Multivariate Stochastic Seismic Inversion With Adaptive Sampling

by

Maryam Hadavand-Siri

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

Doctor of Philosophy

in

Mining Engineering

Department of Civil and Environmental Engineering
University of Alberta

Abstract

Reservoir management tools use geostatistical models to make better decisions and improve hydrocarbon recovery. Developing improved numerical modeling techniques that lead to more accurate and precise geostatistical models will improve flow forecasting, hydrocarbon recovery, reservoir management and contribute to more responsible and sustainable management of our natural resources. The conventional stochastic seismic inversion techniques integrates different source of data to obtain high quality geostatistical models. These techniques initially provide acoustic impedance models that match well and seismic data. Such acoustic impedance models are related to reservoir physical properties via rock physics models. These techniques aim to reproduce the data within the quality of data. However, there is no guarantee that the final petrophysical models will reproduce original seismic data. Fidelity with the original seismic data will be reduced due to an element of randomness at each step. To overcome this issue a new approach is proposed and developed.

The new approach proposes to simulate multiple reservoir physical properties simultaneously. This research study presents a research toward a fully coupled categorical - multivariate continuous reservoir modeling in stochastic inversion context with Petro-Elastic Model and convolution. The application of multivariate geostatistical techniques would improve conventional stochastic inversion approaches. The flexibility of using variogram based techniques and multi point statistic methods to model complex geological features further improves the stochastic inversion approach.

The multivariate stochastic inversion approach combines a trace by trace (column wise) adaptive sampling algorithm with multivariate geostatistical techniques to pick the best physical properties of reservoir that match the actual seismic data. The adaptive sampling method uses an acceptance-rejection approach to condition geostatistical models to well and seismic data. This technique samples the realizations inside the space of uncertainty.

The number of realizations attempted is changed based on the size of space of uncertainty.

In this study, a general framework is presented to calculate the size of the space of uncertainty. This becomes practically relevant when rejection sampling approaches are being used to condition geostatistical models as in the case of stochastic inversion. The size of the space of uncertainty is shown to be the product of exponential entropy values. This is corroborated from information theory, but the application of this in presence of spatial correlation and conditioning data is new.

Modeling multiple reservoir properties simultaneously through the close integration of seismic inversion and multivariate geostatistical techniques leads to high resolution reservoir property models that are suitable for improved reservoir management. A case study with realistic data set is developed to compare the results of multivariate stochastic inversion approach with conventional stochastic method.

To love of my life Reza

Acknowledgements

First and foremost I would like to express my sincere gratitude to my advisor Prof. Clayton Deutsch for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor for my Ph.D study. I admire his integrity, ingenuity, proficiency and diligence as a researcher and as a person. In addition, I thank the Center for Computation Geostatistics and sponsors from industry for the financial support of my research.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Hooman Askari-Nasab, Dr. Larry Bentley, Dr. Jeff Boisvert, Dr. Juliana Leung and Dr. Mauricio Sacchi for their insightful comments and encouragement.

My sincere thanks also goes to the department of civil and environmental engineering staff, Alice Da Silva, Arlene Figley and Lorraine Gran for their support and contribution during my time at university of Alberta.

I am thankful to all my colleagues at Center for Computational Geostatistics. Special thanks to my friends Vahid Dehdari, Filipe Pinto, Mehdi Rezvendehy, Samaneh Sadeghi and Diogo Silva, who have been a source of real friendship and support throughout my PhD time.

Last but not the least, I thank my parents for their endless support, my deepest appreciation to my lovely sweet son and daughter, Soren and Sarina for their great patience and understanding and most of all to Reza, thank you. This thesis is dedicated to you.

Table of Contents

List of Tables

List of Figures

List of Symbols

Symbols	Definition
AI	Acoustic Impedance
ρ	Density
v	Velocity
$R(t)$	Normal incident reflection coefficient
$S(t)$	Seismic trace
$W(t)$	Wavelet
λ	Wave length
$\mathbf{u}$	Location
$Z^*(\mathbf{u})$	Kriging estimation
$Z(\mathbf{u})$	Data value
$m(\mathbf{u})$	Mean of data
$\lambda(\mathbf{u})$	Kriging weight
σ_E^2	Error-variance
$E\{Z(\mathbf{u})\}$	Expected value of Z(**u**)
$C(\mathbf{u})$	Covariance
$F(\mathbf{u}; z)$	Cumulative distribution function
$F(\mathbf{u}; z\|(n))$	Conditional cumulative distribution function
$z(\mathbf{u})$	Unsampled value
$Y(\mathbf{u})$	Random Gaussian function
$y(\mathbf{u})$	Simulated Gaussian value
G^{-1}	Gaussian backtransform
A	Area of interest
m	Model
$E(m)$	Model energy
P, p	Probability

T	Annealing cooling temperature
Q_k	Mismatch in annealing optimization
L	Number of realization
V_p	Compressional wave velocity
V_s	Shear wave velocity
$\vec{\mathbf{u}}$	Location of column
$Z_i^l(\vec{\mathbf{u}})$	Multiple Gaussian variables
n_f	Number of facies variable
n_c	Number of continuous variable
K	Facies category
ρ_f	Fluid density
ρ_r	Rock density
ρ_o	Oil density
ρ_w	Water density
ϕ	Porosity
S_w	Water saturation
S_o	Oil saturation
a	Gardner's parameter
m	Gardner's parameter
Λ	Space of uncertainty
$M(\Lambda)$	Size of space of uncertainty
H	Entropy
$\gamma(h)$	Variogram
h	Distance
$I(z(\mathbf{u}); k)$	Indicator variable
t	Threshold
D	Simulation domain
D^d	Cascading grid over simulation domain
n_d	Nested grid
MSE	Mean square error

List of Abbreviations

Abbrv.	Definition
IEA	International Energy Agency
P-wave	Compressional wave
S-Wave	Shear wave
SGS	Sequential Gaussian Simulation
SGCS	Sequential Gaussian Co-Simulation
AI	Acoustic Impedance
CMP	Common Midpoint
MPS	Multi Point Statistics
PEM	Petro Elastic Model
MSE	Mean Square Error
HZ	Hertz
QSI	Quantitative Seismic Interpretation
AVO	Amplitude Versus Offset
LHS	Left Hand Side
RHS	Right Hand Side
RV	Random Variable
RF	Random Function
Prob	Probability
cdf	Cumulative Distribution Function
ccdf	Conditional Cumulative Distribution Function
PCA	Principle Component Analysis
DSS	Direct Sequential Simulation
USGSIM	Ultimate Sequential Gaussian Simulation
GSLIB	Geostatistical Software Library
McMC	Markove chain Monte Carlo

SIS	Sequential Indicator Simulation
TG	Truncated Gaussian
TGS	Truncated Gaussian Simulation
TPGS	Truncated PluriGaussian Simulation
SNESIM	Single Normal Extended Simulation
VPC	Vertical proportion Curve
CCK	Collocated Cokriging
Fluvsim	Fluvial Simulation

Chapter 1

Introduction

Natural resources such as hydrocarbon reservoirs are an important necessity in the development of modern human society. Over the past century, society has developed a dependence on hydrocarbon reservoirs because they are one of the primary requirements to develop a safe, secure and stable society. The resources in hydrocarbon reservoirs are declining and human population is increasing. Unfortunately, these fundamental resources are not renewable and there is still no reliable replacement for them. According to International Energy Agency (IEA) reports and calculations, the global oil demands was around 75 million barrels per day in the early 2000's which this demand will actually double by 2030 (Birol, 2010; Hamilton, 2015). Therefore, increasing recovery and responsible management of hydrocarbon reservoirs is important.

Geostatistics has been developed over the last forty years as a multidisciplinary field that investigates the spatial distribution of the geological properties to predict the spatial properties of natural resources and quantify the associated uncertainty in the predictions (Journel and Huijbregts, 1978; Chiles and Delfiner, 2009). Two broad classes of geostatistical modeling techniques including estimation and simulation techniques are recognized. Estimation methods provide a single estimate for each location and leads to overly smooth property models. Simulation techniques generate a set of stochastic models of properties that are equiprobable and called realizations. The set of equiprobable realizations represents property models with a reasonable pattern of variation and measure of uncertainty due to data limitation (Deutsch and Journel, 1998).

A high quality geostatistical model consists of a detailed numerical model of reservoir properties including facies, porosity, fluid saturations and permeability. Developing improved numerical modeling techniques leads to more accurate and precise geostatistical

models. Reservoir engineering and management tools such as flow simulation use geostatistical models to make better decisions and optimize hydrocarbon recovery. Therefore, these developed techniques will improve flow forecasting, hydrocarbon recovery, reservoir management and contribute to more responsible and sustainable management of our diminishing natural resources.

1.1 Background

Early techniques for the characterization of spatial variability and heterogeneity of reservoir properties range from deterministic interpolation algorithms (Ripley, 1981; Cressie, 2015) to geostatistical approaches such as Kriging and conditional simulation (Journel and Huijbregts, 1978; Chiles and Delfiner, 2009). These techniques are primarily based on local measurements from wells. Kriging and conditional simulation techniques consider and apply the spatial variability of the well data (Thadani et al., 1987; Pyrcz and Deutsch, 2014). Kriging method provides a single estimate that is overly smooth and does not provide a complete understanding of the spatial variability of reservoir properties. On the other hand, conditional simulation techniques produce a set of equiprobable realizations that reproduce the histogram and spatial covariance of well data (Thadani et al., 1987; Chiles and Delfiner, 2009). Due to the high cost of drilling, few wells are drilled for exploration and well data do not provide a complete understanding of the lateral distribution of reservoir properties. The integration of additional sources of information such as geophysical and production data is required to improve the spatial characteristics of geostatistical models.

Geophysical methods include seismic, gravity, magnetic and geo-electric measurements. Seismic has made a major contribution in hydrocarbon exploration because it responds to multiscale subsurface features (Lines and Newrick, 2004). In the content of this thesis only 2D or 3D surface seismic surveys are considered.

For surface seismic surveys, the acoustic pulses propagate through the subsurface layers with different velocities because of different rock densities, fluid content, shear and bulk

modulus. The reflected acoustic waves from different interfaces are recorded by seismic receivers. A seismic source on the surface generates different waves including compressional (P-wave) and shear (S-wave) waves. These waves expand, attenuate, reflect, refract and change due to geological variability. Seismic data processing methods try to understand this complexity and reveal the underlying geological structures. The outputs of seismic data processing include different seismic attributes that are derived from the reflectivity sequences. To integrate seismic data into geostatistical modeling, the seismic attributes are commonly inverted to acoustic properties like acoustic impedance.

Generally, seismic data have a high lateral coverage and resolution compared to well data while the vertical resolution of seismic data is poor compared to the well data. Therefore, the predictive ability of geostatistical models could be improved by integration of seismic and well data. There are different approaches for the integration of seismic data in geostatistical modeling. Two main approaches have been investigated (Haas and Dubrule, 1994; Hong et al., 2006; Bosch et al., 2010) : 1) Indirect methods, and 2) Direct methods.

Indirect methods consider the large-scale seismic data as secondary variables and combine them with primary small-scale well data through different geostatistical algorithms such as cokriging (Doyen, 1988; Xu et al., 1992; Yao and Chopra, 2000; Grana et al., 2012), indicator coding of seismic data (Thadani et al., 1987; Zhu, 1991; Zhu and Journel, 1993; Moysey et al., 2003; Park et al., 2003), Kriging with external drift (Marechal, 1984; Galli and Meunier, 1987; Bourennane et al., 2000; Bourennane and King, 2003) and block Kriging (Behrens et al., 1998; Lee et al., 2000; Aanonsen and Eydinov, 2006). Most of these indirect integration methods consider a combination of seismic and well data through a probabilistic calibration. These calibrations are critical and difficult to establish (Bortoli, 1992; Bosch et al., 2010).

Direct methods are applied for situations where the data can be considered as a deterministic function of the numerical geological model (Hong et al., 2006). For example, the synthetic seismogram can be obtained from an acoustic impedance model by wavelet convolution. Direct methods combine conventional seismic inversion techniques

with geostatistical approaches to provide high resolution numerical geological models. These direct methods are also referred to as stochastic inversion. The original idea of the direct method was proposed by Bortoli (1992), Haas and Dubrule (1994) in the early 1990's. The original methodology combined a trace-by-trace optimization method with Sequential Gaussian Simulation (SGS) algorithm (Deutsch and Journel, 1998) to provide high resolution geostatistical models.

The original stochastic inversion approach simulates acoustic related properties that are conditioned to the well data. For each realization, a synthetic seismogram is computed and compared with the actual seismic trace. The synthetic seismogram that best matches to the actual seismic data is kept. Under some specific circumstances, the acoustic related variable can be transformed to the units of seismic data such as seismic amplitude (Bortoli, 1992). Based on the reservoir physical properties that are desired to obtain, the related acoustic parameter is chosen to simulated. For example, to obtain a porosity model, the acoustic impedance could be simulated then transformed to porosity. Figure 1.1 shows a flow chart of this methodology.

Compared to the indirect methods of seismic data integration, the original stochastic inversion method does not require a prior probabilistic calibration step like Markove-Bayes calibration (Journel and Zhu, 1990, Zhu and Journel, 1993). Nevertheless, the results are sensitive to the relation between acoustic and physical reservoir properties used to derive the final petrophysical reservoir model (Bortoli et al., 1993).

The Bortoli (1992), Haas and Dubrule (1994) method was modified by Debeye et al. (1996) at Jason Geosystems. This algorithm generates multiple 3D facies realizations consistent with geology and seismic data, then relates the facies to the reservoir physical properties such as porosity. This method is an adaptation of sequential Gaussian simulation (SGS) and sequential Gaussian co-simulation (SGCS) (Deutsch and Journel, 1998) based on simulated annealing (Debeye et al., 1996).

In sequential Gaussian simulation, grid nodes are visited sequentially through a random path and local conditional probability distributions are calculated under a multivariate

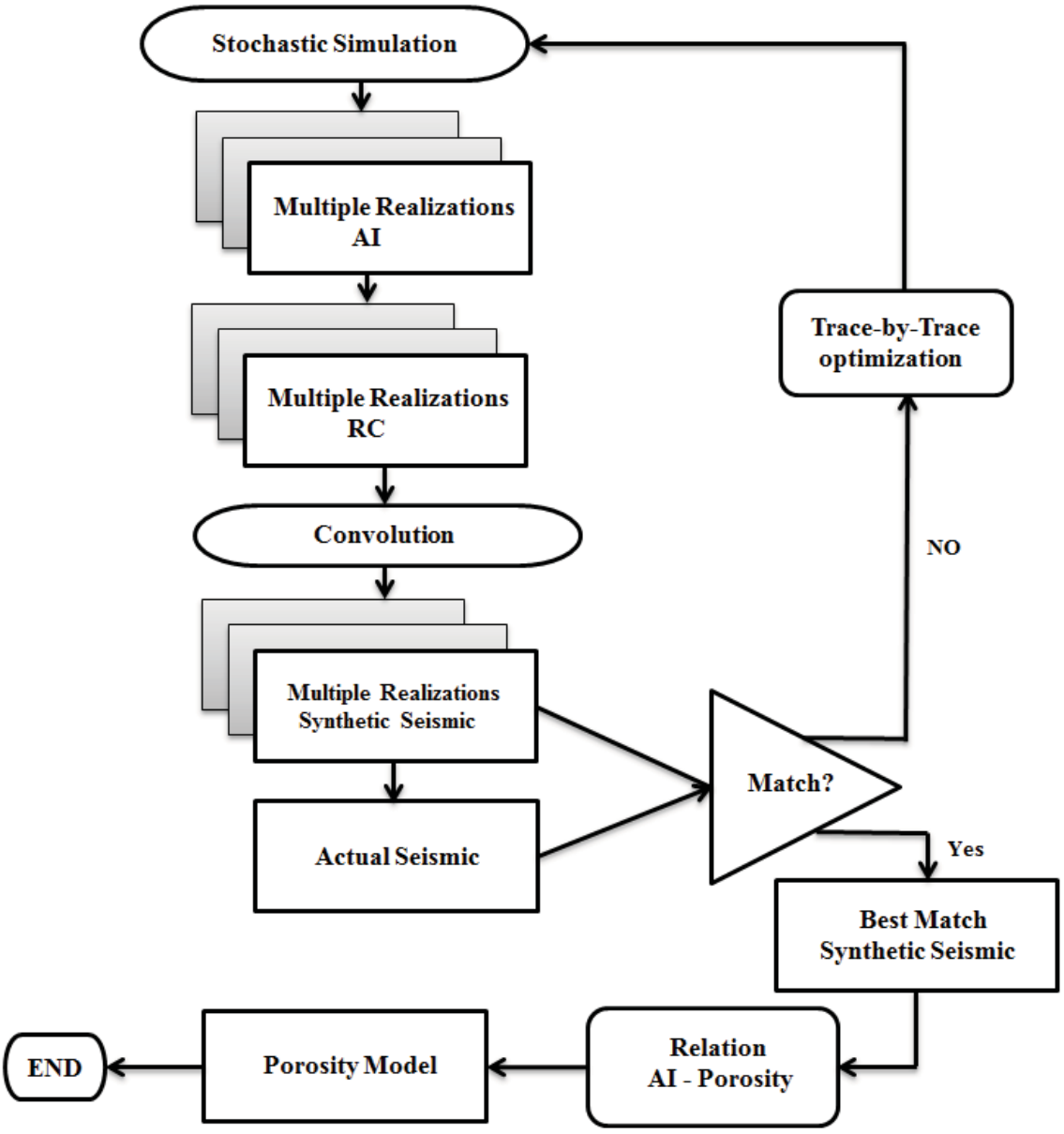

Figure 1.1: Flow chart of original stochastic inversion methodology (Redrawn from Bortoli, 1992).

Gaussian model. Then, a value from the conditional probability distribution is drawn and added to the conditioning data. This procedure is repeated until all the grid nodes are simulated (Deutsch and Journel, 1998). The data are transformed to a Gaussian distribution before the simulation. Sequential Gaussian co-simulation (SGCS) is similar to SGS which multiple reservoir properties that are correlated to each other are simulated simultaneously (Pyrcz and Deutsch, 2014). Simulated annealing is a global optimization method that considers stochastic relaxation and an acceptance rule based on an analogy with annealing. The objective function is the measure of difference between the target spatial characteristics and those of the realization. The optimization algorithm amounts to perturb the specific realization and then accept/reject the perturbation based on how close the realization is to the required properties (Deutsch, 1992)

Other stochastic inversion approaches have been proposed, developed and commercialized in recent years. In the majority of published stochastic inversion methods, around 90% (Bosch et al., 2010), high resolution acoustic impedance models are simulated then they are connected to the physical properties of the reservoir by using a direct relationship or statistical correlations. In some stochastic inversion methods, facies models associated to the acoustic impedance models are simulated directly, then these facies models are connected to the other physical properties such as porosity and fluid saturations.

All stochastic inversion techniques aim to generate high resolution geostatistical models (acoustic impedance or facies) that reproduce the well and seismic data. Then, these high resolution models are related to the reservoir physical properties via rock physics and statistical links, Figure 1.2. There is no guarantee that the physical properties obtained through this sequential linking will match the original seismic data. Fidelity with the original seismic data is likely to be lost due to randomness at each step. This research study proposes and developes a new approach to stochastic inversion to permit generating high resolution geostatistical models with high fidelity with the original seismic traces.

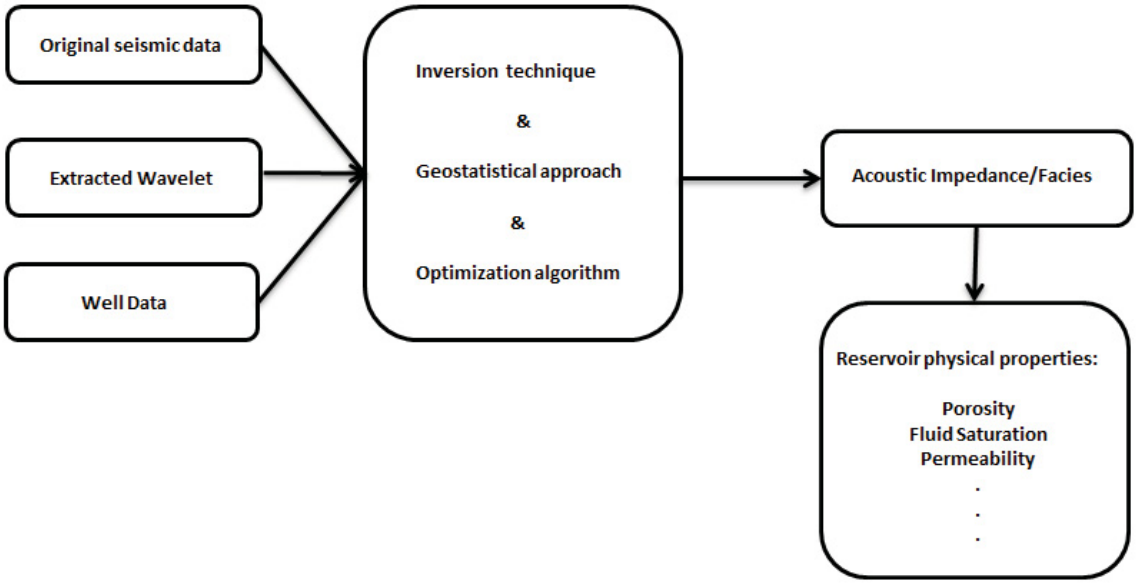

Figure 1.2: Schematic illustration of the sequential connection of properties in stochastic inversion algorithms.

1.2 Problem Solution and Significance of Thesis

Conventional stochastic inversion attempts to reduce the uncertainty to the greatest extent possible while providing fair uncertainty statements. They also aim to reproduce the data within the quality of data. There is no guarantee that the final petrophysical models will reproduce the original seismic data. To address this issue a new approach called multivariate stochastic inversion is proposed and developed. This approach simulates multiple reservoir physical properties simultaneously as a part of stochastic inversion algorithm instead of relating acoustic or facies models to the reservoir physical properties through a set of sequential connections outside the stochastic inversion algorithm. The main purpose of this idea is to simulate multiple reservoir properties by a geostatistical technique and condition them to the seismic data at the same time through a stochastic inversion algorithm. The new method combines a trace by trace (column-wise) multivariate geostatistical technique with adaptive sampling algorithm to overcome the limitation of conventional stochastic inversion methods. The adaptive sampling algorithm considers a different number of realizations through different steps of simulation in this

approach.

In seismic data acquisition, the seismic waves are mechanical perturbations generated by a seismic source. They travel into the earth at a velocity that is governed by the density, fluid content and elastic modulus of the rock they are traveling through. The acoustic impedance of the rock is the product of density and seismic wave velocity of the rock ($AI = \rho v$). When the seismic wave encounters an interface between two geological units with different acoustic impedance, the acoustic impedance contrast causes some of seismic wave energy to reflect and some to refract through the interface (Lines and Newrick, 2004). Figure 1.3 schematically shows the reflected and refracted seismic wave at an interface due to acoustic impedance contrast. The reflected and refracted wave energy are detected, measured and recorded by an array of seismic receivers (geophones) at the surface. The reflected and refracted wave energy from different shot records with a common reflection point such as common midpoint (CMP) at different depth represent a seismic trace, Figure 1.4 . The point on the surface halfway between the source and receiver that is shared by numerous source-receiver pairs is called the common midpoint (CMP). The seismic survey consists of the seismic traces that have been stacked together from different records to reduce noise and improve overall quality of the seismic data (Telford et al., 1990). Figure 1.5 also shows schematically the seismic traces at the different geological interfaces.

A seismic trace represents geological information at different depth in vertical columns. Seismic surveys are inherently column based. As a result, to apply geostatistical modeling as part of stochastic inversion and condition the multiple reservoir properties to the seismic data at the same time in multivariate stochastic inversion approach a full column of multiple reservoir properties should be simulated at each simulation iteration. Therefore, all simulation techniques and algorithms in multivariate stochastic inversion approach must be developed and implemented in column wise manner.

The proposed technique simulates facies and reservoir physical properties at the same time by simulating a set of Gaussian variables. Facies are assigned by truncated (pluri)Gaussian simulation technique. The number of Gaussian variables required for

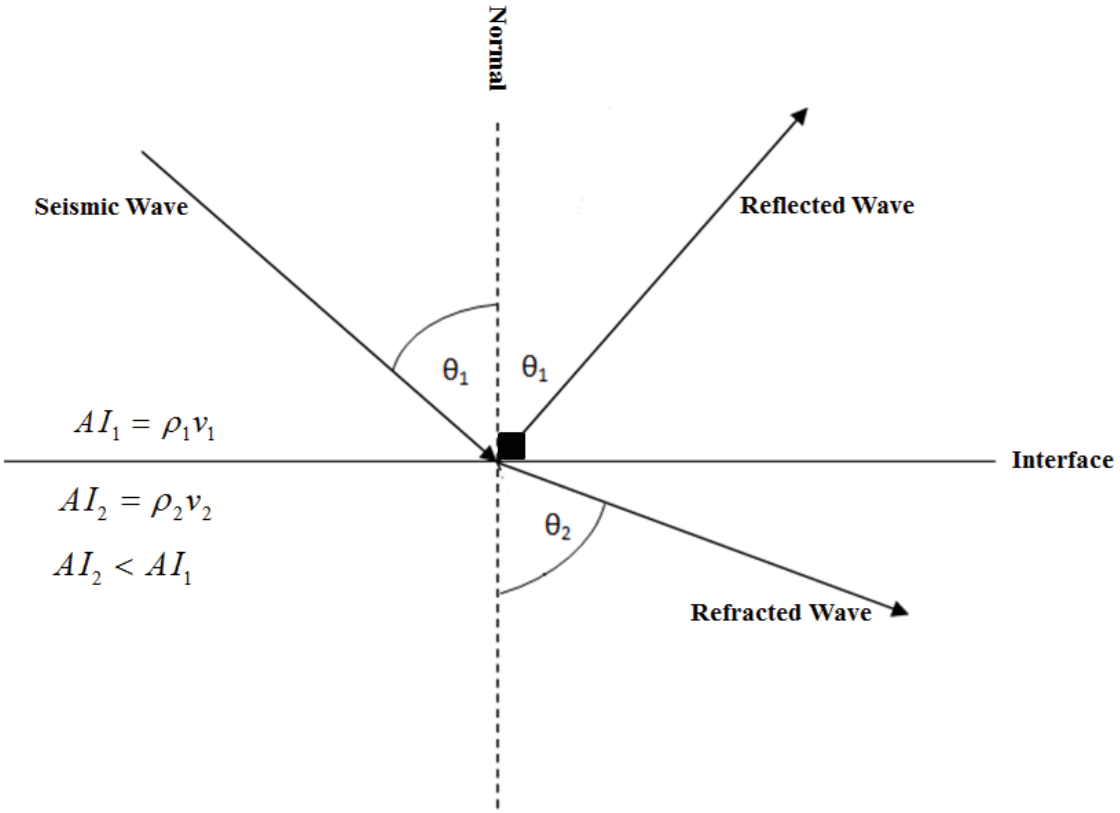

Figure 1.3: Sketch shows the reflected and refracted seismic wave at an interface because of acoustic impedance contrast.

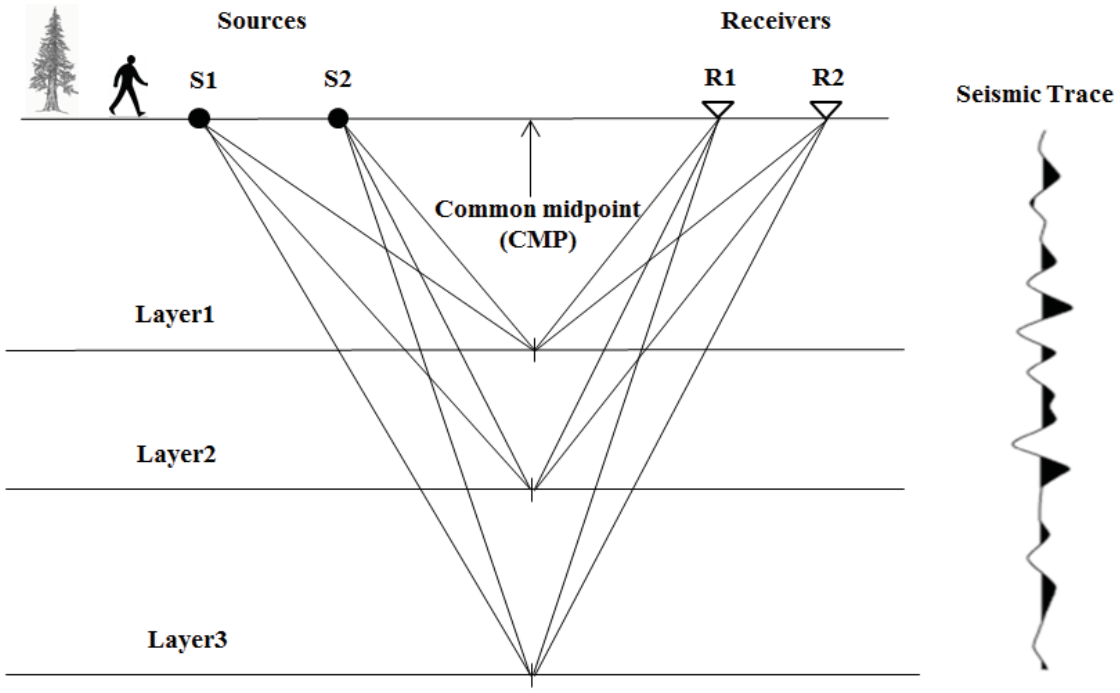

Figure 1.4: An array of two seismic sources/receivers and reflected seismic wave from different layers beside a schematic illustration of a seismic trace with common midpoint (CMP) at different depth.

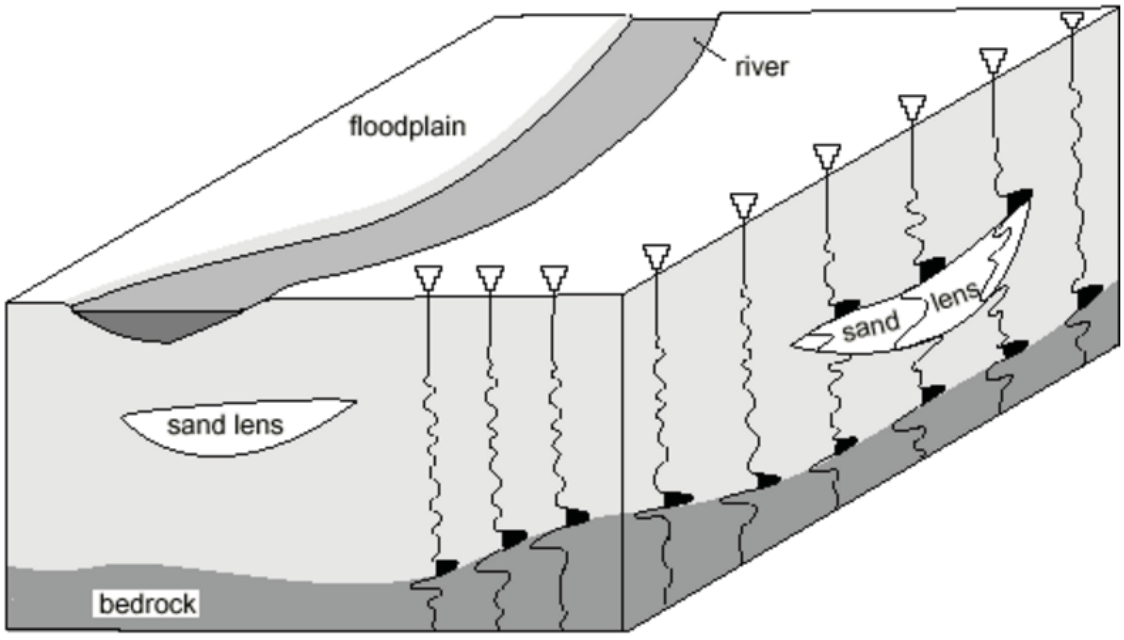

Figure 1.5: Shematic showing of seismic traces at different geological interfaces (WWW.kgs.ku.edu).

facies modeling depends on ordering and transition among the facies categories. To add more flexibility and to handle more complex geological features in reservoir modeling, column wise multi point statistics (MPS) simulation technique is an alternative for facies modeling. Reservoir physical properties (continuous variables) are simulated through the column wise sequential Gaussian simulation (SGS) technique. The number of Gaussian variables required to be simulated by column wise SGS are defined by the number of reservoir physical properties and the number of facies categories.

Each column of multiple reservoir properties are passed to a Petro-Elastic Model (PEM) and a convolution algorithm to derive a synthetic seismogram. The synthetic seismogram is compared with collocated actual seismic trace through the adaptive sampling algorithm. This algorithm selects the acceptable match to the original seismic data based on a stopping criteria. The stopping criteria consists of a target Mean Square Error (MSE) and a reasonable number of realizations that is adapted as the simulation proceeds.

The size of the space of uncertainty is quantified in this research study. The size of the space of uncertainty in presence of different factors is a good indicator of the reasonable number of realizations for each column in multivariate stochastic inversion. The multivariate stochastic inversion approach is explained in detail in Chapter 3

1.3 Delimitation and Limitations

The integration of surface seismic geophysical data in high resolution geostatistical models is studied. Seismic data are originally in the time domain and well conditioning data are in depth. Although time to depth conversion is an important step, it is not in the scope of this project. To obtain high resolution geostatistical models that honor seismic data, the actual seismic traces should be reproduced at each column in the geostatistical model. For this purpose the synthetic seismic traces are computed based on a convolution of reflectivity series calculated through a Petro-Elastic Model (PEM) and an extracted wavelet.

Wavelet extraction is an important step where the synthetic trace along the well is calibrated

to the actual collocated seismic trace. This calibration process is necessary to ensure that the actual seismic data matches the convolution of the extracted wavelet with the reflectivity series along the wells (Bortoli, 1992). A special type of wavelet often used for modeling and inversion purposes is the Ricker wavelet which is defined by dominant frequency. The Ricker wavelet is by definition zero-phase and symmetric. The convolution process can affect not only the amplitude of a seismic traces but its phase as well. The Ricker wavelet is known as a zero phase function which means it does not have a phase effect. Therefore, zero phase Ricker wavelet is a good option for extracted wavelet since it can be easily understood and implemented. This wavelet often reasonably represents the typical earth responses (Yilmaz, 2001, Margrave, 2005). Any other extracted wavelet can be considered and applied in convolution algorithm.

Access to real seismic data would make the study more credible to some practitioners. The case study uses realistic data and compares the results to conventional approach; however, the data was not collected from a real reservoir.

1.4 Thesis Statement

Characterizing multiple reservoir properties simultaneously through the close integration of seismic inversion and multivariate geostatistical techniques leads to improved high resolution reservoir property models. Implementation of column wise multivariate Gaussian simulation technique and quantification of the size of the space of uncertainty adds practical value and theoretical insight.

1.5 Thesis Outline

This thesis consists of eight chapters. Chapter 1 is the introduction. Chapter 2 reviews the relevant literature for integration of seismic data into geostatistical modeling techniques and mainly focuses on conventional stochastic inversion approaches. Chapter 3 provides information and details on the proposed methodology of multivariate stochastic inversion.

Chapter 4 focuses on the quantification of the size of the space of uncertainty and calculation of the space of uncertainty in presence of different parameters. Chapter 5 focuses on column based facies modeling and addresses the practical challenges for implementation of different methods of column based facies modeling. Chapter 6 covers the practical aspects and implementation of the adaptive sampling algorithm. Chapter 7 demonstrates the practical implementation of multivariate stochastic inversion approach to a case study with realistic seismic data. Chapter 8 wraps up the thesis with conclusions and future works. The description of programs written for this research study is provided in an Appendix.

Chapter 2

Literature Review

Reservoir characterization is a collaborative field that relies on different sources of geological and engineering information to describe petroleum reservoirs and the nature of the rocks that contain the hydrocarbons (Lines and Newrick, 2004). Defining the spatial variability of reservoir properties and heterogeneity of reservoir structures are the key problems in reservoir characterization because volume calculation, flow simulation and hydrocarbon recovery are directly affected by these parameters. A high resolution geostatistical model is a gridded numerical model of reservoir properties that provides an assessment of the spatial variability of reservoir properties.

Early techniques to characterize the spatial variability of reservoir properties were primarily based on well log information at few locations. These techniques range from deterministic interpolation algorithms (Ripley, 1981; Cressie, 2015) to geostatistical approaches such as Kriging and conditional simulation (Journel and Huijbregts, 1978; Chiles and Delfiner, 2009; Pyrcz and Deutsch, 2014). Unfortunately, deterministic techniques do not create numerical models that reproduce the spatial variability. Kriging and conditional simulation consider and incorporate the spatial variability inferred from well data. Kriging provides a single, overly smooth model that does not reproduce the spatial variability of the well data. Conditional simulation techniques generate equiprobable realizations that reproduce the histogram and spatial covariance of the well data (Thadani et al., 1987; Chiles and Delfiner, 2009).

The sparsity of well data is a major issue. Well data provide high vertical resolution but do not provide much information related to the lateral continuity of reservoir properties. Therefore, geostatistical techniques that rely only on the well data may provide highly uncertain input to reservoir engineering tools like flow simulation. The integration of

different sources of geological information such as outcrop observation, seismic and production data is important.

2.1 Seismic Data

Among the different geophysical methods, seismic techniques have the greatest role in reservoir characterization because seismic data are capable of detecting small to large scale stratigraphic features.

In reflection seismology, the principles of acoustic theory and wave propagation are used to estimate the subsurface properties. A seismic source sends controlled energy into the earth and reflected seismic waves are recorded by seismic receivers. Compressional waves (P -wave) and shear waves (S-wave) are generally recorded. Compressional and shear wave also called body waves because they propagate in all directions from the seismic source and travel through the earth. Body waves have significant contribution in hydrocarbon exploration (Bjørlykke, 2010). Compressional waves where the particle motion is along to the direction of wave propagation are faster than shear waves where the particle motion is perpendicular to the direction of wave propagation. The propagation of P and S waves through the subsurfaces depends on the acoustic impedance (AI) of the rocks. Acoustic impedance is the product of rock density and wave velocity ($AI = \rho v$). When a seismic wave propagates through the subsurface a portion of the seismic wave energy reflects. The amount of energy that reflects at each interface depends on the contrast of acoustic impedance between the two layers. The acoustic impedance contrast determines the seismic reflection coefficient. The quantities R is known as the normal incidence reflection coefficient, see Equation 2.1. Figure 2.1 shows the propagation of normal incident waves between two subsurfaces. At any interface, the normal incident reflection coefficient (R) can be positive or negative depending on contrast of acoustic impedance across the interface. If the acoustic impedance increases across the two subsurfaces R is positive and if acoustic impedance decreases R is negative. A positive R means that the polarity of the reflected wave will be the same as the incident wave. A

negative R means that the polarity of the reflected wave will be the opposite of the incident wave.

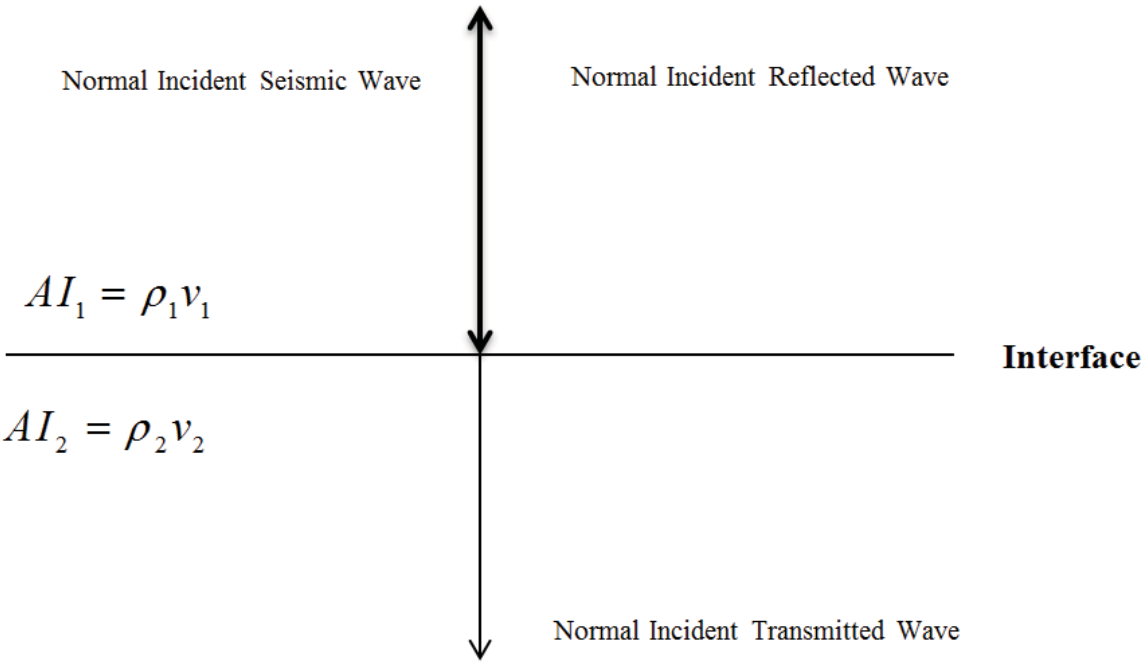

Figure 2.1: Propagation of normal incident waves upon an interface between two subsurfaces.

$$R = \frac{AI_2 - AI_1}{AI_2 + AI_1} = \frac{\rho_2 v_2 - \rho_1 v_1}{\rho_2 v_2 + \rho_1 v_1} \tag{2.1}$$

In a seismic study, there is also the travel time for a wave to arrive at the receiver from an interface. Seismic surveys measure ground motions. The seismic response could be represented as seismic traces that could be interpreted as the convolution of a seismic wavelet with the normal incident reflection coefficient. Equation 2.2 shows the seismic convolution in time domain.

$$S(t) = W(t) * R(t) + N(t) \tag{2.2}$$

Where $S(t)$, $W(t)$, $R(t)$ and $N(t)$ are seismic trace, wavelet, the reflection coefficient of normal incident and noise respectively in time domain. $S(t)$ given by Equation 2.2 is one way to compute synthetic data in exploration geophysics in time domain (Margrave, 2005). The convolution process in depth domain is similar to the convolution in time domain and represents by Equation 2.3.

$$S(d) = W(d) * R(d) + N(d) \tag{2.3}$$

Where $S(d)$, $W(d)$, $R(d)$ and $N(d)$ are seismic trace, wavelet, the reflection coefficient of normal incident and noise respectively in depth domain (Zhang et al., 2016). In case of dealing with depth-migrated seismic data and having reflectivity series in depth domain, convolution in depth domain should be applied to compute synthetic seismic data.

Signal processing methods try to remove noise and obtain a reflectivity sequence that reproduces the measured trace. To integrate seismic data into geostatistical modeling, the seismic data are often processed to compute migration, normal moveout (NMO), reflection coefficients and acoustic impedance. This processing is an inverse problem since we have the result of convolution and not the stacked or migrated sections.

The ability to distinguish different geological features from seismic data depends on the seismic resolution. The vertical resolution of surface seismic surveys is between $\lambda/8$ to $\lambda/4$ (Widess, 1973 and Kallweit and Wood, 1982) where λ is the seismic wavelength. For shallow events with nominal rock velocity of $2000m/s$ and dominant frequency of $50HZ$ the seismic wavelength is $40m$. The seismic wavelength for a deep event with a velocity of $3000m/s$ and frequency of $20HZ$ is $150m$.

Well data have high vertical resolution with limited horizontal coverage. Seismic data with high lateral coverage can be used to map out lateral changes in reservoir properties. Consequently, the integration of different source of geological information such as well and seismic data improved reservoir models.

There are different approaches and techniques to integrate seismic information in reservoir characterizations that are reviewed and explained in the following sections.

2.2 Quantitative Seismic Interpretation (QSI)

Conventional seismic interpretation picks and tracks seismic reflectors that are laterally consistent to map geological structures and the stratigraphy of reservoirs. The fundamental objective of seismic interpretation is to detect hydrocarbon reservoirs, ascertain their extent and calculate hydrocarbon volumes. Generally, conventional seismic interpretation has been fundamentally qualitative that requires professional skill and experience in geology and geophysics (Telford et al., 1990). Seismic interpretation aims to represent and map the geometry of seismic reflectors rather than provide a physical understanding of seismic amplitude variations (Blangy, 1992). Quantitative techniques for seismic interpretation aim to provide additional information for flow forecasting and reservoir characterization (Avseth et al., 2010).

Quantitative seismic interpretation (QSI) represents how rock physics and elastic properties can be applied to model reservoir properties such as facies, porosity, fluid saturation from seismic attributes (Mavko et al., 2009). Integration of quantitative seismic interpretation (QSI) in the early stages of seismic interpretation is an effective methods to reduce the uncertainty in reservoir models. The common and most important quantitative seismic interpretation techniques include: post stack amplitude analysis (bright-dim spot analysis), offset-dependent amplitude analysis (AVO), elastic and acoustic impedance inversion, and forward seismic modeling (Avseth et al., 2010). Seismic attributes such as seismic amplitude represent the contrast in elastic properties between geological layers and contain important information about reservoir physical properties that cannot be obtained from conventional seismic interpretation (Avseth et al., 2010).

2.3 Seismic Data Integration

Creating precise and accurate models of reservoir properties like porosity is an important objective in geostatistical modeling. Consideration of multiple data sources improves the reservoir models. Beyond the traditional application of seismic data for mapping large scale subsurface structures, the integration of seismic data could increase the local accuracy of the geostatistical models. The approaches for the integration of seismic data into geostatistical modeling are considered indirect methods or direct methods.

2.3.1 Indirect Methods

The most common approaches to integrate seismic data into geostatistical modeling are indirect. Indirect methods combine primary small scale well data with secondary large scale seismic data through algorithms such as cokriging and statistical calibration.

Kriging is one of the first methods of spatial estimation that was initially applied for mineral deposits to estimate grades at unsampled locations. Kriging is a linear least-square estimation technique that accounts for the spatial correlation between all data and the location to be estimated. Kriging calculates an estimate at an unsampled location $\mathbf{u}$, $Z^*(\mathbf{u})$, by assigning weights, λ_i, to the set of nearby data (n) (Goldberger, 1962; Matheron, 1962; Pyrcz and Deutsch, 2014), Equation 2.4.

$$Z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{i=1}^{n} \lambda_i(\mathbf{u})[Z(\mathbf{u}_i) - m(\mathbf{u}_i)] \tag{2.4}$$

where $m(\mathbf{u})$ is the mean value at the unsampled location $\mathbf{u}$. In Kriging, the weights minimize the error-variance σ_E^2, which is the expected squared error between the true and estimated value, Equation 2.5.

$$\sigma_E^2 = E[(Z(\mathbf{u}) - Z^*(\mathbf{u}))^2] \tag{2.5}$$

The matrix form of kriging equation system is in the following form, Equation 2.6:

$$\begin{pmatrix} C_{11}(\mathbf{u}) & \ddots & C_{1n}(\mathbf{u}) \\ \vdots & \ddots & \vdots \\ Cn1(\mathbf{u}) & \cdots & Cnn(\mathbf{u}) \end{pmatrix} \begin{pmatrix} \lambda_1(\mathbf{u}) \\ \vdots \\ \lambda_n(\mathbf{u}) \end{pmatrix} = \begin{pmatrix} C_{10}(\mathbf{u}) \\ \vdots \\ C_{n0}(\mathbf{u}) \end{pmatrix} \tag{2.6}$$

where the left hand side (LHS) matrix of C is the data-data covariance matrix and the right hand side matrix (RHS) of C is the covariance between data points and unsampled location, C_{10} to C_{n0}. The weights calculated by the Kriging system consider the spatial variability between the data and unsampled location. The structural function that characterizes the spatial variability among locations is called the variogram. In geostatistics, the variogram provides the linear correlation between two data points as a function of their separation in space. The variogram is used to calculate the covariance values (C) in Equation 2.6

This is a simple and efficient way to relate an unsampled location to nearby data values. Different variants of Kriging such as simple and ordinary Kriging were introduced with different constraints on the mean or transformations of the data (Journel and Huijbregts, 1978; Goovaerts, 1997; Deutsch and Journel, 1998; David, 2012; Cressie, 2015). All Kriging estimation techniques rely on the variogram measure of spatial variability (Pyrcz and Deutsch, 2014).

Multivariate geostatistics was developed as an extension to include several variables. Cokriging is a method of estimation that considers multiple data types. For example, porosity can be estimated by cokriging as a linear combination of nearby porosity values and related seismic attributes like acoustic impedance values. For cokriging, consider the primary well data $\{Z_1(\mathbf{u}_{\alpha_1}), \alpha_1 = 1, ..., n_1\}$ that are complemented by $n_v - 1$ secondary seismic attributes Z_i, $\{Z_i(\mathbf{u}_{\alpha_i}), \alpha_i = 1, ..., n_i, i = 2, ..., n_v\}$. In this situation, the kriging estimator can be extended to accommodate the secondary variables (Journel and Huijbregts, 1978; Carr et al., 1985; Pyrcz and Deutsch, 2014) , Equation 2.7.

$$Z_1^*(\mathbf{u}) - m_1(\mathbf{u}) = \sum_{i_1=1}^{n_1} \lambda_{i_1}(\mathbf{u}).[Z(\mathbf{u}_{i_1}) - m(\mathbf{u}_{i_1})] + \sum_{\alpha=2}^{n_v} \sum_{i_\alpha=1}^{n_\alpha} \lambda_{i_\alpha}(\mathbf{u}).[Z_\alpha(\mathbf{u}_{i_\alpha}) - m(\mathbf{u}_{i_\alpha})] \quad (2.7)$$

Where λ_{i_1} are the weights assigned to the primary data and λ_{i_α} are the weights assigned to the $n_v - 1$ secondary data. To calculate the weights in cokriging the following system of equation must be solved, Equation 2.8.

$$\begin{pmatrix} [C_{11}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1})] & \cdots & [C_{1n_v}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_{n_v}})] \\ \vdots & \ddots & \vdots \\ [C_{n_v1}(\mathbf{u}_{\alpha_{n_v}} - \mathbf{u}_{\beta_1})] & \cdots & [C_{n_vn_v}(\mathbf{u}_{\alpha_{n_v}} - \mathbf{u}_{\beta_{n_v}})] \end{pmatrix} \begin{pmatrix} [\lambda_{\beta_1}(\mathbf{u})] \\ \vdots \\ [\lambda_{\beta_{n_v}}(\mathbf{u})] \end{pmatrix} = \begin{pmatrix} [C_{11}(\mathbf{u}_{\alpha_1} - \mathbf{u})] \\ \vdots \\ [C_{n_v1}(\mathbf{u}_{\alpha_{n_v}} - \mathbf{u})] \end{pmatrix} \quad (2.8)$$

Where the LHS matrix $[C_{ij}(\mathbf{u}_{\alpha_i} - \mathbf{u}_{\beta_j})]$ is the matrix of data-data direct and cross covariances. The RHS matrix $[C_{i1}(\mathbf{u}_{\alpha_i} - \mathbf{u})]$ is the matrix of data-unsampled location direct and cross covariances. The vector of $[\lambda_{\beta_i}(\mathbf{u})]$ are cokriging weights.

