples in the 3D space that are absent in the 2D space, and as vertical continuity increases, the amount of significant information that vertical spatial continuity provides diminishes and makes the 3D samples more similar to a simple averaged 2D sample. This in turn leads to less difference in the uncertainty in those cases. Along with the larger size of the models in the 3D workflow due to the extra dimension present in that workflow, these two are the main observed factors in differences between 2D and 3D residual uncertainty.

Unlike residual tests, the test models conducted on histogram uncertainty are less informative and contain few patterns. The most likely explanation for this difference is how histogram uncertainty is calculated, as subtraction of two different variances, as opposed to residual uncertainty which itself is a variance. This likely leads to more irregular results in terms of histogram uncertainty. Consistent with the rest of the research, this is how histogram uncertainty is measured in this project.

Overall, it appears that 3D histogram uncertainty is higher than 2D, which contradicts the results observed in Chapter 3. This likely further emphasizes the fact that histogram uncertainty is heavily dependent on the nature and configuration of the input data and could be higher or lower in either workflow depending on the factors considered in this chapter.

However, one important takeaway is the results of variable drill hole spacing values. Those results indicate a higher sensitivity towards drill hole spacing in 2D, which is likely due to the smaller size of the 2D data set compared to 3D. This indicates a

higher increase in 2D histogram uncertainty with lower sample density, which is an important note for modelers when considering these two workflows.

Chapter 5

Optimization of sample pairing in the calculation of variograms

The common approach to quantifying spatial correlation in a geostatistical model is through pairing different samples, either in calculating variograms, correlograms, or other methods. Since this pairing process is usually only based on the vector of the distance between a pair of samples, it does not consider the small-scale non-linear geological features that might connect several samples to each other that would not have been connected to one another through typical two-point-based pairings. In the presence of such features, allowing samples to be paired with each other based on the covariance of the petrophysical variables between head and tail samples might lead to a pairing scenario that is based on the geological features that relate samples to each other. In this chapter, a method is developed that finds an optimal pairing scenario between samples of two wells or drill holes.

5.1 Motivation and problem statement

This new method of pairing samples is specifically focused on the geology of the McMurray formation and aims at pairing samples in a way that is aligned with the geology of the formation and captures the correct heterogeneity of the variables by stepping away from the conventional two-point-based methods. While the focus is on the McMurray formation, this method can be applied to many other formations and deposits where the problem of non-linear pairing exists.

Under the assumption of stationarity, the mean and variance of a spatial variable

are constant in the domain and the correlation between any two points is characterized by one of many methods, like variogram, correlogram, or covariance. This two-point correlation is the basis of quantifying spatial correlation for the purpose of geostatistical modeling (Journel and Huijbregts 2003; Pyrcz and Deutsch 2014). The selection of parameters for pairing points is commonly derived from a geological understanding of the domain, as well as the distribution and spatial configuration of samples, and determines which samples are paired together (Chiles and Delfiner 2009). For instance, the understanding of the major and minor directions of anisotropy and the angle of azimuth tolerance dictates which paths to select points along (Wackernagel 2003). Another criterion is the distance between two samples, which associates pairs with specific lag distances (Figure 5.1). The result of the process of selecting pairs of points is the h-scatter plot that is the input to calculating the experimental semi-variogram (Figure 5.2) (Pyrcz and Deutsch 2014; Chiles and Delfiner 2009).

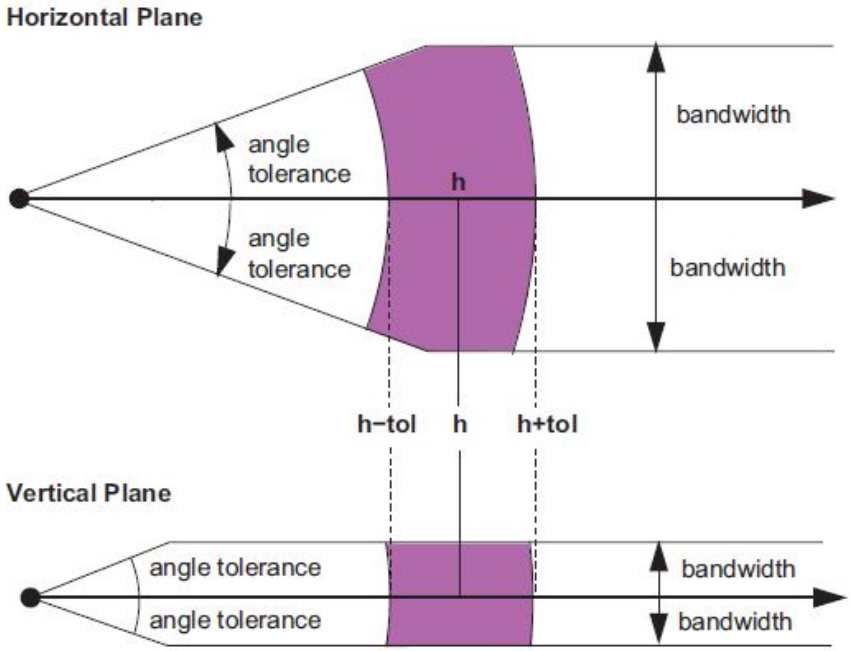

Figure 5.1: Illustration of tolerance parameters and how they define which points are paired to each other (Pyrcz and Deutsch 2014).