In cokriging system of equation with n variables, the covariance matrix needs n^2 variogram/covariance functions which is demanding in terms of data and also the joint model for the covariance matrix can be difficult to model. Therefore, cokriging has not been widely practiced and used in geostatistics (Deutsch and Journel, 1998; Chiles and Delfiner, 2009).

Collocated cokriging is a modification of cokriging that estimates the primary variable at unsampled location from nearby primary data values and the collocated secondary data value at the unsampled location (Doyen, 1988; Xu et al., 1992; Chiles and Delfiner, 2009; Cressie, 2015). By considering a primary variables with n_1 samples and $n_v - 1$ secondary data where $j = 2, ..., n_v$, the collocated cokriging system of equation in matrix form is:

$$\begin{pmatrix} [C_{11}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1})] & [C_{1j}(\mathbf{u} - \mathbf{u}_{\beta_j})] \\ [C_{i1}(\mathbf{u}_{\alpha_i} - \mathbf{u})] & [C_{ij}(\mathbf{u} - \mathbf{u})] \end{pmatrix} \begin{pmatrix} [\lambda_{\beta_1}(\mathbf{u})] \\ [\lambda_{\beta_i}(\mathbf{u})] \end{pmatrix} = \begin{pmatrix} [C_{11}(\mathbf{u}_{\alpha_1} - \mathbf{u})] \\ [C_{i1}(\mathbf{u} - \mathbf{u})] \end{pmatrix} \tag{2.9}$$

where $[C_{11}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1})]$ is the matrix of primary data-data covariances, LHS of simple Kriging equation. $[C_{11}(\mathbf{u}_{\alpha_1} - \mathbf{u})]$ is the vector of primary data-unsampled location covariances, RHS of simple kriging. $[C_{1j}(\mathbf{u} - \mathbf{u}_{\beta_j})]$ and $[C_{i1}(\mathbf{u}_{\alpha_i} - \mathbf{u})]$ are the covariance matrix contains cross covariances between data and collocated secondary variables. $[C_{ij}(\mathbf{u} - \mathbf{u})]$ is the correlation coefficients between secondary variables and $[C_{i1}(\mathbf{u} - \mathbf{u})]$ is the correlation coefficients between primary and secondary variables.

Kriging with external drift (Marechal, 1984; Galli and Meunier, 1987; Bourennane et al., 2000; Bourennane and King, 2003), indicator coding of seismic data (Thadani et al., 1987; Zhu, 1991; Zhu and Journel, 1993; Moysey et al., 2003; Park et al., 2003) and block Kriging (Behrens et al., 1998; Lee et al., 2000; Aanonsen and Eydinov, 2006) are other geostatistical methods to integrate secondary data. These methods will not be pursued in this thesis because they have not proven flexible and consistent with the original seismic data. Kriging with external drift does not provide control on how much the seismic data influence on the final model. In indicator coding of seismic data, multiple thresholds are needed, calibration must be applied and a number of Markov-Bayes parameters are required.

2.3.2 Direct Methods

Direct methods are used in cases where the secondary data can be written as a deterministic function of the geological model (Hong et al., 2006). Direct methods generally integrate the seismic data through the combination of geostatistical approaches and an inversion technique. The procedure of transformation of any geophysical data into physical properties of earth model like elastic properties is considered as seismic inversion process, Figure 2.2.

The principles and mathematics of seismic inversion theory can be found in many

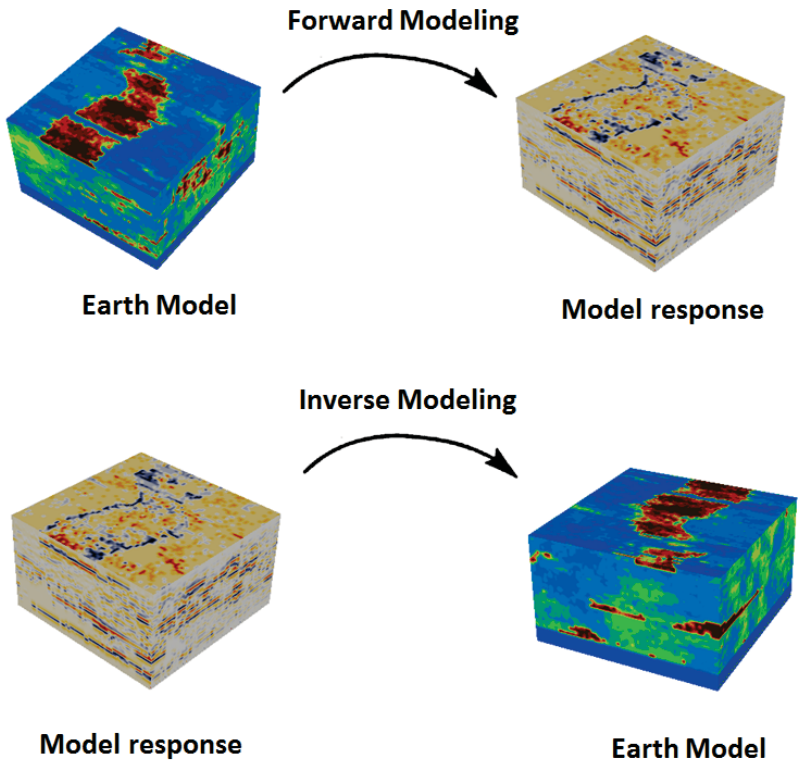

Figure 2.2: Schematic illustration of forward and inverse modeling in seismic data.

Geophysics text books such as Parker (1994) and Oliver et al. (2008). The main objective of applying seismic inversion is to derive physical properties that characterize the rock and fluid properties of the reservoir. In general, inversion does not provide a unique solution due to band limited and noise contaminated seismic data (Bortoli, 1992). Seismic inversion methods are categorized into two major approaches: 1) deterministic, and 2) stochastic.

Deterministic inversion is also known as stratigraphic inversion. Sparse Spike and Model Based deterministic inversion consider optimization techniques (Russell, 1988; Yilmaz, 2001). In most of deterministic seismic inversion methods an initial acoustic model, usually obtained by smoothly interpolating well logs, is constructed from the well. The reflectivity sequence related to the initial acoustic model is then convolved with an extracted wavelet to calculate synthetic seismic. Then, via a process of global optimization the initial acoustic model is modified to minimize the mismatch between the synthetic and actual seismic survey, Figure 2.3. The output of deterministic seismic

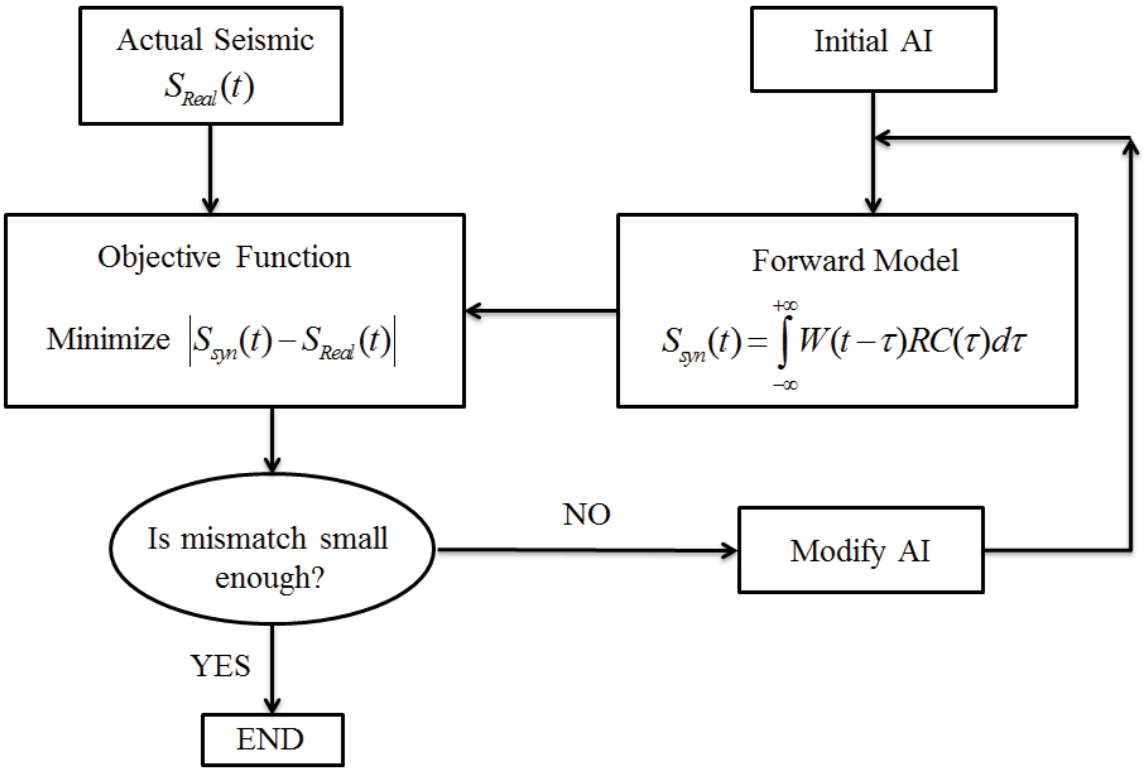

Figure 2.3: Flow chart of deterministic inversion algorithm.

inversion is a 2D or 3D acoustic property that could be used as an input to geostatistical algorithms to model acoustic properties. Finally, a set of statistical relationships are established to relate acoustic properties to the reservoir physical properties (Russell and Hampson, 1991; Ulrych et al., 2001).

Deterministic inversion only works in presence of high quality seismic data (Doyen and Guidish, 1992). The output of deterministic inversion is also relatively smooth and generally unsuitable for direct use in volumetric calculations, estimation of connectivity or fluid flow simulation. Moreover, deterministic inversion does not consider uncertainty or non uniqueness associated with seismic inversion. These techniques only partially account for uncertainty attached to the inversion process (Haas and Dubrule, 1994).

One solution to some of the limitations of deterministic inversion is stochastic seismic inversion. Stochastic seismic inversion combines the inversion process and geostatistical techniques like sequential Gaussian simulation (SGS) to provide high resolution models

that are conditioned to both well and seismic data. The models obtained via stochastic inversion algorithm have higher frequency content than low resolution seismic data. The frequencies outside the seismic bandwidth come from the well data and geostatistical approach.

The basic statistic approaches aim to model the uncertainty about unsampled value (z) as a random variable (RV) Z. The probability distribution is a parameter that characterizes the uncertainty of z. A random variable (RV) is a variable that takes a specified number of outcomes based on probability distribution. A set of random variables (RV) defines the random function (RF) (Deutsch, 1992; Deutsch and Journel, 1998). The cumulative distribution function (cdf) for a continuous random variable (RV) $Z(\mathbf{u})$ at location of $\mathbf{u}$ is:

$$F(\mathbf{u}; z) = Prob\{Z(\mathbf{u}) \leq z\} \tag{2.10}$$

When the cdf is defined based on a set of neighboring data values $Z(\mathbf{u}_\alpha) = z(\mathbf{u}_\alpha)$, $\alpha = 1, \ldots, n$ then the cdf is defined as conditional cumulative distribution function (ccdf):

$$F(\mathbf{u}; z|(n)) = Prob\{Z(\mathbf{u}) \leq z|(n)\} \tag{2.11}$$

Sequential Gaussian simulation approach is one of the most widely used geostatistical techniques in reservoir modeling applications because it is simple, efficient and flexible. The theory underlying sequential Gaussian simulation (SGS) is Kriging (Equation 2.4) to estimate the parameters of local conditional Gaussian distribution. In this method each variable is simulated sequentially based on related normal conditional cumulative distribution function (ccdf) through the kriging estimation system, Equation 2.4 (Deutsch, 1992 and Pyrcz and Deutsch, 2014). The sequential Gaussian simulation of a continuous variable $Z(\mathbf{u})$ followed as (Deutsch, 1992; Deutsch and Journel, 1998; Pyrcz and Deutsch, 2014):

1. Define the univariate cdf $F_Z(z)$ that relates to the entire simulation domain
2. Transform the z data values into y data with a Gaussian cdf
3. Sequentially go through a random path to visit each grid node once
4. At each grid node $\mathbf{u}$, define the conditioning data that consists of original hard data and previously simulated grids in Gaussian unit within a specific neighborhood
5. Determine the mean and variance of the ccdf for the random function (RF) $Y(\mathbf{u})$ by Kriging
6. Draw a simulated value $y^{(l)}(\mathbf{u})$ from the ccdf
7. Add the simulated value $y^{(l)}(\mathbf{u})$ to the data set to use it as a conditioning data in the next iteration
8. Loop over step 3 to 7 until all the grid nodes are simulated
9. Backtransform the simulate Gaussian values $\left\{y^{(l)}(\mathbf{u}), \mathbf{u} \in A\right\}$ into the variables original unit $\left\{z^{(l)}(\mathbf{u}) = G^{-1}(y^{(l)}(\mathbf{u})), \mathbf{u} \in A\right\}$

This procedure generates one realization. More realization can be generated through different random paths $\left\{z^{(l)}(\mathbf{u}), \mathbf{u} \in A\right\}, l = 1, \ldots, L$.

The main and original idea of directly accounting for seismic data in geostatistical modeling with close integration of inversion and geostatistical techniques was proposed by Bortoli (1992), Haas and Dubrule (1994) at ELF in the early 1990's. The original methodology combines a trace-by-trace optimization method with the sequential Gaussian simulation (SGS) algorithm. The vertical variogram is based on the well data and the horizontal variogram obtained from seismic data (Rowbotham et al., 1998; David, 2012). This algorithm is as follows (Bortoli, 1992):

1. Sequentially visit all traces (columns) in a random order
2. At each trace generate multiple realizations of acoustic properties conditioned to the well data and previously simulated values
3. Compute the reflectivity series for the multiple realizations
4. Convolve the reflectivity series with the extracted wavelet to obtain synthetic seismograms

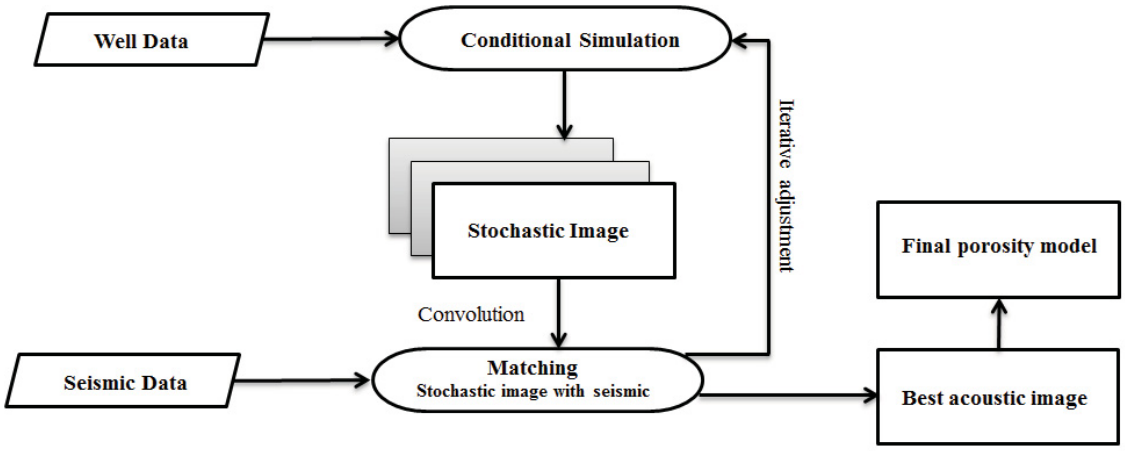

Figure 2.4: Schematically shows the matching approach in stochastic seismic inversion, original seismic data are used for best pick from prior multiple realizations that are conditioned to well data (Redrawn from Bortoli, 1992).

5. Compare the synthetic seismograms with the actual seismic trace and retain the acoustic properties associated to the best one
6. Repeat step 2-5 until all traces are simulated.

The number of realizations performed at each trace as a part of the local optimization must be chosen. Figure 2.4 represent the process of matching approach to the actual seismic data in original stochastic inversion method and Figure 2.5 shows the process of local optimization for this algorithm.

In this algorithm, the process of local optimization for selecting the best match by comparing them with actual seismic data could cause a problem in histogram and variogram reproduction (Deutsch, 2001). Another issue or shortcoming with the algorithm proposed by Bortoli (1992), Haas and Dubrule (1994) is that multiple seismic data are not used explicitly in this approach. However, this issue has been addressed by Hong et al. (2006). They developed a study based on Bortoli (1992), Haas and Dubrule (1994) algorithm to consider multiple-temporal seismic data as well as to integrate secondary seismic data. Multiple-temporal seismic data is considered as time-lapse seismic data that has many practical application in reservoir monitoring (Lumley, 2001).

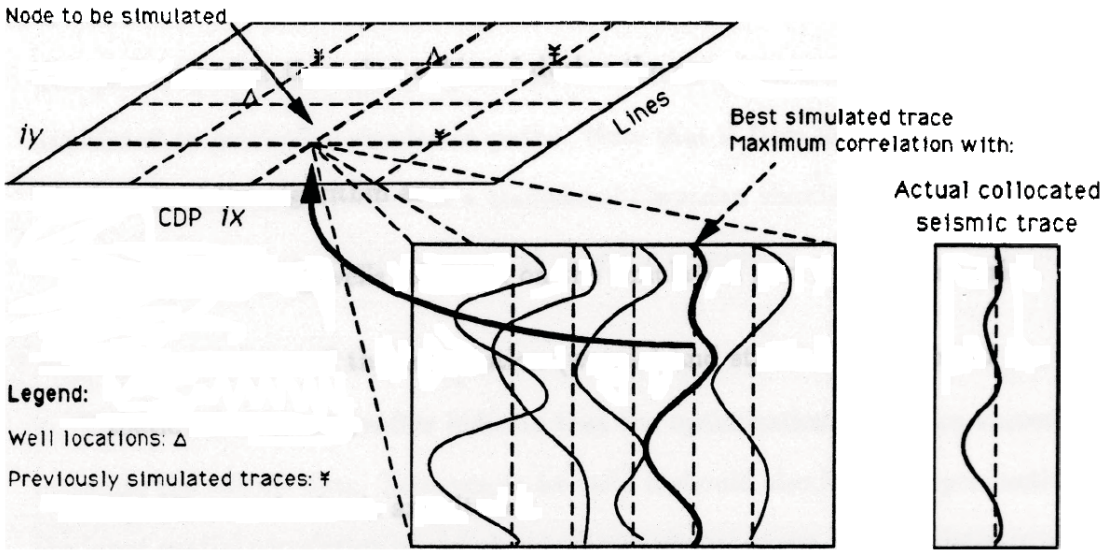

Figure 2.5: Schematic illustration of local optimization process in the original stochastic inversion approach (Bortoli, 1992).

Lamy et al. (1998) developed a study that was based on the Bortoli (1992), Haas and Dubrule (1994) algorithm that focused on different methods of managing and quantifying uncertainty in the stochastic inversion algorithm by focusing on statistical parameters such as the mean, standard deviation, $P10, P50, P90$ and threshold probability.

A different version of stochastic inversion similar to the Bortoli (1992), Haas and Dubrule (1994) method with the addition of a simulation annealing optimization algorithm was proposed by Debeye et al. (1996) at Jason Geosystems. The practical application of this algorithm to generate multiple 3D facies and porosity realizations consistent with geology and seismic data was applied by Sams et al. (1999) in a central Sumatra basin reservoir. This stochastic inversion scheme is an adaption of sequential Gaussian simulation (SGS) or sequential Gaussian co-simulation (SGCS) based on simulated annealing (Debeye et al., 1996; Bosch et al., 2010). The simulated annealing is an iterative optimization method based on trial and error. This method considers an initial model m_0 with associated energy or mismatch $E(m_0)$. It generates a new model m_n and computes the corresponded energy or mismatch $E(m_n)$. If the new mismatch is lower than the initial one, then the new model accepted unconditionally. Although, if it is larger than the initial

one, then the new model accepted with the probability, $P = exp[-(\frac{E(m_n)-E(m_0)}{T})]$ where T is a control parameter that called annealing cooling temperature (Debeye et al., 1996; Ma, 2002; Bosch et al., 2010). The key steps in this stochastic inversion algorithm are (Debeye et al., 1996; Deutsch, 2001; Bosch et al., 2010):

1. Generate an initial model of facies and associated acoustic impedance consistent with geostatistical and geological models and constrained to the well data
2. Compute the synthetic seismic volume from the initial acoustic impedance model and extracted wavelet
3. Sequentially go through a random path and pick a node:
   (a) Check all other possible facies categories ($k = 1, 2, ..., K$) and update acoustic impedance model based on them
   (b) Recalculate the synthetic seismic and the mismatch (Q_k) between synthetic and actual seismic trace for all possible facies categories ($k = 1, 2, ..., K$)
   (c) Pick the best facies categories and retain them as simulated values. The criteria to keep the facies is based on simulation annealing rule. If it reaches to the lower mismatch, the model is accepted unconditionally. Otherwise, the model is accepted with a probability. The probability to keep each category is $P_k = e^{\frac{-Q_k}{T}}$
   (d) Go to the next node and repeat step (a) to (c)until all nodes are simulated
4. Continue the procedure over all nodes of entire simulation domain until the minimum mismatch obtained. Some research studies showed that the number of 10-15 iterations are required to reach the optimal solution (Sams et al., 1999 and Latimer et al., 2000).

There is no clear application of the indicator variogram and facies proportions in this algorithm. Facies must be transformed into acoustic impedance. There is also some ambiguity related to how to define the annealing schedule in this algorithm (Deutsch, 2001).

Kane et al. (1999) developed a seismic-well log inversion method that combines the geostatistical methods for well log interpolation such as kriging with a Monte Carlo search

technique for seismic inversion. This method follows the method originally proposed by Bortoli (1992), Haas and Dubrule (1994). Kriging estimation method is applied to the well data to estimate the velocity and corresponding variance. Then, the velocity estimates and associated variances are used as a prior constraints in the seismic inversion algorithm. A trace by trace inversion algorithm is performed to obtain a complete 2D seismic section. The velocity profiles are derived from previous seismic traces are considered as pseudo well logs in further applications of kriging. This method applies a more efficient Monte Carlo search algorithm in the seismic inversion step and moves away from the wells gradually to minimize the kriging variance at each step.

Another stochastic inversion approach was introduced by Eidsvik et al. (2004). This algorithm formulates the stochastic inversion approach in terms of Bayesian network model. This algorithm accommodates the lateral spatial continuity in the prior distribution of reservoir physical properties through a Markov random field model (Bosch et al., 2010). Later, Larsen et al. (2006) modified this algorithm to incorporate vertical spatial continuity through a stationary Markov-chain prior model.

Francis (2005) introduced a different stochastic inversion approach. This algorithm applies the advantages of Fast Fourier transform-based spectral simulation to generate acoustic properties faster than the sequential simulation method. In this algorithm, conditioning to seismic data is performed by using a generalized linear inversion algorithm to update the primary acoustic impedance model. This method can be used for joint inversion of multiple seismic volumes such as near and far offset volumes or time-lapse seismic data (Bosch et al., 2010).

Another major modification in stochastic inversion was developed by Escobar et al. (2006). This algorithm assumes that both the prior distribution and likelihood are Gaussian and the forward modeling is a linear process. By these assumptions, the high dimensional posterior distribution can be decomposed as a product of Gaussian distributions. Then, Sequential Gaussian Simulation draws realizations from the posterior distribution of elastic properties (Bosch et al., 2010).

Most of the geostatistical algorithms applied in stochastic inversion rely on the variogram (two point statistics) to capture the spatial continuity in reservoir models. Two point statistics methods do not incorporate complex geology features or curvature structures. The solution to this limitation is the idea of Multi Point Statistics (MPS) with training images (Guardiano and Srivastava, 1993).

One of the first research studies to apply MPS method in stochastic inversion and obtain reservoir properties was introduced by González et al. (2007). This approach integrates inversion algorithm with MPS methods to obtain reservoir properties. This technique formulates the inversion problem as an inference problem and combines it with MPS to include prior geological information. It also uses rock physics relationships to characterize the relation between elastic and reservoir properties. This approach incorporates the concepts of sampling from conditional probabilities with concepts of optimization to generate optimal solutions. These solutions are the realizations that reproduce the available seismic data within a certain tolerance given by the rock physics distribution (Bosch et al., 2010).

Stochastic inversion of seismic data using fractal-based initial models was developed by Srivastava and Sen (2010). This approach relies on a fractional Gaussian distribution derived from the statistical parameters of available well data and also fractal theory to generate acceptable and realistic initial models. A global optimization method like simulated annealing algorithm minimizes the mismatch to seismic and well data.

The idea of involving a direct inversion of seismic amplitudes for physical properties of rock such as porosity and fluid saturation that called direct petrophysical inversion was introduced by Doyen (2011) where Petro Elastic-Models (PEM) play an important role in connecting seismic data to reservoir properties.

The application of Principle Component Analysis (PCA) in geostatistical seismic inversion introduced by Jin et al. (2013). PCA provides an orthogonal transformation to project the high dimensional model to a lower dimension by using the correlation between models. The algorithm starts with a large number of training images sampled from a prior probability

distribution, then principle components are computed from the correlation matrix of the training images. The number of principle components are much smaller than the original number of model parameters, so new training images can be obtained efficiently through the posterior probability distribution. The results from this inversion technique show better lateral continuity compared to the trace by trace inversion using the same optimization tool (Jin et al., 2013).

An iterative geostatistical seismic Amplitude Versus Offset (AVO) inversion methodology based on direct sequential simulation (DSS) and co-simulation was proposed by Azevedo et al. (2013) which allows direct inversion of prestack seismic data for density, facies, P-wave and S-wave models (Azevedo et al., 2013). This approach is an iterative procedure based on genetic algorithm. The best models at each iteration will constraint the generation of models in the next iteration. This method initially simulates multiple realizations (L) of density from available well data based on DSS algorithm (Soares, 2001) and for each density model, multiple realizations (L) of V_P/V_S are generated by co-DSS with joint probability distribution that are conditioned to the previously simulated density model. In the next step, multiple realizations of facies (L) are generated conditioned to the previously simulated elastic properties by using classification algorithm like Bayesian classification (Avseth et al., 2010). Then, multiple realizations (L) of V_P conditioned to the well data and simulated facies model by DSS are generated. In the following step, multiple realizations (L) of V_S condition to the well data and previously simulated V_P are modeled by co-DSS. Synthetic angle gather by linear Shuey's approximation (Shuey, 1985) is calculated. In the final step, the mismatch between synthetic and actual seismic data is compared and the best local correlation in a new volume is considered (Azevedo et al., 2013).

Another stochastic seismic inversion similar to the conventional stochastic seismic inversion methods that works based on the calculation of the deviation between the initial guess model and every single elastic property simulated at each simulation iteration was proposed by Pereira et al. (2016). The conventional trace-by-trace correlation coefficient between real and synthetic seismic data along with the deviation to the initial guess model

are integrated as an objective function for optimization algorithm in this method. The combined objective function works based on global optimizer to ensure the convergence of each iteration in the simulation process. This method claims that the computational time in conventional stochastic inversion is decreased and the convergence between synthetic and real seismic data is increased. This method requires many conditioning data to obtain reasonable match to the original seismic data.

2.4 Chapter Summary

High resolution numerical models of reservoir properties have a major role in reservoir characterization as they can improve hydrocarbon recovery, flow simulation and reservoir management. Integration of different sources of geological information such as well data with high vertical resolution along with seismic data with high lateral coverage can significantly improve forecasting of reservoir physical properties. Among the different geostatistical methods for seismic data integration, stochastic seismic inversion appears the most useful and applicable. Stochastic seismic inversion integrates different sources of information including well and seismic data through the close integration of geostatistical approaches and inversion techniques. Stochastic seismic inversion has the advantage of assessing uncertainty associated with inverted elastic properties.

Conventional stochastic inversion approaches simulate either facies or acoustic impedance and then connect them to the reservoir physical properties via a set of sequential connection. However, there is no guarantee that final reservoir models obtained through these sequential connection repreduces the original seismic data. To address the issues related to the conventional stochastic seismic inversion methods, this research study proposes and develops a new stochastic seismic inversion approach. The methodology is explained in the next chapter.

Chapter 3

Methodology

3.1 Introduction

Integration of seismic data into geostatistical modeling is a long-established problem. There are different approaches for seismic data integration in geostatistical modeling. One method is combination of large scale soft seismic data with small scale hard well data through the different geostatistical algorithms such as cokrigging (Doyen, 1988; Xu et al., 1992; Yao and Chopra, 2000; Grana et al., 2012). An alternative to cokriging approach is stochastic inversion which directly account for seismic data into geostatistical modeling (Hong et al., 2006 and Pyrcz and Deutsch, 2014). There are different version of stochastic inversion approaches in literature (Bortoli, 1992; Haas and Dubrule, 1994; Debeye et al., 1996; Sams et al., 1999; Helgesen et al., 2000; Francis, 2005 and Bosch et al., 2010). Most of these approaches connect a high quality facies or acoustic impedance model obtained through close integration of geostatistical modeling and stochastic inversion to the reservoir physical properties via statistical calibrations and petrophysical relationships (Deutsch, 2001; Bosch et al., 2010). Although, integrity to the original seismic data may lost through the sequential connections due to element of randomness at each step. To address the issues related to the conventional stochastic inversion approaches, multivariate stochastic inversion approach is proposed and developed.

3.2 Multivariate Stochastic Inversion Approach

The proposed method, multivariate stochastic inversion, applies a multivariate Gaussian simulation technique as part of stochastic inversion. The new approach simultaneously models all continuous and categorical variables with acoustic properties that they are

conditioned to well and seismic data. To condition multiple reservoir properties to the seismic data at the same time, the synthetic seismogram derived from these properties is compared with the actual seismic trace. As previously discussed, seismic surveys are inherently column based. Therefore, a full column of multiple reservoir properties are required at each step. All simulation algorithms in multivariate stochastic inversion approach are implemented in column wise manner to simulate a full column of physical and acoustic reservoir properties simultaneously at each simulation iteration. Figure 3.2 shows multivariate stochastic inversion algorithm schematically. The new approach works in the following order:

1. Sequentially go through a random path of columns over the simulation domain
2. Create one realization of categorical (facies) and continuous (porosity, fluid saturations, ...) variables simultaneously

    (a) Categorical variables are simulated either by column wise truncated pluriGaussian methods (two point statistics method) or by column wise multi point statistics (MPS) method

    (b) Continuous variables are simulated through the column wise sequential Gaussian simulation (SGS) algorithm

3. Pass the multiple variables to the Petro-Elastic Model (PEM) to calculate elastic properties such as density, velocity, acoustic impedance and reflectivity sequence
4. Convolve the computed reflectivity series with extracted wavelet to obtain synthetic seismogram through the convolution algorithm
5. Compare the synthetic seismogram with collocated actual seismic trace through the selection criteria

    (a) Compute MSE between actual and synthetic seismogram

    (b) Check the stopping criteria: -target MSE and - reasonable maximum number of realizations

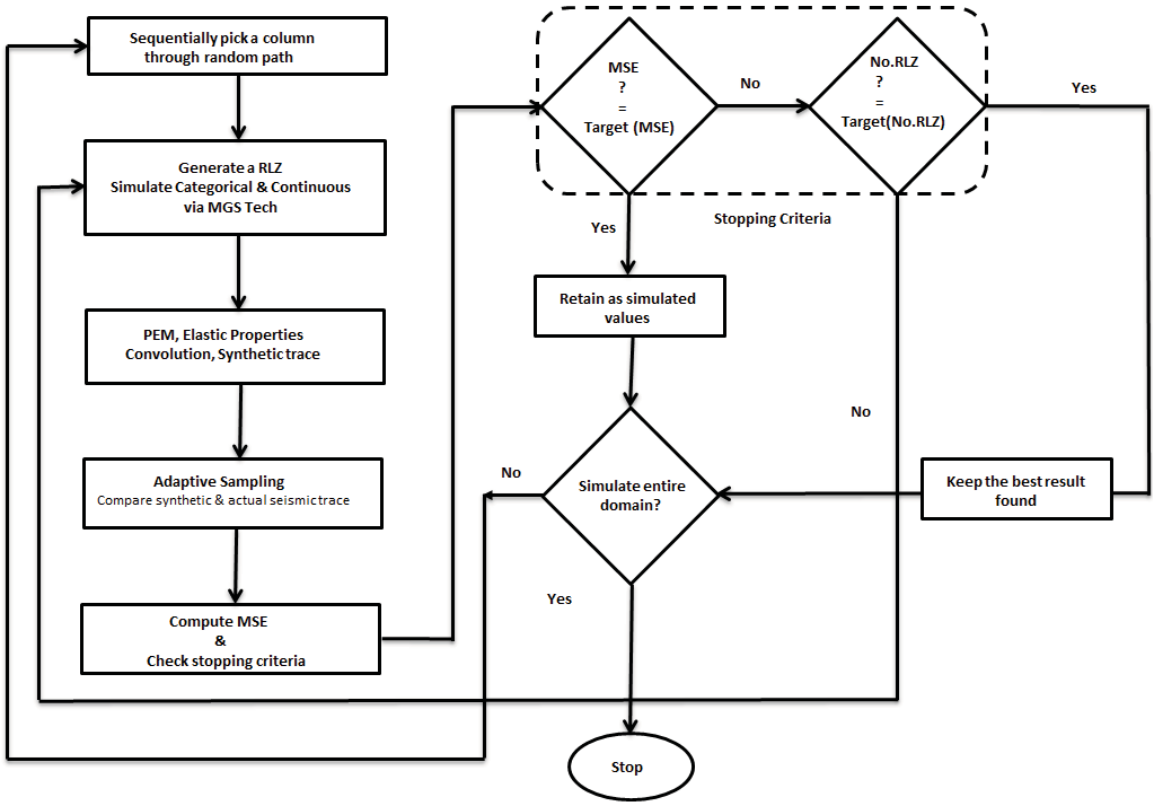

Figure 3.1: Schematic algorithm of multivariate stochastic inversion.

i. If computed MSE reached to target MSE within reasonable number of realizations, go to Step 6. If the maximum number of tries is exceeded go to Step 6

ii. Otherwise go to Step 2 and generate another realization and repeat this procedure until meet the stopping criteria requirements

6. Retain the multiple variables related to the best synthetic trace
7. Go to the next column and repeat Steps 2-6 until all columns are simulated.

One of the major contributions of the new approach is to improve the geostatistical models by applying multivariate geostatistical methods to model multiple reservoir physical properties simultaneously instead of modeling them with sequential connections. This approach models continuous and categorical variables at the same time through the multivariate Gaussian simulation technique that will be explained below.

3.2.1 Multivariate Gaussian Simulation Technique

The multivariate stochastic inversion approach simulates multiple reservoir properties simultaneously via the multivariate Gaussian simulation technique. This technique simulates multiple Gaussian variables for a full column of simulation domain, denoted by $Z_i^l(\vec{\mathbf{u}})$ that i is number of multiple Gaussian variables are simulated by this technique for the column located at $\vec{\mathbf{u}}$ and l number of multiple realizations. Continuous variables are simulated by a column based sequential Gaussian simulation algorithm (Deutsch and Journel, 1998) that performs multivariate geostatistical simulation. Categorical variables like facies are simulated by column based truncated pluriGaussian simulation algorithm as the primary option for facies modeling. To add more flexibility and to handle more complex geological features in reservoir modeling by multivariate stochastic inversion approach, column based multi points statistics (MPS) is an alternative for facies modeling. More details about column based facies modeling are explained on Chapter 5.

The number of Gaussian variables are simulated by the multivariate Gaussian simulation technique is a function of number of continuous variables or reservoir physical properties (n_c), number of Gaussian variables is required to define sequential order among facies categories through the truncated pluriGaussian simulation (n_f) and number of facies categories (K), Equation 3.1:

$$Z_i^l(\vec{\mathbf{u}}) \qquad i = 1, \ldots, (n_f + K \cdot n_c) \qquad l = 1, \ldots, L \tag{3.1}$$

where l is the number of realizations and $\vec{\mathbf{u}}$ the location of the simulated column.

In this technique, continuous variables are simulated for each facies category. At the end, based on the simulated facies value, the corresponding simulated continuous variables are kept for each vertical position. In case of clear sequential order among the facies categories, only one Gaussian variable is required to simulate a facies by truncated Gaussian simulation ($n_f = 1$). Therefore, to simulate a facies with two categories ($K = 2$) like sand and shale

and two reservoir physical properties ($n_c = 2$) such as porosity and water saturation five Gaussian variables are required to be simulated by this technique. One Gaussian variable is required for facies modeling and four Gaussian variables are required for porosity and water saturation associated with sand and shale.

To implement multivariate Gaussian simulation, a sequential Gaussian simulation algorithm (Deutsch and Journel, 1998) that performs multivariate geostatistic simulation called Ultimate Sequential Gaussian Simulation (USGSIM) (Manchuk and Deutsch, 2012) is modified to: (1) perform column based simulation, (2) generate multiple realizations per column, (3) apply multivariate simulation for both categorical and continuous variables simultaneously, (4) couple with Petro Elastic Model (PEM) to generate multiple reservoir elastic properties, (5) convolve the elastic properties with extracted wavelet, and (6) apply a selection criteria to retain one realization. USGSIM is a flexible sequential Gaussian simulation algorithm where multiple variables are cosimulated simultaneously. Multiple secondary data such as seismic and production data can be applied by using locally varying means, Bayesian updating or collocated cokriging. The stepwise transform is integrated in this algorithm as are collocated cokriging, cokriging with a linear model of coregionalization and collocated cokriging with the intrinsic model for the cosimulation of multiple variables (Manchuk and Deutsch, 2012).

Figure 3.2 shows the schematic form of the correlation matrix in the multivariate Gaussian simulation approach for a facies with K categories that simulated by n_f Gaussian variables through the truncated pluriGaussian simulation and n_c continuous variables. This matrix is a block matrix where the top row and first column are related to the facies. The diagonal blocks, second to the K blocks, represent how continuous variables are related to each other within the facies. There is no reason for cross correlation among the continuous variables between facies. Therefore, the remaining off diagonal blocks are zero.

Model validation such as visual assessment and reproduction of statistics such as variogram and histogram is necessary for column based USGSIM algorithm because this algorithm involved many modification steps. For checking, a $2D$ simulation domain

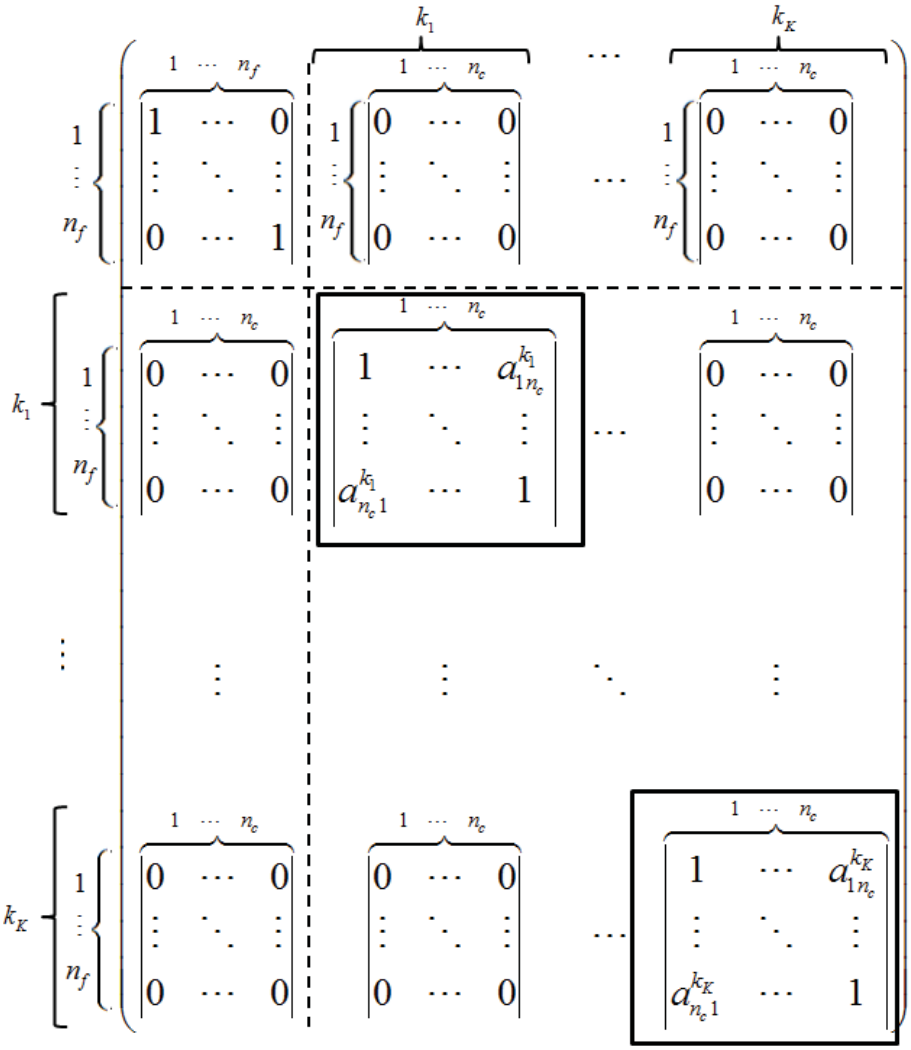

Figure 3.2: Schematic form of correlation matrix of multivariate Gaussian simulation for a facies with K categories and n_c continuous variables which n_f Gaussian variables required for truncated (pluri)Gaussian.

(250×250) are considered and two negatively correlated ($\rho = -0.7$) variables are simulated unconditionally. These simulated variables randomly sampled and considered as conditioning well data for the next steps of model checking. Multiple realizations are generated based on this conditioning data with the USGSIM and column based USGSIM algorithm. For visual assessment, Figure 3.3 and Figure 3.4 display a map of conditioning data and two different realizations of these negatively correlated variables for both implementation of USGSIM. As shown in these figures, the simulation results for column

based USGSIM reproduce the conditioning data with no apparent artifacts. The negative correlation between two variables is reproduced. The correlation coefficient of two different realizations of two negatively correlated variables ($\rho = -0.7$) simulated by column based USGSIM are -0.71 and -0.72. Figure 3.5 and 3.6 represent histogram reproduction of two realizations of simulated variables for USGSIM and column based USGSIM. As shown in these figures, the first row displays histogram of conditioning data and the rest compare the histogram of two different realizations of simulated variables for USGSIM and column based USGSIM. Figure 3.7 displays the omni-directional variogram reproduction of two different realizations of simulated variables for the two algorithms.

Facies and multiple continuous variables are simulated by multivariate Gaussian simulation technique through the column based sequential Gaussian simulation and truncated pluriGaussian simulation algorithm. The simulated multiple reservoir properties are passed to Petro-Elastic Model (PEM) to calculate multiple reservoir elastic properties.

3.3 Petro Elastic Model

The elastic properties of the reservoir are required at the same time as the physical properties. A Petro Elastic Model (PEM) is considered to calculate elastic properties such as bulk density, velocity and acoustic impedance. The Petro Elastic Model (PEM) is coupled into the multivariate geostatistical modeling algorithm. Reservoir physical properties generated via the multivariate geostatistical approach are passed to PEM for calculation of reservoir elastic properties.

There are different Petro-Elastic Model. In the multivariate stochastic inversion algorithm, the bulk density will be calculated by the following equation;

$$\rho = \phi\rho_f + (1-\phi)\rho_r \tag{3.2}$$

where ϕ is porosity, ρ_f and ρ_r are fluid and rock densities, respectively. The fluid density

will be also calculated by following expression;

$$\rho_f = \rho_o S_o + \rho_w S_w \tag{3.3}$$

in Equation 3.3, ρ_o and ρ_w are oil and water phase density and S_o and S_w are oil and water phase saturation respectively.

There are many different petrophysical approaches to calculate velocity. One approach to calculate velocity is to establish a relationship between density and velocity. Bulk density and velocity depend on the mineral composition, the granular nature of the rock matrix, cementation, porosity, fluid content and environmental pressure (Gardner et al., 1974). Gardner's equation is common. Gardner et al. (1974) expressed the relation between density (ρ) and compressional velocity (V_P) as follows, Equation 3.4.

$$\rho = a(V_P)^m \tag{3.4}$$

The constants a and m can be determined from fitting a line to a plot of $log(\rho)$ versus $log(V_P)$, Figure 3.8. Equation 3.4 takes different forms, depending on the unit of density, velocity and also rock types (Gardner et al., 1974). In the proposed research study, Gardner's equation is considered to calculate velocity. The constants of a and m must be determined on a case by case basis. The reflectivity series computed based on acoustic impedance must then be passed to a convolution algorithm to calculate synthetic seismic trace.

For better understanding of PEM algorithm. Let's consider a facies model (250×250) with two categories, shale $(\rho_{shale} = 2.65g/cc)$ and sand $(\rho_{sand} = 2.62g/cc)$, and two related reservoir physical properties such as porosity and fluid saturation as input of PEM algorithm. In this case the pore space consists of two phases; oleic phase $(\rho_{oil} = 0.95g/cc)$ and aqueous phase $(\rho_{water} = 1g/cc)$. The density is calculated by Equation 3.2 and velocity is obtained from Gardner's equation (Equation 3.4) which $a = 0.31$ and $m = 0.25$ for this case. Acoustic impedance is computed as product of

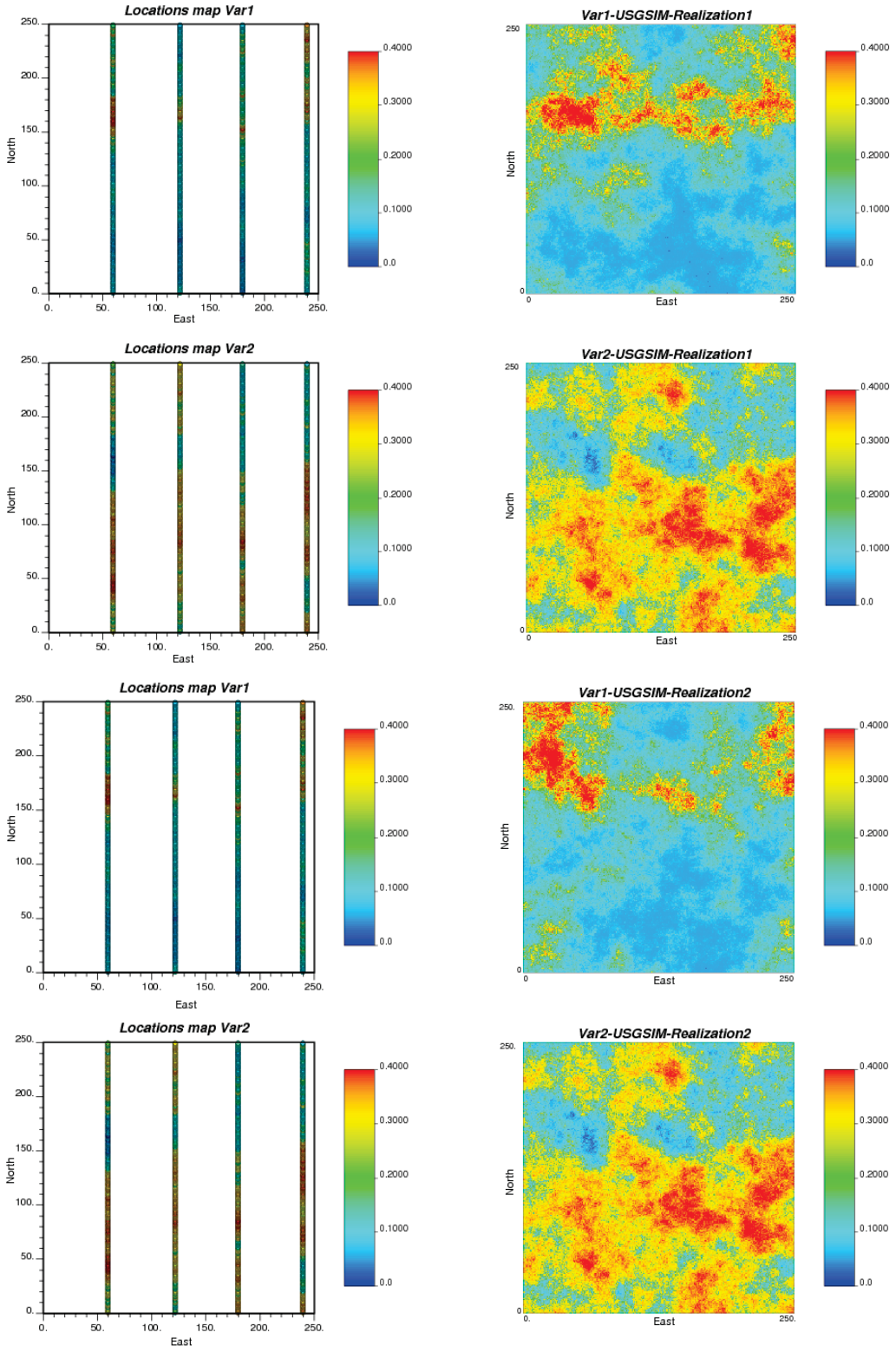

Figure 3.3: Conditioning data and two different realizations of two negatively correlated variables ($\rho = -0.7$) simulated by USGSIM, axes unit (m), color bar unit (m^3/m^3).

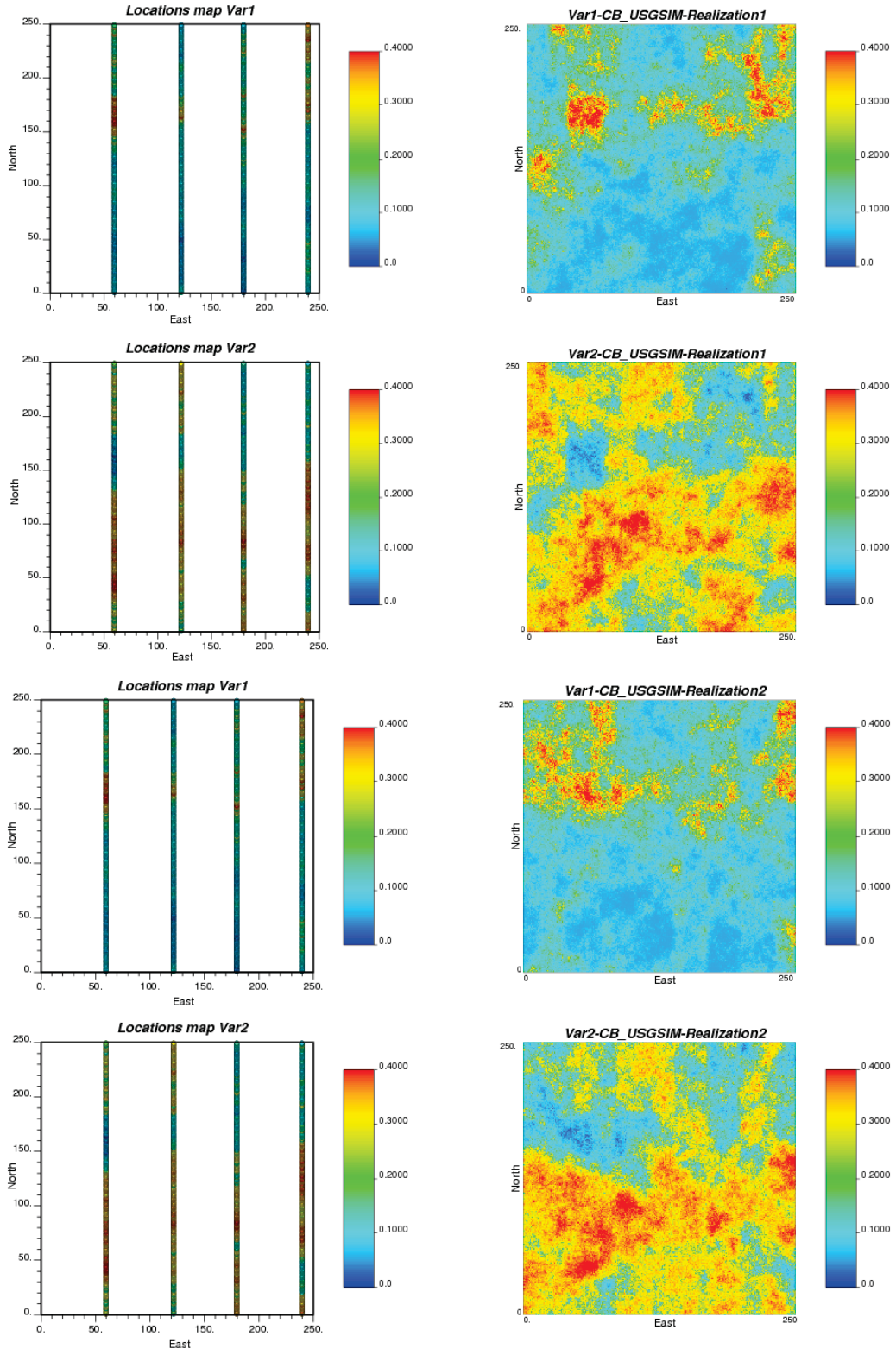

Figure 3.4: Conditioning data and two different realizations of two negatively correlated variables ($\rho = -0.7$) simulated by column based USGSIM, axes unit (m), color bar unit (m^3/m^3).

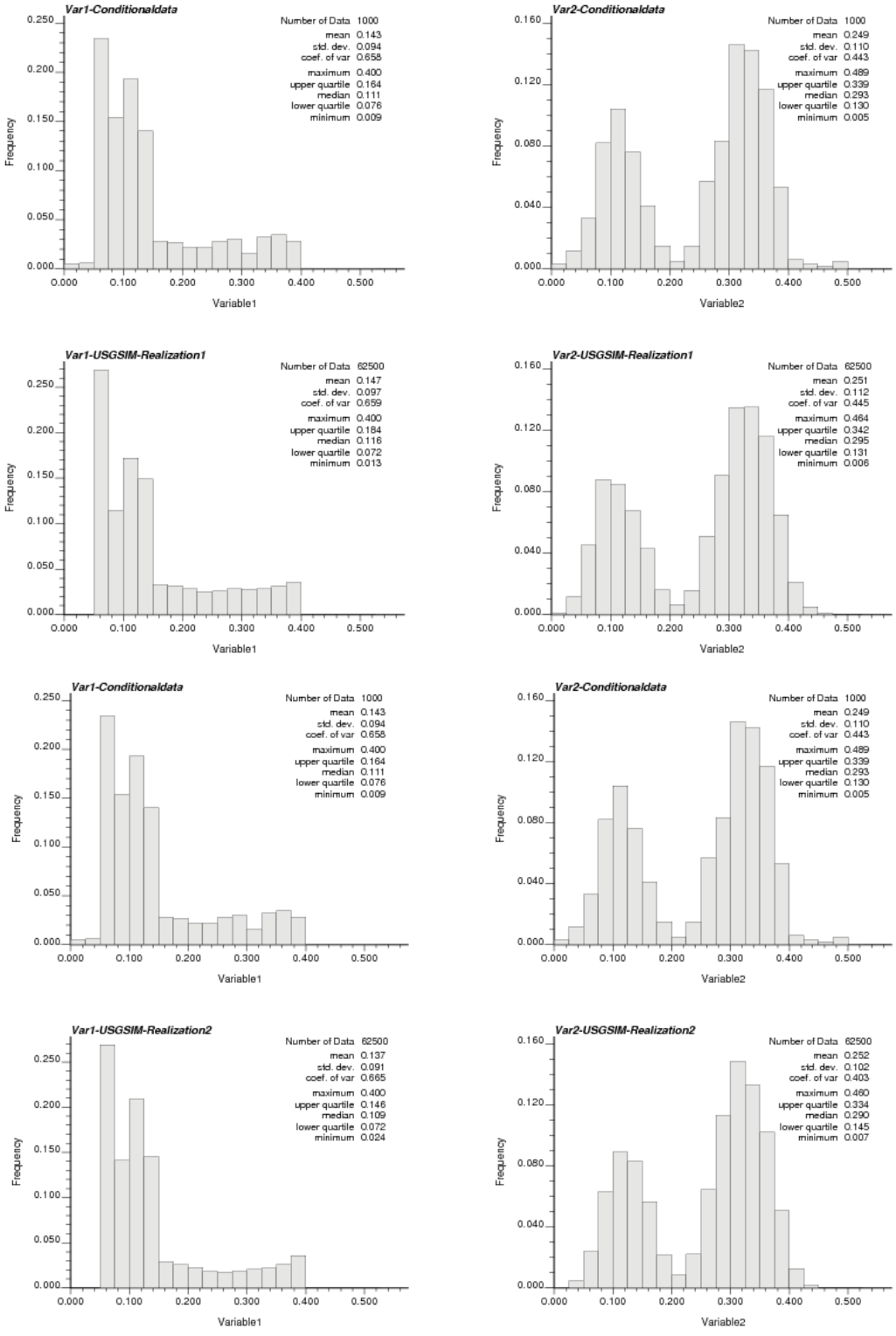

Figure 3.5: Histogram of conditioning data and two different realizations of two negatively correlated variables ($\rho = -0.7$) simulated by USGSIM.

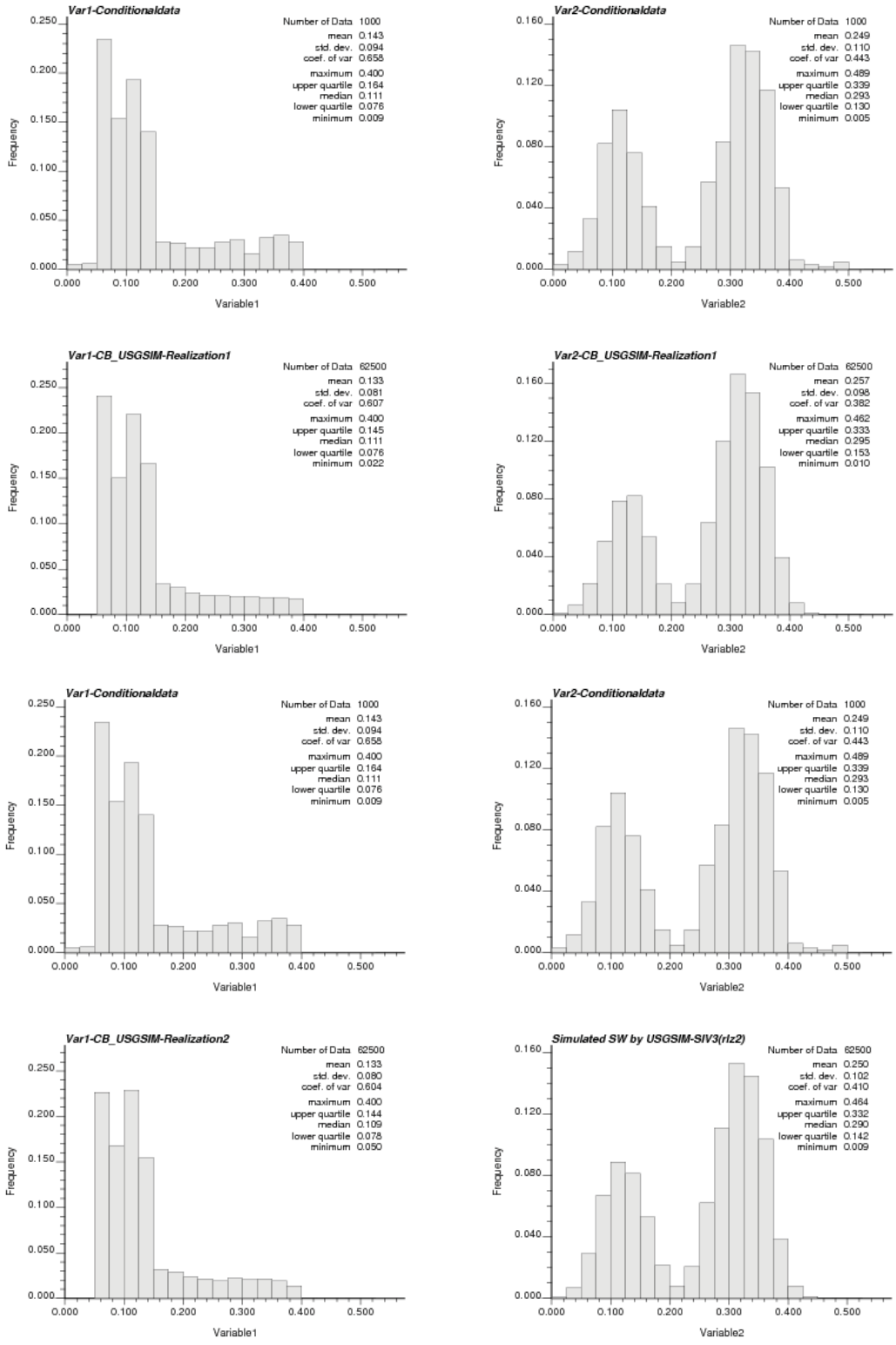

Figure 3.6: Histogram of conditioning data and two different realizations of two negatively correlated variables ($\rho = -0.7$) simulated by column based USGSIM.

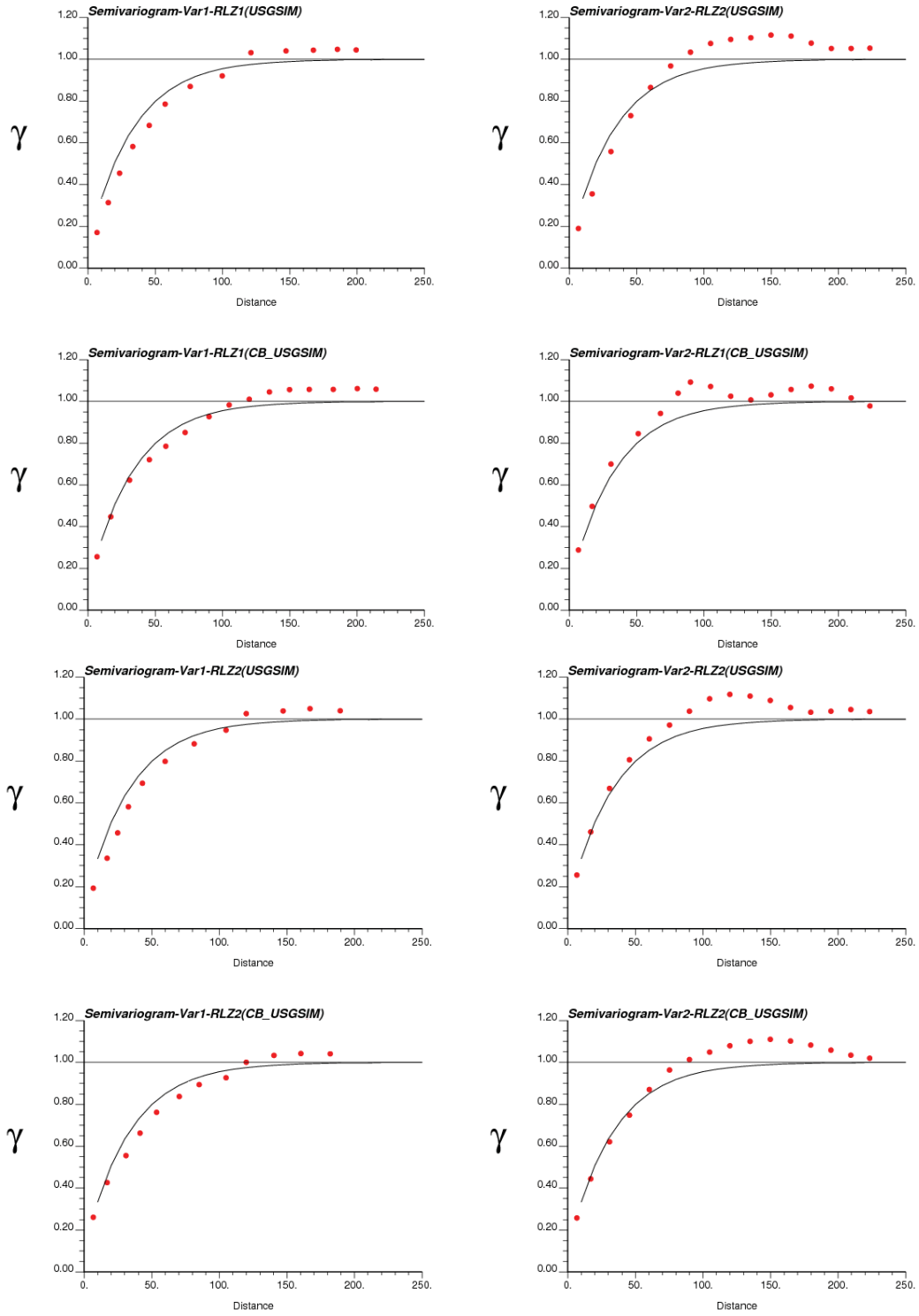

Figure 3.7: Variogram (omni-directional) reproduction of two different realizations of two negatively correlated variables ($\rho = -0.7$) simulated by USGSIM and column based USGSIM (black line variogram model and red dots variogram of simulated values), distance in (m).

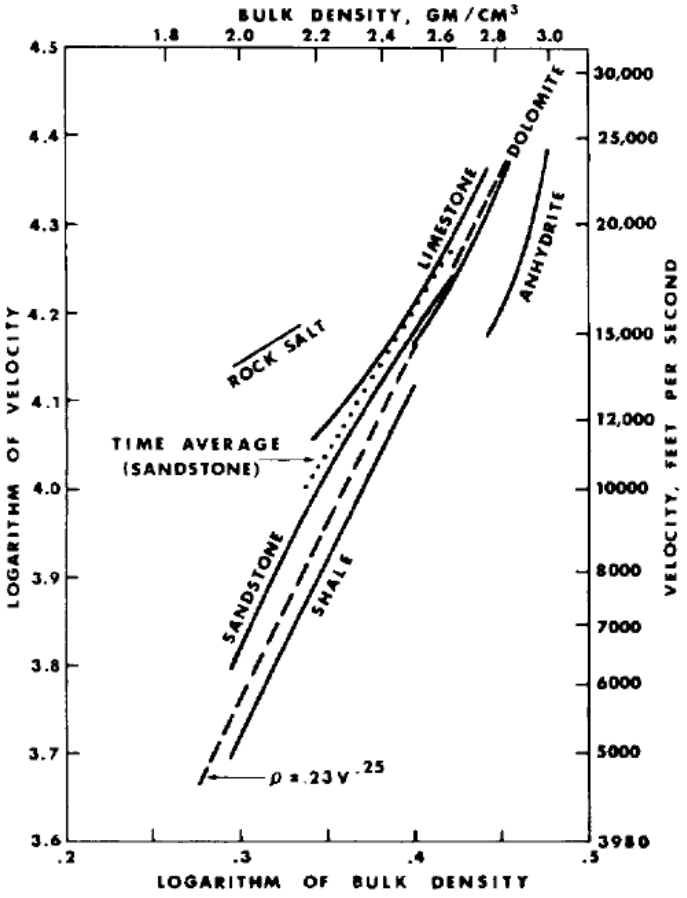

Figure 3.8: Velocity-density relationships in rocks of different lithology (Gardner et al., 1974).

density and velocity and normal incident reflection coefficient (R) by Equation 2.1. Figure 3.9 shows the reservoir properties as input of PEM algorithm and also the reservoir elastic properties obtained from this algorithm. The reflectivity series is used as input to the convolution algorithm to compute the synthetic seismic.

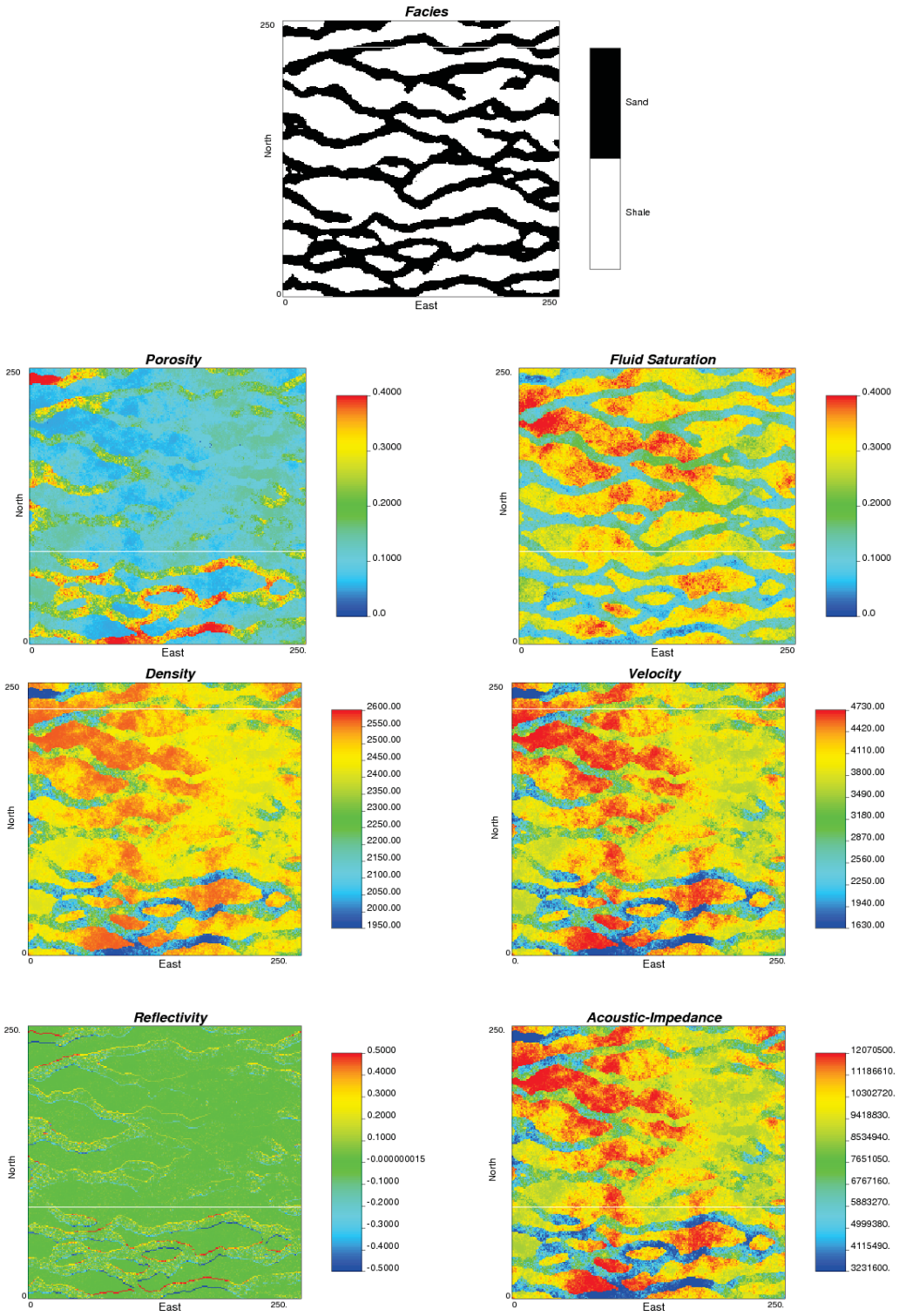

Figure 3.9: Reservoir physical properties as input of PEM and reservoir elastic properties obtained from PEM, axes unit (m), color bar of porosity and saturation (m^3/m^3), density (gr/m^3), velocity (m/sec), and acoustic impedence ($gr/m^2.sec$).

3.4 Convolution and Synthetic Seismogram

The column based multiple realizations of reservoir physical properties must be compared with the collocated actual seismic trace. The one that falls within an acceptance criteria could be chosen. To compute synthetic seismograms at each trace, the reflectivity series obtained via the Petro Elastic Model (PEM) will be convolved with the extracted wavelet.

The convolution of reflectivity sequence, R, with a wavelet, W, yields a seismic trace, S, which usually written as the convolution integral:

$$S(t) = \int_{-\infty}^{+\infty} \mathrm{W}(t-\tau)R(\tau)\mathrm{d}\tau \tag{3.5}$$

When the parameters are discrete, finite length approximations to these quantities, the convolution is written as a summation. Where R_j is reflectivity series with $j = 0, 1, ..., n$ and W_i is the wavelet with $i = -m, ..., 0, ..., m$, Equation 3.6:

$$S_i = \triangle t \sum_{j=i+m}^{i-m} W_{i-j}R_j \tag{3.6}$$

Expanding a few terms of this summation, Equation 3.6, shows that this convolution process can be considered as matrix operation:

$$\begin{aligned} S_0 &= ... + W_0R_0 + W_{-1}R_1 + W_{-2}R_2 + ... \\ S_1 &= ... + W_1R_0 + W_0R_1 + W_{-1}R_2 + ... \end{aligned} \tag{3.7}$$

Convolution is a stationary process where the convolution matrix is the same everywhere. That is, the wavelet that is scaled and used to replace each reflectivity spike does not change with time or location. The convolution process can affect not only the amplitude of the

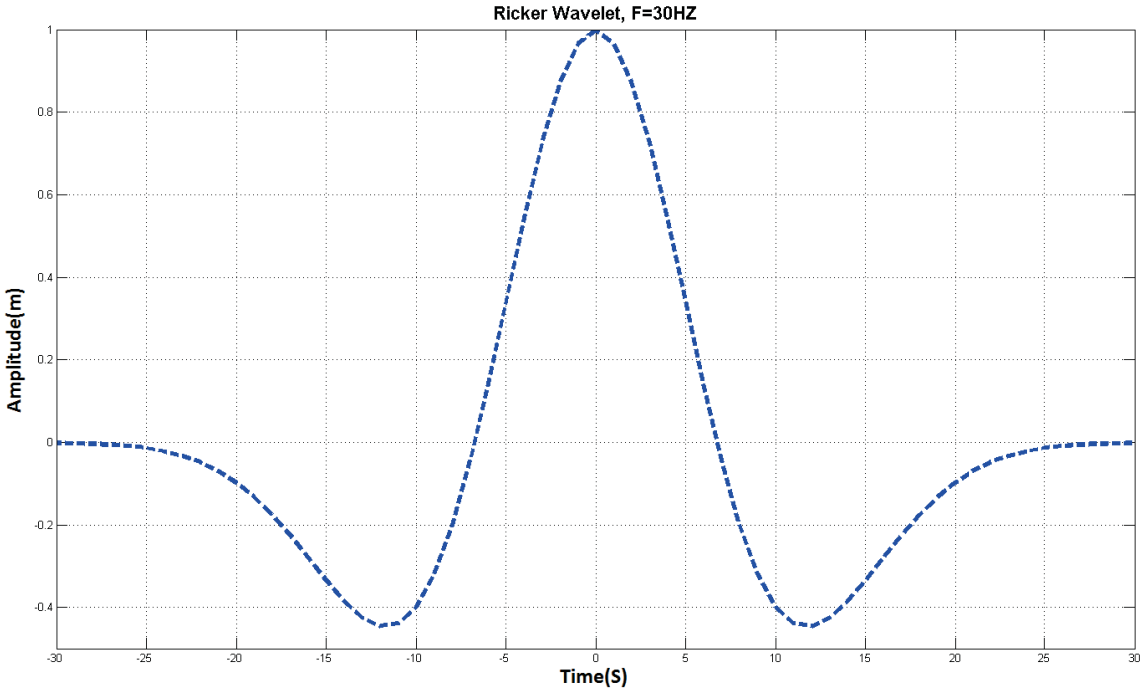

Figure 3.10: Ricker wavelet with central frequency of 30 Hz, amplitude (m), time (sec).

wave but also its phase. The Ricker wavelet known as a zero phase with central frequency of $30HZ$ wavelet does not have amplitude attenuation and phase change under convolution process (Margrave, 2009). Figure 3.10 shows a zero phase Ricker wavelet with central frequency of $30HZ$ in time domain.

It is worth mentioning, the convolution in depth domain can be represent the same as convolution in time domain. In case of dealing with reflectivity series in the depth domain the convolution in depth domain (Equation 2.3) can be applied. In this case the Ricker wavelet in depth domain should be considered. The Ricker wavelet in depth domain follows the same formalism and shape as the time domain, the depth domain Ricker wavelet is formulated with a dominant wavenumber rather than a dominant frequency (Zhang et al., 2016).

The objective of this research study is to consider a reasonable wavelet implemented as a robust and fast convolution algorithm. Wavelet extraction and well tie analysis are beyond the scope of this research study. The convolution algorithm with the specific wavelet

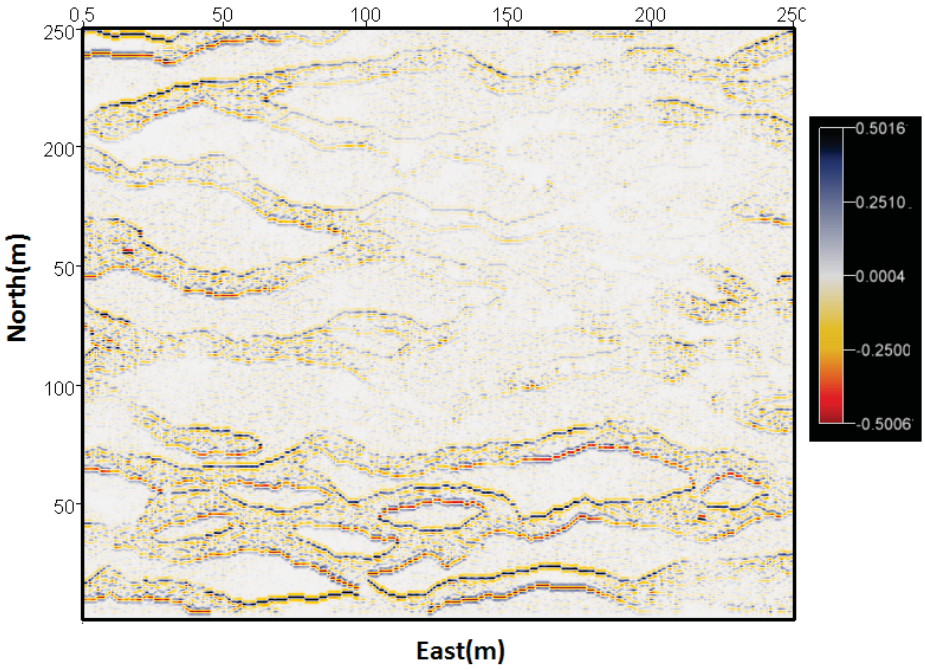

Figure 3.11: Shows 2D seismic that obtained from convolution algorithm, axes unit (m), color bar - amplitude (m).

computes synthetic seismograms based on Petro-Elastic Model(PEM) results from the multivariate sequential Gaussian simulation algorithm. The zero phase Ricker wavelet with a central frequency or wave number are considered in this study. Figure 3.11 shows the 2D synthetic seismic survey obtained from the convolution of Ricker wavelet and R (Figure 3.9) computed by PEM algorithm in the previous example.

Synthetic seismograms are passed to a selection algorithm to pick the best match to the actual seismic data based on acceptance criteria.

3.5 Adaptive Sampling Algorithm

To condition multiple simulated reservoir properties to seismic data, the synthetic seismograms computed by the convolution algorithm are compared to the actual seismic trace through an adaptive sampling algorithm. In multivariate stochastic inversion, adaptive sampling algorithm defines practical stopping criteria to reach an acceptable

match to the original seismic data within a reasonable number of realizations. The mean square error between the synthetic and collocated original seismic trace are computed and compared with the target MSE that is based on the inherent uncertainty in the modeling process. The notion of the space of uncertainty in presence of different parameters is applied to establish a reasonable number of realizations for multivariate stochastic inversion. Adaptive sampling in content of multivariate stochastic inversion means, the simulation process generates large number of realizations at the primary steps to reach an acceptable match and gradually decreases this number by going through the simulation steps as the size of space of uncertainty is reduced. More details on adaptive sampling and stopping criteria can be found in Chapter 6. The quantification of the size of the space of uncertainty is discussed in the next chapter.

3.6 Chapter Summary

Multivariate stochastic inversion applies multivariate geostatistical techniques as part of the stochastic inversion algorithm and simulates multiple reservoir properties simultaneously to provide high resolution reservoir models that reproduce the original seismic data. Multiple reservoir properties, categorical and continuous, are simulated through the multivariate Gaussian simulation technique. For this purpose, the Ultimate Sequential Gaussian Simulation (USGSIM) algorithm is modified. Continuous variables are simulated by column based sequential Gaussian simulation (SGS). The primary option for facies modeling is truncated pluriGaussian simulation method. To be able handle more complex geological features and add more flexibility to the multivariate stochastic inversion approach, column based multi point statistics simulation is an alternative for facies modeling. Reservoir properties are conditioned to the seismic data through a selection criteria that called adaptive sampling. Adaptive sampling algorithm compares the synthetic seismic trace with collocated original one and pick the one that reaches an acceptable match with the original seismic data within a reasonable number of realizations. Size of the space of uncertainty is a good indicator to define reasonable number of realizations in multivariate stochastic inversion approach.

Chapter 4

Size of Space of Uncertainty

4.1 Introduction

Geostatistical approaches, both estimation and simulation, have been extensively applied in reservoir modeling. The estimation methods such as kriging creates an overly smooth numerical model that does not represent the true spatial variability. Simulation methods create a set of realizations conditioned to a variety of information, such as well log data, core data, seismic surveys and production data. These methods reproduce the statistics inferred from the conditioning data and represent the spatial variability. Normally, the set of multiple realizations provide a qualitative and quantitative measure of spatial uncertainty (Goovaerts, 1999). Multiple geostatistical realizations are used for three fundamental purposes (Srivastava, 1996) : 1) assessment of impact of uncertainty, 2) Monte-Carlo risk analysis and 3) spatial variation reproduction.

Reservoir engineers are responsible for flow forecasting and future field development. They search for the most optimistic ($P10$) and pessimistic($P90$) models among the multiple realizations. Therefore, uncertainty assessment requires the generation of many equiprobable realizations. Normally, one hundred realizations is considered to be enough to obtain fairly stable prediction of $P50$ and also choose reliable $P10$ and $P90$. However, the major concern is how many realizations are required in a stochastic inversion context to find one that is acceptably close to the input data. This number may be much larger than one hundred. The set of all possible outcomes is referred to as "*the space of uncertainty*".

The concept of the space of uncertainty has been discussed in many geostatistical studies such as Rossi (1994); Srivastava (1994, 1996); Journel and Xu (1994); Journel (1997), Myers (1994, 1996), Mukerji et al. (2001), Eidsvik et al. (2004), Kjønsberg et al. (2010),

Grana et al. (2012) and Johansen et al. (2013). Some of these studies believe that the space of uncertainty should be defined via the stochastic simulation algorithm by including all possible realizations that generated by the algorithm. Although, few of them believe that the space of uncertainty should be defined outside of stochastic simulation algorithm regardless of the method of simulation (Goovaerts, 1999; Goovaerts, 2006). The term of algorithmically defined the space of uncertainty was suggested later by Deutsch (1994) and its application can be found in different geostatistical modeling research study like Goovaerts (2006).

In addition to the geostatistical studies discussed about space of uncertainty, there are some studies that focus on uncertainty estimation. One recent approach by Zunino et al. (2014) inverts seismic reflection data in the framework of the probabilistic approach to inverse problems by using a Markov chain Monte Carlo (McMC) algorithm. This method aims to directly infer the facies and porosity of reservoir. Therefore, this method combines a rock physics model with seismic data in a single inversion algorithm. For the large data set, the Markove chain Monte Calro (McMC) method is computationally impractical. So, in case of dealing with large data set, this method relies on multi-point based a priori information to quantify geologically plausible models. The solution of inverse problem then represented by a collection of facies and porosity models.

In geostatistical modeling, there is always some degree of uncertainty despite the fact that there is a single and inaccessible truth for the geological structures (Pyrcz and Deutsch, 2014). Uncertainty is due to our limited data relative to the scale of variability. Multiple realizations provide samples from the uncertainty model (Pyrcz and White, 2015). There are different mathematical approaches to model uncertainty described in books including Journel and Huijbregts (1978), Goovaerts (1997) and Chilès and Delfiner (2012).

Many research studies have addressed uncertainty estimation and modeling, little has been published on the actual size of the space of uncertainty and its application in geostatistical modeling. In the early days of simulation, there was a claim that the set of multiple realizations generated by geostatistical simulation may include the truth (Goovaerts,

1999;Mukerji et al., 2001; Leuangthong et al., 2004). In practice, however, we know that the space of uncertainty is very large and the number of realizations is very small. Therefore, the probability that one realization happens to be the truth is infinitesimally small. Although, the size of space of uncertainty is inconceivably large it is also interesting to understand how the space becomes smaller in presence of spatial correlation and conditioning data.

4.2 Size of Space of Uncertainty

To quantify the size of the space of uncertainty, consider a categorical variable that can take one of K outcomes. Consider a set of N locations that define a space of uncertainty denoted Λ. The size of the space of uncertainty is a measure of its extent denoted $M(\Lambda)$. A multivariate distribution of values can be decomposed into a sequence of univariate conditional distributions using the definition of conditional distributions, Equation 4.1:

$$\begin{aligned} f_{Z_1,\ldots,Z_N}(Z_1,\ldots,Z_N) &= f_{Z_2,\ldots,Z_N|Z_1=z_1}(Z_2,\ldots,Z_N) \bullet f_{Z_1=z_1}(Z_1) \\ &\vdots \\ &= f_{Z_N|Z_{N-1}=z_{N-1},\ldots,Z_1=z_1}(Z_N) \bullet \cdots \bullet f_{Z_2|Z_1=z_1}(Z_2) \bullet f_{Z_1=z_1}(Z_1) \\ &= \prod_{i=1}^{N} f_{Z_i|Z_{i-1},\ldots,Z_1=z_1}(Z_i) \end{aligned} \tag{4.1}$$

Therefore, the size of the space of uncertainty for a multivariate distribution could also be considered as the product of the size of the space of uncertainty of univariate conditional distributions, Equation 4.2:

$$
\begin{aligned}
M(\Lambda) &= M(f_{Z_1,\ldots,Z_N}(Z_1,\ldots,Z_N)) \\
&= M(\prod_{i=1}^{N} f_{Z_i|Z_{i-1}=z_{i-1},\ldots,Z_1=z_1}(Z_i)) \\
&= \prod_{i}^{N} M(f_{Z_i|Z_{i-1}=z_{i-1},\ldots,Z_1=z_1}(Z_i))
\end{aligned} \tag{4.2}
$$

Size of the space of uncertainty of a univariate distribution for K completely random categories is K. This size drops to 1 in case of one deterministic category. The size of the space of uncertainty for independent equally probable events, say K categories over N locations, is K multiplied by itself N times, see Equation 4.3. This size dramatically drops to the 1 in deterministic case.

$$
M(\Lambda) = \prod_{i=1}^{N} K = K^N \tag{4.3}
$$

For any probability distribution, there is a quantity called entropy that defines the uncertainty of a single random variable (Cover and Thomas, 2006). In fact, entropy provides a measure of uncertainty associated with the probability density function (PDF) of a random variable. If the random variable is a discrete variable that can take K outcomes values with probability $p_k, k = 1, ..., K$ while $\sum_{k=1}^{K} p_k = 1$, then the entropy can be defined as (Li and Deutsch, 2010):

$$
H = -\sum_{k=1}^{K} p_k ln(p_k) \tag{4.4}
$$

Maximum entropy occurs only when all categories are equally probable, see Equation 4.5. Minimum entropy is achieved when only one category can happen, Equation 4.6.

$$p_k = \frac{1}{K} \quad \Longrightarrow H_{max} = ln(K) \tag{4.5}$$

$$p_k = 1, p_{k'} = 0 \quad \forall k \neq k' \quad \Longrightarrow H_{min} = 0 \tag{4.6}$$

As mentioned above, the maximum size of uncertainty of a univariate distribution is K for K totally random categories. Based on the definition of maximum entropy, this size can be defined as the exponential maximum entropy of the configuration, see Equation 4.7. Minimum size of the space of uncertainty of a univariate distribution is 1 in case of one deterministic category. This size can also be defined as exponential minimum entropy of configuration, Equation 4.8. Therefore, maximum and minimum size of the space of uncertainty of univariate distribution can be characterized by maximum and minimum entropy of the configuration respectively. In fact, the entropy is an indicator of extent or size of the associated distribution and exponential entropy characterizes this size perfectly (Campbell, 1966).

$$H_{max} = ln(K) \Longrightarrow K = e^{H_{max}} = M_{max}(\Lambda) \tag{4.7}$$

$$H_{min} = 0 \quad \Longrightarrow \quad e^{H_{min}} = 1 = M_{min}(\Lambda) \tag{4.8}$$

Size of the space of uncertainty of a univariate distribution in general form can be quantified as the exponential entropy of the configuration, Equation 4.9 . In Figure 4.1, the dash line displays the size of the space of uncertainty of a univariate distribution for maximum and minimum cases while the solid black curve shows exponential entropy of system which is size of the space of uncertainty of univariate distribution in general form.

$$M(\Lambda) = e^H \tag{4.9}$$

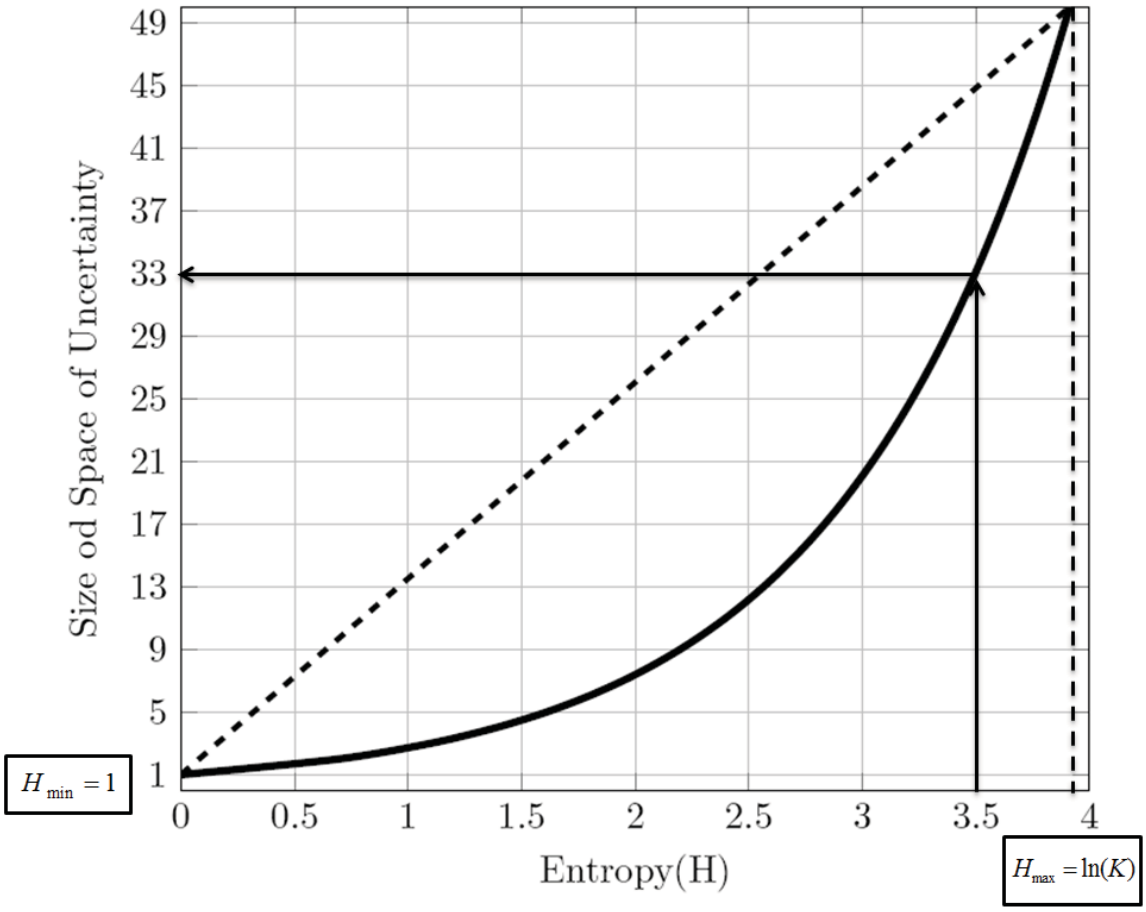

Figure 4.1: Entropy of configuration versus size of the space of uncertainty.

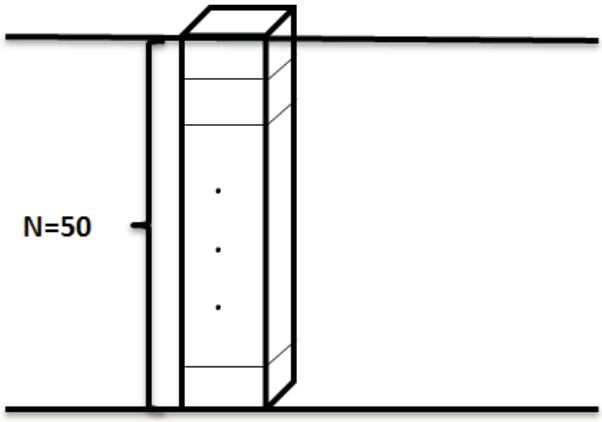

Figure 4.2: Sketch shows a case study, a column over 50 locations, ($N = 50, \Delta Z = 1$).

As a result, size of the space of uncertainty for K categories over N locations can be calculated by Equation 4.10.

$$\begin{aligned} M(\Lambda) &= \prod_{i=1}^{N} K = \prod_{i=1}^{N} e^{H_i} \\ &= \prod_{i=1}^{N} e^{\left(-\sum_{k=1}^{K} (p_k|j=1,\ldots,i-1) \bullet (ln(p_k|j=1,\ldots,i-1))\right)_i} \end{aligned} \tag{4.10}$$

Based on Equation 4.10, size of the space of uncertainty is the product of exponential entropy values. This was proved and applied primarily in information theory (Cover and Thomas, 2006).

The current practice of geostatistical modeling of categorical variables generally considers $K = 2$ to 7 categories over $N = 10^6$ to 10^8 locations. Size of the space of uncertainty in these cases is inconceivably large and cannot be understood from a practical perspective. To have an idea about the size of the space of uncertainty in real practices, consider a case study of a column over fifty locations ($N = 50, \Delta Z = 1$) with $K = 2$ to 4 categories, Figure 4.2.

Table 4.1: Size of the space of uncertainty for $K = 2$ to $K = 4$ categories over fifty locations ($N = 50$, $\Delta Z = 1$).

Number of categories	size of space of uncertainty
$K = 2$	$2^{50} \cong 1.126 \times 10^{15}$
$K = 3$	$3^{50} \cong 7.180 \times 10^{23}$
$K = 4$	$4^{50} \cong 1.268 \times 10^{30}$

As previously discussed, maximum space of uncertainty for K categories over N locations is K^N. As shown in Table 4.1, the maximum size of the space of uncertainty for this case study varies between order of 10^{15} to 10^{30}. It is worth mentioning, one light year is in order of 10^{15} ($9.46 \times 10^{15} m$) and Milky Way radius is in order of 10^{22} ($3.9 \times 10^{22} m$).

Although, size of the space of uncertainty for a case with totally random categories is very big there are different factors that reduce this size significantly including (1) prior global or local proportions for the categories, (2) spatial correlation of the categories, and (3) conditioning data. To investigate the effect of different parameters on the size of the space of uncertainty, the computation of size of the space of uncertainty is implemented.

4.3 Algorithm to implement size of space of uncertainty

Implementation of size of the space of uncertainty must be done through the sequential probability estimation. The probability cannot be estimated without knowing the probability of each category at location. Therefore, the computational process should be done through the sequential simulation algorithm. For this purpose the BLOCKSIS conventional GSLIB code (Deutsch, 2005) was modified to calculate entropy of system, exponential entropy and the size of the space of uncertainty for the simulation domain. The modified BLOCKSIS code performs all the calculations in this order:

1. Sequentially goes through a random path and pick a location
2. Compute p_k for $k = 1, ..., K$ using conditional data and previously simulated data
3. Calculate entropy and exponential entropy; $H = -\sum_{k=1}^{K} p_k ln(p_k)$, e^H
4. Go to the next location, repeat step 2-4 until calculate exponential entropy for entire

simulation domain

5. Calculate size of the space of uncertainty for simulation domain by Equation 4.10

The effects of different parameters on the extension of space of uncertainty will be discussed below.

4.3.1 Unequal Proportion

Size of the space of uncertainty is defined as the product of exponential entropy values. The maximum space of uncertainty is obtained in case of having maximum entropy of configuration. The maximum entropy of system occurs when all categories are equally probable. Therefore, any case of unequal probability that reduces the entropy of system leads to reduction of size of the space of uncertainty. The entropy of system and corresponding size of the space of uncertainty for different proportions are computed through the implemented algorithm on section 4.3 for the case study with two categories $K = 2$ over fifty locations ($N = 50, \Delta Z = 1$). Table 4.2 shows the computed entropy and size of space of uncertainty over the different proportions for this case through the implemented algorithm. Figure 4.3 also shows the entropy and logarithm of size of the space of uncertainty versus proportion respectively. As shown in this figure, the minimum entropy and size of uncertainty obtained in presence of only one category and these parameters increase by closing to the more equal categorical proportion. Maximum entropy and size of uncertainty obtained for equal proportions ($p_1 = p_2 = 0.5$).

4.3.2 Spatial Correlation

Spatial correlation is another parameter that reduces the size of the space of uncertainty. Spatial continuity is generally defined by variogram range and it is an indicator of data dependency. Therefore, a very small variogram range compared to the extension of simulation domain shows completely an independent simulation case and increases the size of the space of uncertainty. By increasing the range of variogram compare to the

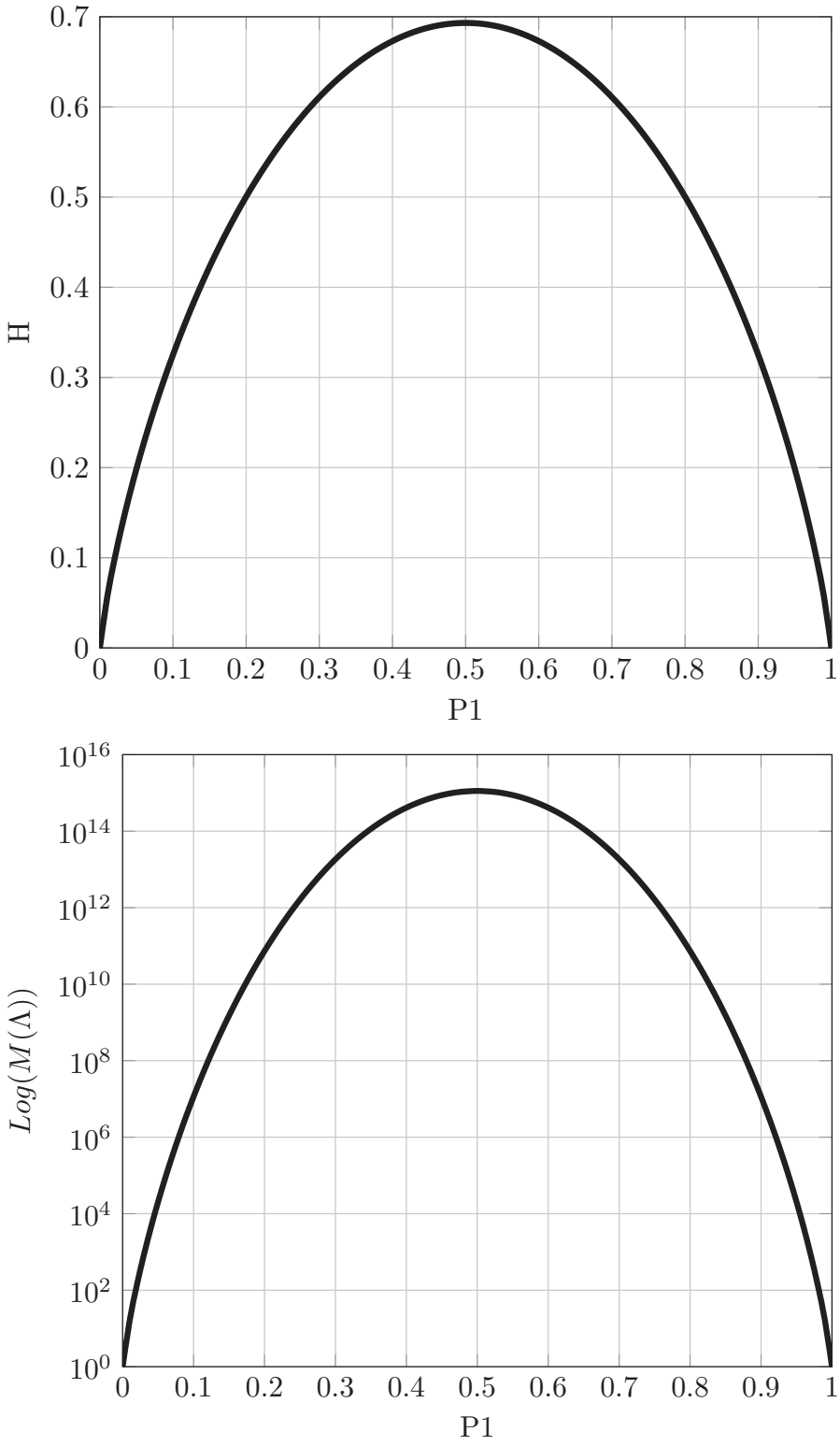

Figure 4.3: shows entropy (H) VS proportion (P_1), the top one and $Log(M(\Lambda))$ VS proportion the bottom one.

Table 4.2: Entropy and size of space of uncertainty for the case study ($N = 50, \Delta Z = 1$) different proportions, results are displayed in Figure 4.3.