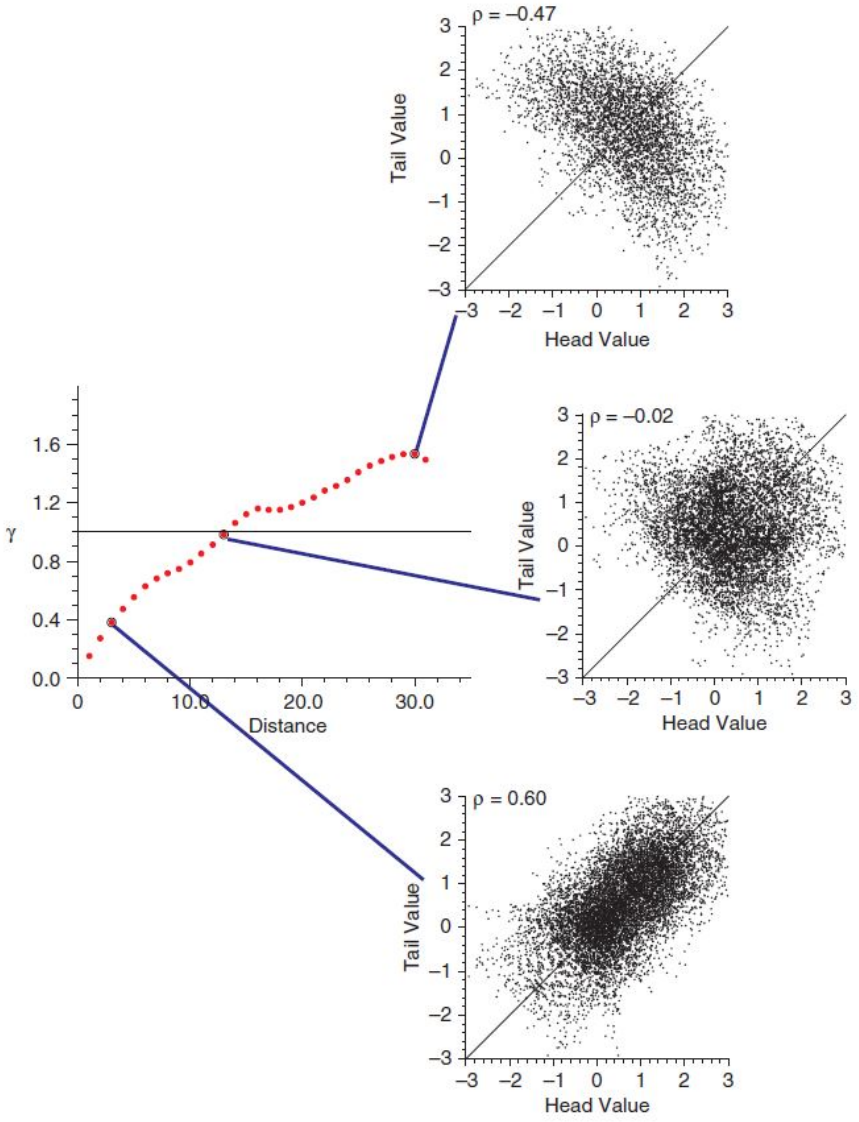

Figure 5.2: The h-scatter plot is the combination of several samples that are paired to each other for a specific lag distance and based on specific tolerance parameters (Pyrcz and Deutsch 2014).

What is often lacking in this approach is a detailed consideration of small-scale non-linear geological features that connect points to each other that might not follow simple distance criteria. This could be the case in certain geological domains, like sandy shales of the McMurray formations of Alberta, which have very small-scale features and petrophysical properties are highly continuous within these features. Such features might not follow linear patterns, and this results in their importance being missed in simple distance vector considerations when pairing samples.

The motivation of this chapter is to develop an algorithm for the pairing of samples that are based on the value of the petrophysical variable of interest and considers the possibility that disrupting geological processes, such as folding or faulting, could lead to a non-linear distribution of variables in space. In this way, samples are paired in a way that maximizes the correlation of a petrophysical variable between the head and tail samples and would lead to an explicit capturing of small-scale non-linear features.

5.2 Methodology

The focus of this research is on capturing and incorporating the effect of small-scale geological features on the pairing process and the basic approach is to compare two drill holes at a time. This means that at every step, two drill holes or wells are compared and the pairing of samples between these two is considered. The calculation of a semi-variogram or covariance between two drill holes is then repeated for many pairs of wells or drill holes. Since the lateral distance between two drill holes is constant for all the samples within them, averaging the covariance or the semi-variogram for pairs of drill holes with the same lag distance will produce the experimental semi-variogram. It is assumed that the pair of wells are parallel and vertical and deviated drill holes are not considered.

As an example, two drill holes are considered in Figure 5.3. Under the conventional methodology, samples in the two drill holes are paired with each other depending on their distance vector, which here means their relative elevation (for the purpose of calculating the horizontal experimental variogram). This scenario is depicted in Figure 5.3, where each green point on the sides represents a sample and each blue line is the connection between two samples on the two drill holes.

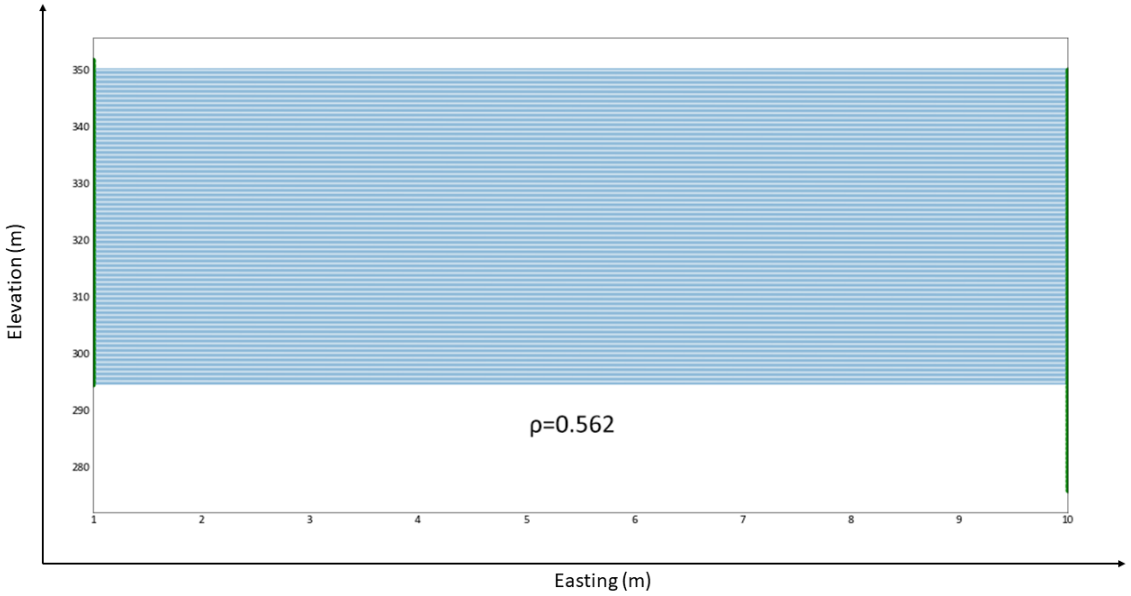

Figure 5.3: Pairing of samples in two drill holes depending on their distance vector, which here means their relative elevation to each other. Green points are samples in a drill hole and blue lines represent the pairing of samples between the two wells. This pairing results in a correlation coefficient of 0.562 for the spatial variable of interest.

To maximize the covariance between the paired samples in the two wells or drill holes, the pairing process must be optimized using an optimization algorithm. In the current paper, a greedy optimization algorithm is implemented (Black *et al.* 2020). To divide the samples in the two wells or drill holes into several distinct geological features that define the pairing process, several control points are randomly chosen on the two drill holes. The samples that fall between these control points are paired with their mirroring group on the other drill hole. For instance, Figure 5.4 shows how two drill holes are divided into three groups using two control points. Each group is paired with its equivalent on the other drill hole.

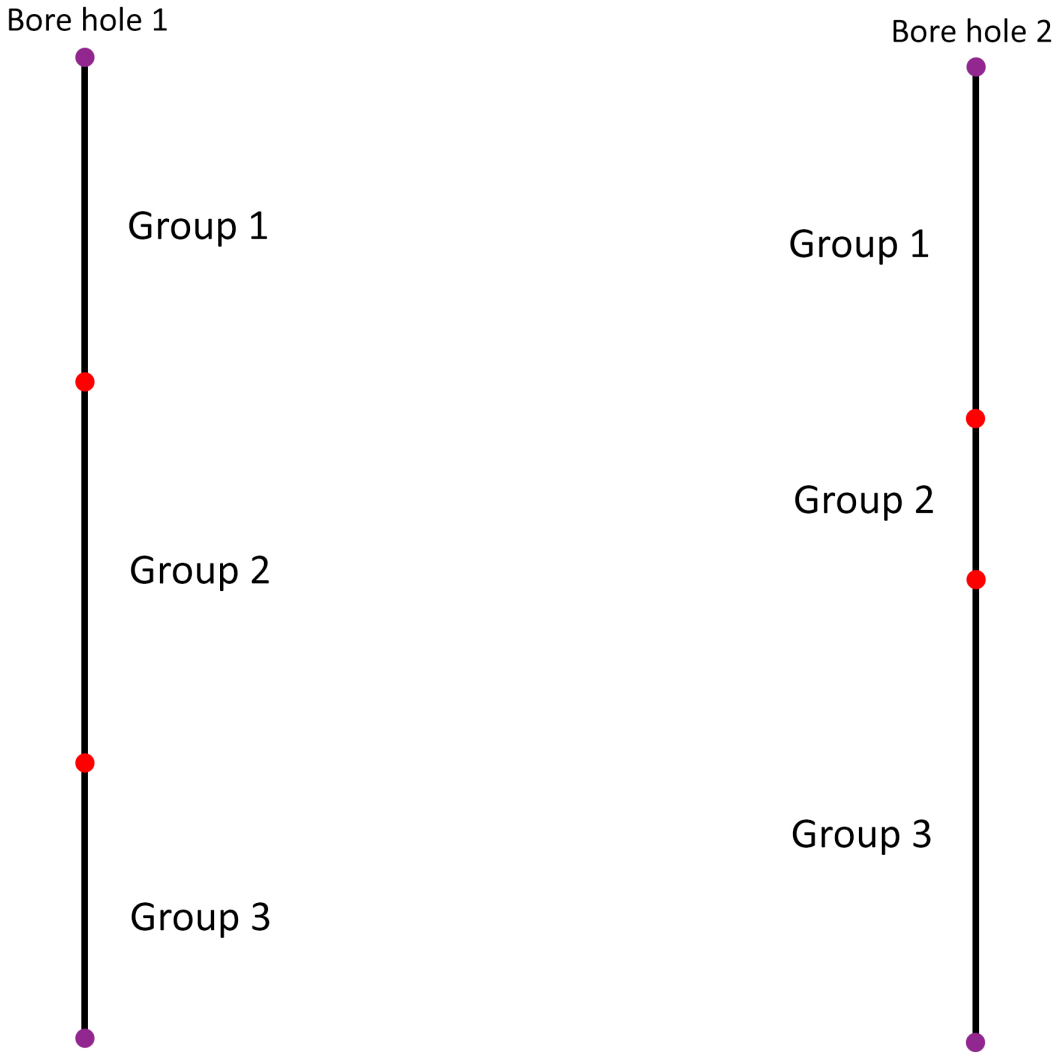

Figure 5.4: An example of two drill holes and how the samples are divided into three groups using two control points. The purple dots mark the top and bottom sample of each drill hole, while the red dots represent a selection of control points.

As said above, the pairing process allows for multiple points to be paired with each other. When an unequal number of samples exist in the same pairing group in the two drill holes, samples from the smaller group will have multiple samples from the larger group paired with them. The only limit is that any sample must be paired with at least one other sample. The algorithm goes through the following loop:

1- Randomly select a given number of control points from between the samples in the two drill holes.

2- Match all the points in the same groups from the two drill holes.

3- For each of the control points, go through a given number of iterations. Each time, randomly changing the control point from among samples between the neighboring bottom and top control points. Each time the control point is changed, recalculate the covariance of the new pairs. Accept the change in the iteration if the overall covariance of pairs increases.

4- Once the algorithm goes through all the control points, a random restart is initiated, where the control points are randomly reselected. The same process of step 3 is repeated for the new set of control points. By the end, if the results are improved, the new set of control points replace the previous ones. This process is repeated for a given number of random restarts.

By the end of the algorithm's implementation, the experimental semi-variogram value is calculated for the optimal group of pairs between the two drill holes. Figure 5.5 shows the result of such a pairing process. Here, the deciding factor was the covariance of effective porosity between groups of pairs from the two wells.

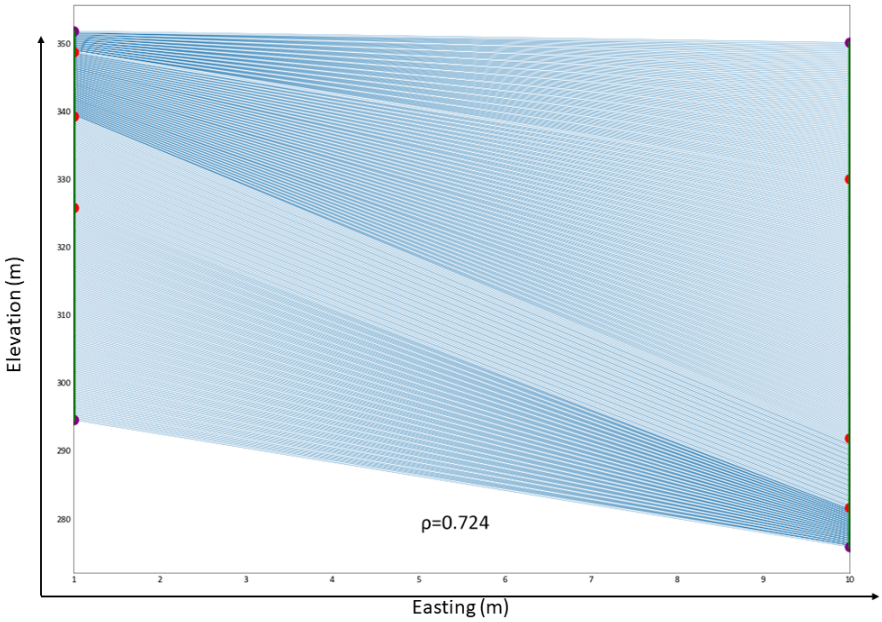

Figure 5.5: Pairing of samples in the two wells according to the optimization algorithm. The red dots represent the final control points, as optimized by the program. The purple dots are the top and bottom samples of the drill holes, which are treated as control points for the purpose of dividing samples into different groups, but remain constant through the algorithm. The correlation of head and tail samples in this pairing scenario is 0.724, which is higher than the result from the conventional pairing.

The control points can be seen as a method of detecting the footprint of a set of complex geological features within the two drill holes. The optimal set of control points results in certain groups of samples being paired from the two drill holes that might have very different elevations and might never have been paired in a conventional pairing process. But since the target is the maximum covariance of a spatial variable between the group of pairs, the result will point to groups of samples that have the greatest amount of similarity in terms of petrophysical properties. As a result, these control points and how they group the samples together are meant to replicate the geological features that exist in the space between the two drill holes. Referring to the example in Figure 5.4, Figure 5.6 illustrates how the grouping of samples is meant to infer the spread of geological features in the space between the two drill holes.