P_1	P_2	H	$M(\lambda)$
0	1	0	1
0.1	0.9	0.3251	11457426
0.2	0.8	0.5004	7.35×10^{10}
0.3	0.7	0.6109	1.84×10^{13}
0.4	0.6	0.6730	4.11×10^{14}
0.5	0.5	0.6931	1.13×10^{15}
0.6	0.4	0.6730	4.11×10^{14}
0.7	0.3	0.6109	1.84×10^{13}
0.8	0.2	0.5004	7.35×10^{10}
0.9	0.1	0.3251	11457426
1	0	0	1

extension of simulation domain, as more information is provided over the simulation domain, the level of uncertainty decreases and size of the space of uncertainty reduces. Figure 4.4 shows different variogram ranges that applied to calculate size of the space of uncertainty for the case study with two categories. Figure 4.5 also shows logarithm of size of the space of uncertainty versus the corresponding variogram ranges that computed through the implemented algorithm. As shown in this figure, for variogram range $a = 1$ which is very small compared to the extension of simulation domain ($N = 50, \Delta Z = 1$) size of the space of uncertainty is very large. By increasing the variogram range compared to the extension of simulation domain size of the space of uncertainty dramatically decreases and this reduction continues until variogram range reaches to the half of simulation domain ($a = 25$). The variogram ranges greater than half of simulation domain have little effect on size of the space of uncertainty.

4.3.3 Conditioning Data

Conditional data generally provide valuable information over simulation domain that reduce the level of uncertainty. Therefore, size of the space of uncertainty is decreased by more conditioning data. Distance of conditioning data from simulated nodes compare to the range of spatial continuity is the key parameter that governs size of the space of uncertainty. To investigate the effect of conditioning data on the size of space of uncertainty, let's consider

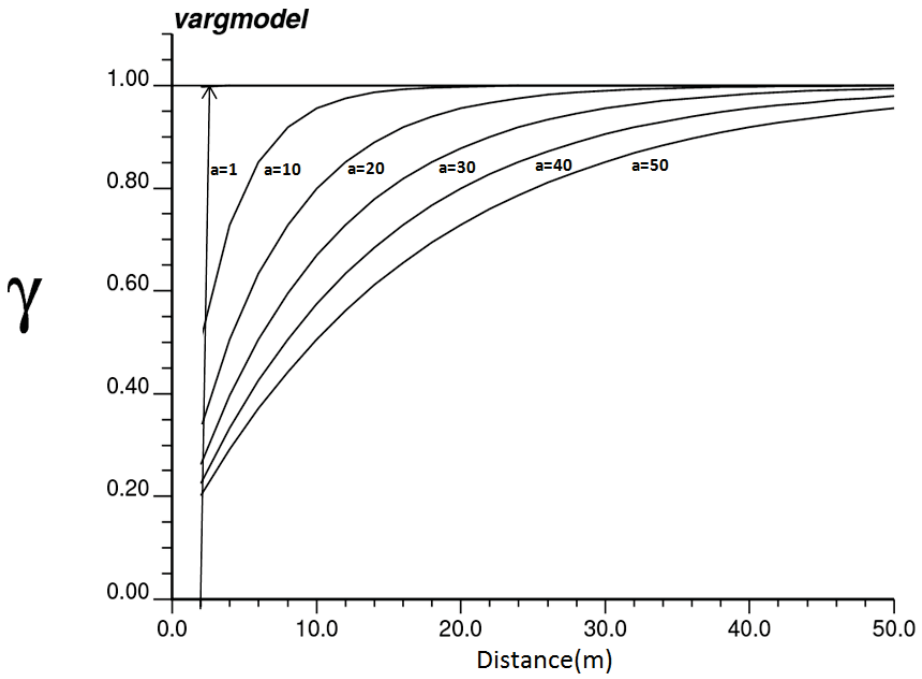

Figure 4.4: 2D omni-directional variogram model with different ranges, distance (m).

a 2D model of 50×50 nodes which size of space of uncertainty is calculated over one column ($N = 50$) while the position of the column of conditioning data changes over simulation domain through the implemented algorithm. Figure 4.6 schematically shows the 2D model and the distance of the column of conditional data from the simulated column compare to variogram range. Decreasing the ratio of this distance to the variogram range directly reduces size of the space of uncertainty. Figure 4.7 shows the logarithm of size of the space of uncertainty versus this ratio is computed via the implemented algorithm. Minimum size of the space of uncertainty obtained when the column of conditioning data is located adjacent the simulation column. The maximum one happened when the distance of conditioning data to the simulated column is half of variogram range (ratio = 0.5). The ratio greater than 0.5 doesn't affect the size of space of uncertainty.

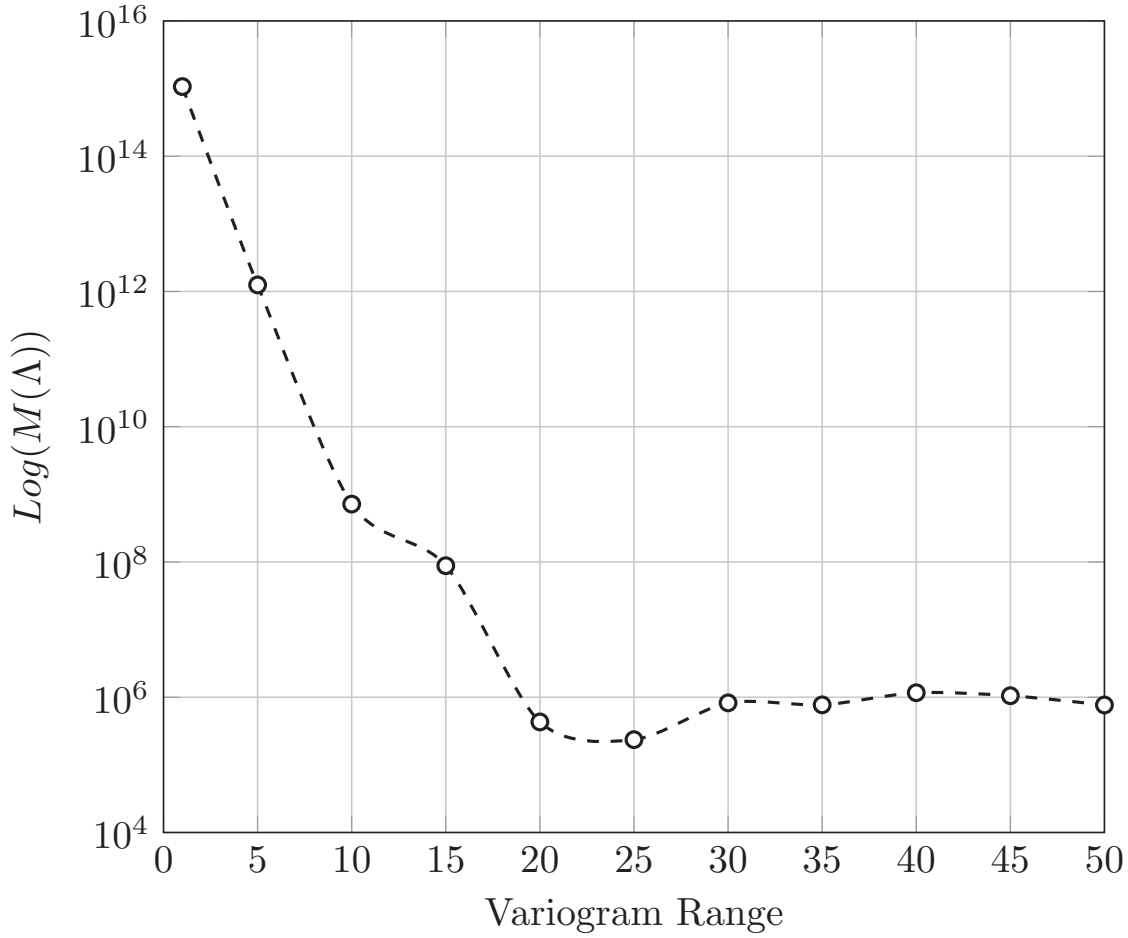

Figure 4.5: Size of the space of uncertainty versus variogram range (m).

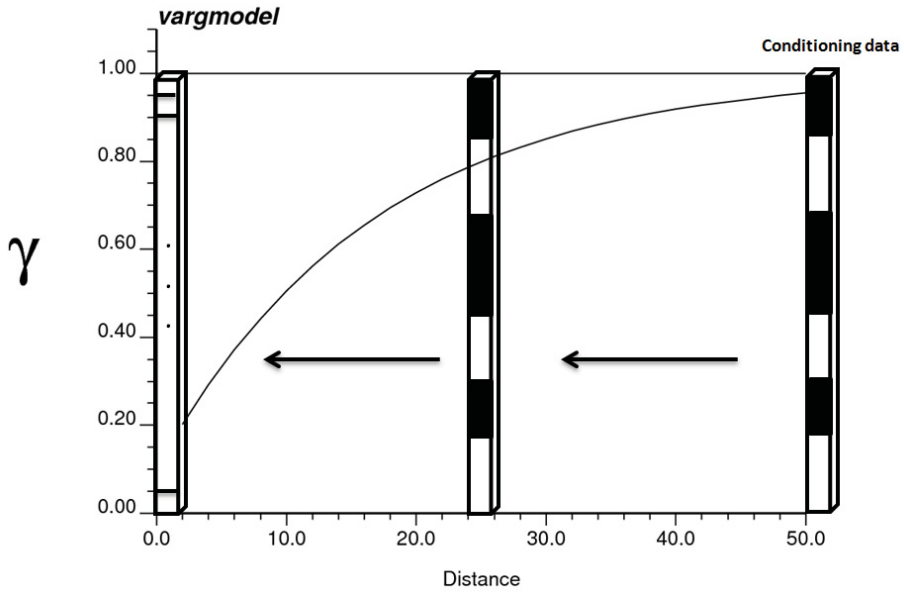

Figure 4.6: Sketch shows distance of conditioning data compare to variogram range.

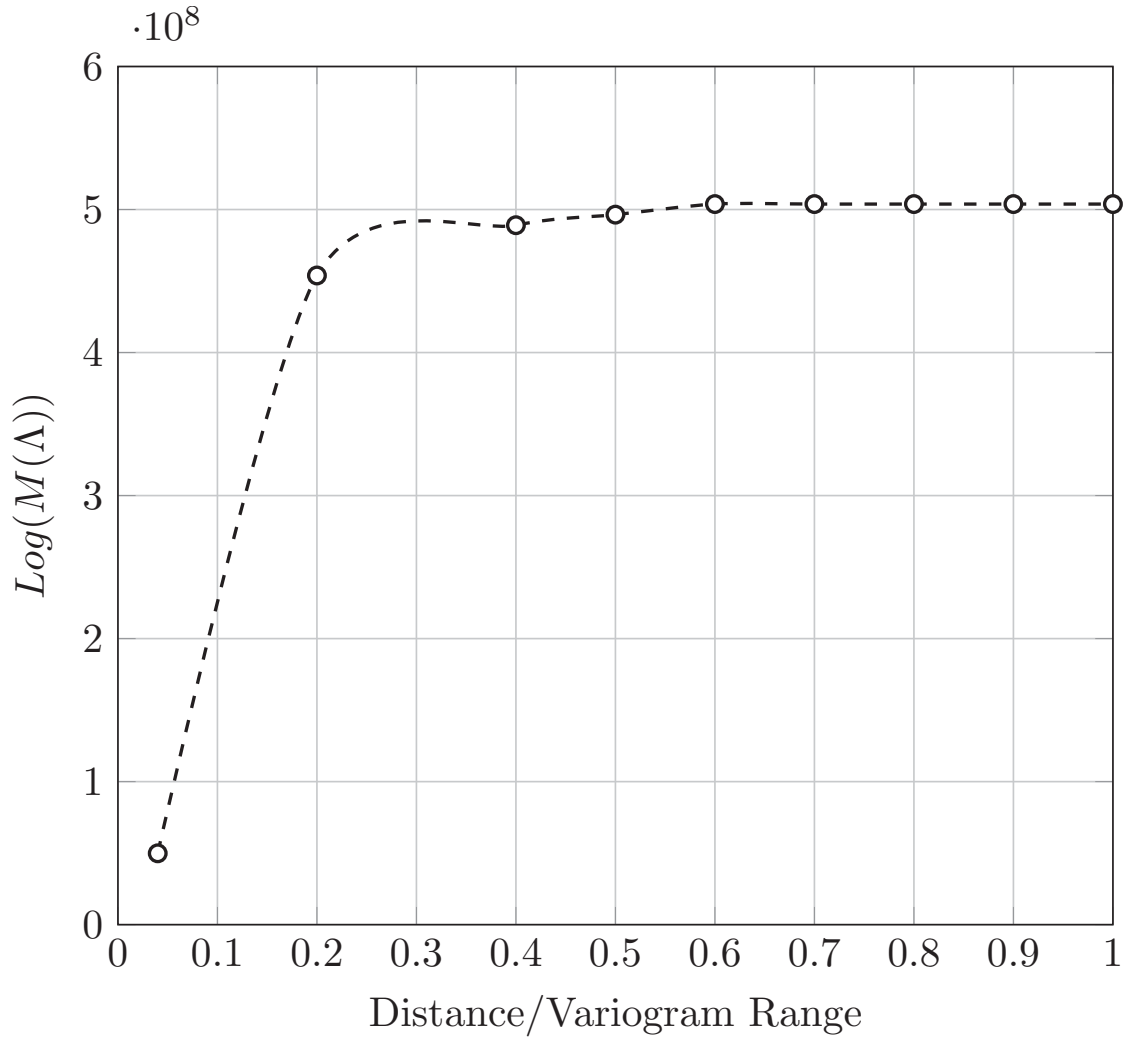

Figure 4.7: Size of the space of uncertainty versus ratio of distance to variogram range.

4.4 Discussion

As previously discussed, the maximum possible size of the space of uncertainty is immensely large and cannot be understood from a practical perspective. However, this space becomes much smaller and calculable in presence of different factors. Let's consider the size of the space of uncertainty for K categories over N locations in presence of total randomness with equal probabilities as A_0 which is the maximum size of space of uncertainty and equal to K^N. Applying more limitation to approximate proportion, size of the space of uncertainty reduces to A_1. Further limitation to approximate spatial correlation ($\gamma(h)$), size of the space of uncertainty becomes smaller (A_2). Additional restriction to approximate conditioning data, size of the space of uncertainty becomes very small, A_3. Figure 4.8 schematically shows the size of space of uncertainty in presence of different factors.

The basic promise of the algorithms that use acceptance-rejection sampling approaches to condition geostatistical models such as stochastic inversion is to sample A_3 and then reject the results, if it is not in the target zone, A_4. In stochastic inversion algorithms, the size of target zone (A_4) is defined by seismic data.

Multivariate stochastic inversion approach applies geostatistical modeling as part of stochastic inversion algorithm to provide high resolution reservoir models by conditioning them directly to the seismic data. For this purpose, the synthetic seismogram is computed based on simulated reservoir physical and elastic properties then compared with collocated actual seismic trace. This procedure is repeated by generating multiple realizations until reach to the acceptable match to the original seismic data. The main purpose of quantifying the size of the space of uncertainty in this thesis is to compute size of the space of uncertainty in presence of different parameters such as unequal proportion, spatial correlation and conditioning data (A_3) and then obtain the reasonable number of realizations at each simulation iteration of multivariate stochastic seismic inversion based on the extension of space of uncertainty. It would be an interesting idea instead of running constant number of realizations per column through the simulation process, define the

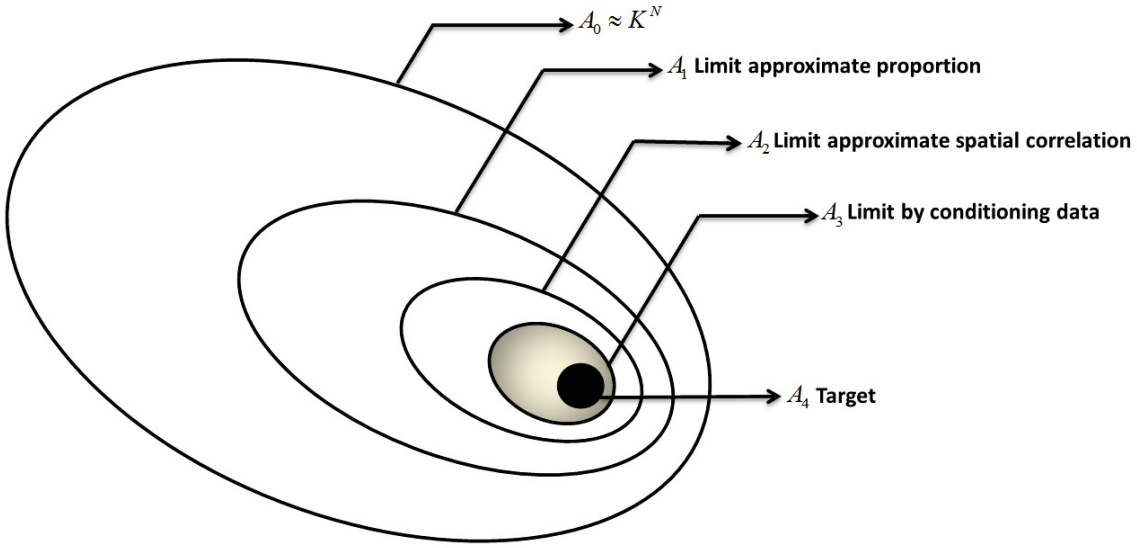

Figure 4.8: Schematically displays the size of space of uncertainty by applying some limitations.

number of realization according to the extension of space of uncertainty at each step of simulation process in multivariate stochastic inversion.

One of the parameters that reduces the size of space of uncertainty is conditioning data. More conditioning data over simulation domain leads to less space of uncertainty on the domain. Figure 4.9 shows this concept schematically. The first 10% to 20% of simulation domain (first dash line) is critically important. For this part of domain, as less conditioning data is provided over the simulation domain then the size of space of uncertainty is large and consequently more number of realizations is required to reach to the truth. The next 30% of simulation domain (between the dash lines) is mildly important because more conditioning data is provided and size of the space of uncertainty is gradually decreased. It is inevitable, when half of domain (50%, second dash line) is simulated then for the rest of domain there is at least one node beside them which is already informed. Therefore, for this part of domain size of the space of uncertainty reduces rapidly and number of realizations dramatically drops.

To demonstrate this logic, let's consider a numerical example with the previous case study

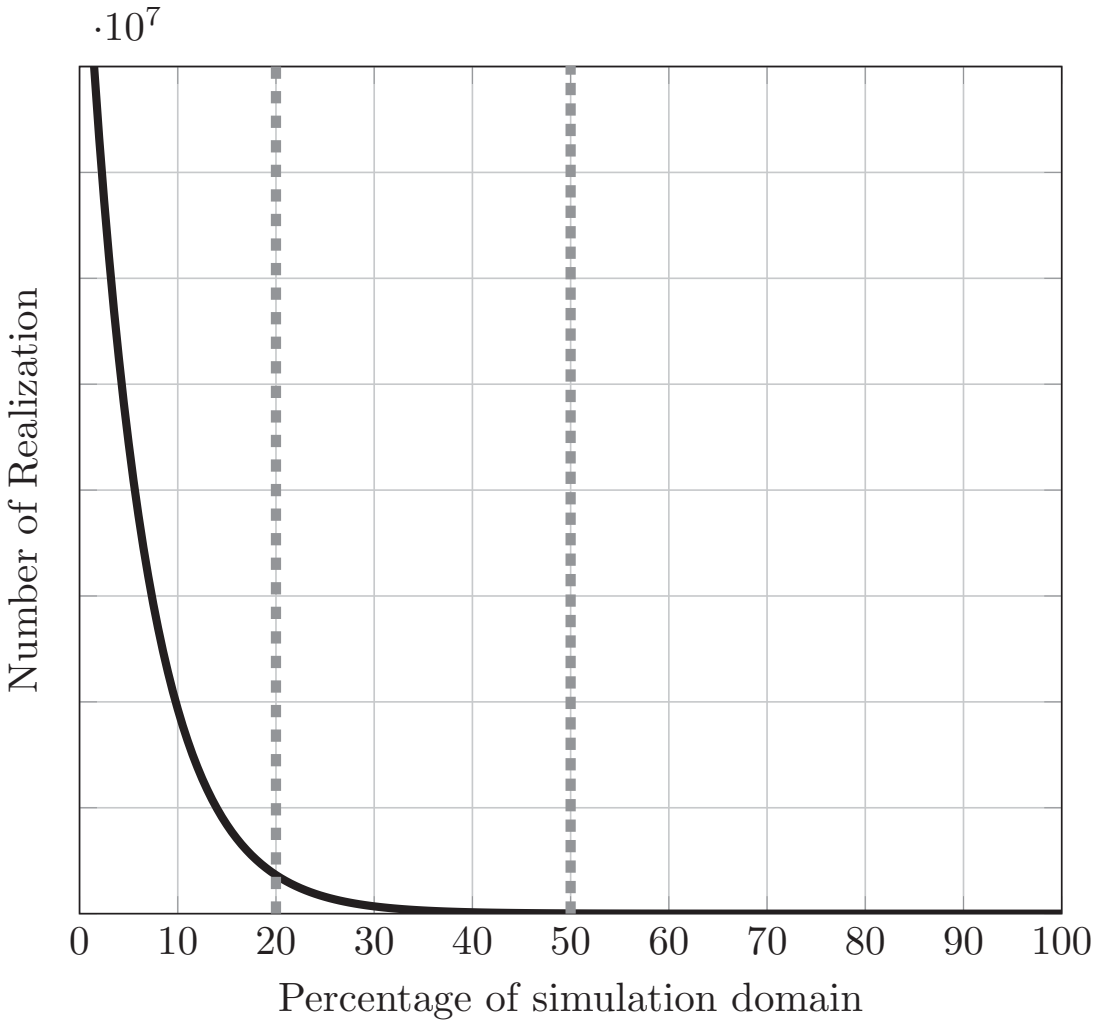

Figure 4.9: Schematically displays number of realizations versus size of domain and how this number is large for the first 20% of domain and gradually drops by going through the simulation process as more conditioning data is available.

as a simulation domain. We aim to simulate this case study under different scenarios to obtain the match to a truth. Figure 4.10 schematically shows the simulation domain and the truth which is a fully informed column with two categories ($K = 2$) over fifty location ($N = 50$). The first scenario is an unconditional simulation which there is no conditioning data over the simulation domain. In this case, size of space of uncertainty is maximum (2^{50}) and the possibility to obtain the match to the truth among few hundreds realizations is very small. Therefore, we need to generate multiple realizations in order of 10^{15} to obtain a perfect match to the truth which is huge and impractical. For the next scenarios, gradually

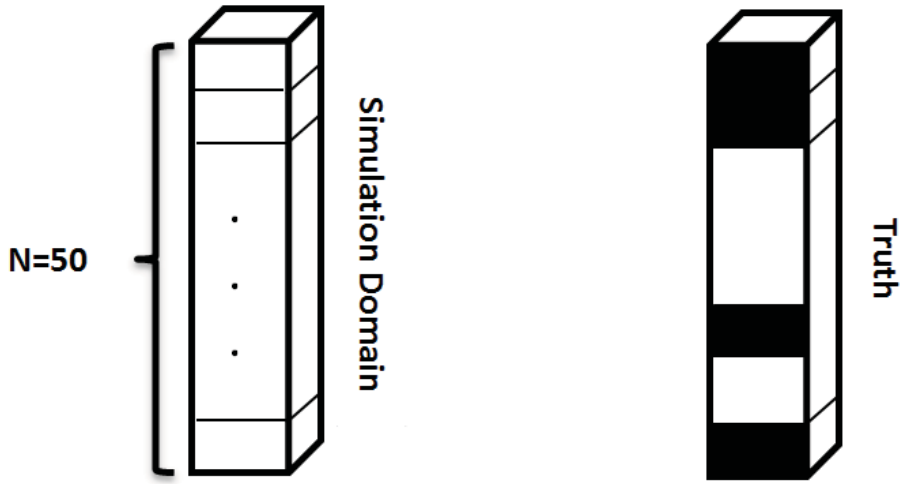

Figure 4.10: Schematically displays a column of fifty location as simulation domain and the truth with two categories.

conditioning data is added to the domain and size of the space of uncertainty is calculated by implemented algorithm. It is trivial, more conditioning data provide more information over simulation domain which reduces the size of space of uncertainty and consequently the number of realizations. When half of the domain is already informed (25 conditioning data) size of the space of uncertainty dramatically drops to order of 10^3 and by reasonable number of realizations the perfect match to the truth cab be obtained.

In another example to demonstrate this logic, consider a 2D case study (20×20) as simulation domain and also consider a fully informed 20×20 domain with two categories as a truth. In this example, a full column of the 2D domain ($N = 20$) is simulated through a random path at each simulation step. Multiple realizations are generated per column and compared with corresponded column of the truth until reach to the acceptable match. The acceptable match is the one with maximum 10% mismatch with the truth. Figure 4.11 shows the number of realization over the simulation path for the 2D case. As shown in this figure, for the first portion of simulation domain (first dash line) the number of realizations to obtain acceptable match to the truth is large. Then, for the second portion of simulation domain (between dash lines) number of realizations quickly drops to around 1000. While the number of realizations for the last part of simulation domain even

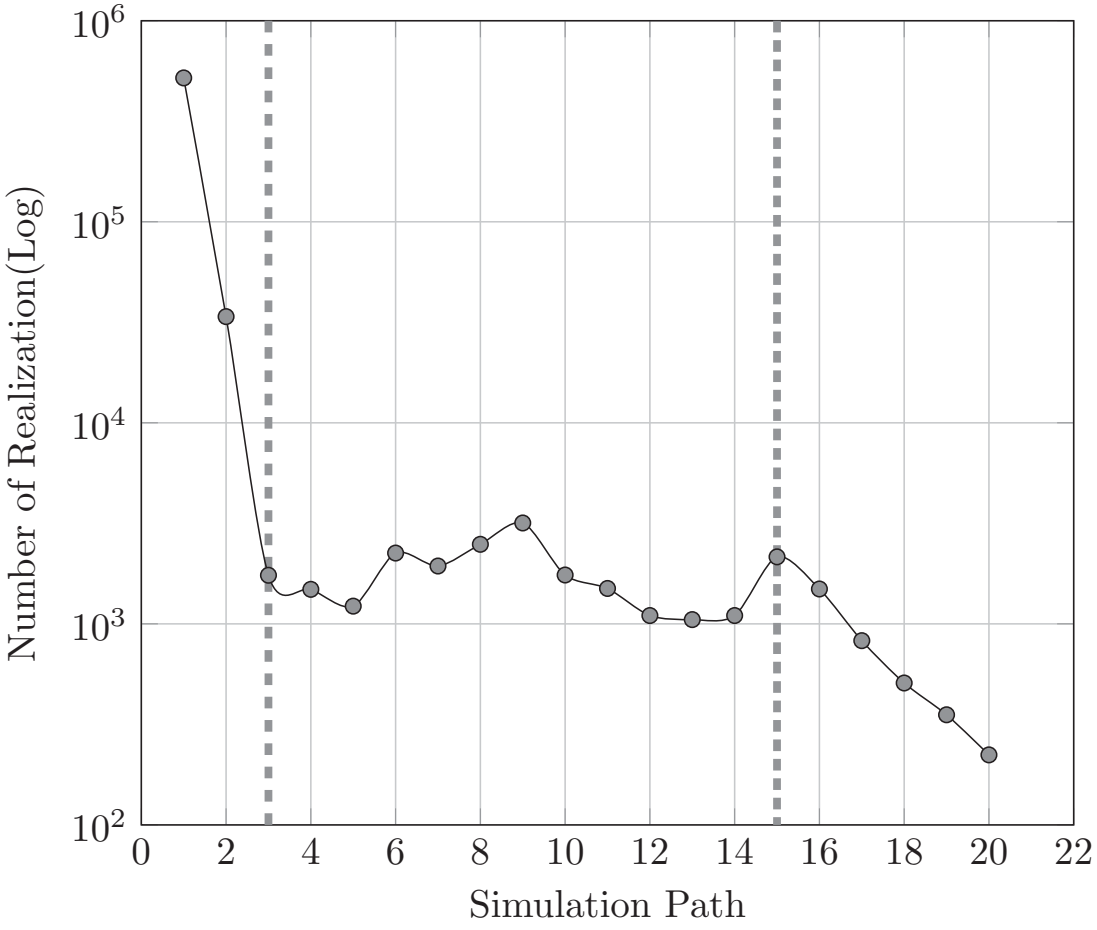

Figure 4.11: Number of realizations over the simulation path.

becomes smaller by around 100 realizations. In hence, we can claim this concept in multivariate stochastic inversion process and apply dynamic number of realizations for simulation process. More details is discussed in Chapter 6.

4.5 Chapter Summary

Size of space of uncertainty is only considered for categorical variables. Continuous variables are not considered because they can take a value continuously between over a range of values. The number of values that can be taken depends on the number of significant digits that are preserved. The best approach would be to consider classes of the continuous variable and, perhaps, take the limit as the number of classes approach a large

number. A reasonable expression is presented to compute the space of uncertainty in categorical variable modeling. The size space of uncertainty is the product of exponential entropy of the configuration. This size in case of totally random with no correlation is very large. However, size of the space of uncertainty becomes minimum in case of completely deterministic. It means size of the space of uncertainty decreases by adding more information over the configuration. Size of the space of uncertainty in some sense provides number of realizations over the simulation domain to reach the truth.

Chapter 5

Column Based Facies Modeling

Subsurface geological models involves modeling of facies or rock type within stratigraphic zones. Facies are important in reservoir modeling because reservoir physical properties such as porosity, permeability and fluid saturation are similar within facies and different between facies. The spatial structure of facies is different from continuous variables; the structure is usually better defined due to larger scale geological controls. Facies are generally categorized into discrete integer codes for the purpose of modeling. Facies can be defined as indicators or transformed values of zero or one corresponding to the K different facies values:

$$I(z(\mathbf{u});k) = \begin{cases} 1, & \text{if} \quad z(\mathbf{u}) = k \\ 0, & \text{otherwise} \end{cases} \tag{5.1}$$

Indicator values are exclusive and exhaustive which means each location only has one non-zero indicator values;

$$\sum_{k=1}^{K} I(z(\mathbf{u});k) = 1 \tag{5.2}$$

$$I(z(\mathbf{u});k) \cdot I(z(\mathbf{u});k^{'}) = 0 \quad k \neq k^{'} \tag{5.3}$$

There are three broad classes of facies modeling; object based modeling, variogram based methods and multiple-point statistics (MPS) that will be explained below.

5.1 Object Based Modeling

Object based simulation algorithms aim to model geological objects with their geometries at the same time (Haldorsen et al., 1984; Deutsch and Wang, 1996). Object based approaches distribute geometric objects in the simulation domain based on some probability rules by using marked point processes. Conditioning the simulation results to the well data is often a trial-and-error approach by moving objects around, adding/removing objects and modifying the size of objects. The main advantage of the object based framework is the preservation of geological features and direct control on the shape, size and orientation of these features (Honarkhah, 2011). In terms of conditional simulation, it is hard to condition object based models to the local data particularly when there are many well data or the geological objects are large. Considering seismic or production data as conditioning data makes the simulation more challenging. There is no explicit control of spatial correlation with object based modeling (Galli et al., 2006). Different objects require custom algorithms to be able model complex geological features.

5.2 Variogram Based Methods

variogram based algorithms easily account for a variety of data such as well, seismic and production data. Variogram or cell based simulation algorithms based on two-point statistics (variogram models) include Sequential Indicator Simulation (SIS) and Truncated PluriGaussian Simulation (TPGS). The spatial continuity of the facies could be measured by the indicator variograms:

$$2.\gamma(k;\mathbf{h}) = E\left\{[I(z(\mathbf{u});k) - I(z(\mathbf{u}+\mathbf{h});k)]^2\right\} \qquad k = 1, ..., K \tag{5.4}$$

Sequential Indicator Simulation (SIS) was developed by Alabert (1987) and Journel (1988) to simulate categorical variables. SIS considers the set of indicator variograms and draws the categories from estimated local probability distributions to build a categorical

image (Journel, 1988 ; Alabert et al., 1990). Indicator simulation techniques are able to consider various types of data and different indicator variograms for each facies (Zhu and Journel, 1993 ; Deutsch and Journel, 1998). The simulation process visits each grid node sequentially in a random path. The conditional distribution at each location considers well data and previously simulated values. Once the conditional distribution is constructed, a simulated value for the facies is drawn from the set of probabilities. The SIS method does not reproduce the cross-correlations and transitions between the different facies (Galli et al., 2006). These limitations motivated Matheron et al. (1987) to propose a truncated Gaussian random function.

Truncated Gaussian simulation (TGS) was developed by Matheron et al. (1987) to simulate ordered categorical variables with locally varying proportions. Later, the algorithm was extended to pluriGaussian simulation by Galli et al. (1994). The multivariate stochastic inversion algorithm in this study considers this approach as one approach for facies modeling when simulating categorical and continuous variables simultaneously. This method and its application in the proposed stochastic inversion approach will be explained below.

5.2.1 Truncated PluriGaussian - TPG approach

In truncated Gaussian simulations (TGS), the facies are simulated indirectly via the simulation of a stationary Gaussian random function. The simulated Gaussian values are then transformed into facies variables by truncations. For example to simulate a facies with two categories, 1 and 2, first simulate a Gaussian variable and then if the simulated Gaussian value is lower than the specific number t_1 is facies category one (1) otherwise is category two (2). The value of t_1 is called a threshold or cut-off, see Figure 5.1.

To define thresholds in TG approach, let $\mathbf{u}$ be any location over the simulation domain, $z(\mathbf{u})$ be the indicators of facies and $y(\mathbf{u})$ the corresponding simulated Gaussian function at location $\mathbf{u}$. The threshold mathematically defined by:

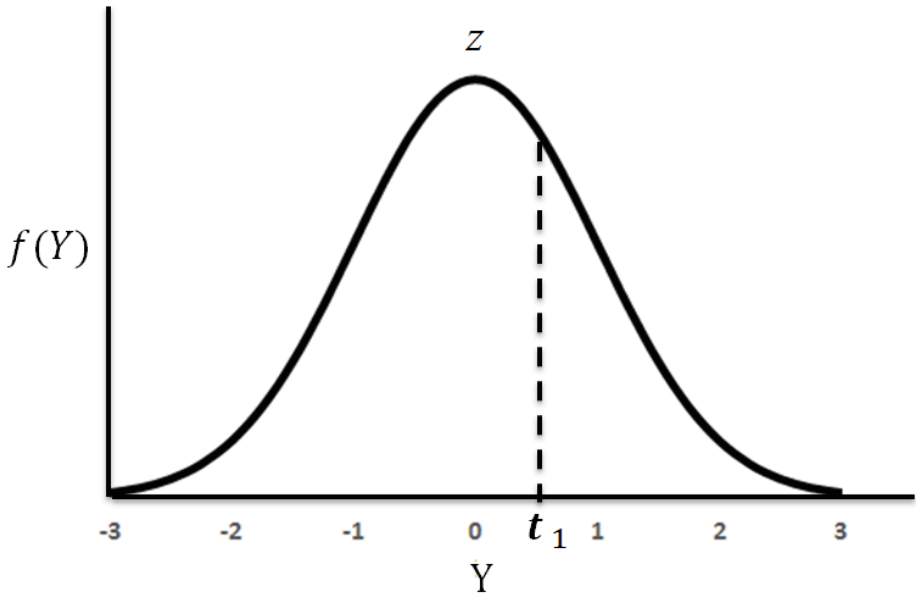

Figure 5.1: Schematic illustration of truncation of Gaussian probability density function

$$z(\mathbf{u}) = 1 \quad \Longleftrightarrow \quad -\infty \leq y(\mathbf{u}) < t_1 \tag{5.5}$$

In general, for more than two facies the thresholds defined by:

$$\mathbf{u} \in z_k \quad (k = 1, ..., K) \quad \Longleftrightarrow z(\mathbf{u}) = k \quad \Longleftrightarrow t_{k-1} \leq y(\mathbf{u}) < t_k \tag{5.6}$$

that $t_0 = -\infty$ and $t_K = +\infty$. For K possible facies, there are $K-1$ thresholds in increasing order that lead to facies that are exclusive and exhaustive:

$$t_1 \leq t_2 \leq \cdots \leq t_{k-1} \leq t_k \leq t_{k+1} \leq \ldots t_{K-1} \tag{5.7}$$

Although, the truncated Gaussian approach reproduces the ordering between facies there is only one variogram of the Gaussian variable. Truncated pluriGaussian (TPG) simulation is a generalized form of truncated Gaussian approach that has more flexibility to overcome the limitations related to transitions and anisotropies in TG approach (Galli et al., 2006).

Truncated pluriGaussian (TPG) considers more than one Gaussian variable.

The idea is to use rules to transfer the multiple Gaussian values to facies types. The proportion and ordering of facies are controlled by truncation rules and the multivariate Gaussian distribution controlled by the variograms of the Gaussian variables (Deutsch and Deutsch, 2013).

In truncated pluriGaussian simulation the main and initial step is defining a model type that depends on the pattern of relations and transitions between facies. The two key parameters that controls the truncated pluriGaussian simulation results are the thresholds that define the truncation of Gaussian variables and variogram models of the underlying Gaussian variables. The proportion of facies categories, facies rules and the correlation between the Gaussian variables determine the thresholds. The facies rules are represented by a diagram that determines the contacts and transitions between facies, see Figure 5.2. Moreover, the mathematical relation between indicator variograms and the variogram of Gaussian variables defines suitable variogram models for the underlying Gaussian variables. In case of conditional truncated pluriGaussian simulation the conditioning data need to be transformed into the Gaussian values. The last step in truncated simulation approaches is to transform the simulated Gaussian values into facies categories by the rules and threshold.

Although variogram based methods are practical for geostatistical modeling, they cannot handle complex geological features. Multiple-Point Statistics (MPS) Simulation technique is an option to overcome these limitations and handle complex geological features with straightforward conditioning.

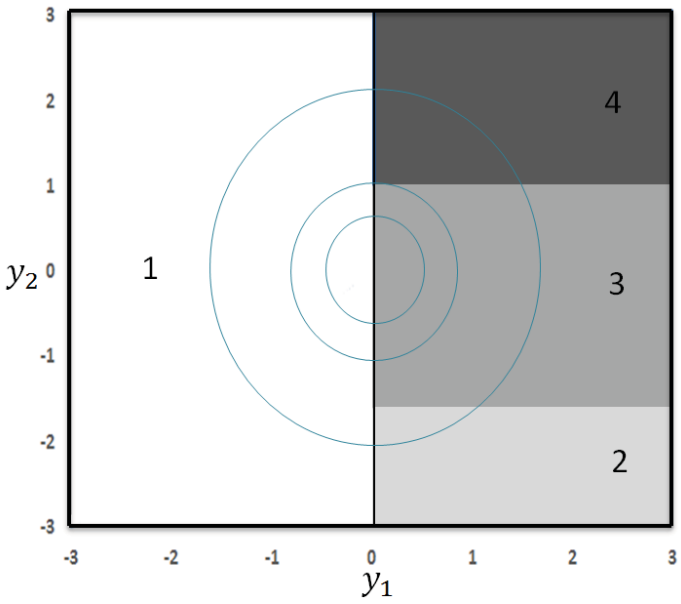

Figure 5.2: Schematic illustration of truncation of bivariate Gaussian probability density function, facies 2, 3, 4 are ordered while facies 1 crosses all other facies.

5.3 Multiple-point Geostatistical approach

The heterogeneity of subsurface reservoirs may be too complex to be captured by two-points statistics, or variogram based techniques. Multiple-points statistics (MPS) is an alternative to model complex and curvilinear geological features. The idea of using non-Gaussian models to consider spatial continuity in facies modeling was proposed by Alabert and Journel (1989). A multiple-point statistics approach (MPS) to model more complex geological features was proposed by Guardiano and Srivastava (1993). The multiple-point method accounts for a configuration of more than two points which helps reproduction of non-linear spatial continuity. Figure 5.3 shows the two points configuration that infer the statistics from variogram (left) and an example five point configuration illustrative of MPS method (right).

The frequency of different multiple-point configurations come from a conceptual training image. The conceptual image characterizes the geometry and complexity of geological features deemed relevant to the reservoir under consideration. The spatial relationships

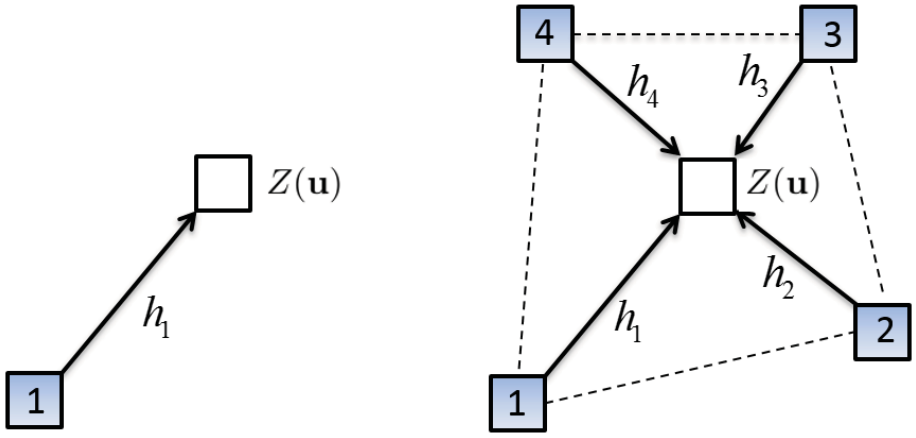

Figure 5.3: Schematic illustration of two point statistics (Left) and Multi Point Statistics (Right).

that come from a training image represent more information than variograms because they consider many points at the same time. Object based simulation algorithms are sometimes used for generating training images.

The original MPS algorithm of Guardiano and Srivastava (1993) applies the concept of multiple-point statistics by considering an $n + 1$ configuration template of $(0, \mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n)$, similar to the shown in Figure 5.3, but in 3D with more points. In the next step, the training image is scanned based on the template to compute the frequency of the specific $n + 1$ pattern. The n data are fixed and the unsampled location could take 1 to K outcomes with different frequency. The empirical conditional distribution of $Z(\mathbf{u})$ at grid node $\mathbf{u}$ can be obtained from these replicates. Then, a value can be sampled from this distribution and assigned to the simulation grid. The sequential approach used in this method to simulate the grid nodes is similar to variogram-based geostatistics (Honarkhah, 2011). The original MPS algorithm was impractical in terms of computational time because of scanning the entire training image for each grid node. Since then, many multiple-points statistics algorithms have been developed due to the importance of modeling complex geological features in reservoir management.

The first practical multiple-point statistics algorithm called ”single normal (extended) simulation (SNESIM)” was proposed by Strebelle (2000). To overcome the computational

limitations of the original approach, this method scans the training image once and stores probabilities in a search tree. The SNESIM algorithm follows a probabilistic approach similar to the original algorithm proposed by Guardiano and Srivastava (1993). In this approach the probability of each facies given the nearby conditioning multiple-points data event and the training image. The SNESIM approach will be explained below.

5.4 SNESIM Approach

As previously mentioned, the MPS method considers multiple-point configurations and infers spatial statistics from a training image to go beyond bivariate moments to handle more complex geological features. The SNESIM algorithm is a robust MPS technique that handles the computational issues of the original approach by the use of a search tree and multiple grid. The search tree is an effective dynamic algorithm that allows the fast retrieval of conditional probabilities during the simulation process. The search tree represents a chain of nodes that connect the multiple data events based on the location of the data in the template (Deutsch and Silva, 2014). A template is an arrangement of 2D or 3D nodes with regard to a central location that is used for estimation of conditional probability values in MPS modeling. The SNESIM algorithm works in a sequential manner where each grid node is simulated and added to the conditioning data for subsequent grid nodes. SNESIM algorithm considers multiple-point data events and simulates the categorical variables by inferring conditional probabilities from the training image. The multigrid approach is essential to reproduce large scale features. The main steps of the SNESIM algorithm are briefly explained below (Deutsch and Silva, 2014):

Build/Store Search Tree

The training image is scanned based on a maximum template size and the multiple-points data event frequencies stored in the search tree.

Set up Multigrid Simulation Path

All the conditioning data are relocated to the the closest grid node. The purpose of this relocation is computational convenience and to maximize the effect of the conditioning data on the final model.

Create a Pseudo-Random Path

Similar to sequential indicator simulation, all the nodes are simulated sequentially through a random path but in the manner where the most informative nodes are visited first. This means an informative factor is integrated to the random path generation which the most informative nodes are visited first without significant change in the randomness of the simulations. The process generates the random path by assigning a random value from a uniform distribution to all the nodes in the simulation domain. Based on the search parameters, all the nodes are then visited to count and store the number of conditioning data within the search ellipsoid. After visiting all nodes, the number of conditioning data in the neighborhood are added to the random values of the path. At the end, the pseudo-random path is obtained by sorting the updated random sequence in a descending order.

Select Data Event

In case the data event is not found in the search tree, the data event will be updated by dropping the most distant conditioning data.

Computation of Conditional Probability

Based on the proportion obtained from the search tree, the conditional probability of each facies at the unsampled central grid node is computed.

Draw Simulation Value

The facies is simulated by Monte Carlo Simulation from the conditional probabilities.

More details about the MPS approach can be found in geostatistical text books like "Geostatistical Reservoir Modeling" by Pyrcz and Deutsch (2014) and "Guide to MPS Simulation" by Deutsch and Silva (2014).

5.5 Multi-Grid Approach

The concept of a multi-grid was introduced by Gomez-Hernandez (1991). Later, Tran (1994) applied the concept in sequential simulation algorithms for dense simulation grids with long-range variograms. The concept of the multi-grid permits reproduction of large-scale features with small templates.

In multi-grid framework, the simulation domain (D) contains of a set of cascading grids, D^d. The cascading grids discretize the simulation domain into increasingly nested finer grids and vary from 1 to n_d ($1 \leq d \leq n_d$) where D^1 relates to the final simulation grid and n_d is the total number of cascading grids or multi-grids. The grid nodes of multi-grid d is separated by 2^{d-1} nodes of the finest grid which is the final simulation grid. An example of four nested multi-grids is displayed in Figure 5.4.

The template of each multi-grid is a scaled version of the original template and will be rescaled by vector of 2^{d-1}. For each multi-grid, all stages of the simulation process including scanning of the training image, computation of conditional probabilities and drawing simulation value of the nodes will be performed independently. The simulated nodes of previous multi-grid are considered as conditioning data for the next multi-grid. In this way, the large scale geological patterns condition the short scale ones to facilitate the reproduction of spatial geological features of the training image (Deutsch and Silva, 2014).

Facies modeling is required in the developed multivariate stochastic inversion approach.

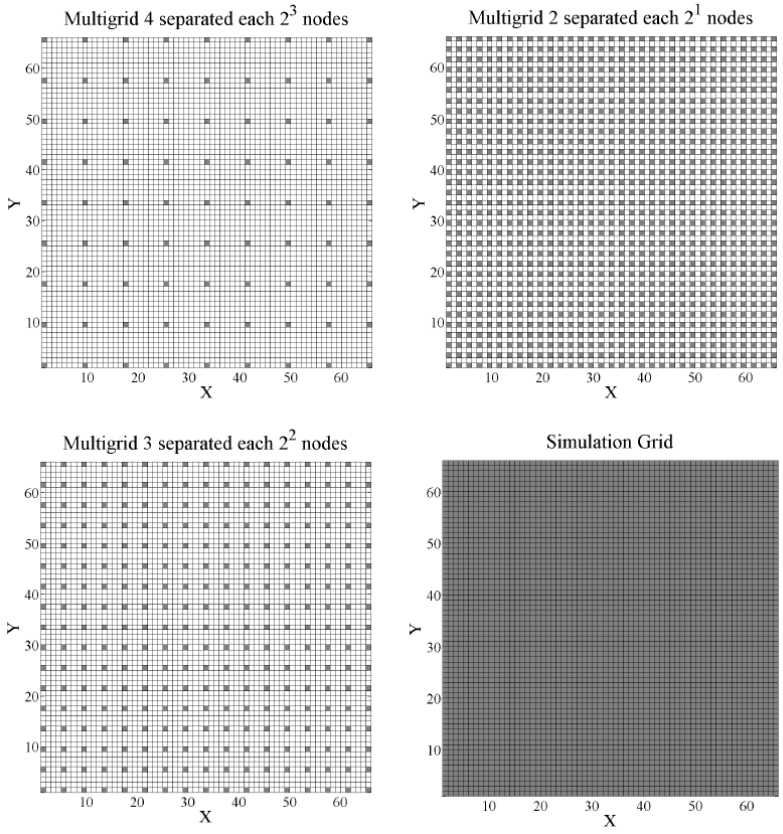

Figure 5.4: Illustration of four nested 2D multi-grid (Deutsch and Silva, 2014).

Seismic surveys consist of seismic traces that are the convolution of an extracted wavelet and a vertical column of corresponding reflectivity series. Therefore, the framework of multivariate stochastic inversion developed in this thesis is inherently column based. This means that facies and property modeling algorithms must be implemented in a column-wise manner instead of a node based order.

5.6 Column-Wise Facies Modeling

Conventional stochastic seismic inversion techniques initially provide acoustic impedance model that is conditioned to both well and seismic data. Later, the acoustic impedance model is related to reservoir properties through a rock physics models. Although, there is no guarantee that the physical properties obtained from such a model matches the actual seismic data (Bortoli, 1992 ; Bosch et al., 2010). To provide high quality reservoir models and overcome this limitation, the multivariate stochastic inversion approach simulates multiple reservoir properties (continuous and categorical) simultaneously as a part of stochastic inversion algorithm instead of relating acoustic impedance or facies models to reservoir physical properties through the sequential connection outside the stochastic inversion algorithm. In the multivariate stochastic inversion approach, the multiple variables are simulated via the multivariate Gaussian simulation technique. As seismic data depends on the full vertical column through convolution, the multivariate Gaussian simulation technique must proceed in a column-wise manner where multiple variables for a full column are simulated at each iteration.

The multivariate Gaussian technique simulates a full column of multiple Gaussian variables at the same time, denoted by $Z_i^l(\vec{\mathbf{u}})$, see Equation 3.1. The number of Gaussian variables are simulated by multivariate Gaussian simulation approach is defined by number of Gaussian variables is required to define the sequential order among the facies categories through the truncated pluriGaussian approach (n_f), number of reservoir physical properties or continuous variables (n_c) and the number of facies categories (K). In general, the n_f Gaussian variables simulated and turned to indicator vector of facies by

truncated pluriGaussian approach and each reservoir physical property is simulated per facies categories by sequential Gaussian simulation (Deutsch and Journel, 1998) approach. Based on the simulated facies value of each cell, the corresponding simulated continuous variables are retained. When there is a clear sequential order among the facies, only one Gaussian variable is required ($n_f = 1$) and facies are simulated by the truncated Gaussian approach. Therefore, the number of Gaussian variables are simulated by multivariate Gaussian technique drops to $1 + (K \cdot n_c)$. Figure 5.5 is a schematic of the correlation matrix in the multivariate Gaussian simulation approach. The matrix is a block matrix that represents three facies categories ($K = 3$) simulated by truncated Gaussian approach ($n_f = 1$) and n_c reservoir properties. The first column and top row are related to the facies. The second, third and forth diagonal blocks represent how the continuous variables are related to each other in the first, second and third facies category, respectively.