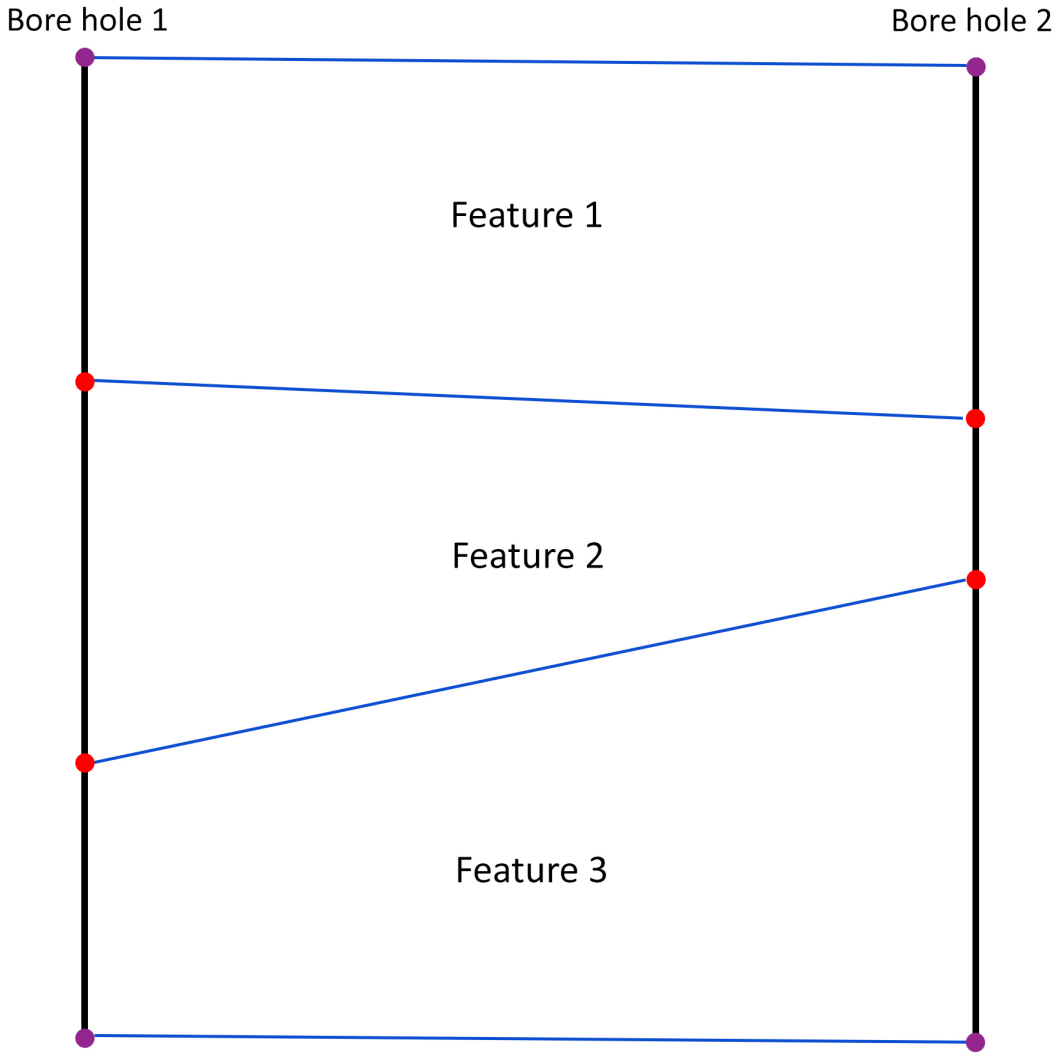

Figure 5.6: The spread of control points in the optimal model is an interpretation of the presence of geological features in the space between the two drill holes.

One additional feature in this method is a tolerance parameter that is meant to allow the user to determine the degree of linear behavior in the optimization workflow. Based on a preexisting understanding of the geological domain, the user could define how linear the pairing of samples should be. A high degree of tolerance in the optimization workflow leads to a possibility of highly non-linear pairing, while a low degree of tolerance leads to a more linear result.

This tolerance is implemented as a range of values that the elevations of control points can take. The length of each borehole is divided into equal parts based on the

number of control points, and the center of each part is the predetermined elevation of each control point. Higher tolerances lead to a larger range of allowed elevations for each control point around these fixed elevations. This definition means that even a full 100 percent tolerance also has some degree of linear behavior. These tolerances are illustrated in Figure 5.7 and the formula for finding the ranges can be seen in Equations 5.1 and 5.2. In these Equations, $R1$ and $R2$ refer to the bounding limits of the elevation of control points in each borehole, CP refer to the number of control points in each borehole, $LBH1$ and $LBH2$ refer to the length of the array of samples in each borehole, and Tol is tolerance parameter defined by the user from 0 to 1.

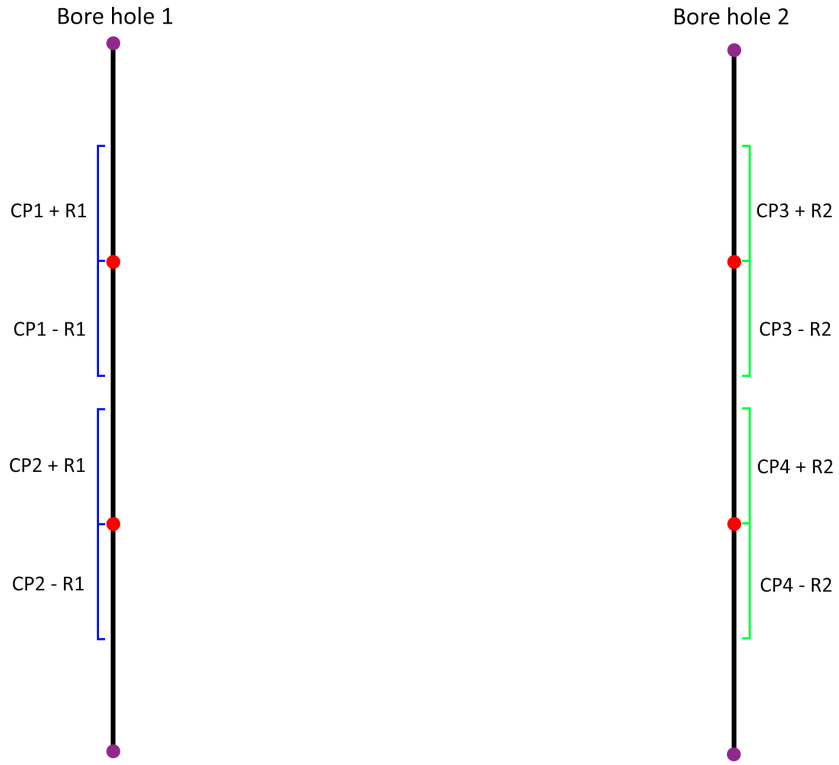

Figure 5.7: In the current revised workflow, a tolerance parameter is used to define the bounding limits of elevations of control points (R1 and R2) in each drill hole. These ranges are calculated using Equations 5.1 and 5.2.

$$R1 = ((LBH1/(CP+1))/2) * Tol \tag{5.1}$$

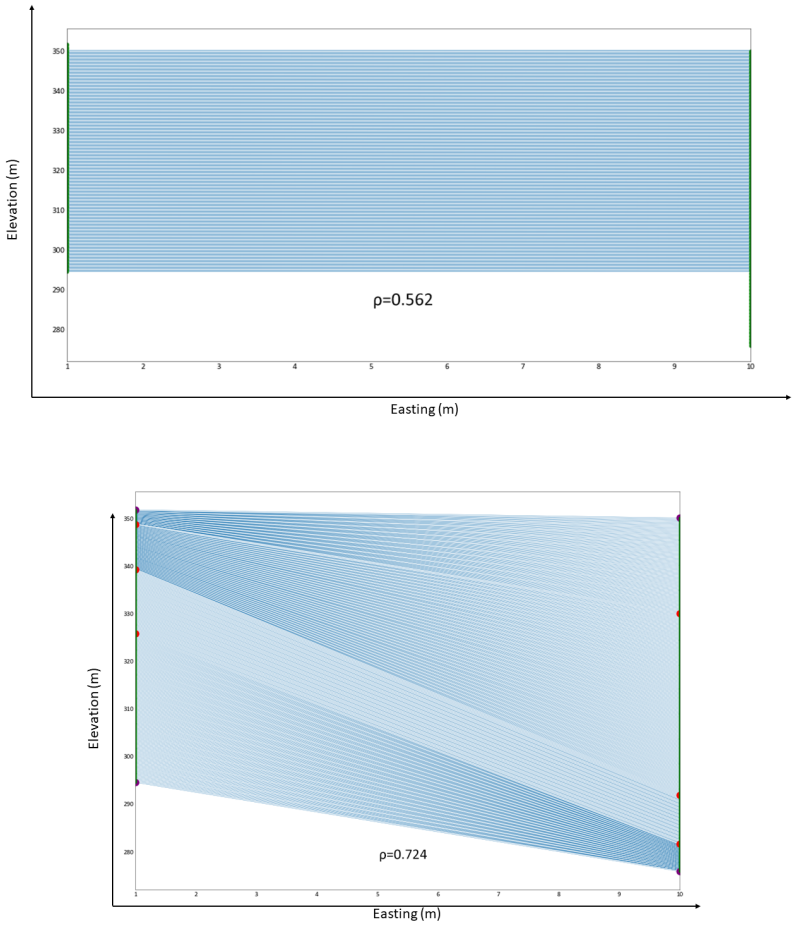

Figure 5.8: An example of the optimal sample pairing algorithm implemented on two wells, with results in a simple pairing scheme (top) and the optimal pairing scheme (bottom) showing a higher correlation in the results from the workflow

$$R2 = ((LBH2/(CP+1))/2) * Tol \tag{5.2}$$

5.3 Results

In this section, the results of this algorithm for a 14-well data set located at the McMurray formations will also be reviewed. The McMurray oil formations are known for their small-scale shale features that make it an excellent choice for implementing the optimization algorithm. In all the following results, the variable oil saturation is used to determine the optimal pairing set.

The results in Figure 5.8 demonstrate a typical performance by the optimal algorithm. Here, three control points are given to the program and the samples in the two wells are divided accordingly. A noticeable improvement in the correlation of the paired samples from 0.562 to 0.724 shows how the algorithm matches samples to improve their geological similarity.

Figure 5.9 shows similar results in another well pair along with the geological surfaces in those two wells. The results here show that although the algorithm does not exactly detect or predict the correct elevation of surfaces, it does divide the two wells into sections of similar behavior in terms of the spatial variable (here, oil saturation) (Figure 5.9).

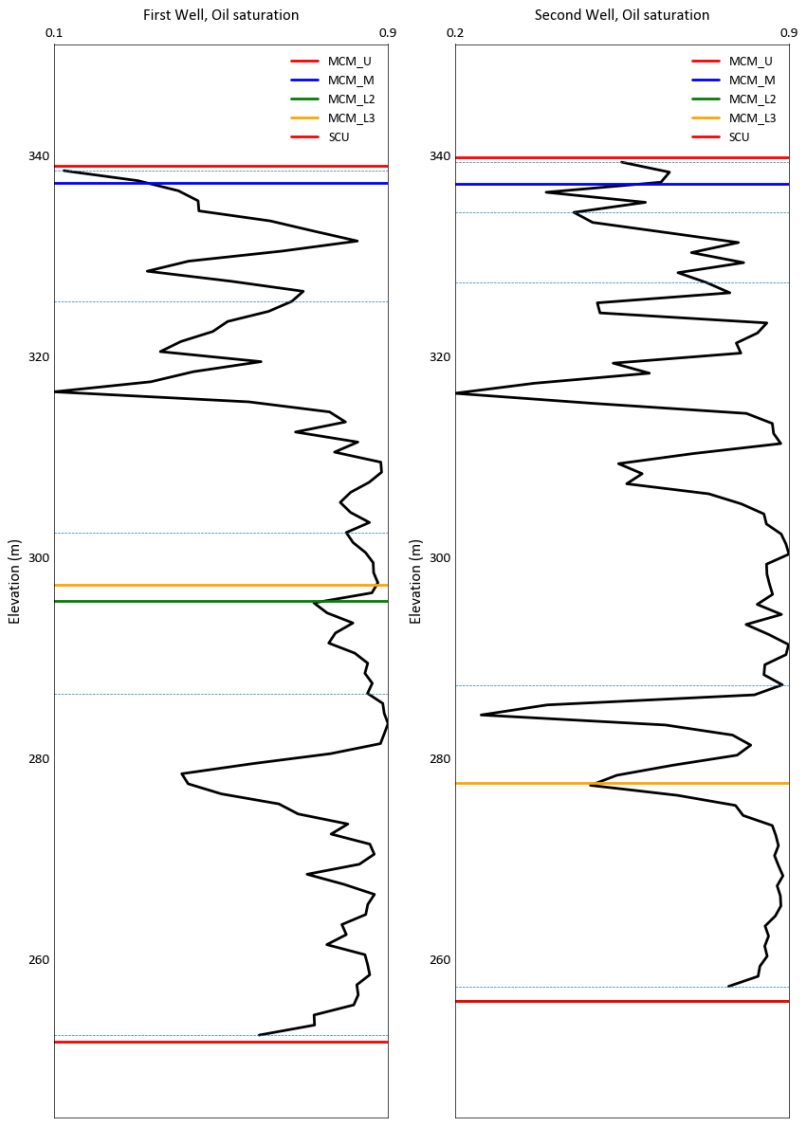

Figure 5.9: The results of optimizing the pairing scheme using three control points in two wells. The colored lines refer to the elevation of geological surfaces in these two wells, while the light blue lines refer to the elevation of control points. The black lines are the drill hole plots illustrating the value of oil saturation in the two wells.

Figure 5.10 shows the impact of tolerance in the optimal correlation found in the algorithm in different pairs of wells. Each set of colored points refers to one pair of wells that have been used as input in the pairing algorithm. Although in general the higher the tolerance the optimal correlation also increases, this trend has different levels of intensity between the different well pairs. This shows that when a pair of wells have a high degree of linear behavior, it is highly probable that the global optimal pairing scheme is found in the low tolerances. In contrast, when the formation is

highly non-linear within the well pair, the true optimal set is only revealed at high tolerances.

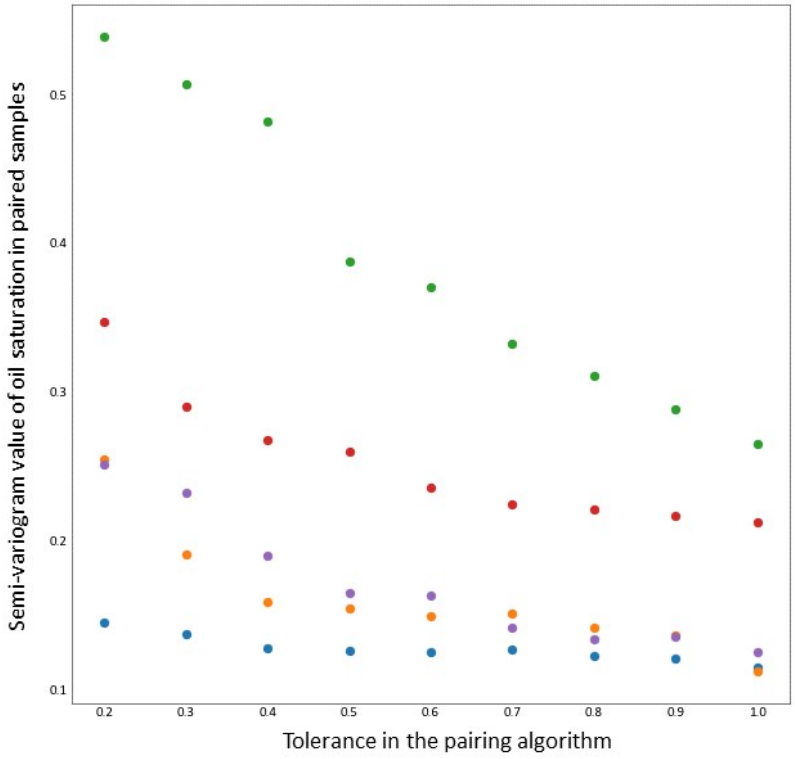

Figure 5.10: The relation between tolerance of the pairing algorithm and the semi-variogram value of the paired samples show a general higher correlation (lower variogram value) with higher tolerance but with different degrees of intensity

In general, the pairing algorithm leads to a more continuous variogram compared to conventional two-point pairing schemes (Figure 5.11).