To add more flexibility to the multivariate stochastic inversion approach and be able to handle more complex and curvilinear geological features in reservoir models, the SNESIM approach is an alternative to truncated Gaussian (TG) or truncated pluriGaussian (TPG) simulation of facies. For this purpose the SNESIM algorithm is modified to perform column based MPS simulation. An issue for implementation of column wise SNESIM is the multi-grids approach.

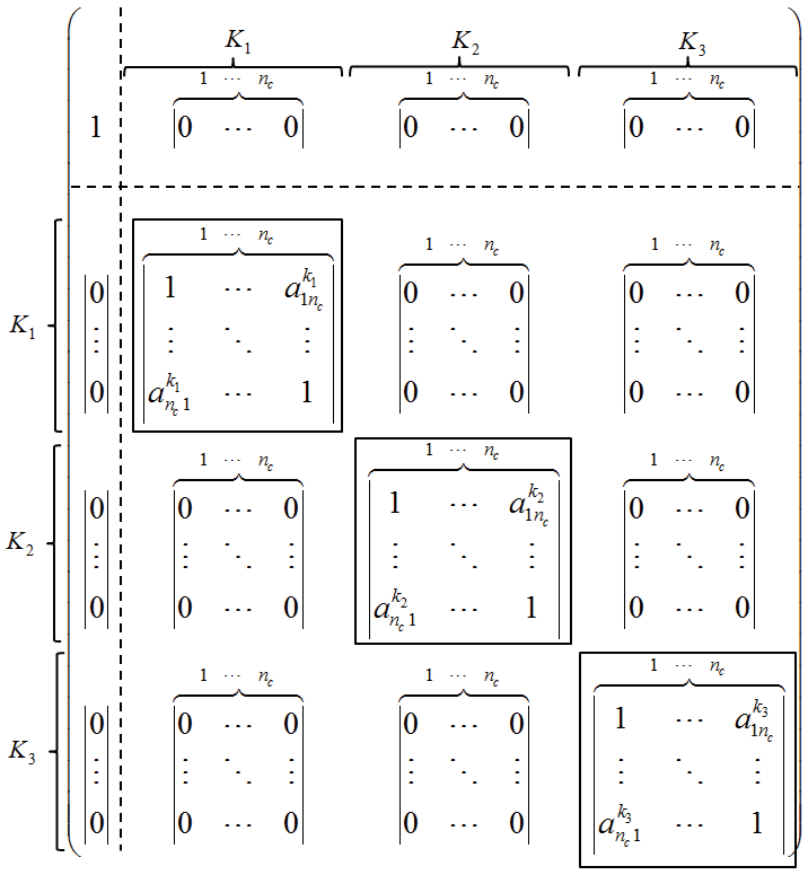

Figure 5.5: Correlation matrix of multivariate Gaussian simulation for a facies with three categories and n_c continuous variables.

5.7 Column-Wise SNESIM Algorithm

The nature of seismic data is inherently column based. Therefore, the framework of stochastic inversion developed in this thesis is column wise: the multivariate stochastic inversion approach simulates a full column of multiple reservoir properties (categorical and continuous variables) simultaneously with multivariate geostatistics techniques. The simulated column of multiple reservoir properties passed to a petro-elastic model to compute reservoir elastic properties and a seismic reflectivity series. Then, seismic reflectivity is convolved with the extracted wavelet to compute synthetic seismic traces and pick the best match to the corresponding actual seismic trace. A realization is ultimately chosen for subsequent conditioning. As a result, facies and property modeling

algorithms must be implemented in a column wise manner instead of a completely random node based order. For this purpose, the original SNESIM algorithm is modified to perform column wise multi point statistic facies modeling. However, the original multi-grid approach is not based on columns. Therefore, the original multi-grid approach was modified to be applied in column wise fashion. The modified multi-grid approach called directed multi-grid.

5.7.1 Directed Multi-Grid

A directed multi-grid performs the concept of multi-grid approach in different dimensions or directions individually to satisfy the requirement of column based MPS simulation for multivariate stochastic inversion. Two types of directed multi-grid can be considered for MPS simulation: 1) $X-Y$ multi-grid, full column of Z and 2) $X-Y$ multi-grid plus Z multi-grid.

5.7.1.1 $X-Y$ Multi-Grid, Full column of Z

In this approach, a $2-D$ multi-grid in $X-Y$ direction is applied and at every step of $X-Y$ multi-grid a full column of Z is simulated with a random path. For example consider a $3D$ simulation domain by $128 \times 128 \times 64$ ($n_x \times n_y \times n_z$) grid nodes and 6 nested finer grid ($d = 6, 1 < n_d < 6$). Based on this directed multi-grid approach, the grid nodes are simulated by the sequence of $(4 \times 4 \times 64), (8 \times 8 \times 64), \ldots, (128 \times 128 \times 64)$. In this approach reproduction of large scale geological features in Z direction may be an issue.

5.7.1.2 $X-Y$ Multi-Grid, Z Multi-Grid

This directed multi-grid approach is similar to the previous one. However, in this approach a $2-D$ multi-grid in $X-Y$ direction is applied and at every step of $X-Y$ multi-grid a full column of Z is simulated with another set of multi-grid. In fact, at one column there is a multi-grid that is embedded in a $2-D$ multi-grid. This means in a $3D$ grid of $128 \times 128 \times 64$ ($n_x \times n_y \times n_z$) by $n_{d_{X-Y}} = n_{d_Z} = 6$, for the first stage of the multi-grid in $X-Y$

direction which is (4×4), each full column is simulated by another set of multi-grid that is $(4, 8, 16, 32, 64)$. Then, it goes to the next level of multi-grid in $X - Y$ that is (8×8) and applies another set of multi-grid for each full columns. This procedure is replicated until reach to the last step of $X - Y$ multi-grid which is (128×128).

The SNESIM algorithm is modified to perform the column wise MPS simulation. Both directed multi-grid methods that explained above are available in column based SNESIM algorithm.

The column wise SNESIM algorithm consists of many independent modification steps including different multi-grid approaches. Therefore, model validation is necessary to make sure all concepts are applied and implemented correctly.

5.8 Model Checking

There are alternatives to check simulation results and judge different multi-grid approaches including, visual assessment, reproduction of lower order statistics such as the variogram and reproduction of higher order statistics such as the frequency of multiple point data events.

Visual assessment is qualitative, but very useful. To evaluate column wise SNESIM algorithm an example of $3D$ grid by $128 \times 128 \times 64$ ($n_x \times n_y \times n_z$) nodes with two categories is considered, category one (K_1) with 0.7 proportion ($P_1 = 0.7$) and category two (K_2) with 0.3 proportion ($P_2 = 0.3$). The XY and XZ slices of training image and node based SNESIM as a reference (Figure 5.6) compared with column based SNESIM with different multi-grid approaches; no multi-grid, $X - Y$ multi-grid-full Z and $X - Y$ multi-grid plus Z multi-grid (Figure 5.7). As shown in these figures, column based SNESIM with no multi-grid shows a very noisy result that emphasizes the importance of multi-grid in reproduction of large scale geological features in the MPS simulation process. The MPS simulation algorithm with $X - Y$ multi-grid and full column of Z approach shows a better result on $X - Y$ slice while the Z slice is noisy. Although, the

MPS with $X - Y$ multi-grid plus Z multi-grid represent better reproduction of large scale geological features in three dimensions.

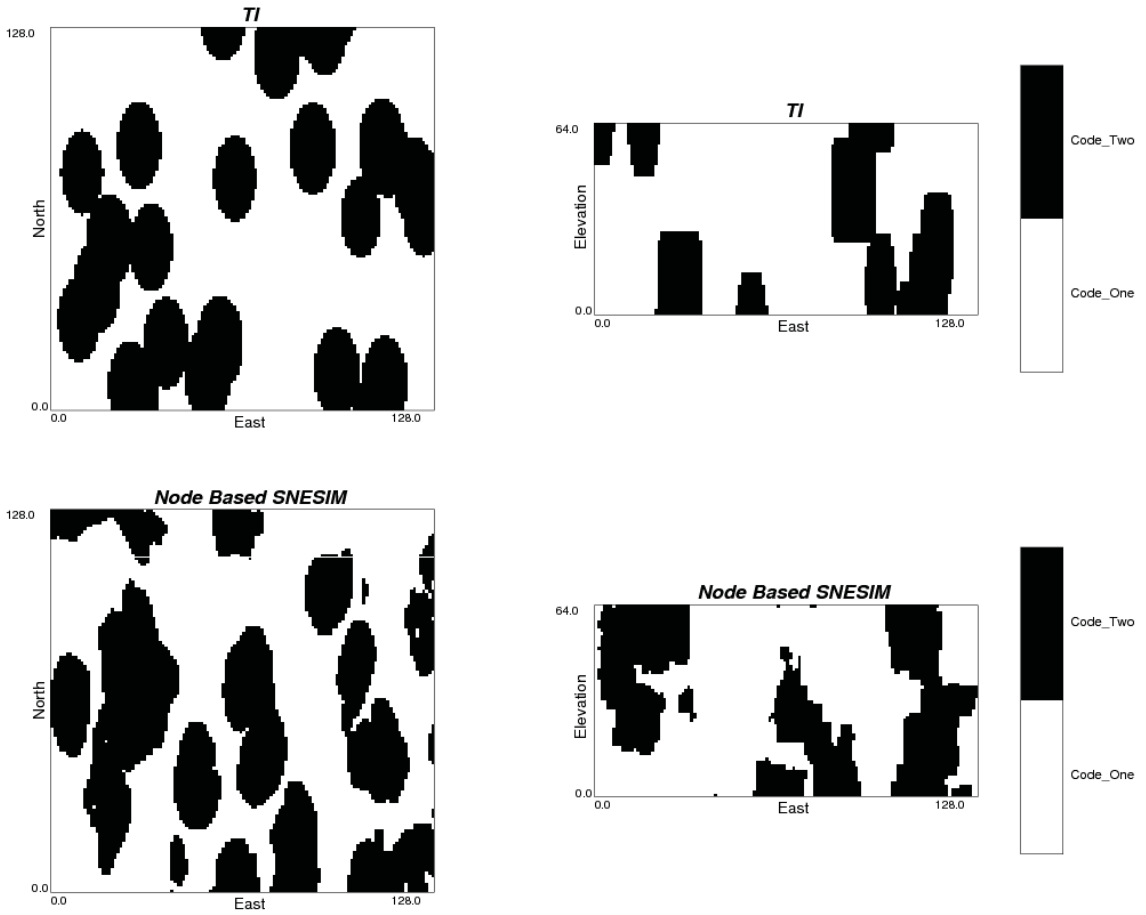

Figure 5.6: XY (left) and XZ (right) slices of training image and node based MPS simulation with original multigrid approach, axes unit (m).

Visual assessment is the first step for quality control and model checking. Using the variogram to check MPS results is not convincing since the variogram is deemed inadequate in the first place.

Another option for checking the MPS simulation results is comparing the frequency of some multiple point patterns. If N (10 to 20) points are chosen in a reasonable configuration- for K categories, there are K^N configurations. A compact configuration is considered to be the most informative. The patterns can be indexed by $i = 1, \ldots, K^N$ and the proportion

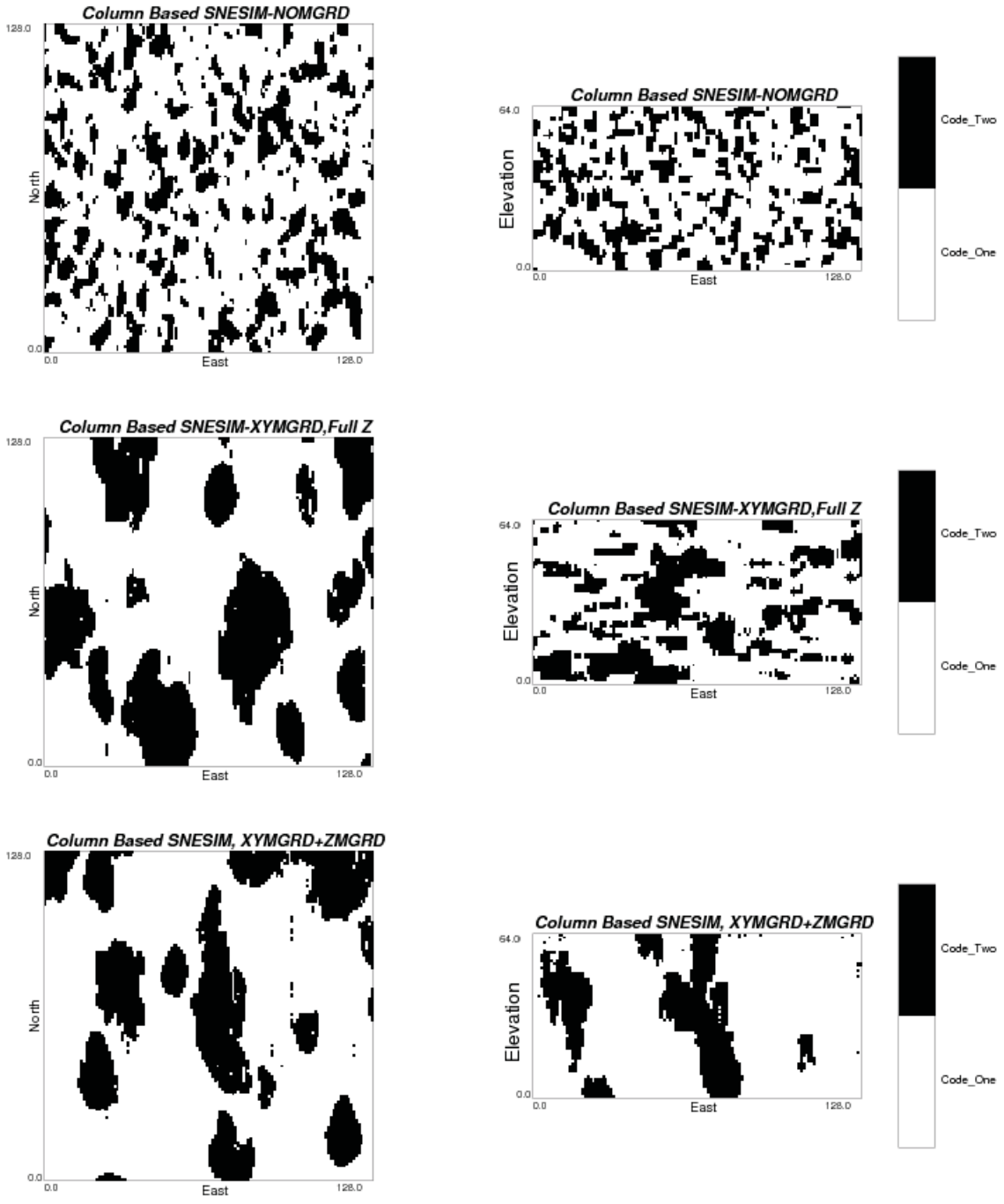

Figure 5.7: XY (left) and XZ (right) slice for column based with no multi-grids and column based with different type of multi-grids SNESIM, axes unit (m).

or frequency of each configuration denoted by P_i. The multiple point patterns should be reproduced better with the multigrid.

A configuration by $3 \times 3 \times 2$ nodes ($N = 18$ locations) is considered and the proportion calculated and compared. When the frequency of the patterns (P_i) is computed, the first and the last one are pure facies with no mixing (either 0 or 1 category) but the rest are different mixed patterns (both 0 and 1 category). By keeping all the patterns together, pure and mixed one, the results will be dominated by the pure ones since they have a much higher proportion. Therefore, they will be kept separate and the Mean-Square Error (MSE) computed for the others, Equation 5.8, where n is the number of patterns, P_i^* is the frequency of patterns in the realization being checked and P_i^{Ref} is the frequency of patterns in the reference training image.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i^* - P_i^{Ref})^2 \tag{5.8}$$

One option is to consider a relative of Mean-Square-Error, MSE_{Rel} (Equation 5.9). However, zero frequency patterns cause numerical instability. Zero proportion features may be implausible and could be important. An observed zero may also be due to a small simulation domain. The only way to know the difference between real and relative zero is with a very large model.

$$MSE_{Rel} = \frac{1}{n}\sum_{i=1}^{n}(\frac{P_i^* - P_i^{Ref}}{P_i^{Ref}})^2 \tag{5.9}$$

Figure 5.8 to Figure 5.11 show the frequency of multiple point data events of a training image versus node based and column based SNESIM with different multi-grid approaches. As these charts show, the simulation results for node based SNESIM is similar to the training image data (close to 45° line). The SNESIM with directed multi-grid approaches show better results compare to the SNESIM without multi-grid and also their trends are closer to the node based SNESIM. Table 5.1 also shows the MSE of pure and mixed data events. These results demonstrate that the column wise SNESIM with $X - Y$ multi-grids plus Z multi-grid approach works better than $X - Y$ multi-grid and full column of Z approach and they both represent better reproduction of large scale geological features compare to

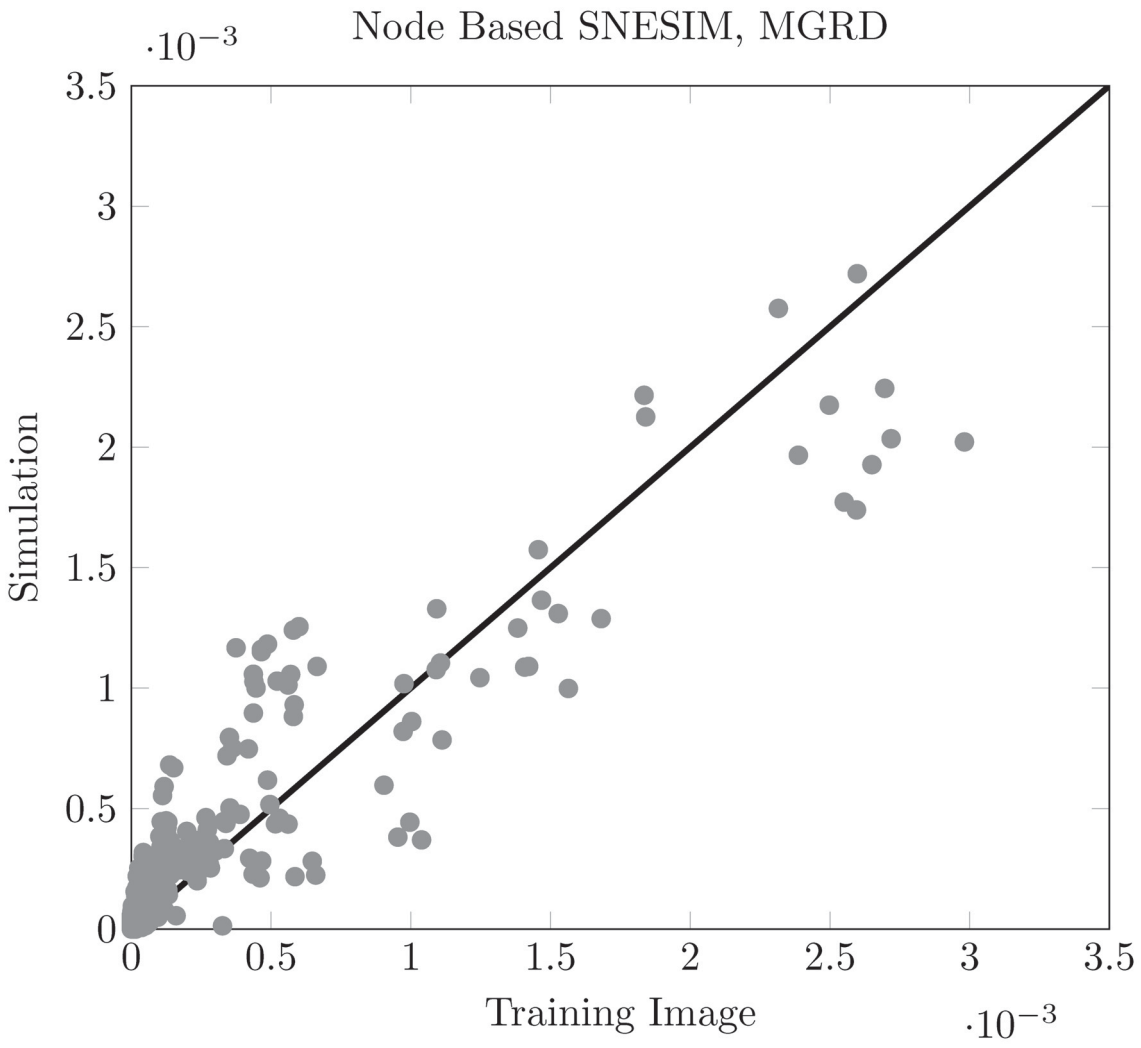

Figure 5.8: Frequency of patterns of training image versus simulation results for node based SNESIM with original multi-grid approach.

the SNESIM without multi-grid.

Figure 5.12 and Figure 5.13 show another example of column based SNESIM with different multi-grid approaches for three categories. As shown in these figures, the column based SNESIM with directed multi-grid shows better results that are closer to the node based full multigrid SNESIM than no multi-grid.

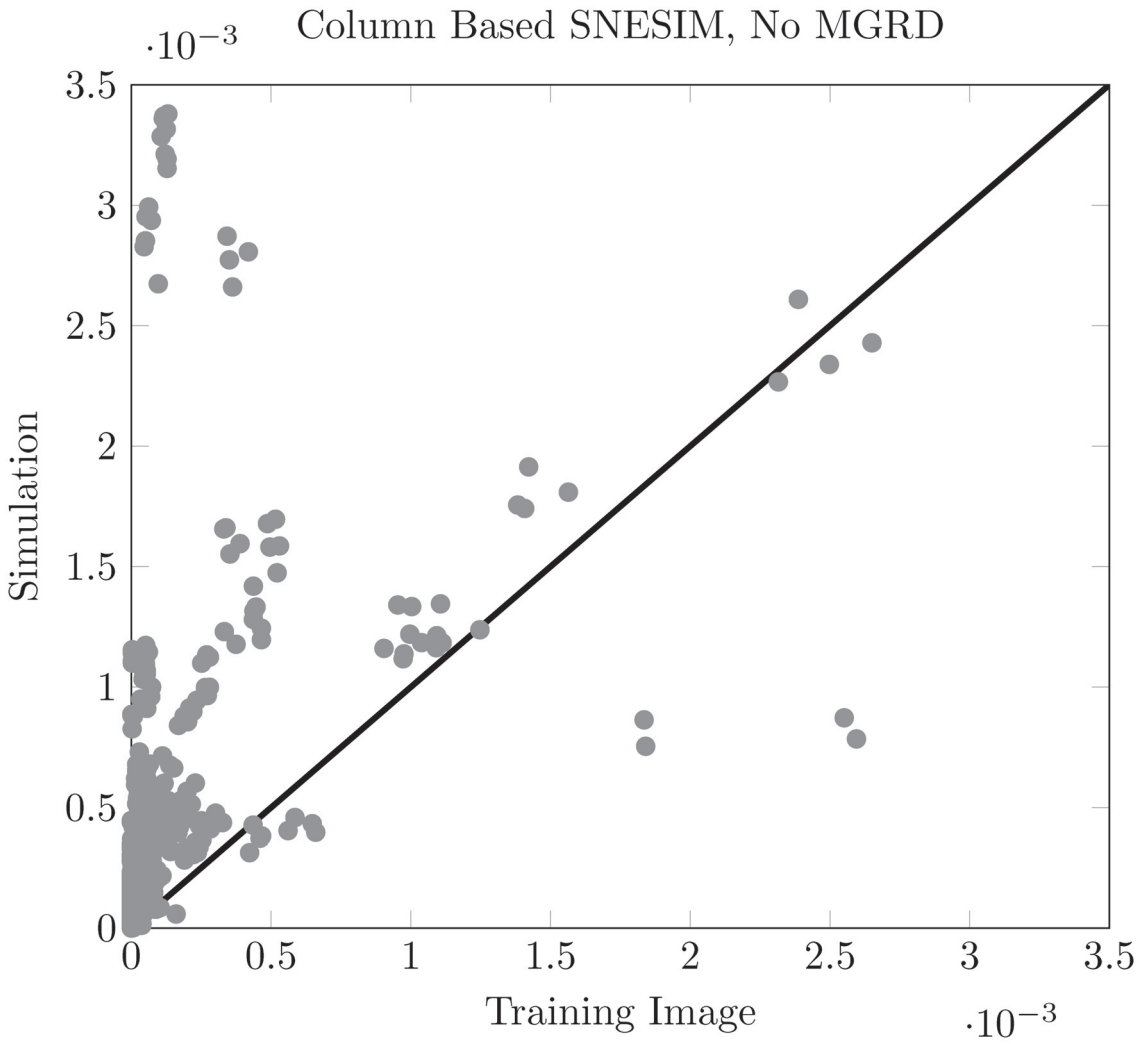

Figure 5.9: Frequency of patterns of training image versus simulation results for node based SNESIM with no multi-grid approach.

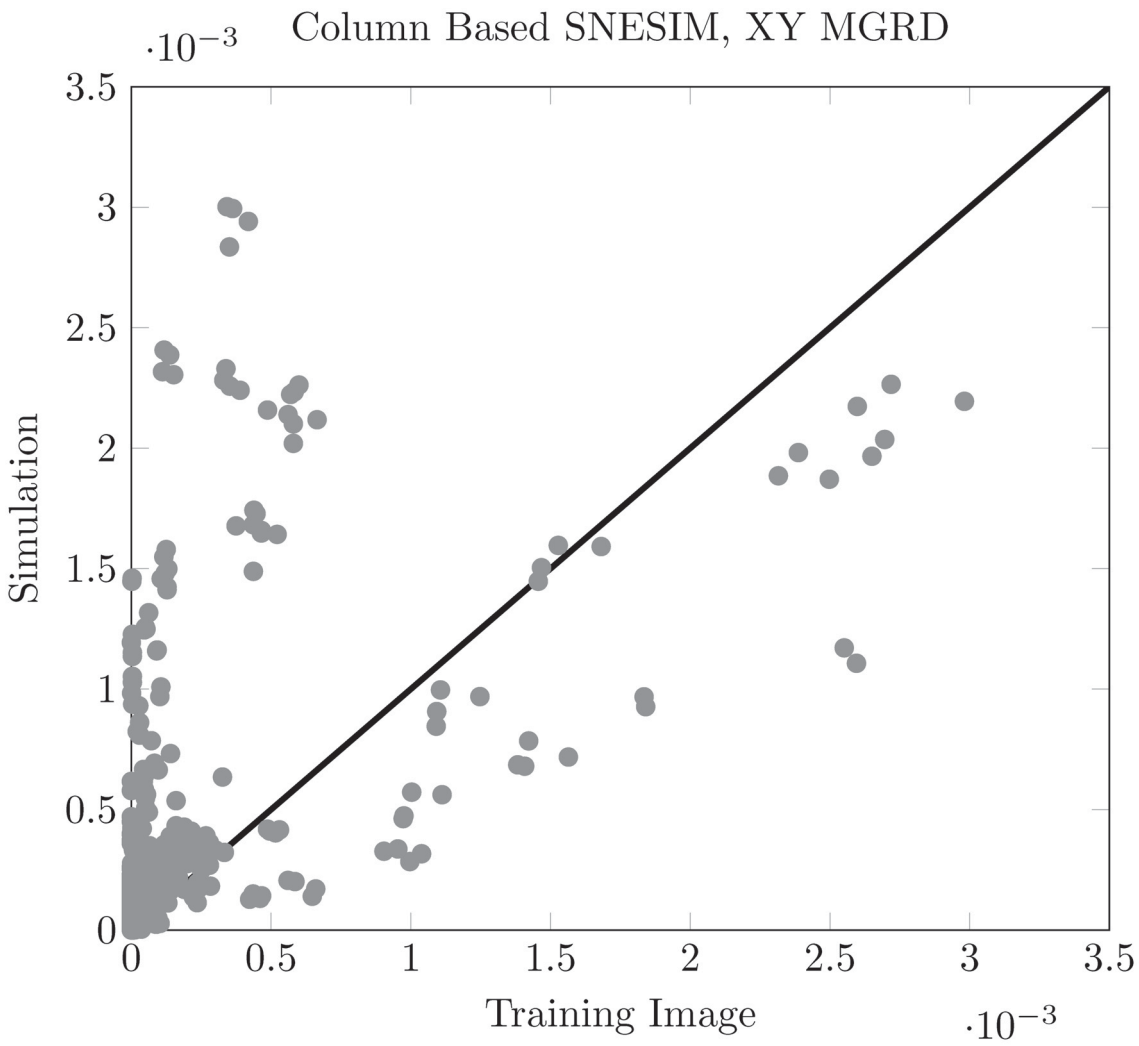

Figure 5.10: Frequency of patterns of training image versus simulation results for column based SNESIM with $X - Y$ multi-grid approach.

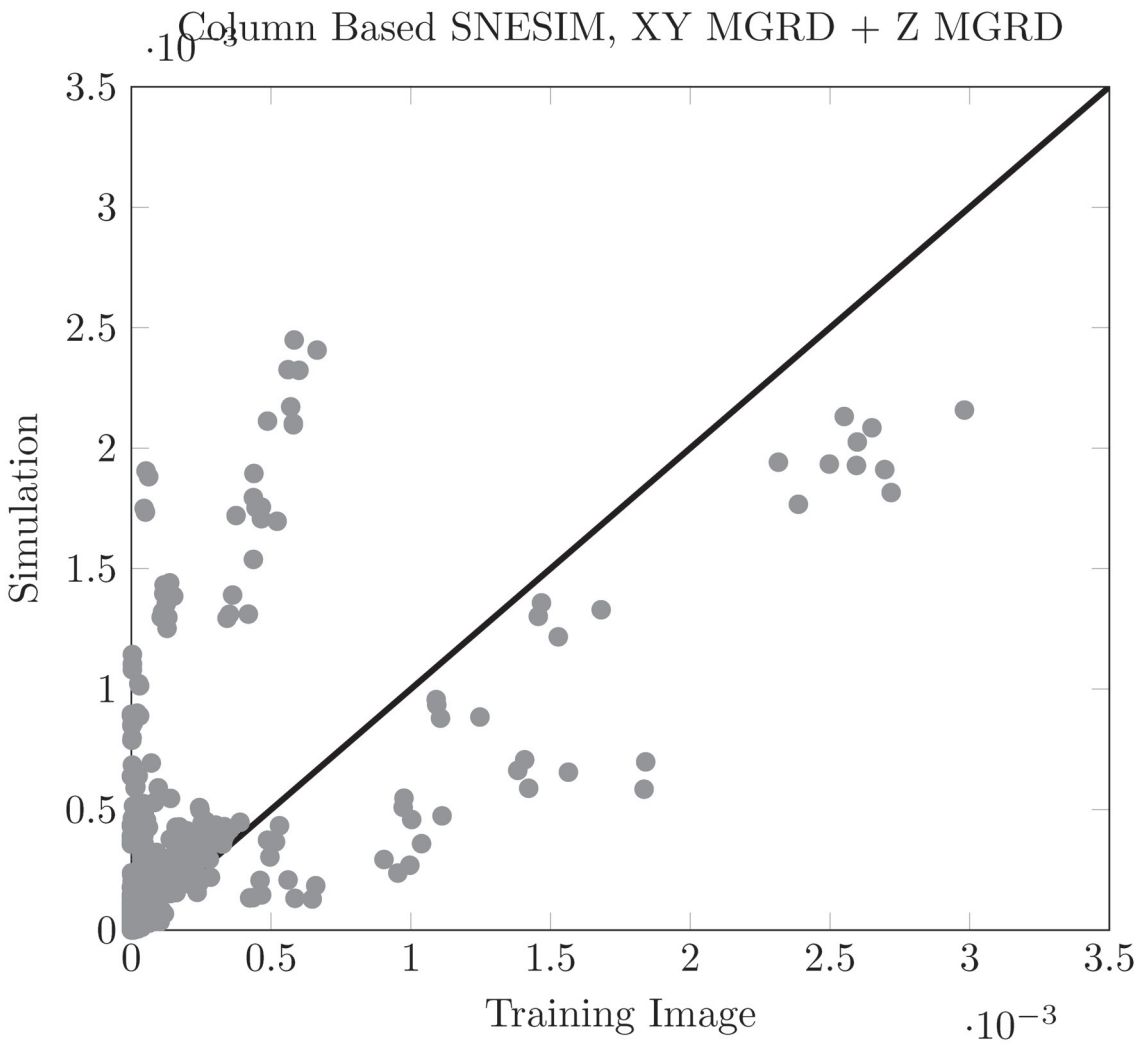

Figure 5.11: Frequency of patterns of training image versus simulation results for column based SNESIM with $X - Y$ plus Z multi-grid approach.

Table 5.1: Frequency, MSE of pure and mixed (MSE_p & MSE_m) patterns for training image, node based (NB) SNESIM and column based (CB) SNESIM with different multi-grid approaches.

	% White	% Black	$MSE_p(\times 10^{-3})$	$MSE_m(\times 10^{-10})$
Training Image	61.37	22.67	–	–
Node Based	52.55	28.66	5.63	2.70
CB,No MGRD	36.91	28.48	49.70	27.40
CB,XYMGRD-FullZ	46.58	17.76	12.20	8.80
CB,XY&ZMGRD	48.27	23.15	8.59	8.07

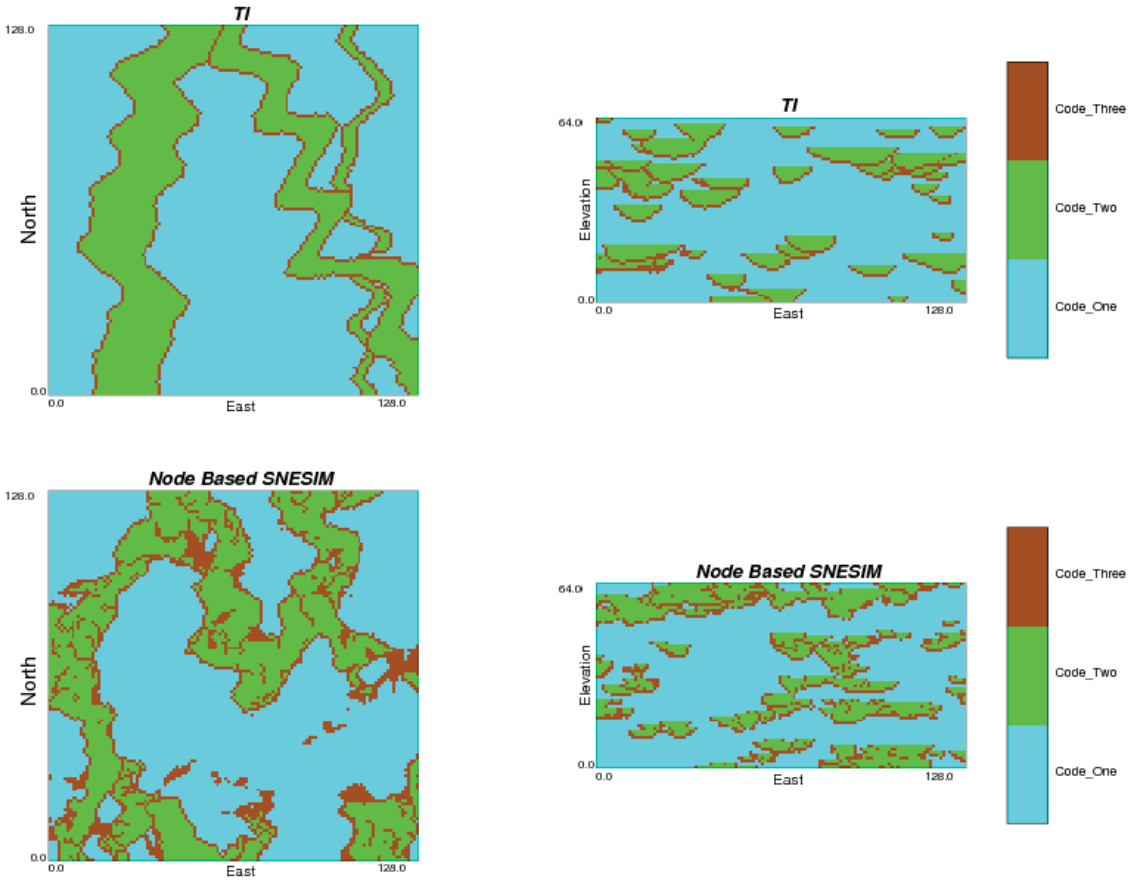

Figure 5.12: XY (left) and XZ (right) slice of training image and node base MPS simulation, axes unti (m).

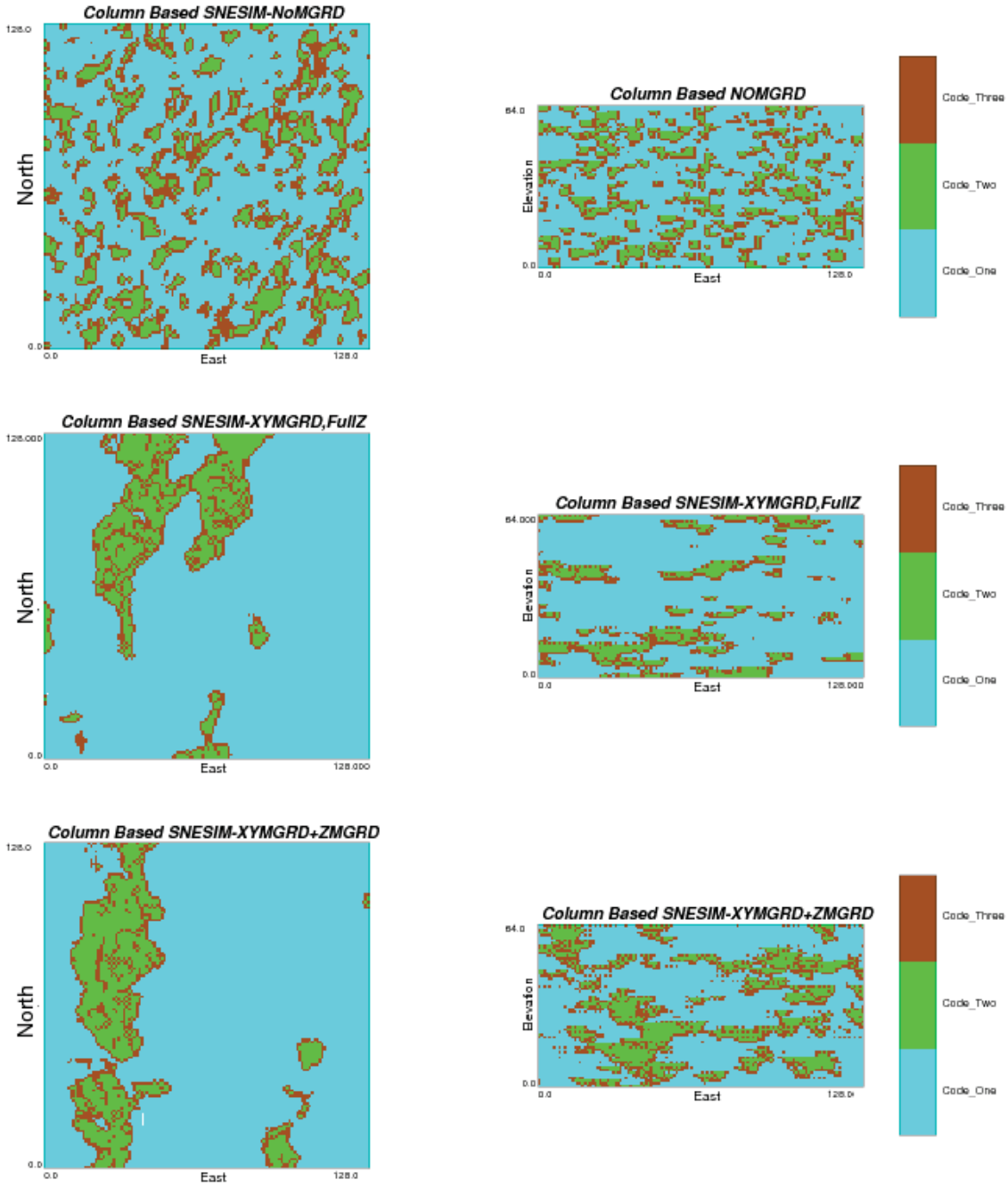

Figure 5.13: XY (left) and XZ (right) slice for column base with no multi-grids and column based with different type of multi-grids SNESIM, axes unti (m).

5.9 Chapter Summary

Different facies modeling techniques have been reviewed. Variogram based and multi-point statistic (MPS) simulation are appropriate for column based stochastic inversion. Among the different variogram based methods, truncated (pluri) Gaussian is a good candidate for facies modeling in multivariate stochastic inversion approach to simulated categorical and continuous variables simultaneously. Multi-point statistics (MPS) simulation techniques have seen a lot of popularity recently as they can handle more complex geological features on reservoir modeling. Therefore, column wise truncated (pluri) Gaussian and column wise multi-point statistics simulation technique have been implemented in this research study and considered as different options for facies modeling.

In multivariate stochastic inversion approach, facies and reservoir physical properties could be simulated by multivariate Gaussian simulation techniques simultaneously. In this method, facies is simulated by truncated (pluri) Gaussian technique and reservoir properties are simulated by sequential Gaussian simulation technique. The number of Gaussian variables are required for facies modeling via the truncated (pluri) Gaussian simulation technique depends on the number of Gaussian variables that describe the sequential ordering among the facies categories. The number of Guassian variables for simulation of reservoir properties is also defined by the number of reservoir properties and facies categories.

To be able to handle more complexity in reservoir modeling, column wise multi-point statistics is an alternative for facies modeling in multivariate stochastic inversion. For this porpuse, the directed multigrid approach is implemented in this thesis to be able to apply column wise multi-point statistics simulation.

Chapter 6

Adaptive Sampling

Multivariate stochastic inversion simulates a full column of facies and multiple reservoir physical properties simultaneously through the multivariate Gaussian simulation technique. To condition the simulated reservoir properties to the seismic data, the synthetic seismic trace related to the simulated column of multiple properties is computed through the forward modeling process in the PEM and convolution algorithms. The synthetic seismic trace is compared with the corresponding actual seismic trace in the adaptive sampling algorithm. Multiple realizations are generated per column through the described procedure to reach to an acceptable match with the original seismic data. The multiple simulated reservoir properties related to the acceptable match is kept as simulated values for the column. This process is replicated for all the columns to simulate the entire domain and obtain a facies, porosity and fluid saturations that reproduce the original seismic data.

The set of multiple realizations is used for uncertainty assessment and reservoir management. Uncertainty assessment requires many equiprobable realizations (Goovaerts, 1999). In the early days of simulation, there was a claim that the set of multiple realizations generated by geostatistical simulation may include the truth. Although, in reality, the space of uncertainty is vast and the number of realizations is very small. Therefore, the probability that one realization happens to be the truth is infinitesimally small (Goovaerts, 1999; Mukerji et al., 2001; Leuangthong et al., 2004). Also, considering the computational time and CPU usage, it is not practical to generate many realizations through the geostatistical simulation algorithm. Besides, uncertainty in the upstream petroleum industry is typically summarized by a P_{10} and P_{90} that only requires hundreds of realizations to infer reliably (Pyrcz and Deutsch, 2014).

The space of uncertainty is very large. The maximum size of space of uncertainty is K^N

for K categories over N locations. For only three categories ($K = 3$) over fifty locations ($N = 50$) the size of space of uncertainty is 3^{50} or about 10^{23}. To have a sense of how big this is consider that the radius of Milky Way is about 10^{22} meters. The current geostatistical practice deals with $K = 2$ to 7 categories and $N = 10^6$ to 10^8 locations; therefore, the size of the space of uncertainty is inconceivably large and it is impossible to sample through the simulation process.

As mentioned above, sampling the full space of uncertainty to obtain a perfect match to the original seismic data in multivariate stochastic inversion requires an unreasonably large number of realizations. Moreover, it may not even be possible to generate a true model due to modeling assumptions. A practical stopping criteria is required for the simulation of multivariate properties simultaneously conditional to the well and seismic data.

6.1 Practical Stopping Criteria

The multivariate stochastic seismic inversion approach simulates multiple reservoir physical properties simultaneously as a part of stochastic inversion algorithm to obtain high resolution reservoir models that recreate the seismic data through an adaptive sampling algorithm. A practical stopping criteria is defined two factors: 1) acceptance/rejection principle for the column based realizations, and 2) a reasonable maximum number of realizations.

6.1.1 Acceptance - Rejection Principle

In the proposed algorithm, the synthetic seismogram is compared to the collocated original seismic trace and the Mean Square Error (MSE) is calculated to pick the best match to the original seismic data. Multiple realizations are considered until a reasonable match to the original seismic data is attained. One basis for the target MSE is based on the inherent uncertainty in the model. There are different levels of uncertainty in the modeling process.

One source of uncertainty is the Petro-Elastic Model (PEM). A number of parameters

influence the reservoir elastic properties such as rock density and velocity. Facies and porosity can be related to the velocity empirically by the time-average equation. This equation is the most reliable when the rock is under significant pressure, saturated by brine and contains fine cemented grains (Gardner et al., 1974). Rock density and velocity depend upon the granular nature of the rock matrix, cementation, mineral composition, porosity, fluid content and pressure. Depth of burial and geological age also have some effect (Gardner et al., 1974). There are different empirical and practical connections between different reservoir elastic properties. A popular relation between rock density and velocity is Gardner's equation ($\rho = 0.31{V_p}^{0.25}$) that connects the density ($\rho(g/cc)$) to compressional velocity ($V_p(m/sec)$) in rocks based on different lithology, see Figure 3.8. This empirical relationship between density and velocity is based on laboratory and field observations for different sedimentary rock types. Reservoir elastic properties are related to each other through the empirical relations. Therefore, there is some level of error or uncertainty related to the elastic properties such as density and velocity that are computed by the Petro-Elastic Model (PEM) in multivariate stochastic inversion approach.

Inherent uncertainty in multivariate stochastic inversion is also related to seismic procedures like wavelet extraction. Seismic wavelets are essential inputs for many seismic processes such as inversion algorithms because they link geology to seismic amplitudes. Since they are generally unknown, an accurate extraction of the wavelet is a prerequisite for reliable seismic data analysis and any error or variation could lead to inaccurate results. Quantification of uncertainty related to seismic analysis has been documented by, for example, (Buland and Omre, 2003) Generally, there are two different approaches for wavelet extraction: 1) analytical wavelet extraction, and 2) data - driven optimization (Buland and Omre, 2003).

Analytical wavelets such as Ricker or Ormsby wavelet (Ryan, 1994) are often applied in quantitative seismic interpretation, seismic modeling and seismic inversion. The analytical wavelets are easy to define as they are determined by a limited number of parameters such as skewness, amplitude and polynomial order. The shape of the wavelet is predefined and adjustment to a more realistic wavelet shape is challenging (Skauvold

et al., 2015). The parameters that define the shape of the extracted wavelet vary over simulation domain because the geology and consequently seismic amplitudes are changing. The wavelet may be non-stationary with its shape varying over the area of interest.

A data- driven optimization method may be applied if many well data are available. The most common methods are based on the Walden and White (1998) approach but alternative methods have been developed by Buland and Omre (2003) and Gunning and Glinsky (2006). These methods are based on tying amplitude of seismograms in the neighborhood of one or more wells to the synthetic seismograms calculated by well data. The goal of this method is to extract the spectral properties of seismic data which is the reflectivity series by removing the geological contribution to the seismic data. There is also an uncertainty related to the seismic well-tie process. There is inevitable uncertainty in stochastic inversion due to the wavelet extraction procedure.

A zero phase Ricker wavelet with a central wavenumber is considered here for the multivariate stochastic inversion approach. Uncertainty in the parameters of this wavelet will be considered. Alternatives could be considered in practice with specific well and seismic.

An example of the uncertainty that propagates into the final model via the Petro-Elastic Model (PEM) algorithm in the multivariate stochastic inversion is developed. A column with fifty locations (1×50) with different reservoir physical properties, elastic properties and seismic data is displayed in Figures 6.1 and 6.2 respectively. In the Petro-Elastic Model (PEM) algorithm, the velocity model can partially cause the uncertainty in multivariate stochastic inversion results. In this study, the velocity is calculated by Gardner's equation ($\rho = aV_p{}^m$) via the Petro-Elastic Model (PEM) which $a = 0.31$ and $m = 0.25$ are the fitting factors obtained through the empirical relationships and are uncertain. To quantify the uncertainty related to the velocity model, two Gaussian distribution $Y_1(mean, std) = Y_1(0.31, 0.05)$ and $Y_2(mean, std) = Y_2(0.23, 0.03)$ are assigned to the fitting factors. The standard deviation is small, yet considered

representative given some of the experimental published results (Mukerji et al., 2001; Goovaerts, 2006). Multiple realizations of velocity generated over the range of fitting factors were simulated. Corresponding elastic properties such as acoustic impedance (AI), reflection coefficient (R) and seismic data are calculated for the multiple realizations. Figures 6.3 and 6.4 show the range of velocity, R and seismic trace for multiple realizations. The average Mean-Square Error (MSE) between seismic data obtained from multiple realizations of velocity and the original one ($MSE_{ave} = 9.4 \times 10^{-4}$) over the variance of original seismic data ($\sigma^2 = 1.6 \times 10^{-2}$) represent around 6% relative error in the seismic data. This helps define the acceptable mismatch between original seismic data and simulated one via the multivariate stochastic inversion in presence of uncertainty in the Petro-Elastic Model (PEM). The adaptive sampling algorithm could stop when the MSE reaches to the order of 10^{-4}.

As discussed above, some part of the cumulative uncertainty in the final model is due to wavelet extraction. To get closer to the cumulative uncertainty of velocity modeling and wavelet extraction, a random noise is added to the Ricker wavelet in the previous example. In this situation, the average Mean-Square Error (MSE) between multiple realizations of seismic data and the original one ($MSE_{ave} = 1.53 \times 10^{-3}$) over the variance of original seismic data ($\sigma^2 = 1.6 \times 10^{-2}$) represent around 10% error in the seismic data. Figure 6.5 shows the range of seismic data for this situation. The uncertainty in velocity modeling and extracted wavelet process leads to higher mismatch between synthetic and actual seismic data and increases the target MSE. Therefore, the relative error of the final model is increased as well. Considering the other source of uncertainty in the model may increase the target MSE or mismatch more than this number.

The discrepancy between reality and model obtained through the multivariate stochastic inversion is mainly because there are many steps and algorithms involved in this approach. The uncertainty in the Petro-Elastic Model (PEM) is not only about the parameters that define the elastic properties. The fundamental physics may not be completely or correctly represented in the PEM. The PEM may also change over the area of interest. In the convolution process, the sources of uncertainty include the nature of the

extracted wavelet, time-depth relationship, edge effects, fundamentals of physics and stationarity. There may also be a scale effect, that is, using grid blocks with the constant properties when reality is variable at all scales. There are many sources of uncertainty in the final model that cannot be considered and processed. Some sensitivity analysis has been done on the PEM and wavelet extraction.

In general, each step and algorithm embedded in multivariate stochastic inversion approach should be analyzed and its associated uncertainty quantified. The cumulative uncertainty defines the acceptable target MSE and relative error for seismic data matching. These parameters support the primary stopping criteria in the adaptive sampling algorithm.

6.1.2 Reasonable Number of Realizations

In multivariate stochastic inversion approach, multiple realizations are generated through the multivariate Gaussian simulation technique. These multiple realizations are processed and compared to the actual seismic data. The question being considered here is how many realizations should be generated to obtain an acceptable match with seismic data. The acceptable match principle and target Mean Square Error (MSE) have been discussed above. In this part, the focus is on the practical and reasonable number of realizations for the multivariate stochastic inversion approach.

The size of the space of uncertainty is very large and cannot be exhaustively sampled. It is not necessary to obtain a perfect match with the actual seismic data as there is some level of error and uncertainty associated with data and models. Nevertheless, a reasonable number of realizations for simulation process must be chosen. The size of the space of uncertainty will be reduced in presence of conditioning data. Therefore, as the simulation proceeds the number of realizations will be reduced. An acceptable match or target MSE with the actual seismic data would likely be reached with less realizations in presence of strong conditioning data. More realizations are required to reach the target MSE early in the simulation process far from conditioning data. When the conditioning data is close to

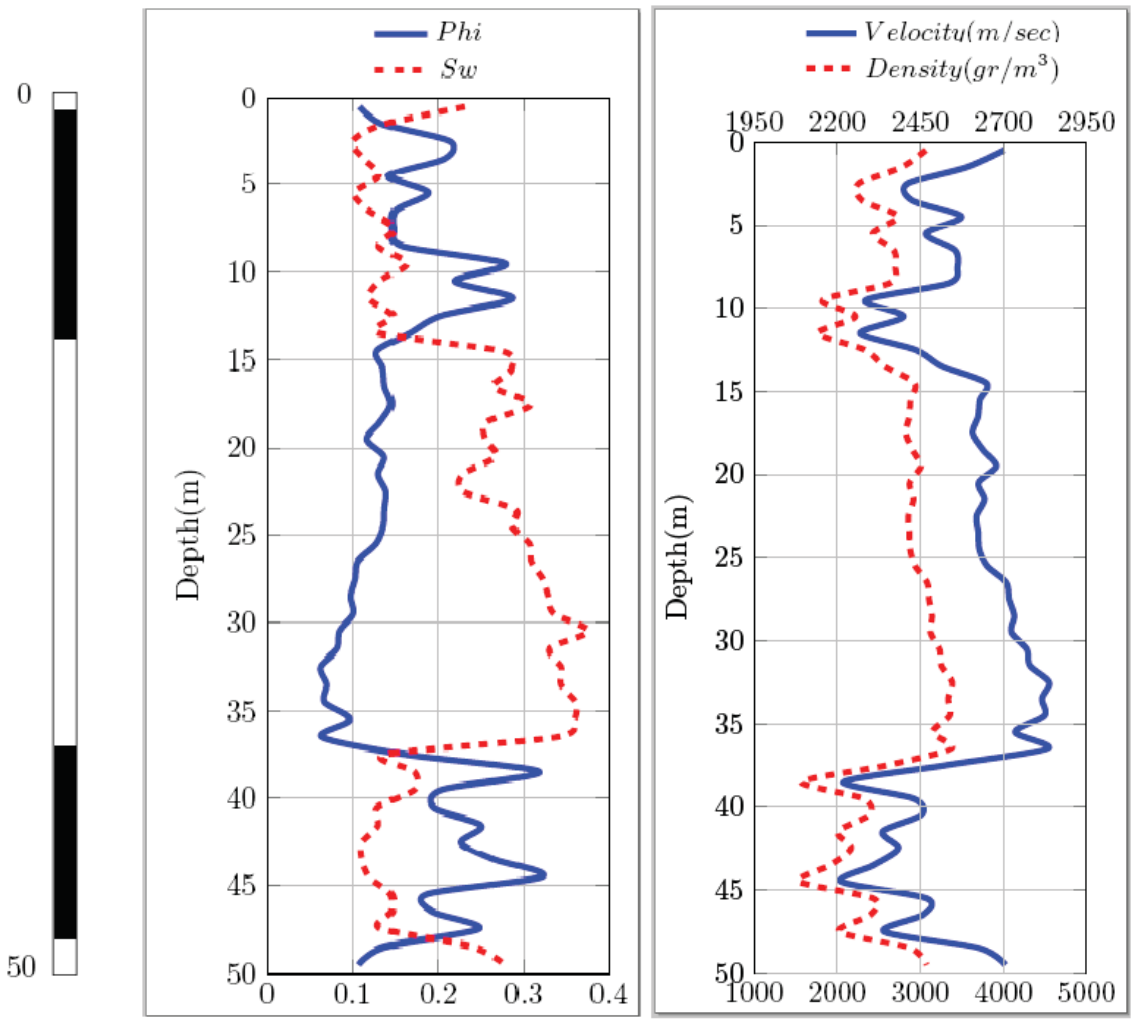

Figure 6.1: Facies and physical properties of the example of uncertainty propagation via PEM.

the simulation location, the size of the space of space of uncertainty is reduced and number of realization dramatically drops. Figure 6.6 schematically represents this concept.

To demonstrate this concept, a 2-D model (20×20) is considered. A column of this 2-D model is simulated unconditionally and conditionally through the multivariate stochastic inversion approach. For the conditional simulation, three different positions are considered with weak, mild and strong conditioning. Figure 6.7 illustrates these scenarios. Ten thousand realizations are generated to simulate multiple reservoir properties simultaneously through the multivariate Gaussian simulation algorithm. The multiple

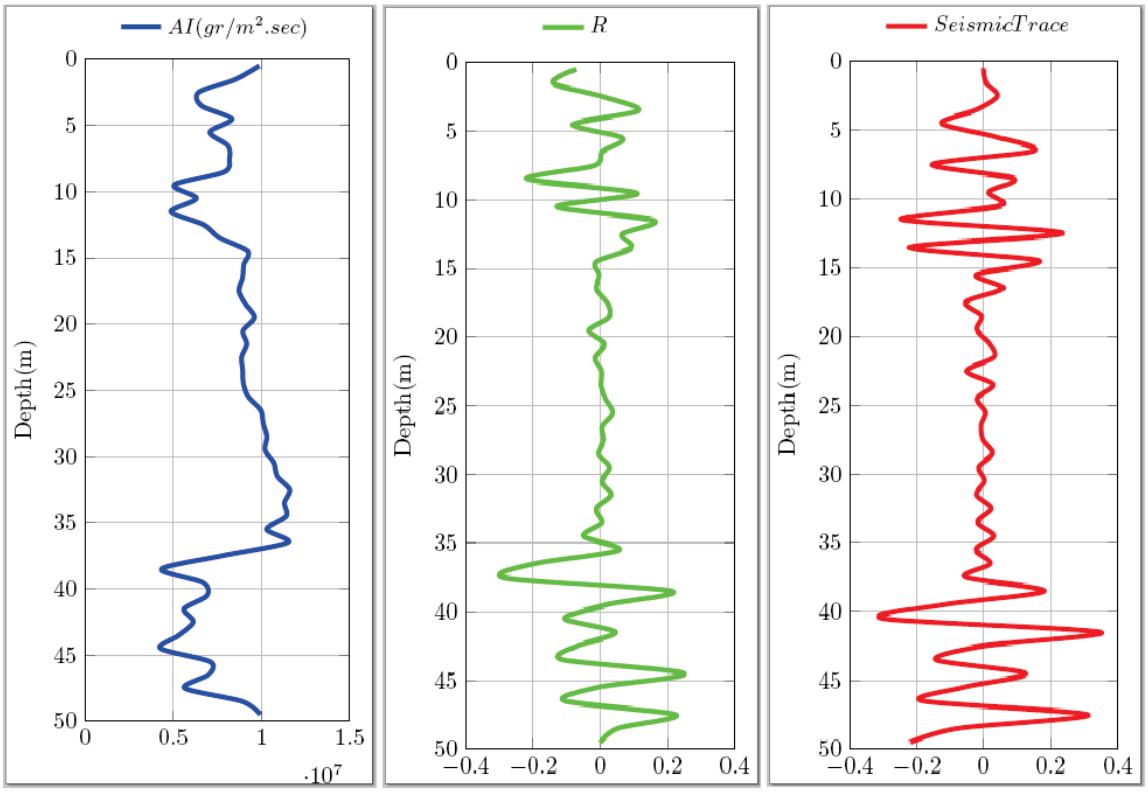

Figure 6.2: Elastic properties and seismic trace of the example of uncertainty propagation via PEM.

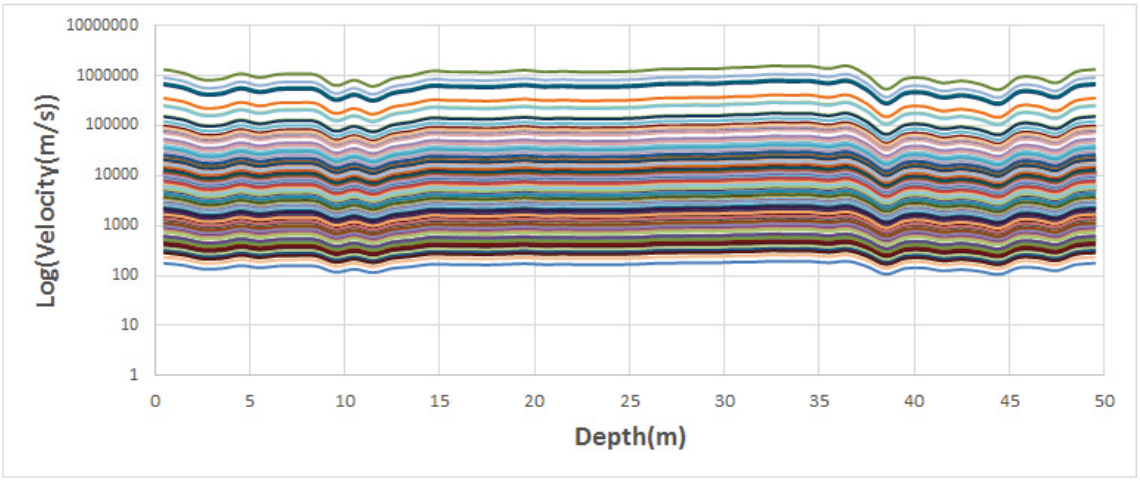

Figure 6.3: Range of velocity for multiple realizations of velocity.

realizations of different reservoir properties are passed to the PEM and convolution algorithms to calculate synthetic seismograms. Calculated synthetic seismograms are

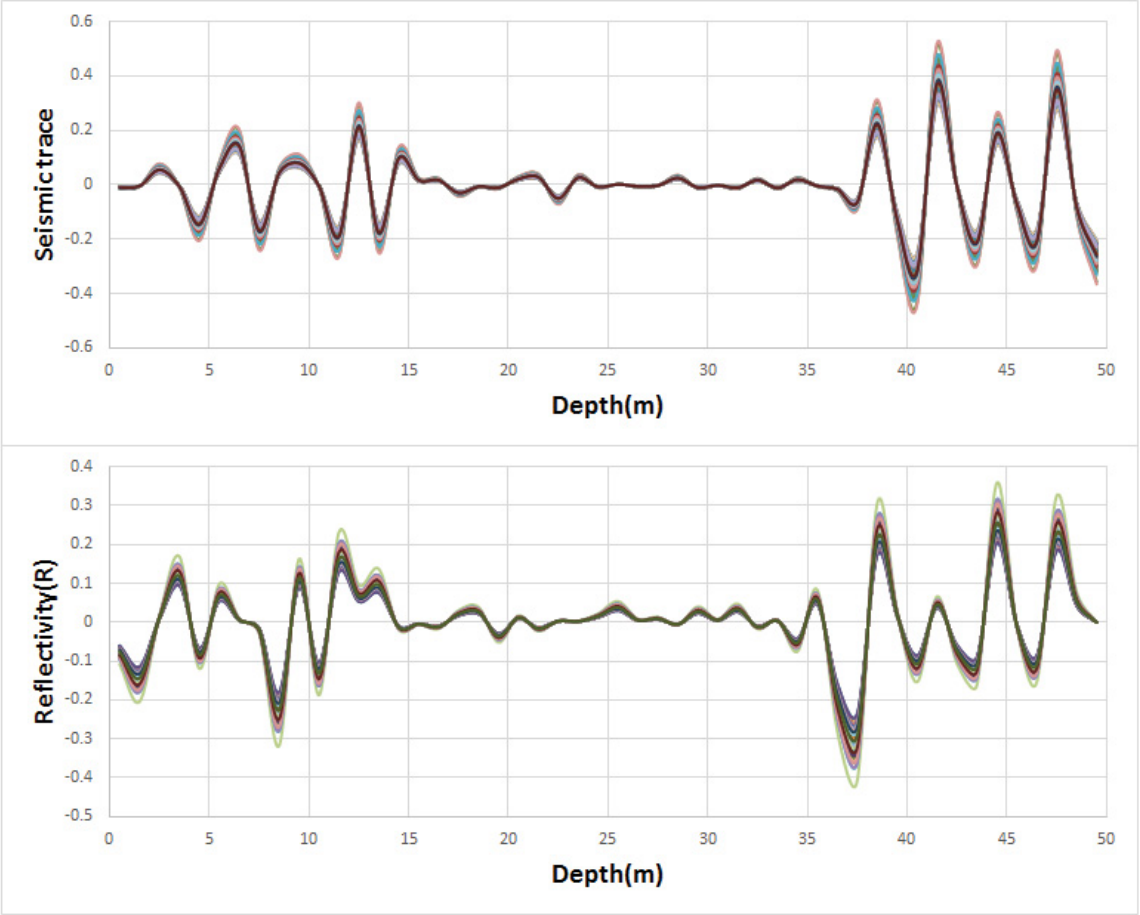

Figure 6.4: Range of reflectivity (R) and seismic trace for multiple realizations of velocity.

compared with actual seismic trace and the MSE is computed with the best fit. Figure 6.8 shows the comparison of synthetic seismogram related to the minimum MSE and actual seismic trace for unconditional and different conditional simulation scenarios. As shown in this figure, unconditional and weak conditional simulations reach a higher MSE or mismatch with original seismic data within a fixed number of realizations (10000) compared to the mild and strong conditional simulation. Strong conditional simulation almost reaches to a perfect match with actual seismic data.

In another example, a single column is simulated as above but this time trying to reach a specific target MSE. The realization with the lowest MSE is retained as simulation proceeds. Figure 6.9 displays the number of realizations versus MSE for this example. The MSE

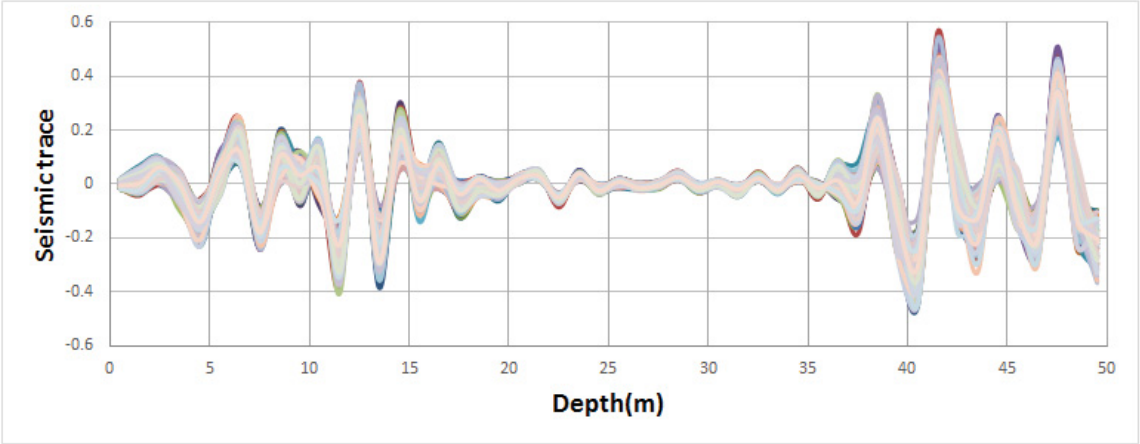

Figure 6.5: Range of seismic trace for multiple realizations based on cumulative uncertainty of PEM and wavelet extraction.

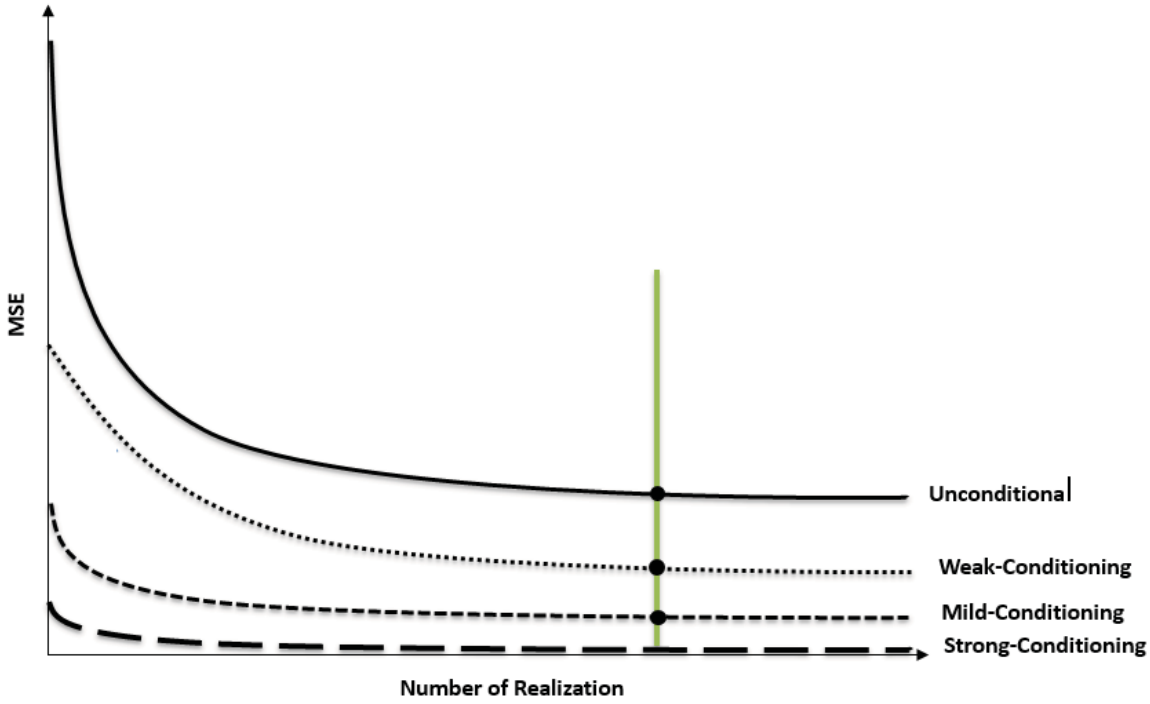

Figure 6.6: Schematic illustration of number of realization versus Mean Square Error.

of the first realization for unconditional simulation is large but the target MSE (10^{-4}) is reached within 1000 realizations. The MSE of the first realization for the conditional case is very low compared to the unconditional one. The weak conditional simulation reaches to the target MSE within 500 realizations while mild and strong simulation reach to the

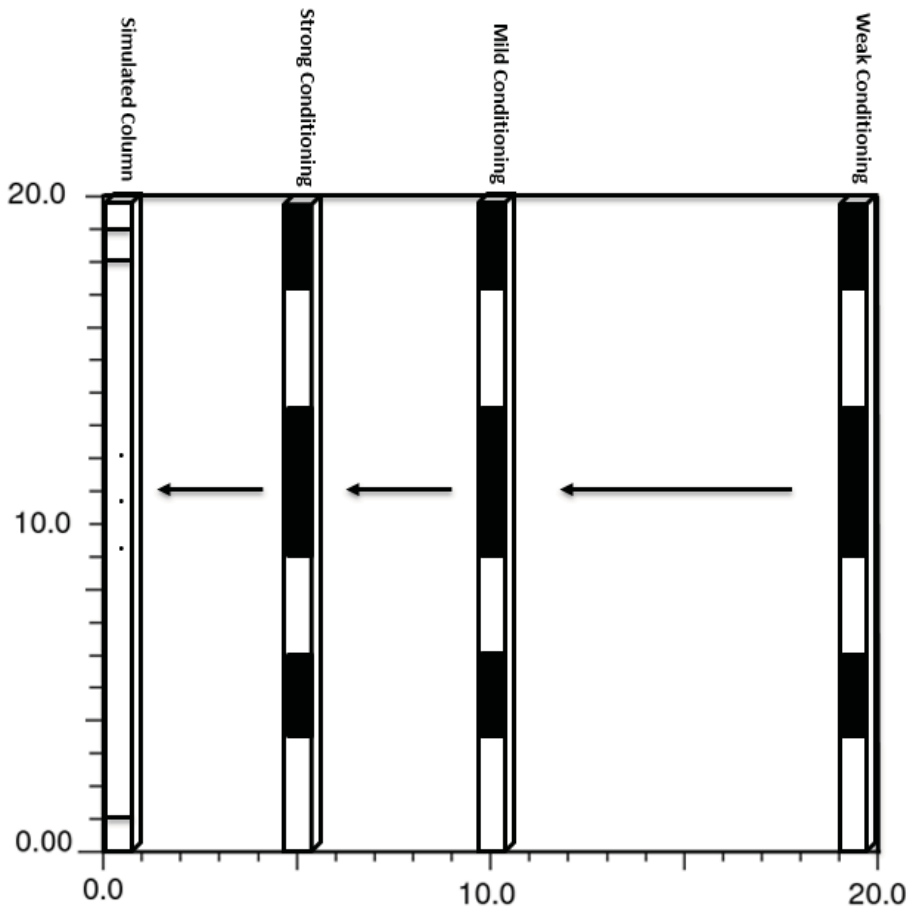

Figure 6.7: Schematic illustration of 2-D model for unconditional and conditional simulation.

target MSE within 100 and 10 realizations, respectively. Figure 6.9 demonstrates that the number of realizations drops when going through the simulation path as more information is provided from previously simulated locations, that is, a better match or lower MSE can be reached within less realizations as the simulation proceeds. The issue being considered is how the different stages for conditioning data should be defined through the simulation steps in multivariate stochastic inversion to define and generate different number of realizations on the different stages of conditioning data.

Multigrid approach is a solution to this issue. The multigrid approach uses a structured sequence to visit the grid nodes in the simulation domain. Two or more grids are defined over the simulation domain. The sequence starts with the coarsest multigrid. At this stage

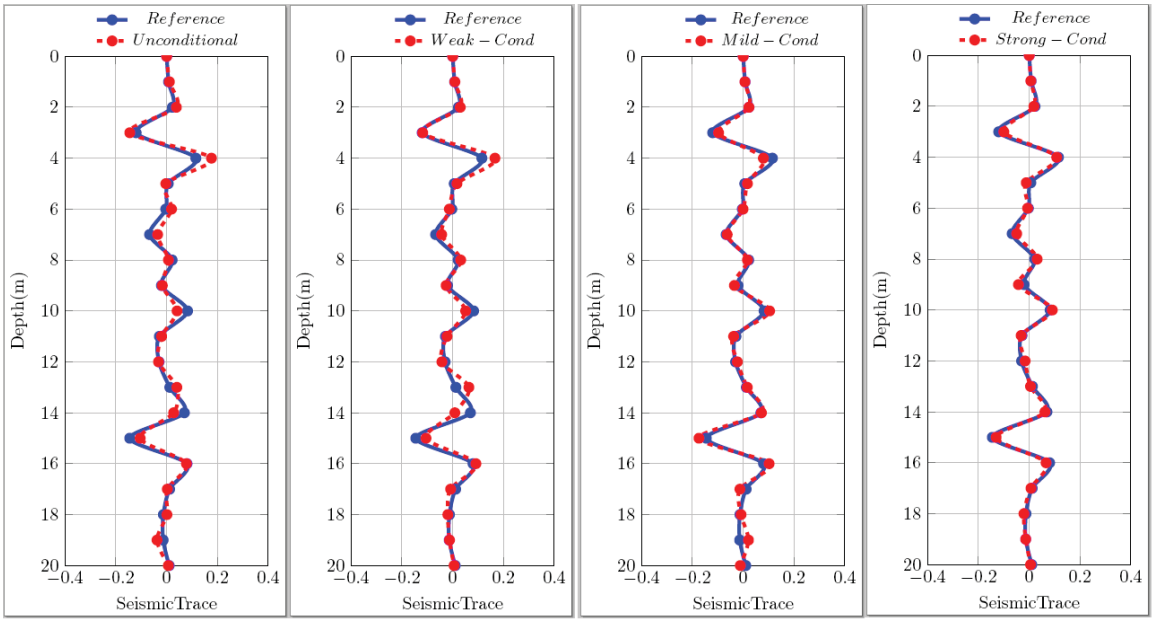

Figure 6.8: Comparison of reference seismic trace with synthetic seismogram corresponded to unconditional, weak - mild - strong conditional simulation at 10000 realizations.

of multigrid, since the spacing between grid nodes is large, the simulation is not strongly conditioned. Since the simulated nodes at each level of multigrid are added to the conditioning data, these levels of multigrid can be considered as weak, mild and strong conditional simulation. Therefore, a different number of realizations can be considered for each stage of the multigrid. Figure 6.10 schematically represents this concept. The multigird approach is important in simulation as it can facilitate the local conditioning and reproduction of large scale features.

The different levels of multigrid are considered as different stages for conditioning data in multivariate stochastic inversion in that simulation process starts with large number of realizations in the first level of multigrid and gradually decreases the number of realizations by going through the different levels of multigrid. The multivariate stochastic inversion approach considers a directed multigrid, $X - Y$ multigrid with full column of Z. The number of realizations per multigrid stage and the target MSE must be defined for the adaptive sampling algorithm. The number of realization within each multigrid stages

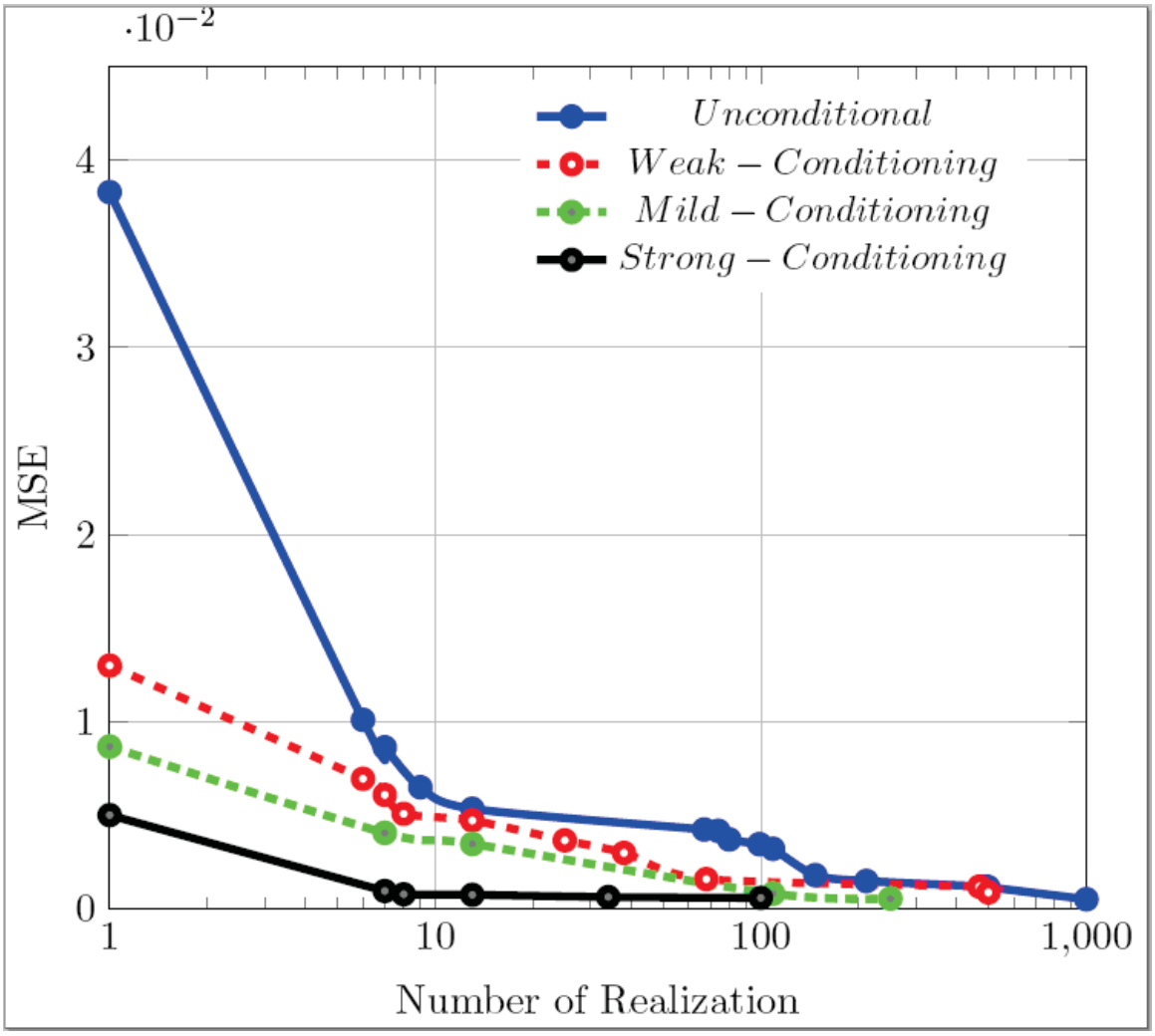

Figure 6.9: Mean Square Error versus number of realizations (Log scale) for unconditional and conditional scenarios for case study, 2-D model (20×20) .

will be kept the same for simplicity.

The multivariate stochastic inversion approach simulates multiple reservoir properties simultaneously through the multivariate Gaussian simulation in column wise manner. These simulated properties are passed to PEM and convolution algorithm to compute reservoir elastic properties and corresponding synthetic seismic. The synthetic seismogram is compared with actual seismic trace and MSE calculated through the adaptive sampling algorithm. This procedure is continued by generating multiple realizations until the column either reaches the target MSE or the maximum number of

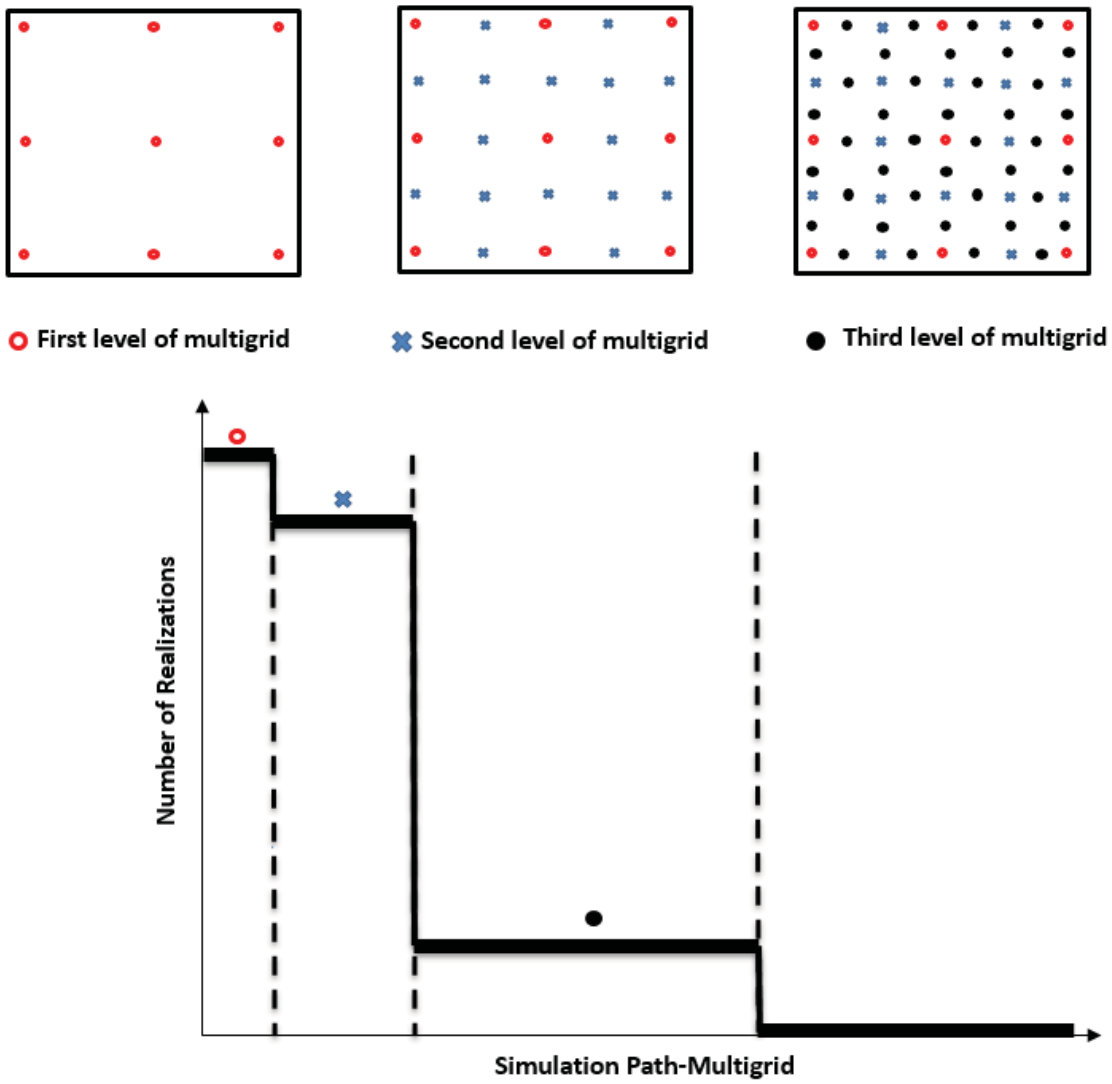

Figure 6.10: Schematic representation of the concept of multigrid and dynamic number of realizations for each level.

realizations for the level of multigrid. A 2-D example demonstrates the work flow and shows how the different number of realizations are defined for each level of multigrid.

6.2 2-D Example

A 2-D example is considered to demonstrate the performance of the multivariate stochastic inversion approach. A 40×40 model of facies, porosity, water saturation and seismic survey are simulated as a reference model. Column 20 is sampled and considered as well data for conditioning. The objective is to reproduce the reference model through the different scenarios with different stopping criteria.

As a first scenario, the reservoir properties are only conditioned to the well data. The synthetic 2-D seismic model related to the different realizations are calculated through the forward model. Figure 6.11 and Figure 6.12 show these simulated properties for two different realizations beside the conditioning data and reference model respectively. Figure 6.13 to 6.15 show the reference seismic survey and the synthetic seismic surveys obtained by forward modeling for two different realizations.

In the second scenario, the reservoir properties are simulated simultaneously through the multivariate stochastic inversion approach conditioned to the well and seismic data. In this scenario, five levels of multigrid is considered for simulation process. The number of multiple realizations for each multigrid level is displayed in the second column of Table 6.1. The target MSE is defined as 10^{-4}. The adaptive sampling algorithm, at each column, either reaches to the target MSE within the defined number of realizations for each multigrid level or stops on the defined number of realizations and considers the simulated values related to the minimum MSE as simulation results for the corresponded column. Figure 6.16 and 6.17 show these simulated properties for two different realizations beside the related conditioning data and reference models. Figure 6.18 and 6.19 also show the synthetic seismic for two different realizations of this scenario. In this scenario, only the finest level multigrid ($mmult = 1$) converges to the target MSE. The rest of multigrid levels stop at the maximum number of realizations before reaching the target MSE. As

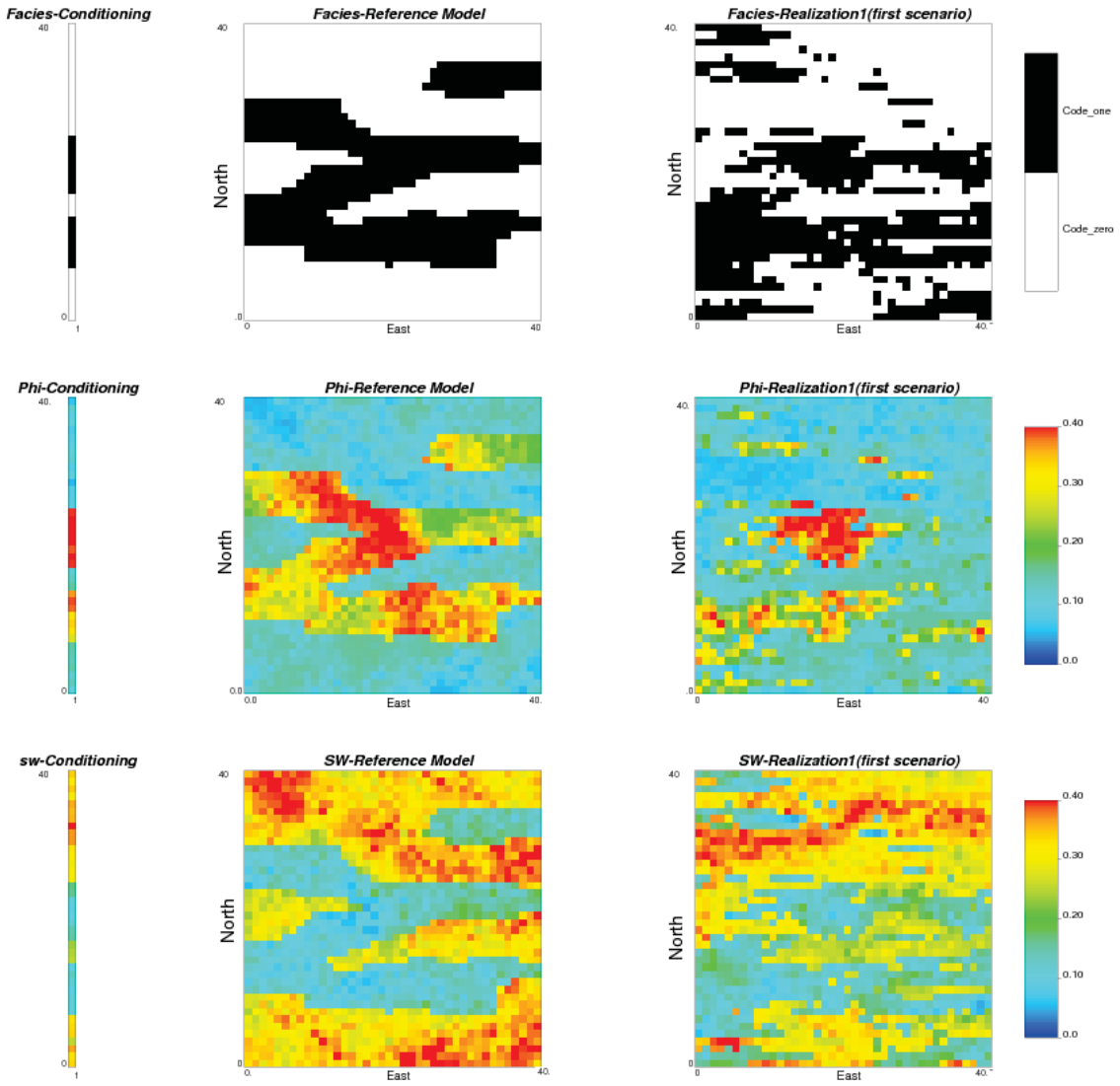

Figure 6.11: Comparison of conditioning data (first column) with different properties of reference model (second column) and first realizations of simulated properties (third column) for first scenario, axes unit m, color bar unit m^3/m^3.

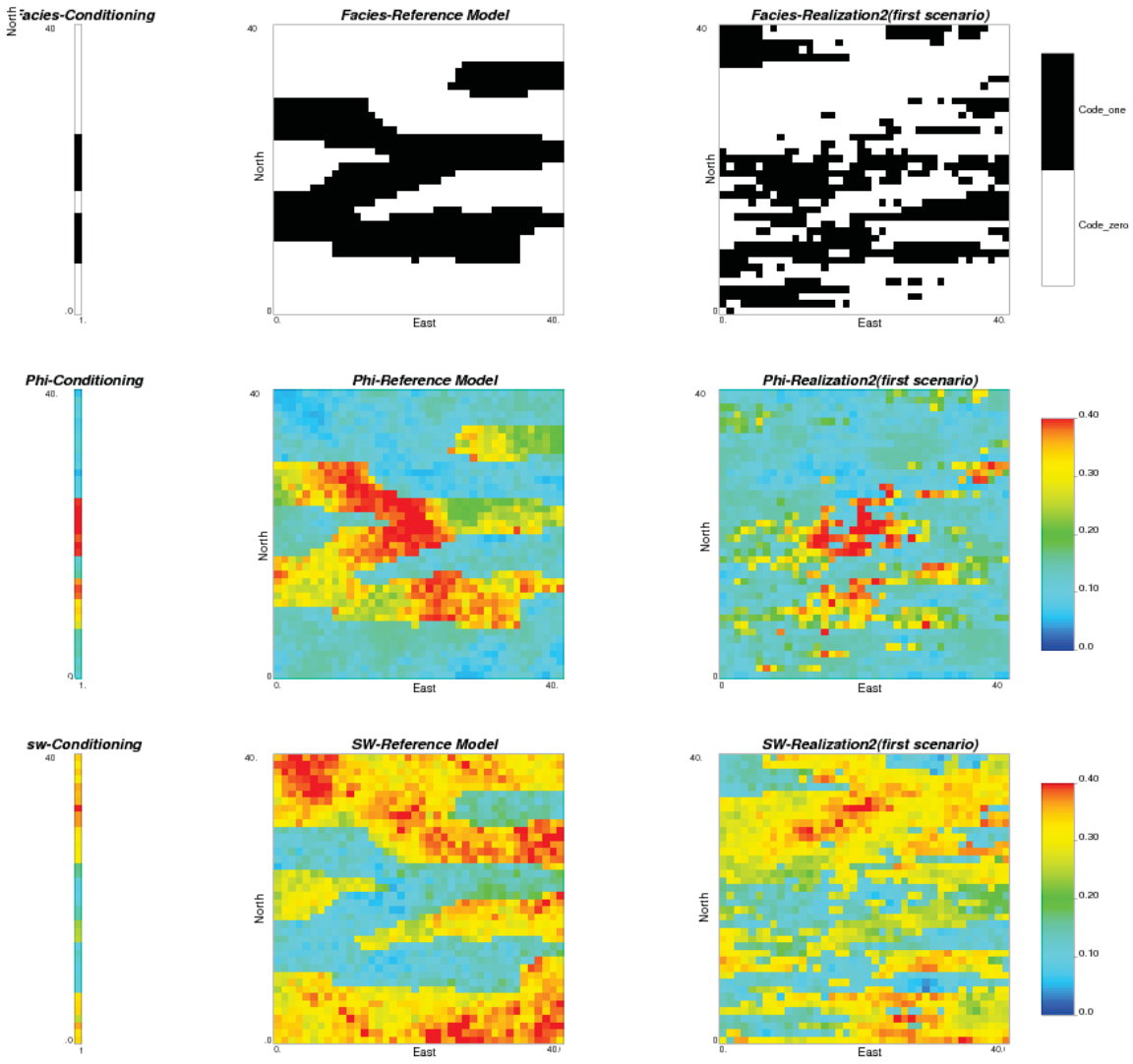

Figure 6.12: Comparison of conditioning data (first column) with different properties of reference model (second column) and second realizations of simulated properties (third column) for first scenario, axes unit m, color bar unit m^3/m^3.

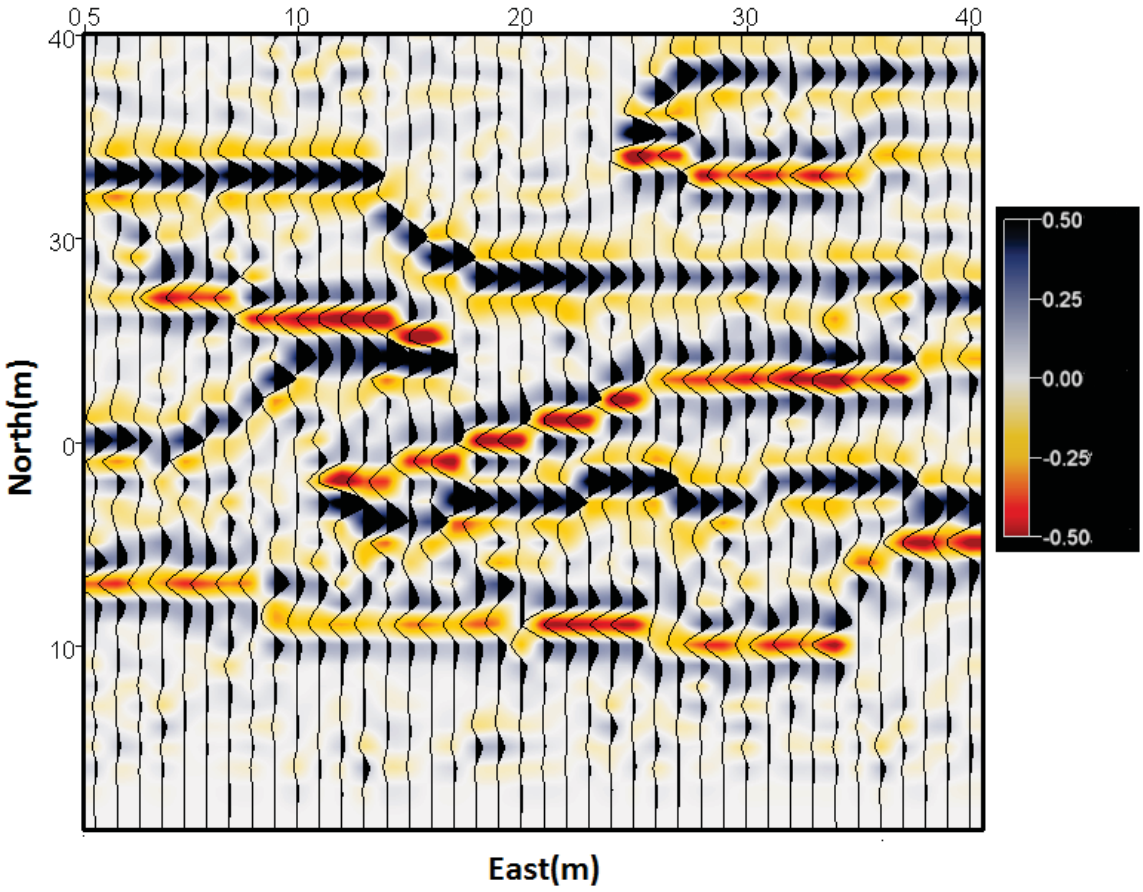

Figure 6.13: 2-D seismic survey considered as the reference model, axes unit (m) and color bar - amplitude (m).

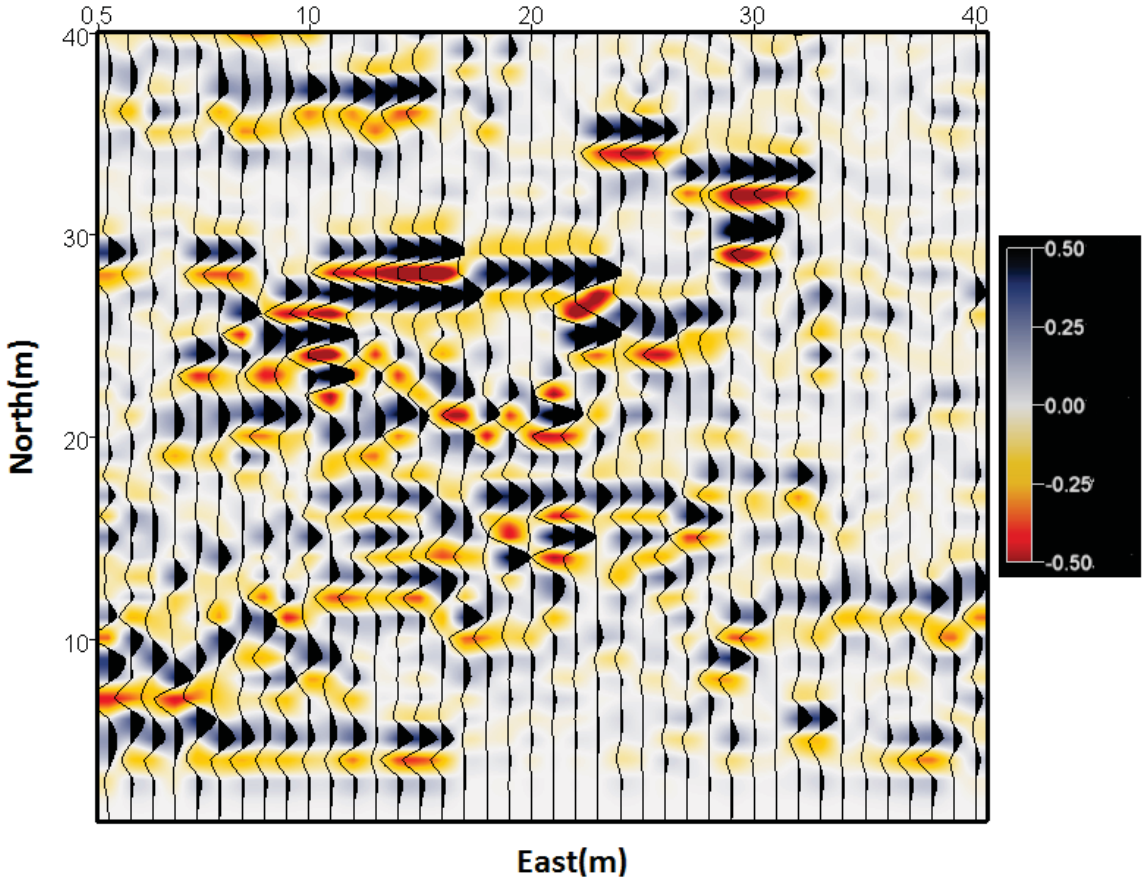

Figure 6.14: The synthetic 2-D seismic survey of first realization of first scenario conditioned only to the well data, axes unit (m) and color bar - amplitude (m).