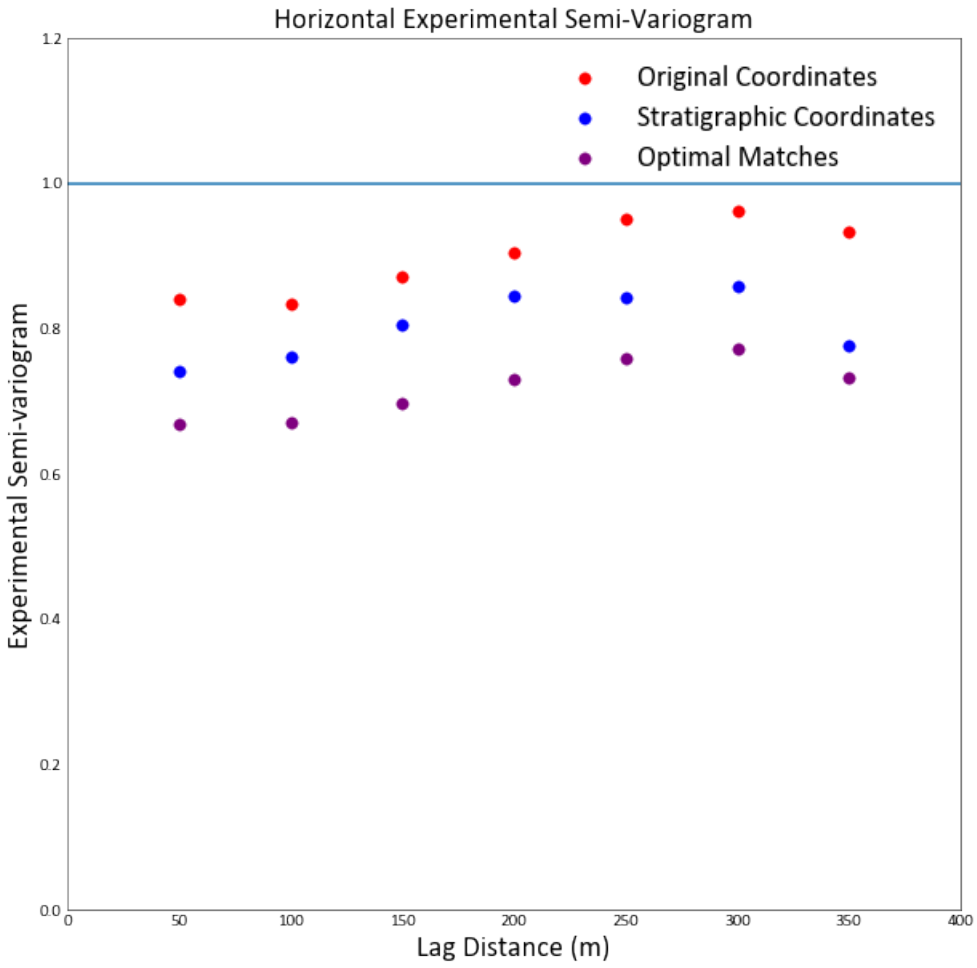

Figure 5.11: The experimental variogram calculated on the same data set with three different schemes: original elevations (simple pairing), stratigraphic elevations (simple pairing), and the optimal scheme. The results show a more continuous variogram in the optimal scheme.

5.4 Conclusion

The pairing of samples is a fundamental question in any quantification of spatial correlation in a geostatistical model. This is usually done based on the vector of the distance between any pair of samples, which pairs them based on lag distances, angle tolerances, and other variogram calculation parameters. But this pairing scheme does

not consider the role of geological features which might relate certain samples to each other that might not have a clear and straightforward connection in terms of their distance vectors. In other words, the geological background of points and the features to which they belong could have a major impact on the actual spatial correlation in the domain. This is not properly captured in a normal pairing scheme.

An alternative method for pairing samples was proposed which was based on grouping samples from two drill holes or wells based on the geological features that occupy the space between those two wells or drill holes. This method optimizes the pairing based on the covariance of the spatial variable of interest in tail and head samples. In this method, a greedy optimization algorithm is proposed, whereby defining a certain number of control points that effectively divide samples into groups and optimizing the placement of these control points in the two drill holes or wells, the pairing of points is optimized for a given number of iterations and random restarts.

As was seen in the results, this workflow improves the spatial correlation between samples and yields a horizontal experimental variogram that has higher continuity than the pairing made in original or stratigraphic coordinates. The algorithm manages to divide the two wells/bore holes into sections of similar patterns in the value of the spatial variable of interest. The algorithm also shows a higher correlation in results when the tolerance increases.

For access to the source code used in the algorithm described in this chapter, readers are encouraged to contact the authors to provide access to the source code.

Chapter 6

Conclusion

6.1 Summary

As per the thesis statement, the purpose of this research is to implement probabilistic resource modeling in 2D and 3D modeling workflows and discuss possible differences between the two workflows and the contributing factors to those differences.

Towards this goal, the main methodology used to generate probabilistic resource modeling in the case study is described through a step-by-step analysis in Chapter 2. Chapter 2 highlights steps in the modeling process where uncertainty is quantified in the final results and how this is done differently in the 2D and 3D workflows, which helps better explain uncertainty in final results in later chapters.

Following a description of the methodology, Chapter 3 details the results of implementing the methodology in the case study of the McMurray formation of Northern Alberta, in addition to multiple checks and validations of the modeling results. Results in Chapter 3 point to similar global uncertainty in estimations of OOIP in both 2D and 3D without facies workflows, although higher uncertainty in the 3D with facies workflow due to additional uncertainty in the modeling of lithofacies. The contribution of each of the three main categories of uncertainty is also measured separately in Chapter 3, which points to the high impact of uncertainty in the geometry of the domain, followed by residual uncertainty and uncertainty in model parameters, dominated by histogram uncertainty. Additional analyses on the impact of well spacing density and SAGD pad uncertainty are also considered in Chapter 3.

Chapter 4 focuses on residual and histogram uncertainty, with the goal of identifying what factors lead to these two categories of uncertainty being higher or lower

in 2D and 3D workflows using several test models. Results of residual uncertainty tests generated by unconditional simulation point to two critical factors in differences of residual uncertainty in 2D and 3D workflows: the vertical size of the domain or thickness of the domain, and the degree of vertical spatial continuity. While higher thickness leads to more differences between the two workflows, lower spatial continuity has the same effect. While histogram uncertainty is observed to show few patterns and appears to be dependent on the specific data set and domain, the test results show higher impact on 2D histogram uncertainty as a result of sparse drill hole density, which is likely due to the smaller size of the 2D data set.

Finally, Chapter 5 reviews a method of pairing samples in the process of variogram modeling in a stratigraphic domain. This method relies on maximizing the correlation between paired groups of samples using a greedy optimization algorithm. The method relies on a geological understanding of the McMurray formation, or other stratigraphic formations, where undulating processes have disturbed the horizontal continuity of structures and simple pairing schemes might not capture the full spatial continuity in the domain. Using this method, maximizing the correlation of paired samples within a given number of groups could lead to the detection of those structures and an improvement in the spatial continuity of variograms. The results show that while the algorithm cannot fully detect existing layers, it can provide a variogram with more continuity compared to the original variogram.