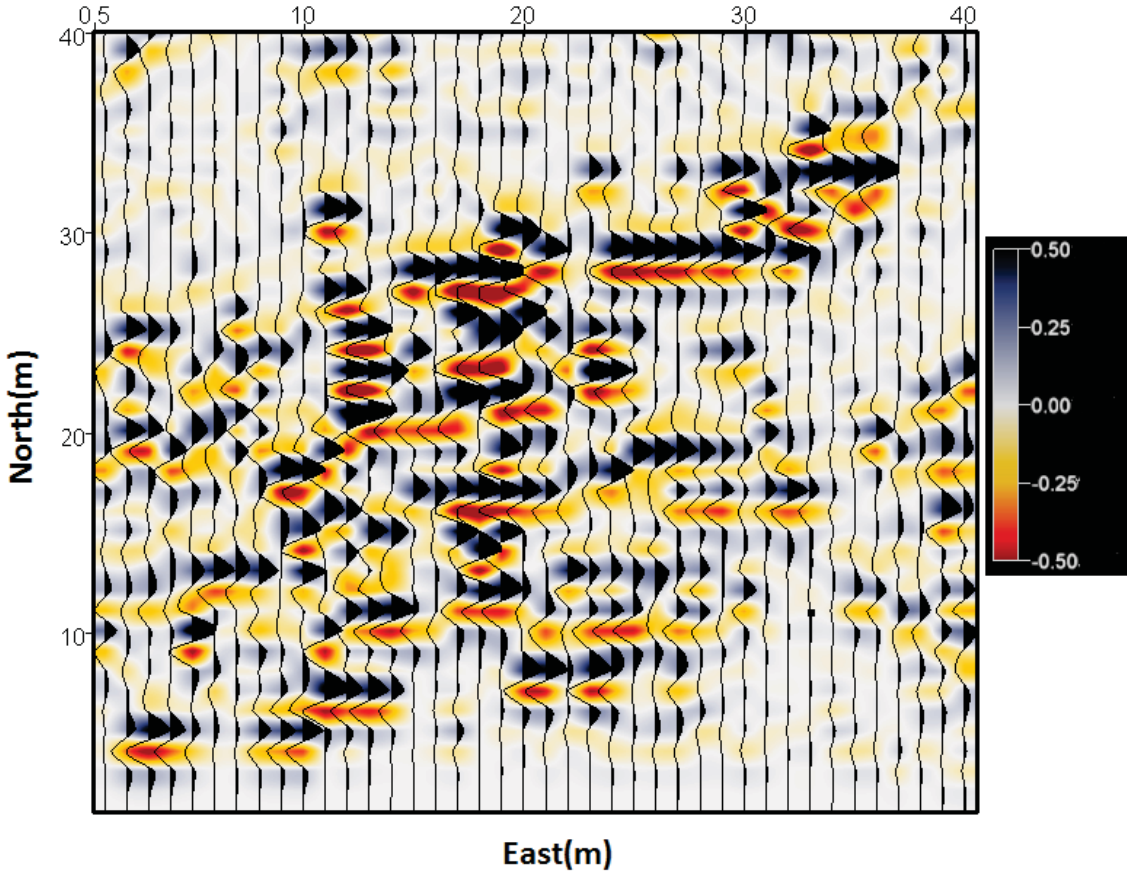

Figure 6.15: The synthetic 2-D seismic survey of second realization of first scenario conditioned only to the well data, axes unit (s) and color bar - amplitude (m).

mentioned before, defining the stopping criteria in the adaptive sampling algorithm is a trade off, fewer of realizations requires less computational time yet leads to a higher mismatch between synthetic and original seismic data.

Table 6.1: Multigrid level, number of columns times dynamic number of realization per multigrid level for the second and third scenarios.

Level of Multigrid	No.Columns $\times$ No.RLZ(SC 2)	No.Columns $\times$ No.RLZ(SC 3)
5	3×100	3×500
4	5×50	5×250
3	10×20	10×100
2	20×10	20×50
1	2×5	2×25

The third scenario increases the number of realizations to improve the mismatch between original and synthetic seismic. The number of multiple realizations for this scenario displayed in the third column of Table 6.1. Figure 6.20 and 6.21 show these simulated properties for two different realizations beside the related conditioning data and reference models. Figure 6.22 and 6.23 show the synthetic seismic survey of two different realizations for this scenario. In this scenario, the number of multiple realizations at each multigrid level is five times the second scenario and the majority of simulation iterations converge to the target MSE. The computational time is also increased linearly by the number of realizations. However, a better match with original seismic data is obtained with this scenario.

Figure 6.24 shows the conditional variance of multiple realizations of different properties for these three scenarios. Figure 6.25 shows the conditional variance of the seismic survey for the first scenario that is only conditioned to the well data and also the conditional variance of seismic surveys related to the second and third scenario that are conditioned to well and seismic data through the multivariate stochastic inversion. As shown in these figures, the conditional variance of reservoir properties and seismic for the first scenario are very low in the middle of simulation domain which is well location. However, the variance increases away from the well. The conditional variance of different properties and seismic data for the second scenario are lower compared to the first scenario.

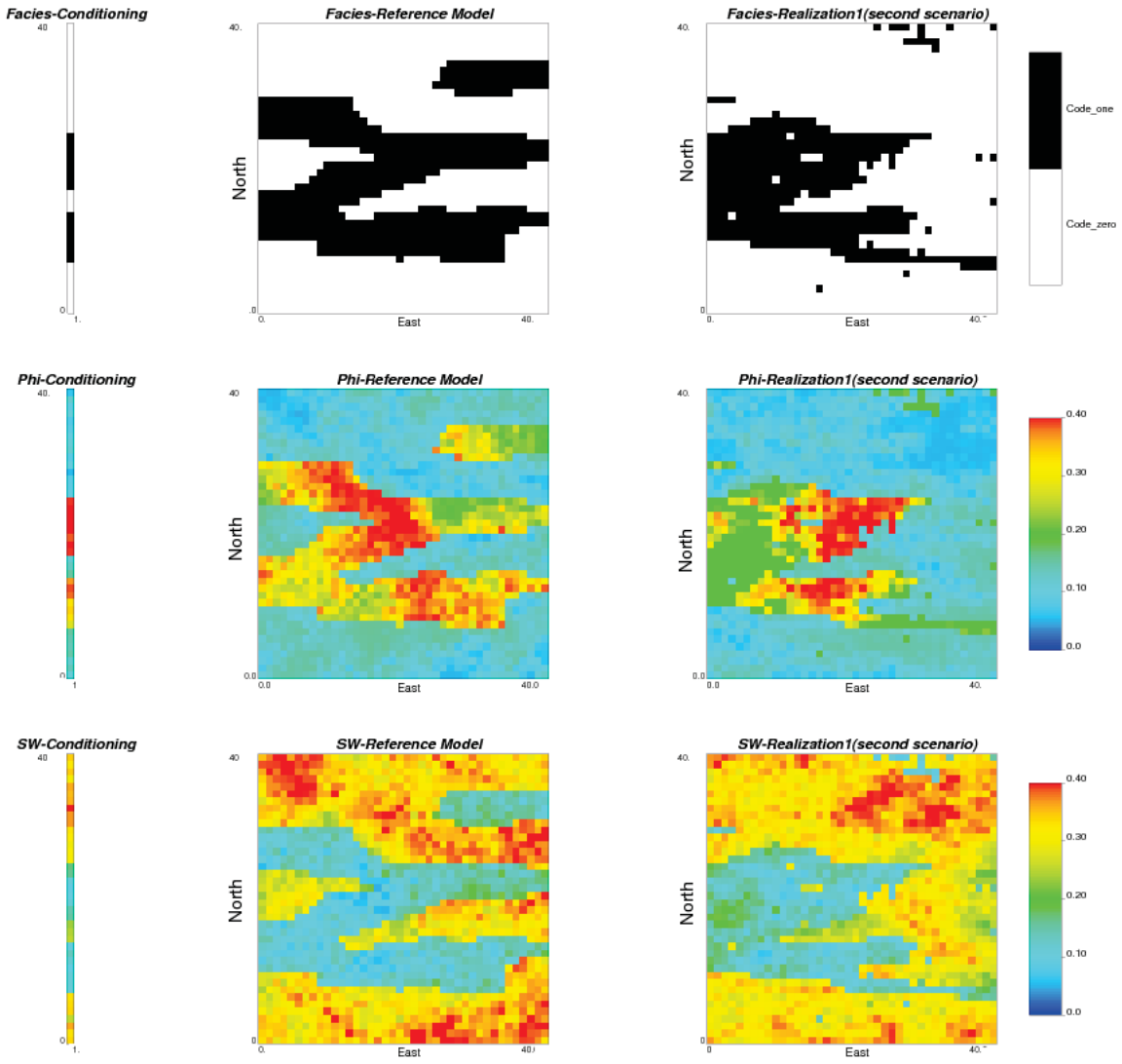

Figure 6.16: Comparison of conditioning data (first column) with different properties of reference model (second column) and first realizations of simulated properties (third column) for second scenario, axes unit m, color bar unit m^3/m^3.

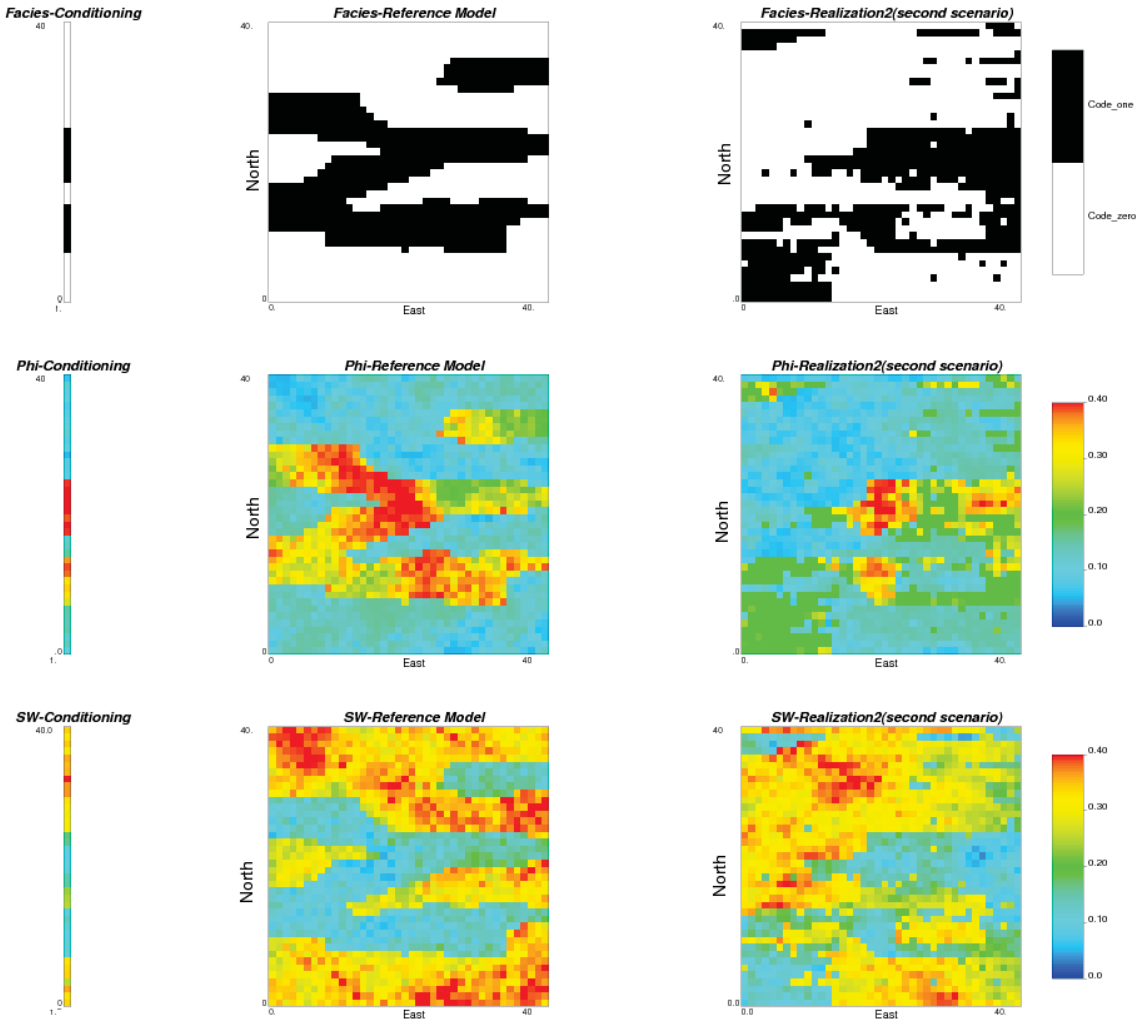

Figure 6.17: Comparison of conditioning data (first column) with different properties of reference model (second column) and second realizations of simulated properties (third column) for second scenario, axes unit m, color bar unit m^3/m^3.

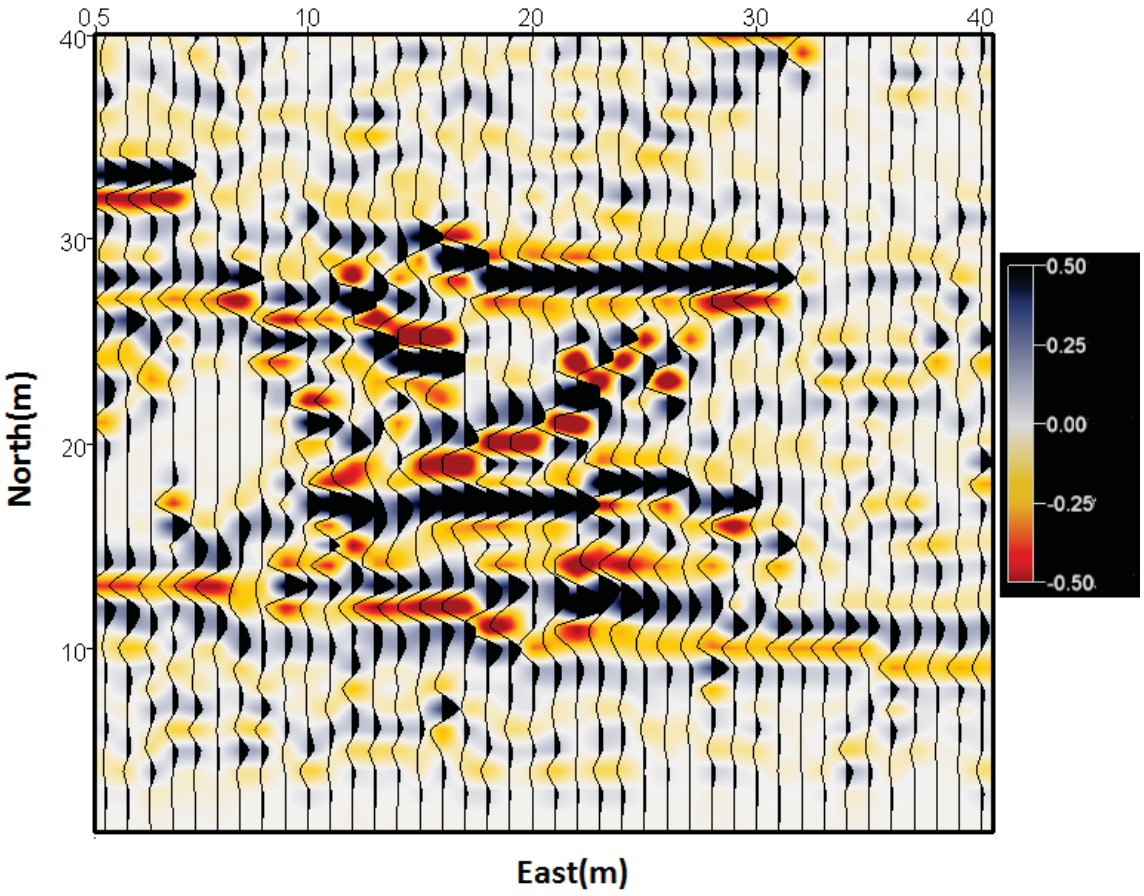

Figure 6.18: The synthetic 2-D seismic survey of first realization of second scenario conditioned to well and seismic data through the multivariate stochastic inversion, axes unit (m) and color bar - amplitude (m).

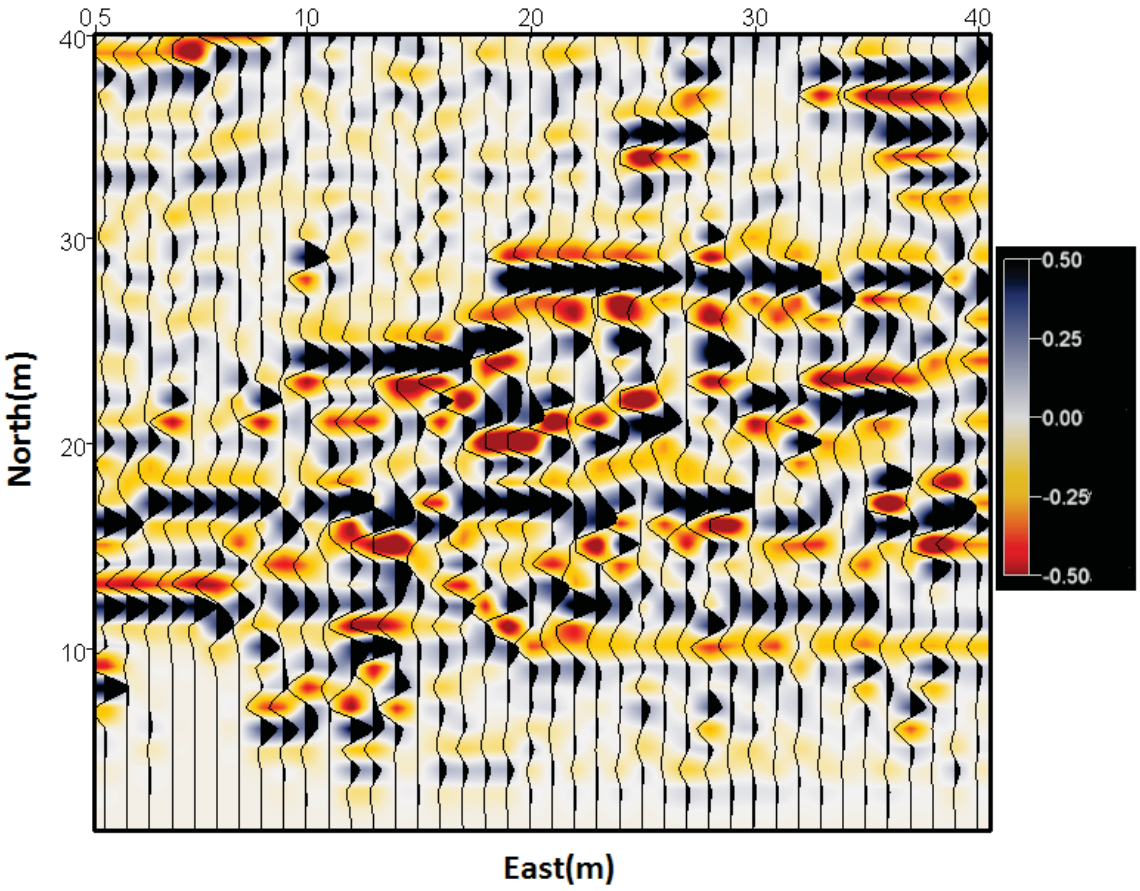

Figure 6.19: The synthetic 2-D seismic survey of second realization of second scenario conditioned to well and seismic data through the multivariate stochastic inversion, axes unit (m) and color bar - amplitude (m).

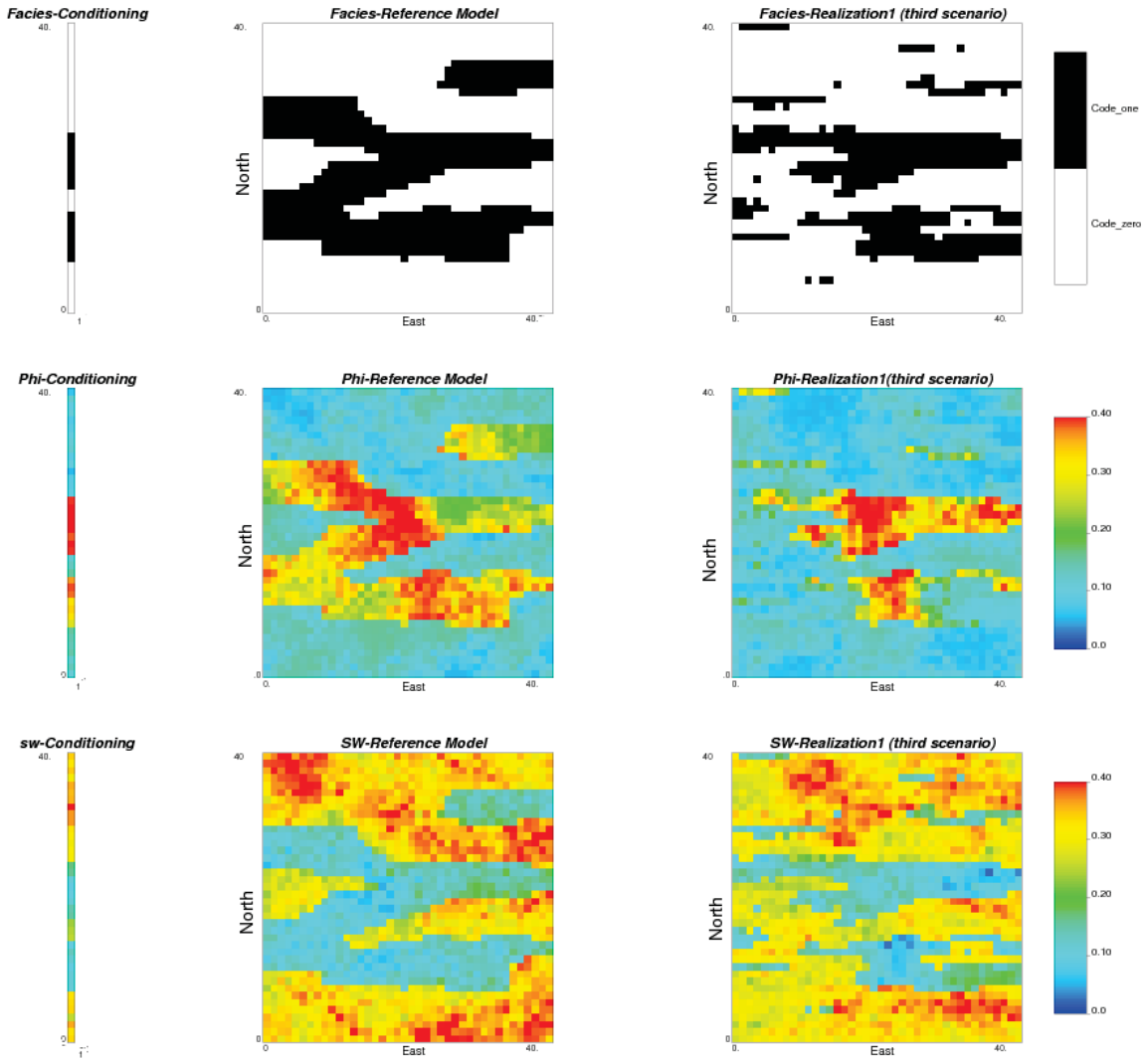

Figure 6.20: Comparison of conditioning data (first column) with different properties of reference model (second column) and first realizations of simulated properties (third column) for third scenario, axes unit m, color bar unit m^3/m^3.

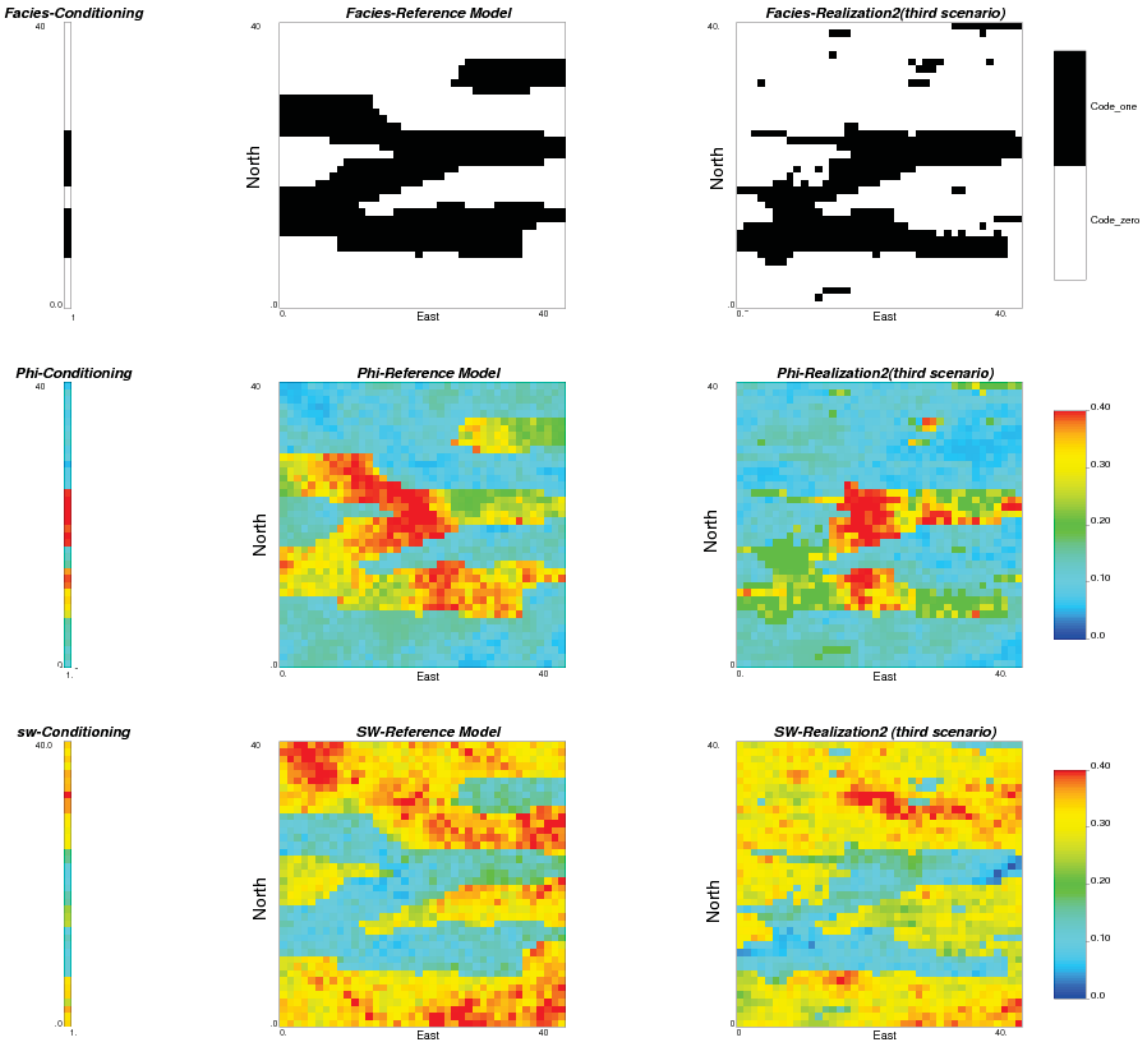

Figure 6.21: Comparison of conditioning data (first column) with different properties of reference model (second column) and second realizations of simulated properties (third column) for third scenario, axes unit m, color bar unit m^3/m^3.

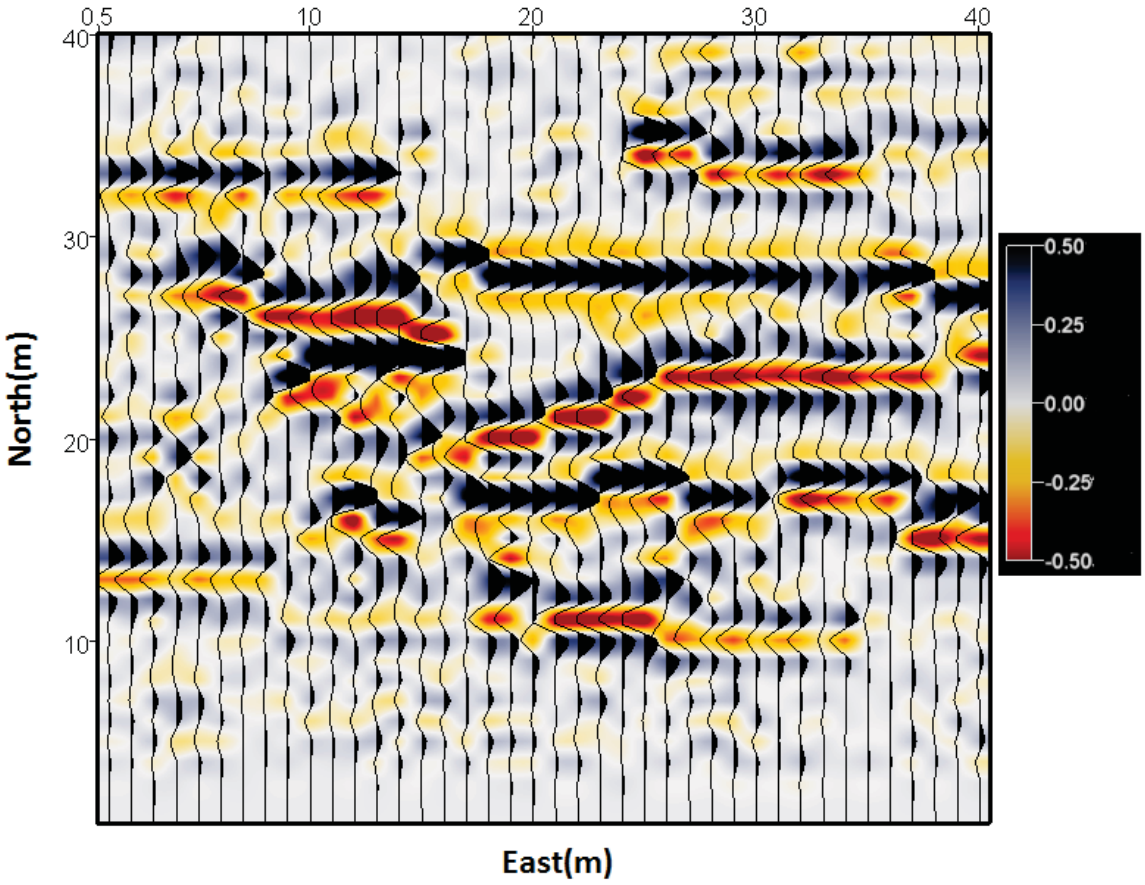

Figure 6.22: The synthetic 2-D seismic survey of first realization of third scenario conditioned to well and seismic data through the multivariate stochastic inversion, axes unit (m) and color bar -amplitude (m).

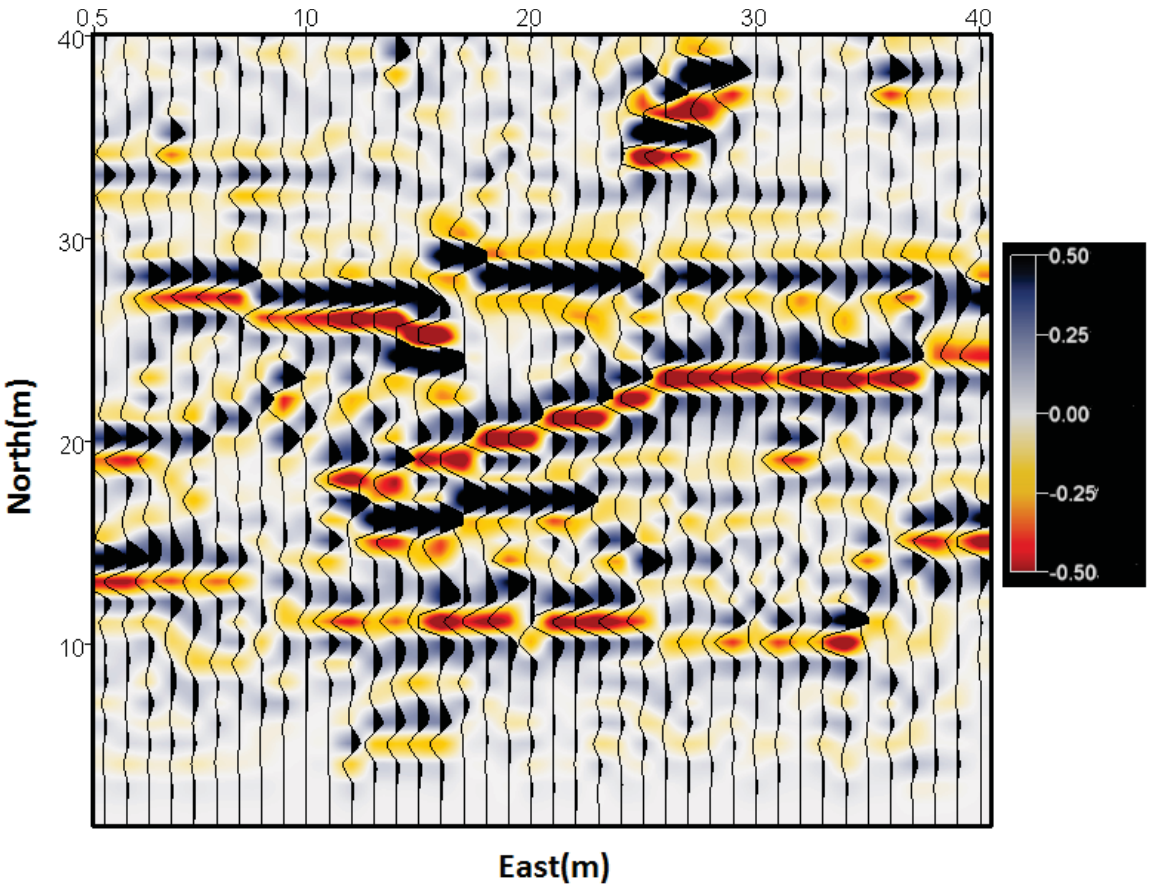

Figure 6.23: The synthetic 2-D seismic survey of second realization of third scenario conditioned to well and seismic data through the multivariate stochastic inversion, axes unit (m) and color bar -amplitude (m).

Conditioning to the seismic data via the multivariate stochastic inversion reduces the conditional variance and provides improved reservoir models. Moreover, the third scenario with more realizations shows lower conditional variance compared to the other scenarios.

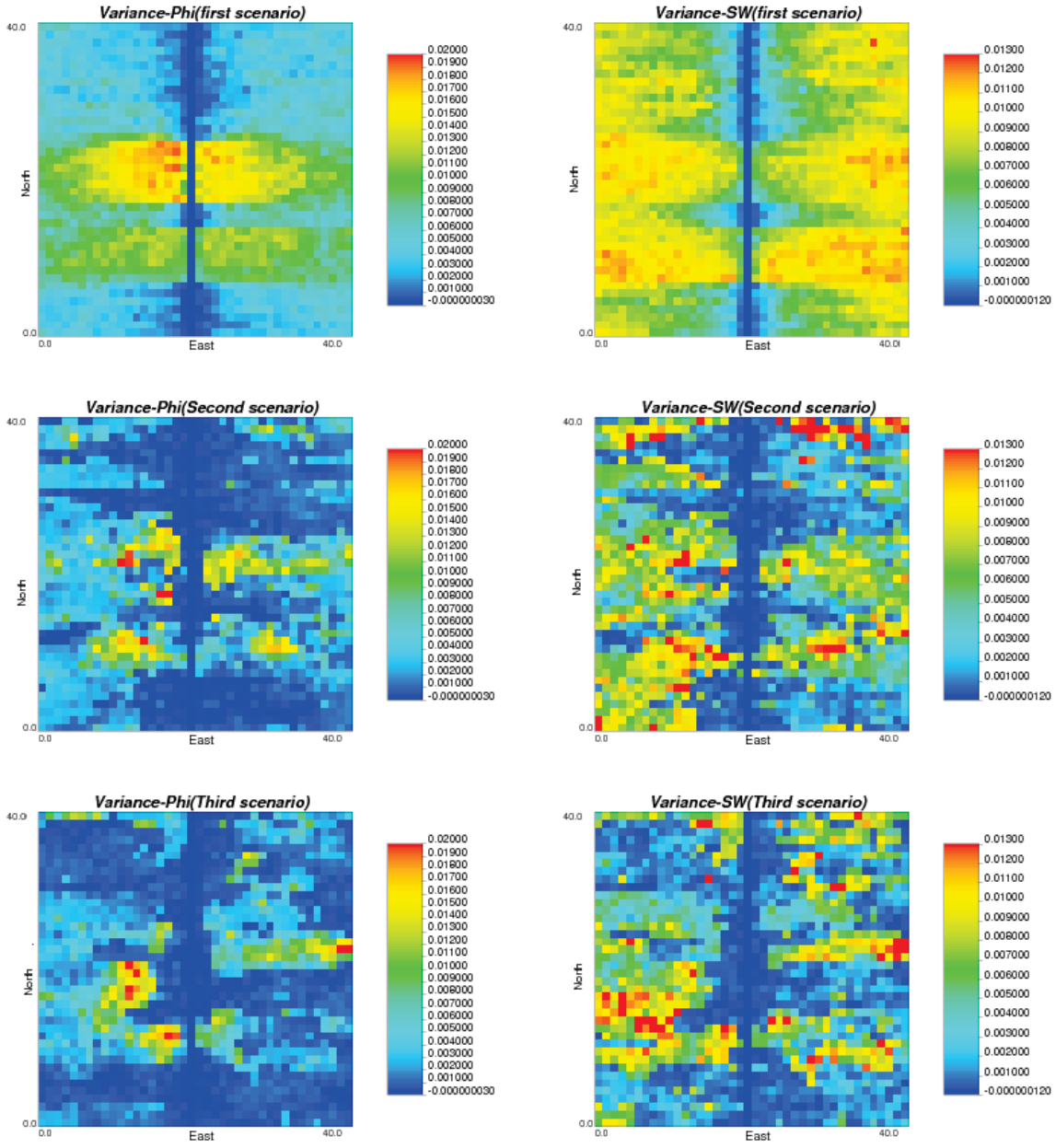

Figure 6.24: Comparison of conditional variance of multiple realizations of three scenarios for different reservoir physical properties, axes unit m.

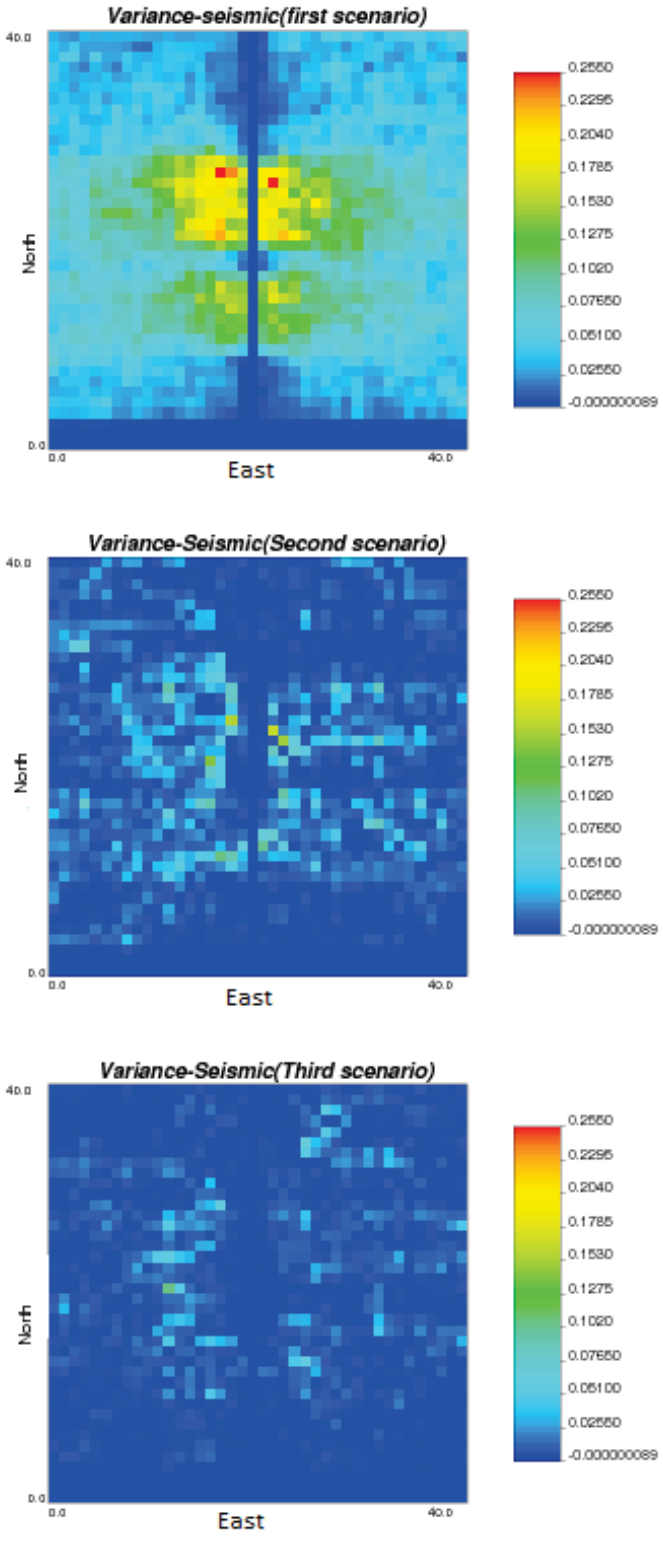

Figure 6.25: Comparison of conditional variance of multiple realizations of three scenarios for seismic data, axes unit (m).

6.3 Chapter Summary

The multivariate stochastic inversion approach applies multivariate geostatistical modeling as part of stochastic inversion technique to condition the reservoir properties to the seismic data directly. For each column of simulation domain, multiple reservoir physical and elastic properties are simulated and synthetic seismic trace related to these simulated properties is compared with collocated actual seismic trace through the adaptive sampling algorithm. The adaptive sampling algorithm selects the acceptable match with original seismic data and retains the corresponded simulated multiple reservoir properties as simulated values of the column. This selection is based on the acceptance-rejection criteria which consists of target MSE and reasonable number of realizations.

The target MSE is defined based on the cumulative uncertainty in the final model. The cumulative uncertainty of the model is attributable to different sources such as geostatistical modeling process, PEM, wavelet extraction, seismic convolution, etc. Recognition and quantification of all sources of uncertainty in the final model is difficult but would be required to accurately define the target MSE.

To obtain the acceptable match to the original seismic data and reach to the target MSE at each simulation iteration a reasonable number of realizations should be generated for each column. As the conditioning data increases over the simulation domain, the size of space of uncertainty reduces. Therefore, by adaptive sampling means more realizations generates in the early stage of simulation and the number of realizations gradually decreases through the simulation process. The dynamic number of realizations per multigrid level adapts to this observation.

Chapter 7

Case Study

The developed multivariate stochastic inversion is applied to a case study. Access to well and seismic data from a real reservoir would add value to study and make it more persuasive to some practitioners; however, access to processed 3-D seismic data is quite challenging in the petroleum industry. Moreover, Time-Depth conversion and well-tie analysis are not in the scope of this research study. A realistic data set consisting of a 3-D model of facies, reservoir physical properties such as porosity and water saturation and related 3-D seismic survey are generated and considered as the reference models for the case study. Although, the data set is not collected from a real reservoir. Multiple reservoir property models and synthetic seismic surveys obtained through the different approaches including multivariate geostatistical modeling, multivariate stochastic inversion and conventional stochastic inversion approach are generated and compared in this case study.

7.1 Reference Model

The reference model is a 3-D model of $64 \times 64 \times 32$ which each grid node is $20m \times 20m \times 1m$. Initially, a 3-D facies model with two facies categories (shale and sand) is generated by the fluvsim program and considered as reference facies model. Fluvsim is a program for object-based stochastic modeling of fluvial depositional systems (Deutsch and Tran, 2002). For porosity, two unconditional realizations are generated by sequential Gaussian simulation. Two uniform distributions between $(0.1, 0.4)$ and $(0.05, 0.15)$ are generated and transformed to the normal score units. Then, the two unconditional realizations are transformed back to the original units. The purpose is to generate porosity models for the two facies categories. The uniform distribution between $(0.1, 0.4)$ is an indicator of sand porosity and the one between $(0.05, 0.15)$ is an indicator

of shale porosity. These two porosity models are merged based on reference facies model to obtain the reference porosity model. For water saturation, two unconditional realizations that are negatively correlated ($\rho = -0.7$) with the two previous unconditional realizations for the porosity are generated by SGS algorithm. Two uniform distributions $U_1(mean, std) = U_1(0.15, 0.05)$ and $U_2(mean, std) = U_2(0.3, 0.05)$ are generated and transformed to normal score units. Then, the two unconditional realizations are transformed back to original units. The uniform distribution, $U_1(mean, std) = U_1(0.15, 0.05)$ is for sand water saturation or net water saturation. The uniform distribution, $U_2(mean, std) = U_2(0.3, 0.05)$ is for shale water saturation or non-net water saturation. The two water saturation models are merged based on the facies categories of each cell in the facies reference model to generate water saturation reference model.

The facies, porosity and water saturation models are passed to the Petro-Elastic Model to calculate reservoir elastic properties. In the reference model, the pore space consists of two phases; oil phase ($\rho_{oil} = 0.95g/cc$) and water phase ($\rho_{water} = 1g/cc$). The density is calculated by Equation 3.2 and velocity by Gardner's equation, Equation 3.4. The Gardner's parameters are considered as $a = 0.31$ and $m = 0.25$ for the reference model. The acoustic impedance is calculated as product of density and velocity. The normal incident reflection coefficient is computed by Equation 2.1. The reflectivity series is convolved with zero phase Ricker wavelet to obtain a 3-D seismic survey. This 3-D seismic survey is considered as original seismic data for this case study. For well data in this case study, three different columns of facies, porosity and water saturation model are sampled and considered as well log information. Well data are located on $(14, 56)$, $(32, 39)$ and $(40, 54)$ grid cells. Figure 7.1 shows the location of well data and axes represent grid cells.

Figure 7.2 shows the 3-D visualization of the seismic survey that is considered as the original seismic data. Three different slices of facies, porosity, water saturation reference model and original seismic survey in different orientations are shown in Figure 7.3 and 7.4. Slice $XY = 5$, $XZ = 54$ that is located close to the well 1 and $YZ = 23$ is located between well 1, 2 are displayed. The same slices of facies, reservoir physical properties and seismic

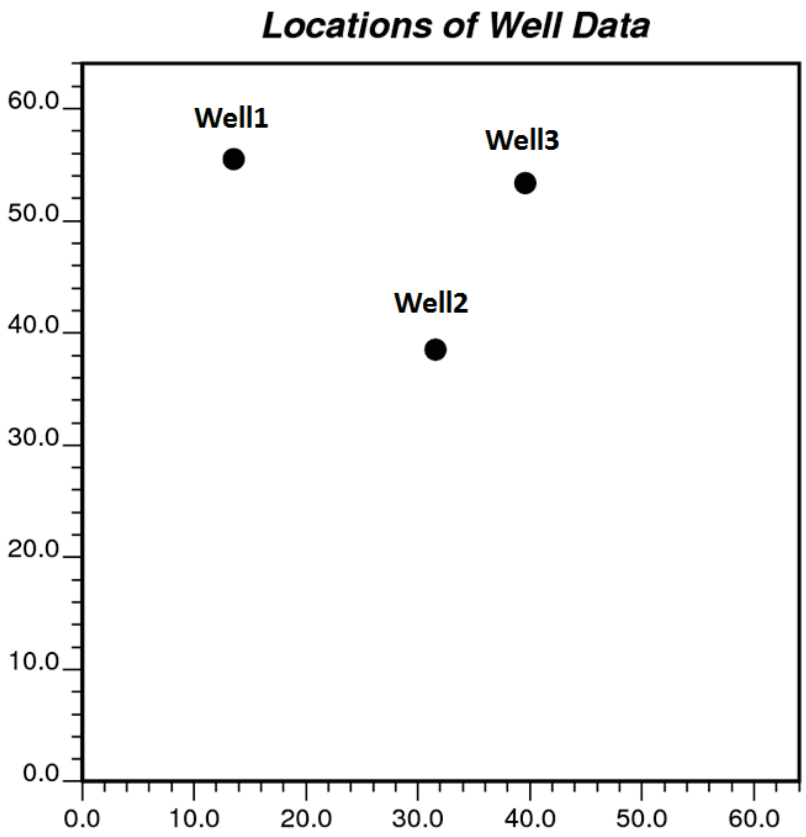

Figure 7.1: Location of well data in the case study, axes represent the grid cells.

survey that computed through the different approaches will be compared with these slices of reference models and original seismic data in the following sections. The axes unit of all slices of facies, porosity, water saturation and seismic that are displayed in this chapter for different approaches and settings represent the grid cells. The color bar unit for porosity and water saturation are m^3/m^3. Color bar unit for seismic amplitude is m.

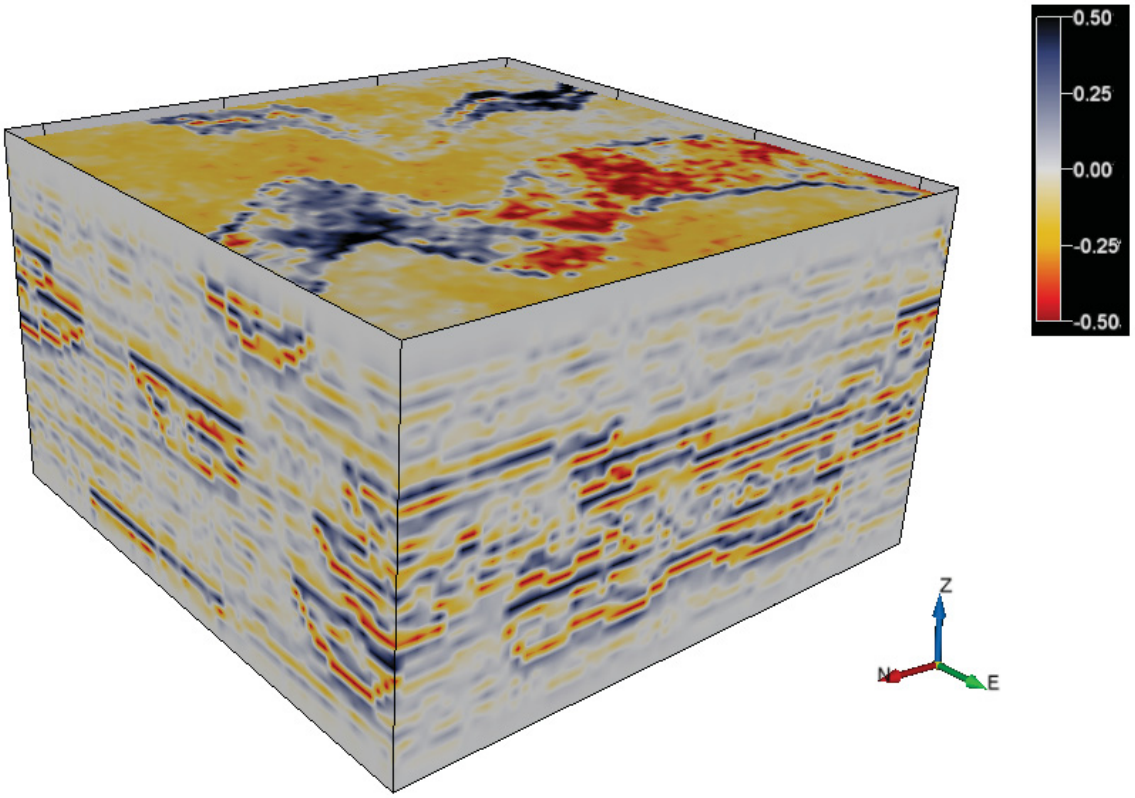

Figure 7.2: 3D visualization of seismic amplitude, considered as original seismic data for the case study, axes represent grid cells, color bar - amplitude (m).