6.2 Contributions

This thesis provides a number of insights on the key question in the thesis statement. Regarding the results from the case study, the key takeaway from results in Chapter 3 is the almost equal uncertainty quantified by 2D and 3D without facies modeling workflows when all aspects of parameter uncertainty are considered, as opposed to higher residual uncertainty in the 2D workflow in the absence of any parameter uncertainty. Once facies are incorporated into the 3D workflow, the results show higher uncertainty compared to 2D. This overall comparison could be used as a default presumption in any domain within the Athabasca oil sands, and to some degree in other stratigraphic domains with similar characteristics. This could help modelers across

the industry to shape some expectations before building a more detailed model of these domains.

The higher residual uncertainty in 2D is one of the main motivations for this research and is also observed in the results. Common industry practices in probabilistic resource estimation frequently use fixed modeling input parameters and a fixed geological boundary model (no uncertainty in model parameters and no uncertainty in geometry), which leads to the entire global uncertainty in results to be residual uncertainty, hence the observation of higher 2D uncertainty shown in Figure 1.3. Modelers should have an expectation of higher 2D uncertainty in the absence of parameter or geometry uncertainty.

These results reinforce the practice of incorporating all aspects of uncertainty for a meaningful assessment of resources. Given the assumptions behind those results and the fact that these results come from a single geological formation, the conclusions and recommendations can be applied to any other domain within the Athabasca oil sands (given the sufficient similarity with the McMurray formation) and to some degree to other formations where 2D-like characteristics can be observed, such as tabular coal deposits or other stratigraphic formations.

The results from Chapter 3 also indicate the importance of including uncertainty in surfaces and boundaries of stratigraphic formations as a best practice and how critical they could be in the global uncertainty of results. Approaches to capturing uncertainty in surfaces and boundaries are diverse and could lead to different quantification of uncertainty, and depending on the geological context behind the models, it may be preferred to assume a fixed boundary. It must also be pointed out that while all the uncertainty in variable boundaries is considered as a single category, this uncertainty itself comes from a variety of sources, such as histogram and variogram uncertainty in modeling the surfaces.

Similar to results from the case study, results from test models in Chapter 4 show that residual uncertainty is always higher in the 2D workflow, primarily due to the smaller size of the modeling domain in that workflow (due to the lack of a vertical dimension in the modeling grid). Higher thickness leads to a larger disparity in the size of the 2D and 3D modeling grids and leads to lower 3D uncertainty and thus more disparity between the two workflows. On the other hand, higher vertical spatial

continuity leads to lower 3D uncertainty which leads to lower differences between 2D and 3D results. A higher degree of vertical geological structure which appears in the form of higher vertical spatial continuity increases the amount of information lost in the transition of the data set from 3D to 2D (in this study through averaging of 3D samples) and leads to more significant differences in the uncertainty generated in the two workflows.

The test results point to these two factors to be major controls on the degree of differences in residual uncertainty between 2D and 3D workflows, which appears to be higher in 2D in almost all cases. This means that modelers should pay close attention to these two factors in a domain when determining the degree of disparity between residual uncertainty in 2D and 3D.

6.3 Assumptions and Limitations

One of the key assumptions in this research is the focus on tabular deposits and features observed within these domains. While many tabular and stratigraphic formations share features like the ones in the McMurray and many of the results in this research can be extended to other tabular formations, there are unique geological characteristics within each stratigraphic formation that has the potential of altering the results of probabilistic modeling. This also extends to types of samples and input data that could also have a significant impact on the results. Focusing on a single formation and a single data set might have prevented a broader understanding of 2D and 3D resource estimation in different contexts.

Additionally, some specific modeling workflows and tools such as SGS, SIS, and spatial bootstrap are also used in this research which could be seen as a limitation since other geostatistical methods and their corresponding results were not explored. Other geostatistical tools and workflows could potentially lead to different results, but the methods used here are commonly used by geostatistical modelers and are well-established and well-tested.

Regarding uncertainty in geometry, some important assumptions and limitations exist. Firstly, uncertainty in the aerial boundaries of the domain is excluded due to the geological nature of the McMurray formation which has no natural aerial

boundary within the lease examined. Including uncertainty in aerial boundaries could be important in other domains. Additionally, the approach to quantify uncertainty in geometry through thickness and surface modeling is only one way to approach this problem and other geological or geophysical workflows could lead to a different measurement of uncertainty in the geometry of the domain. Also, different types of formations and domains might require different approaches. As a result, analyzing the consequences of different modeling approaches in different domains on the uncertainty in the final results requires another study that is out of the scope of this research.

6.4 Future works

While the current research sheds some light on how 2D and 3D modeling of a tabular or stratigraphic formation could have major consequences on the uncertainty driven by those models, there are some important gaps that must be further explored.

Firstly, uncertainty in geometry is largely left unexplored in terms of how different approaches and methods could lead to different results. Additionally, exploring the contribution of aspects of uncertainty could also be further analyzed in this case.

Secondly, applying probabilistic resource estimation in other tabular and stratigraphic formations with different data sets and geological characteristics could further validate the findings and provide confidence in generalizing the findings to other geological settings. While this research provides many domain-specific observations for the Athabasca oil sands, those research projects could do the same in other domains.

Bibliography

[1] M. Pyrcz and C. Deutsch, *Geostatistical Reservoir Modeling, 2nd Edition, Oxford University Press, New York, p. 448.* May 2014, ISBN: 978-0199731442.

[2] M. Rossi and C. Deutsch, *Mineral Resource Estimation*. Jan. 2014, ISBN: 978-1-4020-5716-8. DOI: 10.1007/978-1-4020-5717-5.

[3] P. Darling, *SME mining engineering handbook*. SME, 2011, vol. 1.

[4] A. Journel and C. Huijbregts, *Mining Geostatistics*. Blackburn Press, 2003, ISBN: 9781930665910. [Online]. Available: https://books.google.ca/books?id=Id1GAAAAYAAJ.

[5] W. Ren, J. Mclennan, O. Leuangthong, and C. Deutsch, "Reservoir characterization of mcmurray formation by 2d geostatistical modeling," *Natural Resources Research*, vol. 15, pp. 111–117, Jan. 2006. DOI: 10.1007/s11053-006-9011-8.

[6] J.-L. Mallet, "Space–time mathematical framework for sedimentary geology," *Mathematical Geology*, vol. 36, pp. 1–32, Jan. 2004. DOI: 10.1023/B:MATG.0000016228.75495.7c.

[7] C. Langenberg, F. Hein, D. Lawton, and J. Cunningham, "Seismic modeling of fluvial-estuarine deposits in the Athabasca oil sands using ray-tracing techniques, Steepbank River area, northeastern Alberta," *Bulletin of Canadian Petroleum Geology*, vol. 50, no. 1, pp. 178–204, Mar. 2002, ISSN: 0007-4802. DOI: 10.2113/50.1.178. eprint: https://pubs.geoscienceworld.org/cspg/bcpg/article-pdf/50/1/178/3312188/178.pdf. [Online]. Available: https://doi.org/10.2113/50.1.178.

[8] K. Khan and C. Deutsch, "Practical incorporation of multivariate parameter uncertainty in geostatistical resource modeling," *Natural Resources Research*, vol. 25, May 2015. DOI: 10.1007/s11053-015-9267-y.

[9] S. Mondal, L. Zeite, and M Mallik, "2d geostatistical modeling and volume estimation of an important part of western onland oil field, india," Dec. 2015.

[10] D. G. Krige, "A statistical approach to some basic mine valuation problems on the witwatersrand," *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, no. 6, pp. 119–139, 1951.

[11] M. E. Hohn, "An introduction to applied geostatistics: By edward h. isaaks and r. mohan srivastava, 1989, oxford university press, new york, 561 p., isbn 0-19-505012-6, isbn 0-19-505013-4 (paperback), 55.00*cloth*,35.00 paper (us)," *Computers & Geosciences*, vol. 17, no. 3, pp. 471–473, 1991.

[12] M. David, *Geostatistical ore reserve estimation*. Elsevier, 2012.

[13] G Matheron, "The theory of regionalised variables and its applications," *Les Cahiers du Centre de Morphologie Mathématique*, vol. 5, p. 212, 1971.

[14] G Matheron and F Blondel, "Traité de géostatistique appliquée. tome 1. editions technip, paris," *Á. Martínez-del-Pozo et al*, 1962.

[15] A. Journel, "Resampling from stochastic simulations," *Environmental and Ecological Statistics*, vol. 1, no. 1, pp. 63–91, 1994.

[16] A. G. Journel, "Modelling uncertainty and spatial dependence: Stochastic imaging," *International Journal of Geographical Information Systems*, vol. 10, no. 5, pp. 517–522, 1996. DOI: 10.1080/02693799608902094.

[17] O. Bertoli, J. Mukangai, J Vann, and S. Dunham, "Two-dimensional geostatistical methods — theory, practice and a case study from the 1a shoot nickel deposit, leinster, western australia," *Proceedings Fifth International Mining Geology Conference*, Jan. 2003.

[18] D. Harris and R. Perkins, "A case study of scaling up 2d geostatistical models to a 3d simulation model," SPE Annual Technical Conference and Exhibition, Oct. 1991. DOI: 10.2118/22760-MS. [Online]. Available: https://doi.org/10.2118/22760-MS.

[19] L. Wang *et al.*, "A mps-based novel method of reconstructing 3d reservoir models from 2d images using seismic constraints," *Journal of Petroleum Science and Engineering*, vol. 209, p. 109 974, 2022, ISSN: 0920-4105. DOI: https://doi.org/10.1016/j.petrol.2021.109974. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920410521015886.

[20] P. Ballin, A. Journel, and K. Aziz, "Prediction of uncertainty in reservoir performance forecast," *Journal of Canadian Petroleum Technology*, vol. 31, no. 04, Apr. 1992, ISSN: 0021-9487. DOI: 10.2118/92-04-05. eprint: https://onepetro.org/JCPT/article-pdf/doi/10.2118/92-04-05/2169296/petsoc-92-04-05.pdf. [Online]. Available: https://doi.org/10.2118/92-04-05.

[21] A. G. Journel, "Modeling uncertainty: Some conceptual thoughts," in *Geostatistics for the Next Century: An International Forum in Honour of Michel David's Contribution to Geostatistics, Montreal, 1993*, R. Dimitrakopoulos, Ed. Dordrecht: Springer Netherlands, 1994, pp. 30–43, ISBN: 978-94-011-0824-9. DOI: 10.1007/978-94-011-0824-9_5. [Online]. Available: https://doi.org/10.1007/978-94-011-0824-9_5.

[22] J.-P. Chiles and P. Delfiner, *Geostatistics: modeling spatial uncertainty*. John Wiley & Sons, 2009, vol. 497.

[23] F. Wang and M. M. Wall, "Incorporating parameter uncertainty into prediction intervals for spatial data modeled via a parametric variogram," *Journal of agricultural, biological, and environmental statistics*, vol. 8, no. 3, pp. 296–309, 2003.

[24] O. Babak and C. V. Deutsch, "Accounting for parameter uncertainty in reservoir uncertainty assessment: The conditional finite-domain approach," *Natural resources research*, vol. 18, no. 1, pp. 7–17, 2009.

[25] P. Dowd and E Pardo-Igúzquiza, "The incorporation of model uncertainty in geostatistical simulation," *Geographical and Environmental Modelling*, vol. 6, no. 2, pp. 147–169, 2002.

[26] P. J. Diggle and P. J. Ribeiro Jr, "Bayesian inference in gaussian model-based geostatistics," *Geographical and environmental modelling*, vol. 6, no. 2, pp. 129–146, 2002.

[27] A. G. Journel and A. Bitanov, "Uncertainty in n/g ratio in early reservoir development," *Journal of Petroleum Science and Engineering*, vol. 44, no. 1-2, pp. 115–130, 2004.

[28] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.

[29] H. Putter and G. A. Young, "On the effect of covariance function estimation on the accuracy of kriging predictors," *Bernoulli*, pp. 421–438, 2001.

[30] A. Journel and P. Kyriakidis, *Evaluation of Mineral Reserves: A Simulation Approach* (Applied Geostatistics). Oxford University Press, 2004, ISBN: 9780199771264. [Online]. Available: https://books.google.ca/books?id=ujSHhABUYNwC.

[31] J. Masliyah, J. Czarnecki, and Z. Xu, "Handbook on theory and practice of bitumen recovery from athabasca oil sands - volume 1: Theoretical basis," *Kingsley Knowledge Publishing*, pp. 115–116, Jan. 2011.

[32] A. M. Latifi and J. Boisvert, "Stratigraphic coordinate transformation," *GeostatisticsLessons. Retrievedfromhttp://www. geostatisticslessons. com/lessons/stratcoords*, 2022.

[33] C. Deutsch and J. Deutsch, "Cell declustering parameter selection," *Geostatistics Lessons*, 2015.

[34] J Manchuk and C. V. Deutsch, "A short note on trend modeling using moving windows," *Centre for Computational Geostatistics, University of Alberta, Edmonton, Canada, CCG Paper (403)*, 2011.

[35] J. A. McLennan, "The decision of stationarity," 2007.

[36] J. Qu and C. V. Deutsch, "Geostatistical simulation with a trend using gaussian mixture models," *Natural Resources Research*, vol. 27, no. 3, pp. 347–363, 2018.

[37] H. Wackernagel, *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media, 2003.

[38] B. Harding and C. V. Deutsch, "Trend modeling and modeling with a trend," *GeostatisticsLessons. Retrievedfromhttp://www. geostatisticslessons. com/lessons/trendmodeling*, 2021.

[39] J. Vincent and C. V. Deutsch, "The multivariate spatial bootstrap," *Geostatistics Lessons*, 2019.

[40] C. V. Deutsch, A. G. Journel, *et al.*, "Geostatistical software library and user's guide," *Oxford University Press*, vol. 8, no. 91, pp. 0–1, 1992.

[41] J. J. Gómez-Hernández and A. G. Journel, "Joint sequential simulation of multigaussian fields," pp. 85–94, 1993.

[42] A. G. Journel and P. C. Kyriakidis, *Evaluation of mineral reserves: a simulation approach*. Oxford University Press, 2004.

[43] C. V. Deutsch, "A sequential indicator simulation program for categorical variables with point and block data: Blocksis," *Computers & Geosciences*, vol. 32, no. 10, pp. 1669–1681, 2006.

[44] R. M. Butler, "Steam-assisted gravity drainage: Concept, development, performance and future," *Journal of Canadian Petroleum Technology*, vol. 33, no. 02, pp. 44–50, 1994.

[45] P. E. Black *et al.*, "Dads: The on-line dictionary of algorithms and data structures," *NIST: Gaithersburg, MD, USA*, 2020.