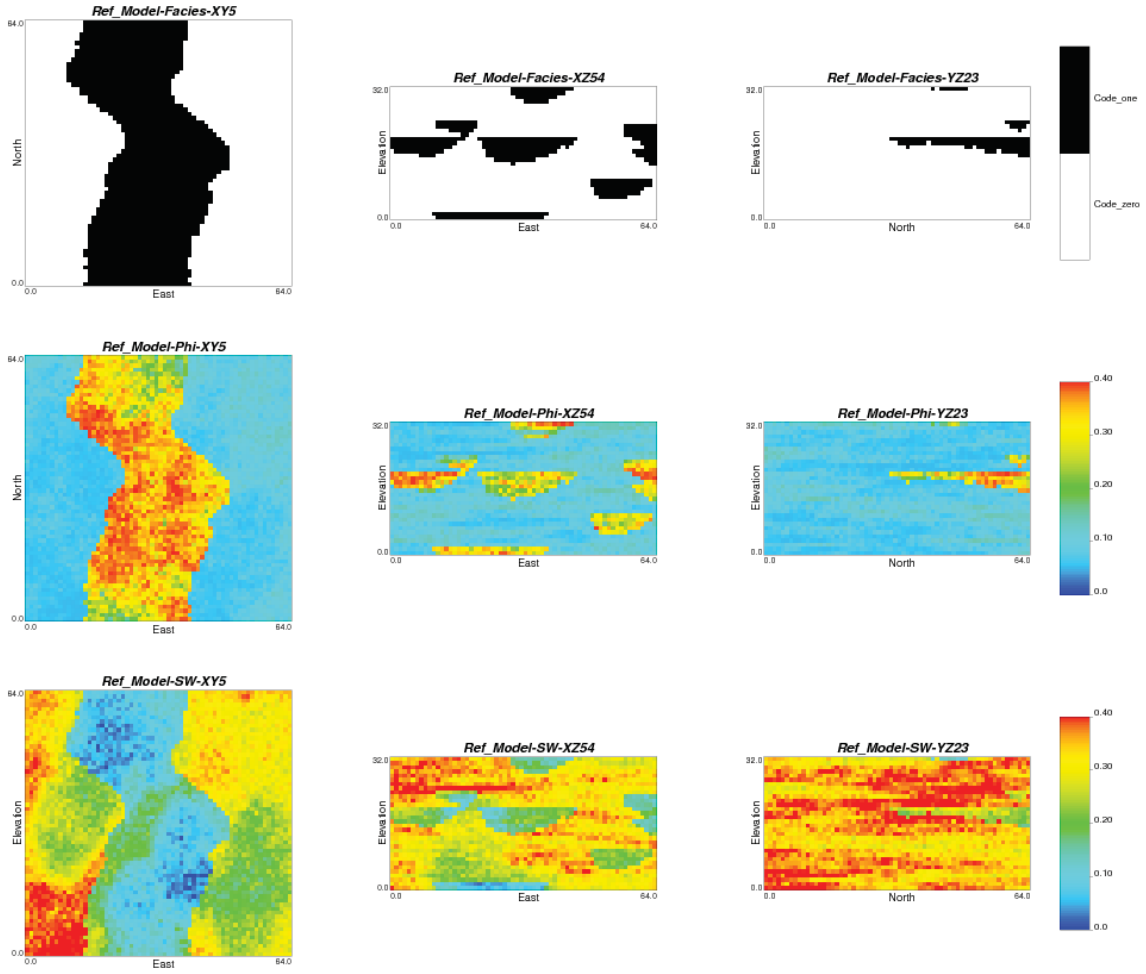

Figure 7.3: Different slices of facies, porosity and water saturation of reference model in different orientations, axes represent grid cells, color bar unit (m^3/m^3).

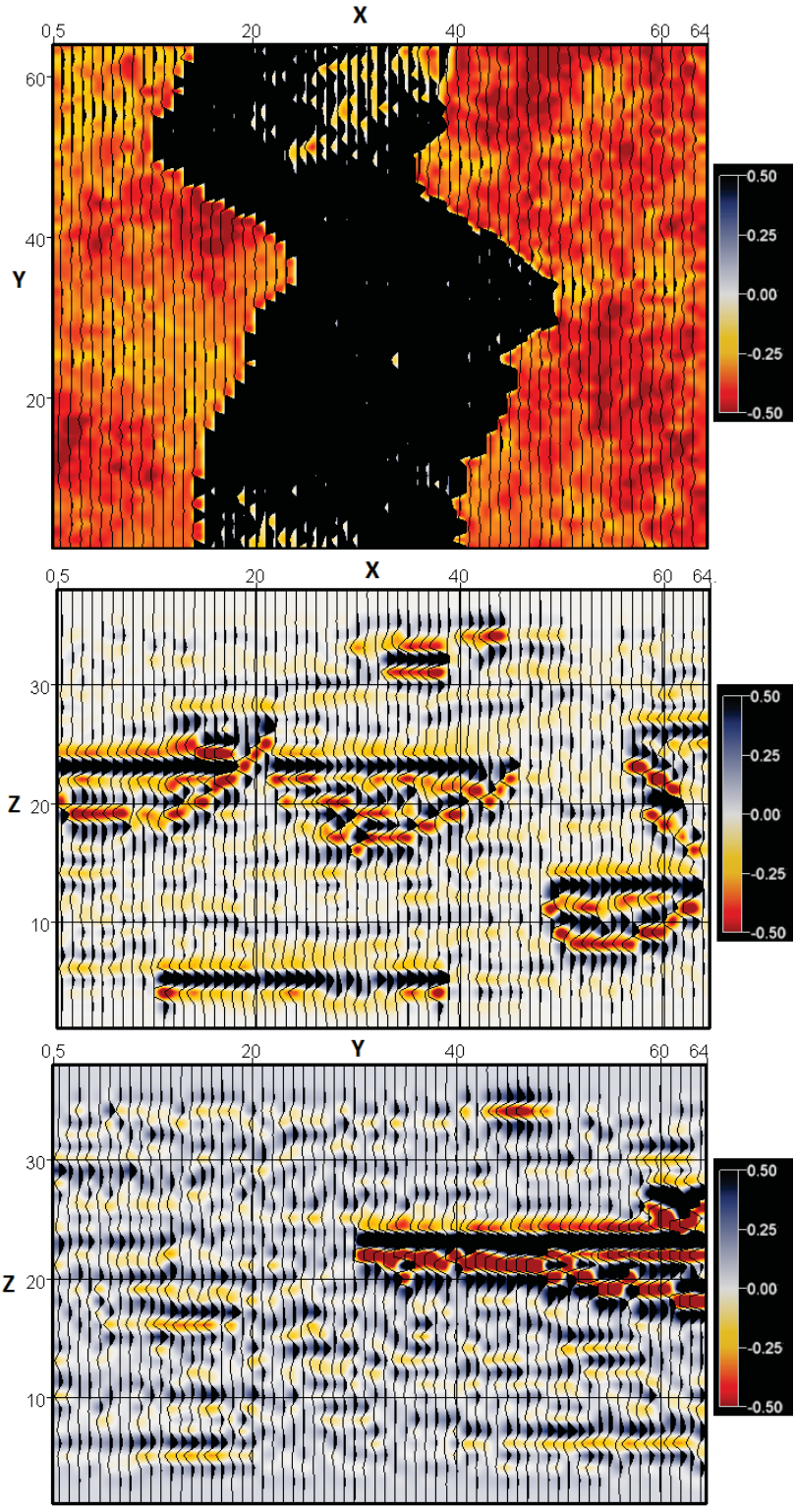

Figure 7.4: Different slices of the 3-D seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom), axes represent grid cells, color bar - amplitude (m).

7.2 Multivariate Geostatistical Method

In the first setting of the case study, facies and multiple reservoir physical properties like porosity and water saturation are simulated simultaneously condition to the well data by the multivariate Gaussian simulation technique through the column based ultimate sequential Gaussian simulation (USGSIM) algorithm. Facies simulated by TGS, porosity and water saturation are simulated by SGS. In this case the constant number of realizations per column, 100, are generated. Multiple realizations of facies, porosity and water saturation are passed to PEM and convolution algorithm to compute the synthetic seismic cubes through the forward modeling process. The synthetic seismic cubes are compared to the original seismic survey to compute the mismatch. Figure 7.5 and 7.6 show the best two realizations of facies, porosity and water saturation. Figure 7.7 and 7.8 also show the synthetic seismic cubes of these realizations that have lower mismatch with the original seismic data compared to the other realizations. In general, for case of $64 \times 64 \times 32$ with 100 realizations per column, 409600 column of realizations are generated.

As shown in Figures 7.5 and 7.6 and also comparing them with the reference models, the reservoir property models that are conditioned only to the well data do not capture the spatial variability and heterogeneity of reservoir. Well data do not provide much information on lateral variability of reservoir properties due to low areal resolution. According to the Figure 7.7 and 7.8 and comparing them with the original seismic data, these models also could not reproduce the original seismic data. In these figures, the results of slice $XZ = 54$ are better compare to the other slices as this slice is close to the well 1 data. However, the slices that are away from the wells are more noisy compare to the slice is close to the well location.

Integration of seismic data with high areal coverage and well data with high vertical resolution into geostatistical modeling significantly improves the quality of reservoir models. Stochastic seismic inversion is the most common method that directly accounts for the seismic data integration into geostatistical modeling. In the next two settings for the case study, multiple reservoir properties and synthetic seismic cubes are generated by developed multivariate stochastic inversion and conventional stochastic inversion

approach and the results will be compared.

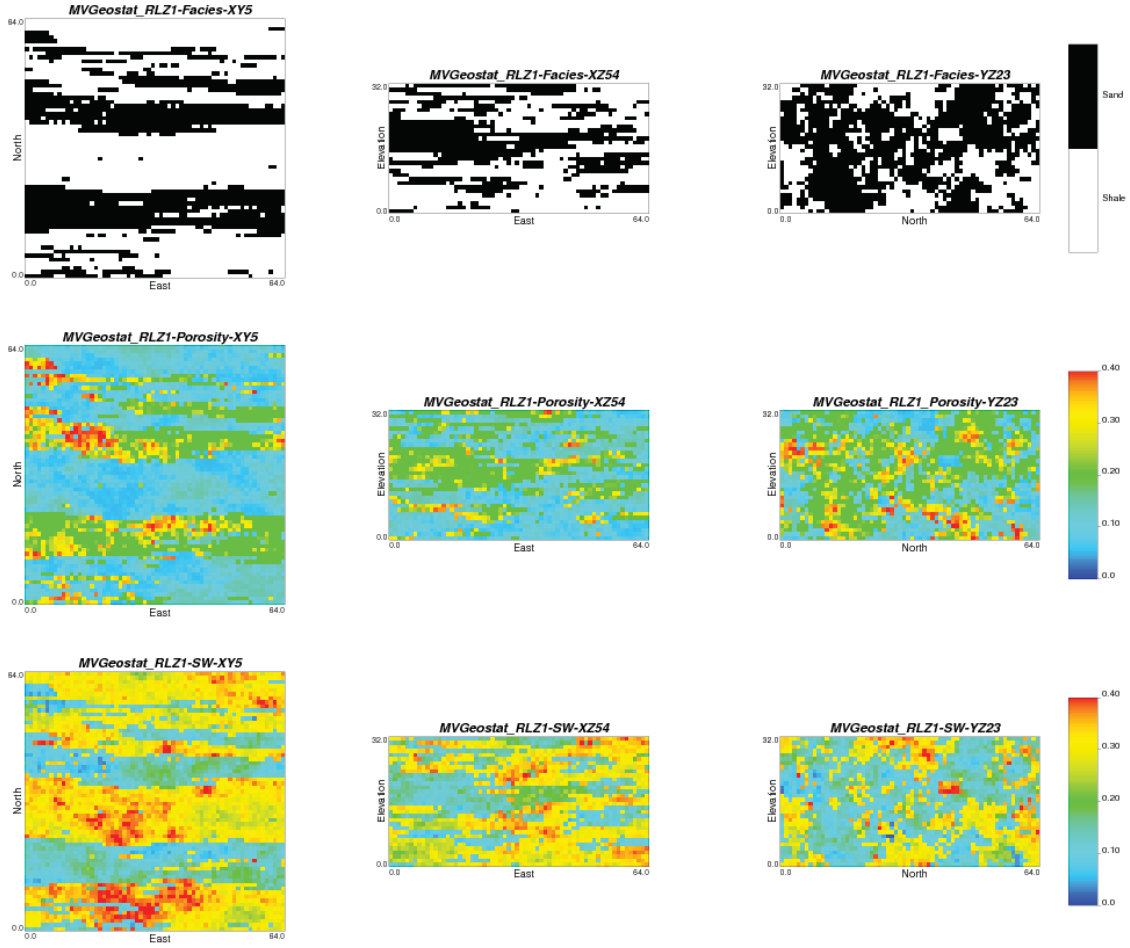

Figure 7.5: Different slices of facies, porosity and water saturation of first realization in different orientations for multivariate geostatistical modeling that are conditioned to the well data, axes represent grid cells, color bar unit (m^3/m^3).

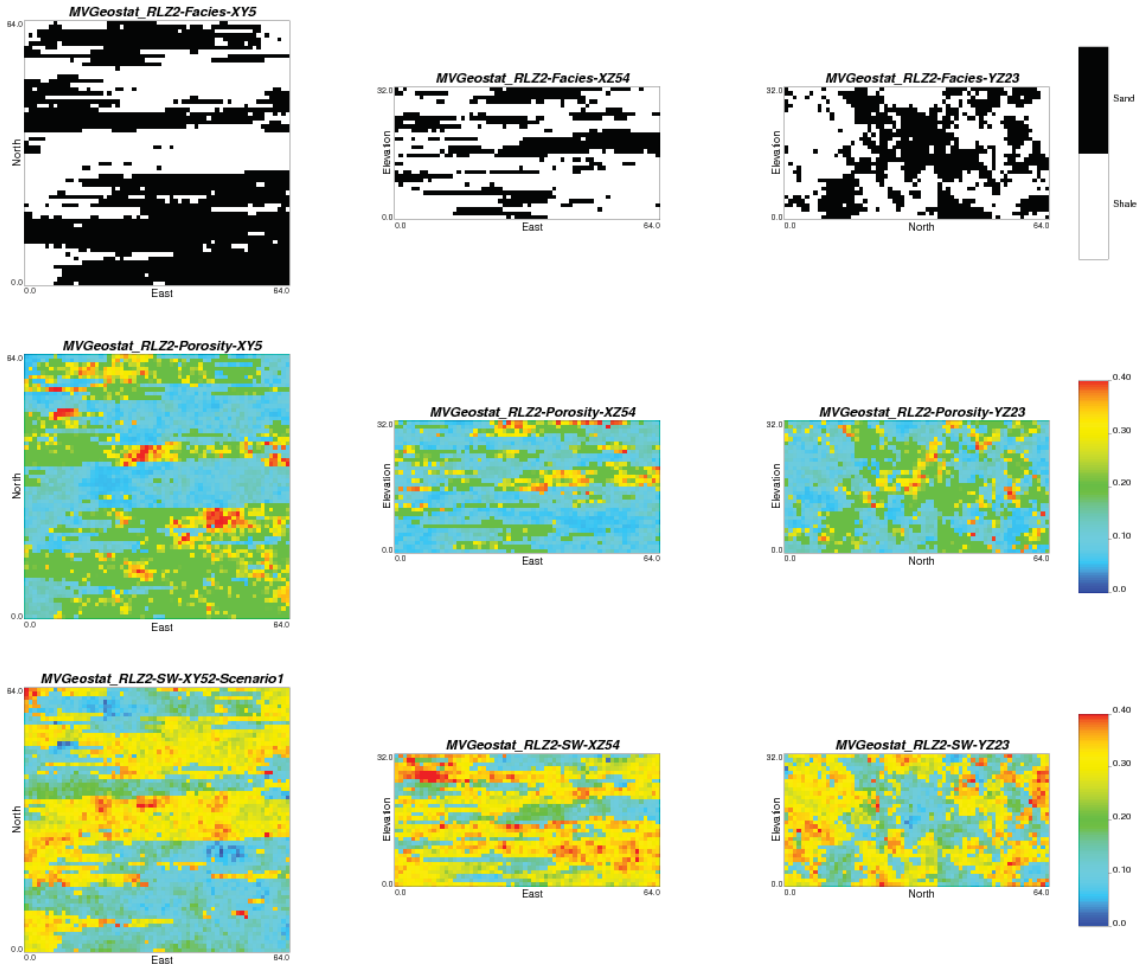

Figure 7.6: Different slices of facies, porosity and water saturation of second realization in different orientations for multivariate geostatistical modeling that are conditioned to the well data, axes represent grid cell, color bar unit (m^3/m^3).

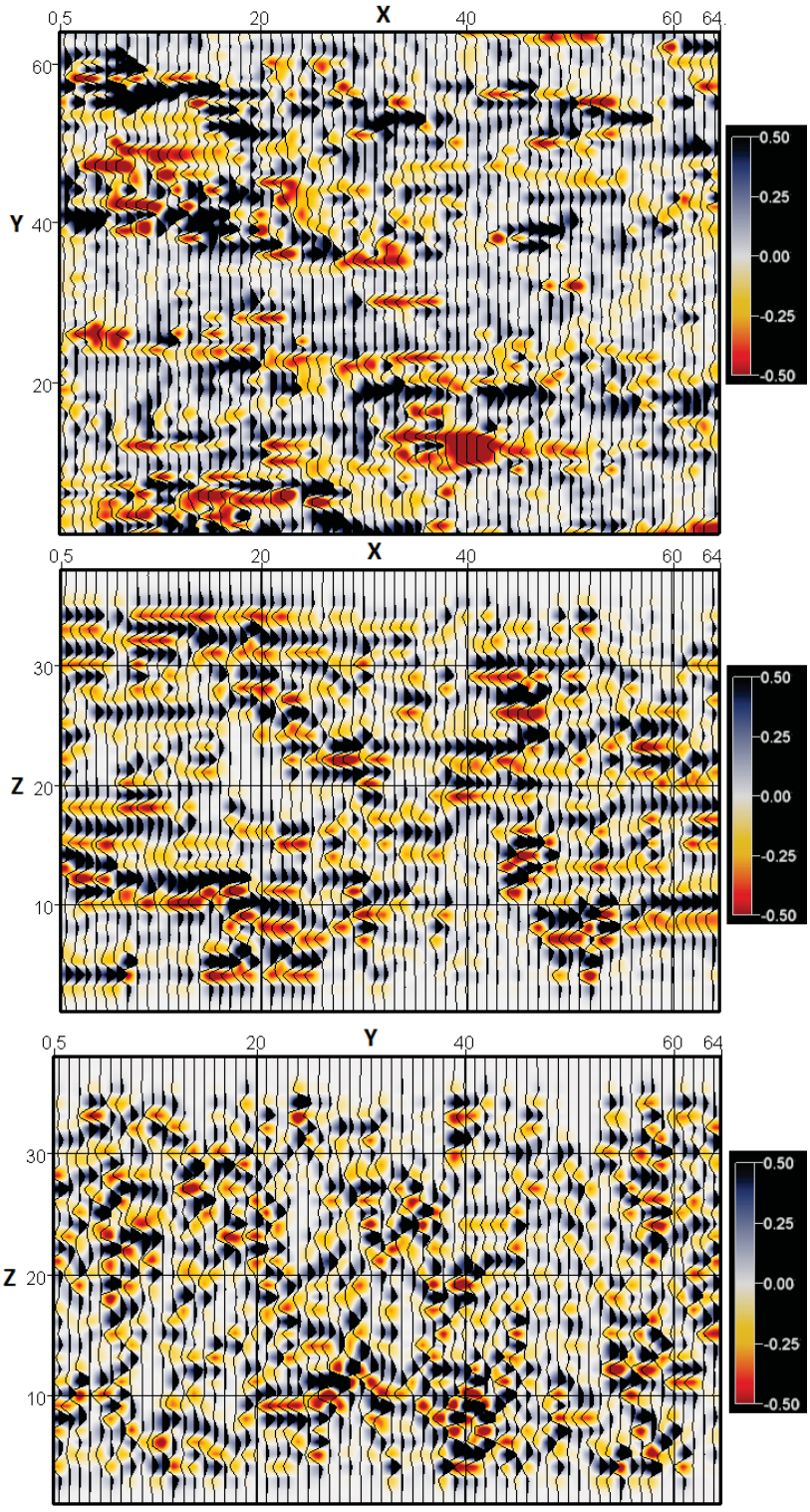

Figure 7.7: Different slices of the synthetic seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom) for the first realization of multivariate geostatistical modeling, axes represent grid cells, color bar - amplitude (m).

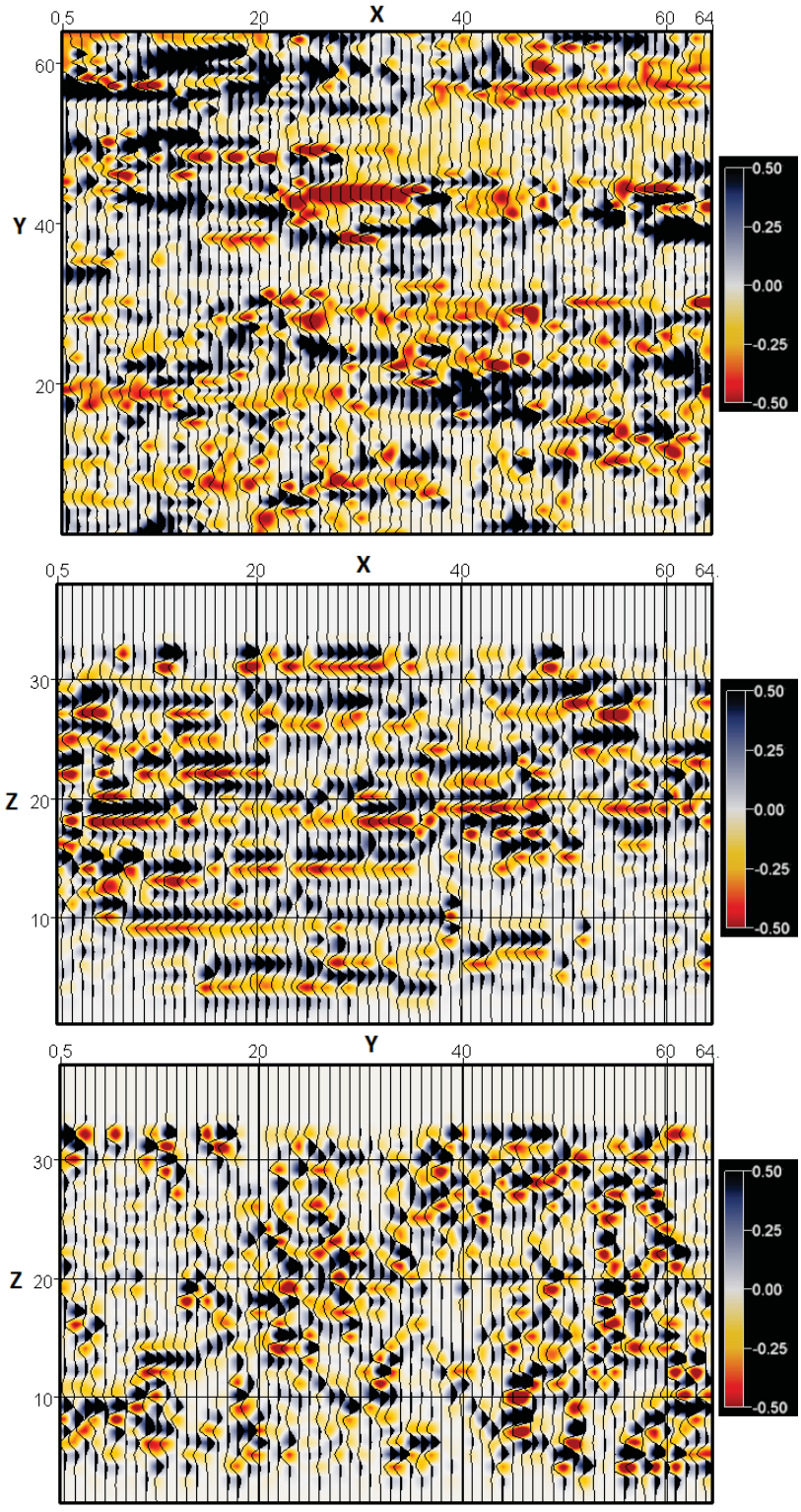

Figure 7.8: Different slices of the synthetic seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom) for the second realization of multivariate geostatistical modeling, axes represent grid cells, color bar - amplitude (m).

7.3 Multivariate Stochastic Inversion Method

Now the facies, porosity and water saturation are simulated simultaneously conditioned to the well and seismic data through the developed multivariate stochastic inversion approach. The stopping criteria in this setting is considered as follows. The target MSE is in order of 10^{-3} that defined based on inherent uncertainty of PEM model and extracted wavelet. Five levels of multigrid are considered for the simulation process, $mmult = 5$. A dynamic number of realizations per multigrid level is considered. To help choose how many realizations are required per multigrid level, a few columns of each multi-grid level are simulated. A different number of realizations are generated for each column and minimum mismatch between synthetic and original seismic data (minimum MSE) for each set of multiple realizations is recorded. This procedure is replicated for each column until a relative stable minimum MSE is reached. Figure 7.9 shows the minimum MSE versus different number of realizations for third level of multigrid ($mmult = 3$). As shown in this figure, for most of the columns after 100 number of realizations the MSE changes slightly over more number of realizations and the curves reach to plateau. Therefore, 100 number of realizations is considered for this level of multigrid. The number of realizations per multigrid level for this setting is displayed in Table 7.1. As shown in this table, the simulation process starts with a large number of realizations for the first level of multigrid and gradually reduces this number for lower levels of multigrid that consider more conditioning data. The total number of realizations for all the columns are generated is 38272 for this process which is less than the number of realization for the case of generating a constant number of realizations, 100, per column which is 409600.

Two different realizations of facies, porosity and water saturation for this setting are displayed in Figure 7.10 and 7.11. The corresponded synthetic seismic cubes of these realizations are also shown in Figure 7.12 and 7.13. As shown in these figures and also comparing them with reference models, multivariate stochastic inversion approach by applying geostatistical modeling as part of stochastic inversion technique to condition reservoir model to the seismic data directly provides reservoir models that reproduce the

seismic data.

Table 7.1: Multigrid level, number of columns simulated, number of realization and total number of realizations per multigrid level.

Multigrid level	No.Columns	No.Realization	Total NO.Realization
5	$[4 \times 4]$	1000	16000
4	$[(8 \times 8) - (4 \times 4)]$	200	9600
3	$[(16 \times 16) - (8 \times 8)]$	100	1920
2	$[(32 \times 32) - (16 \times 16)]$	10	7680
1	$[(64 \times 64) - (32 \times 32)]$	1	3072
			$38272 < 409600$

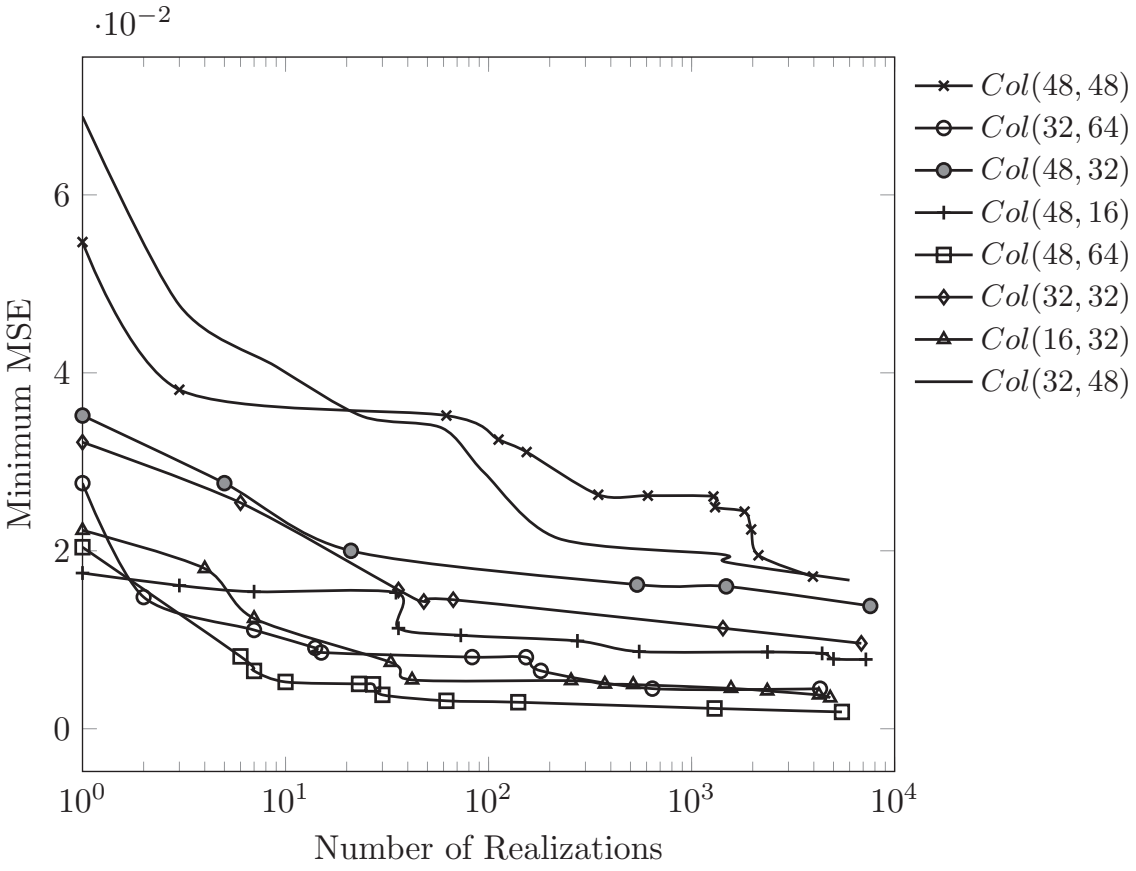

Figure 7.9: Display minimum MSE versus different sets number of realization for few columns of third level of multigrid ($mmult = 3$) in multivariate stochastic inversion approach for the case study.

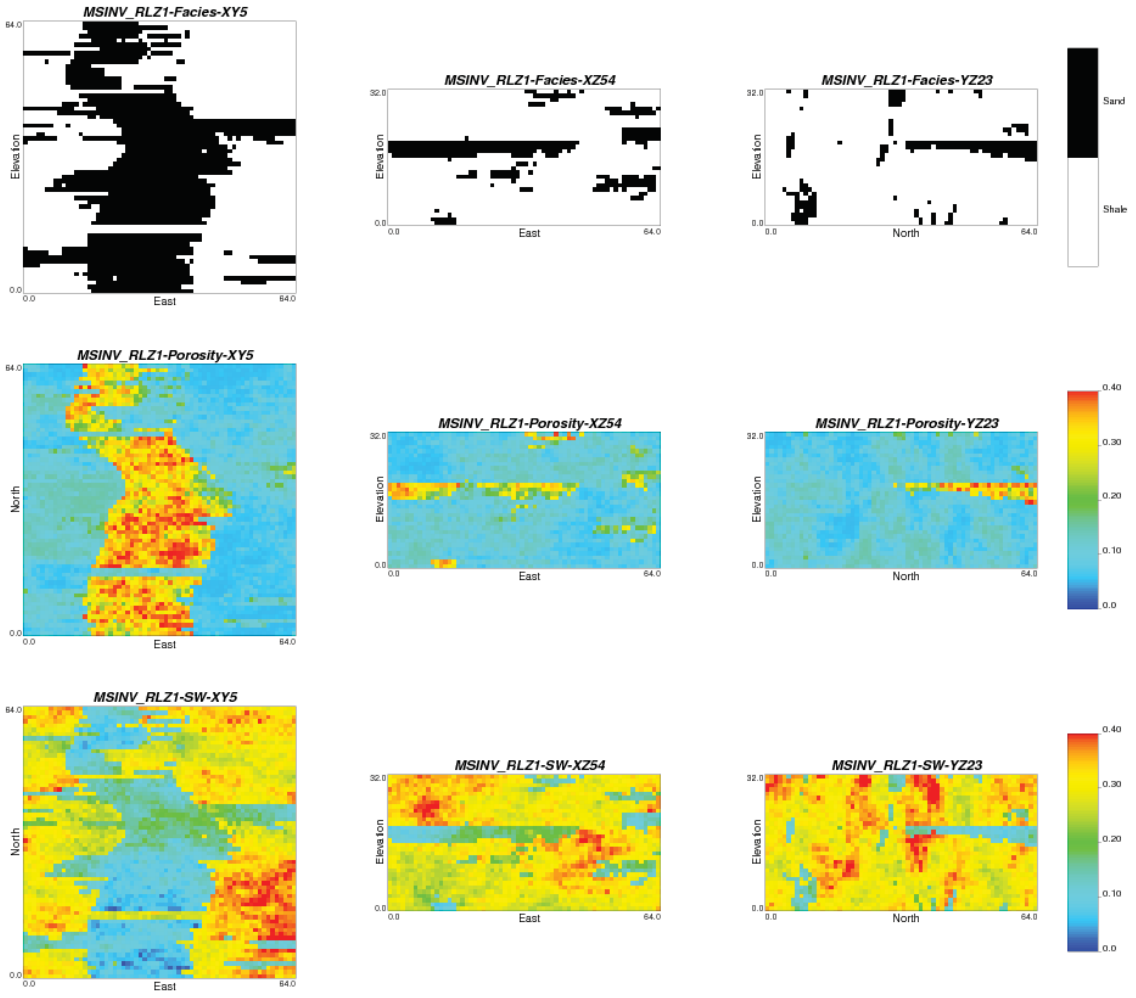

Figure 7.10: Different slices of facies, porosity and water saturation of first realization in different orientations for multivariate stochastic inversion approach that are conditioned to the well and seismic data, axes represent grid cells, color bar unit (m^3/m^3).

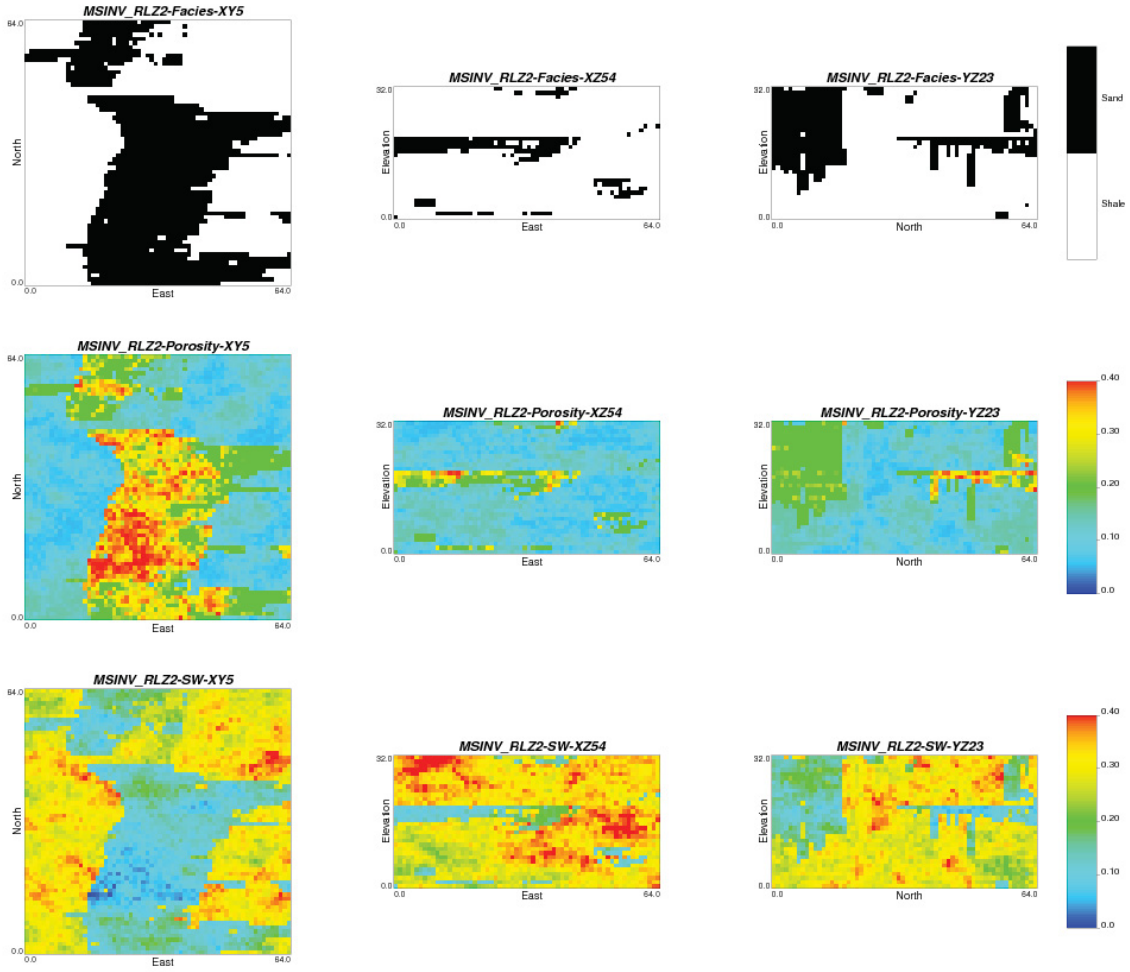

Figure 7.11: Different slices of facies, porosity and water saturation of second realization in different orientations for multivariate stochastic inversion approach that are conditioned to the well and seismic data, axes represent grid cells, color bar unit (m^3/m^3).

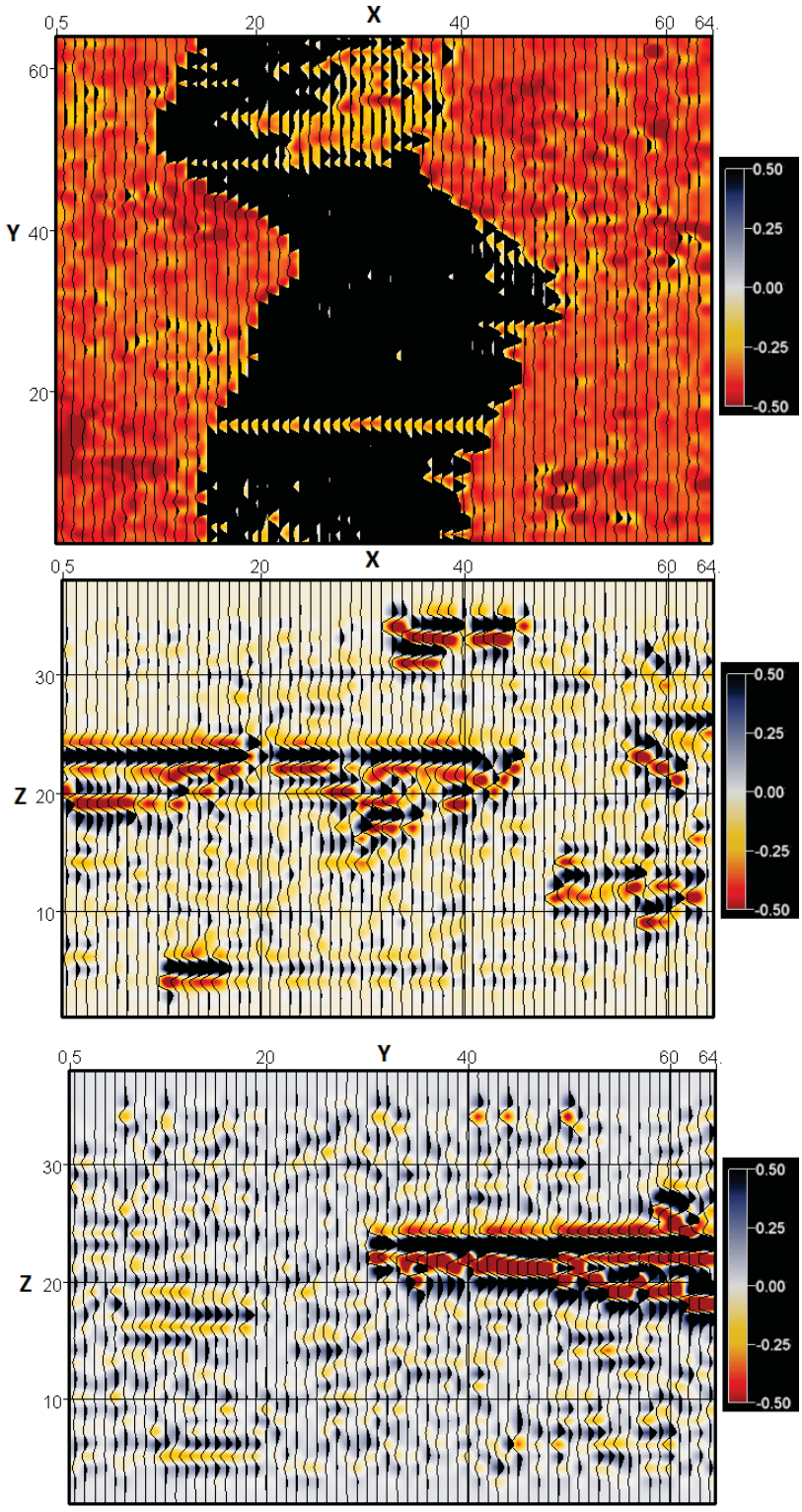

Figure 7.12: Different slices of the synthetic seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom) for the first realization of multivariate stochastic inversion approach, axes represent grid cells, color bar - amplitude (m).

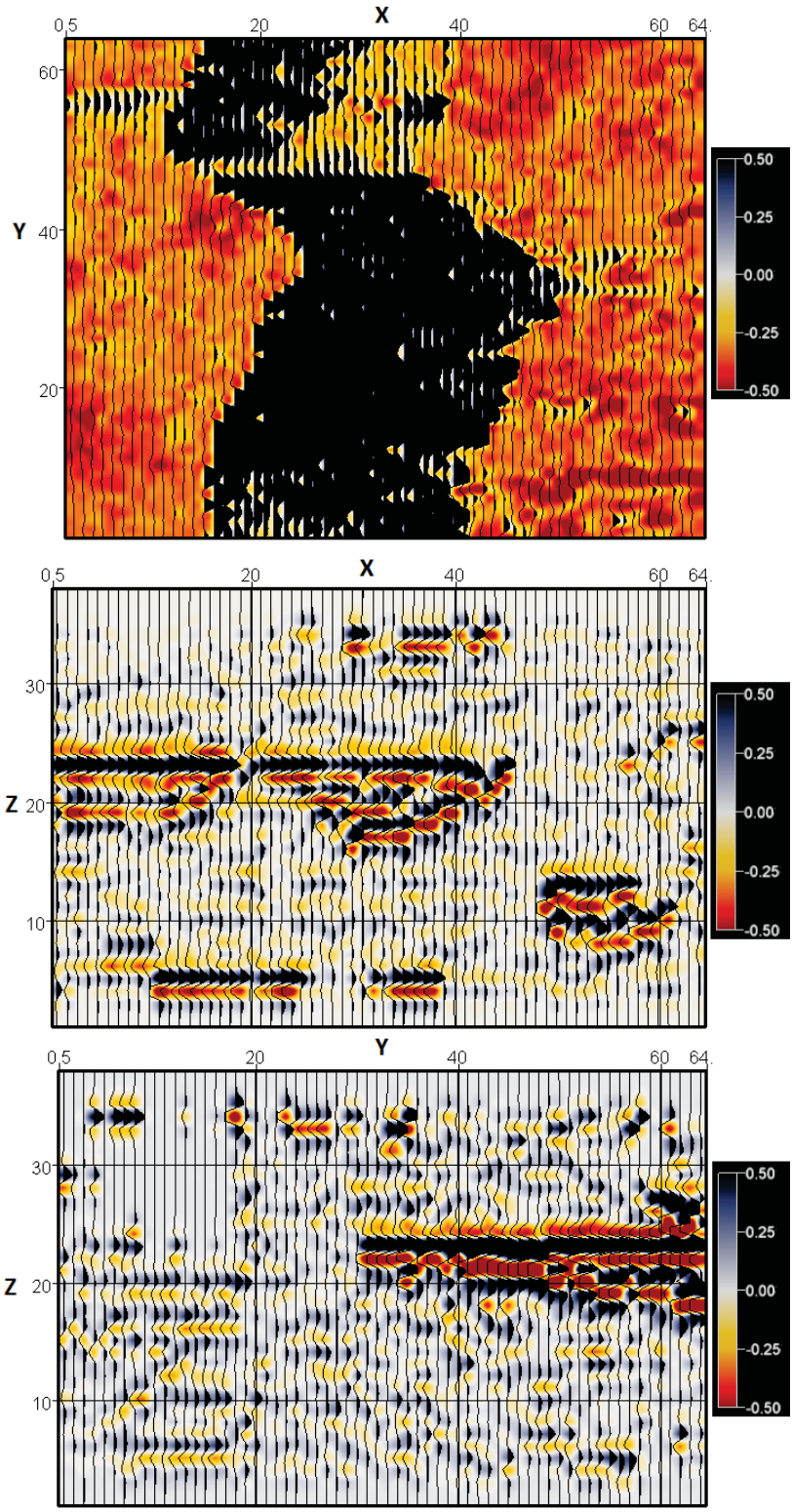

Figure 7.13: Different slices of the synthetic seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom) for the second realization of multivariate stochastic inversion approach, axes represent grid cells, color bar - amplitude (m).

7.4 Conventional Stochastic Inversion Approach

In this section the facies, porosity and water saturation are simulated conditioned to the well and seismic data by conventional stochastic inversion. In conventional stochastic inversion, a high quality acoustic impedance or facies is simulated and connected to the reservoir physical properties such as porosity through a set of sequential connections. The methodology of conventional stochastic inversion applied in the case study followed as:

1. Calculate proportion of facies categories on an 8 cell moving average at well locations
    - In general the vertical resolution of seismic data is coarser than vertical resolution of well data, so moving average is applied
    - Proportion or probability of shale is denoted by P_{sh} and for sand is $P_s = 1 - P_{sh}$
2. Extract acoustic impedance (AI) at well locations
3. Cross plot AI versus proportion at well location to obtain vertical proportion curve (VPC), Figure 7.14
4. Convert AI to probability model cell by cell
5. Extract proportions of facies categories at well locations and convert them to the Gaussian units
    - $y_{sh} = G^{-1}((P_{sh})/2)$ and $y_s = G^{-1}((1 + P_{sh})/2)$
6. Apply truncated Gaussian simulation (TGS) with local proportion to simulate facies
    - Take care of conditioning data to consider the area with low and high AI which means preparing the conditioning data for each facies categories separately
7. Cross plot porosity of shale and sand versus AI to obtain the correlation coefficient per facies categories
8. Run sequential Gaussian simulation (SGS) by facies considering collocated cokriging (CCK) to simulated porosity of shale and sand
9. Merge two porosity models based on facies categories of each cell at simulation domain

10. Simulate water saturation for shale and sand by SGS correlated with porosity ($\rho = -0.7$) and merge them based on facies categories of each cell at simulation domain
11. Pass facies and reservoir physical properties to the PEM to calculate reservoir elastic properties
12. Convolve reflectivity series with extracted wavelet to obtain synthetic seismic cube via the forward modeling

Figure 7.15 and 7.16 show different slices of the facies, porosity and water saturation for two different realizations obtained by conventional stochastic inversion approach. By comparing the conventional results with reference models, the conventional approach generated high quality facies. Although, the resolution of reservoir physical properties that obtained through the sequential connections are reduced gradually. Figure 7.17 and 7.18 display the synthetic seismic cubes related to the two realizations of conventional stochastic inversion approach. According to these figures and based on visual assessment, the reservoir properties model generated by conventional stochastic inversion approach do not reproduce the original seismic data within the quality of the data. Moreover, the MSE between the seismic reference model and synthetic seismic cube obtained via the multivariate stochastic inversion is 7.8×10^{-3} while the MSE for conventional stochastic inversion is 6.4×10^{-2}.

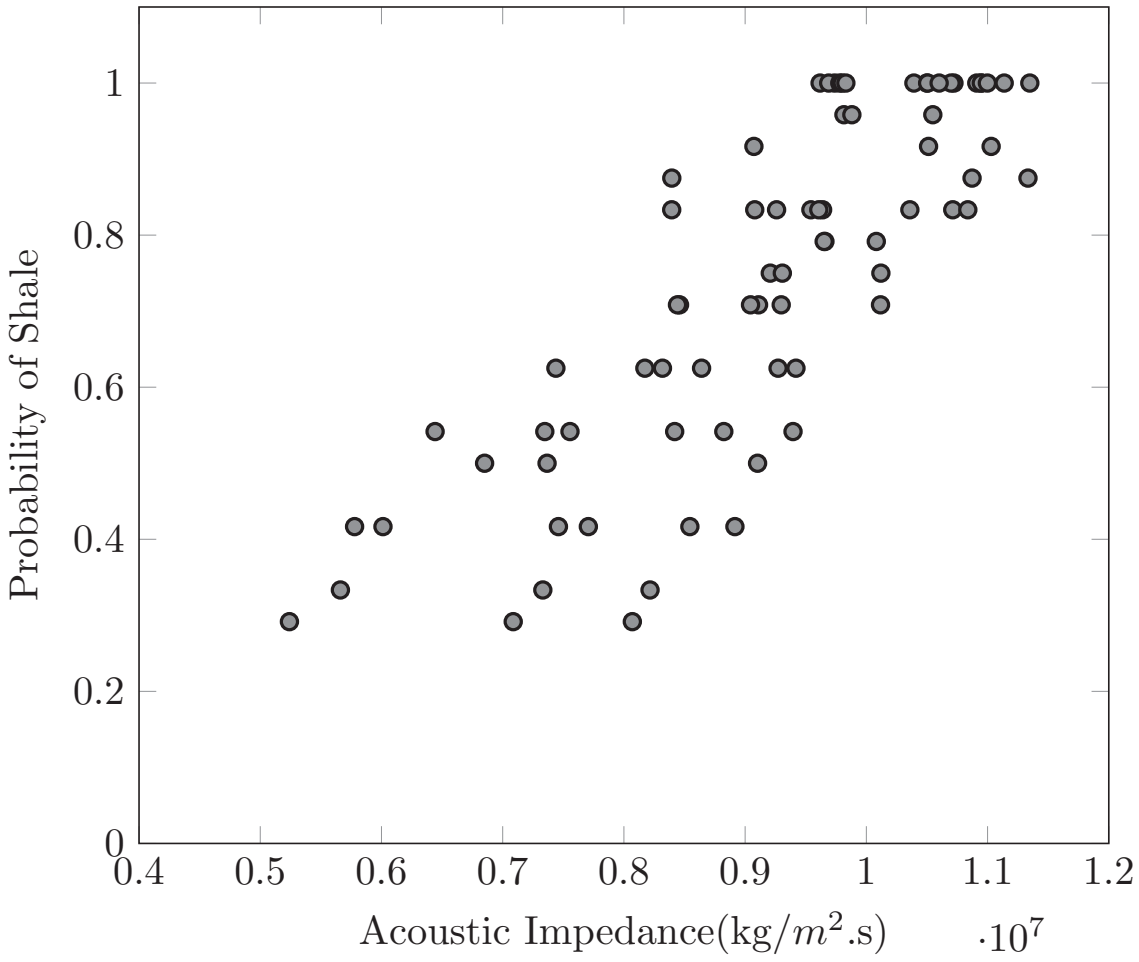

Figure 7.14: Shows the relationship between probability of shale versus acoustic impedance for probability model.

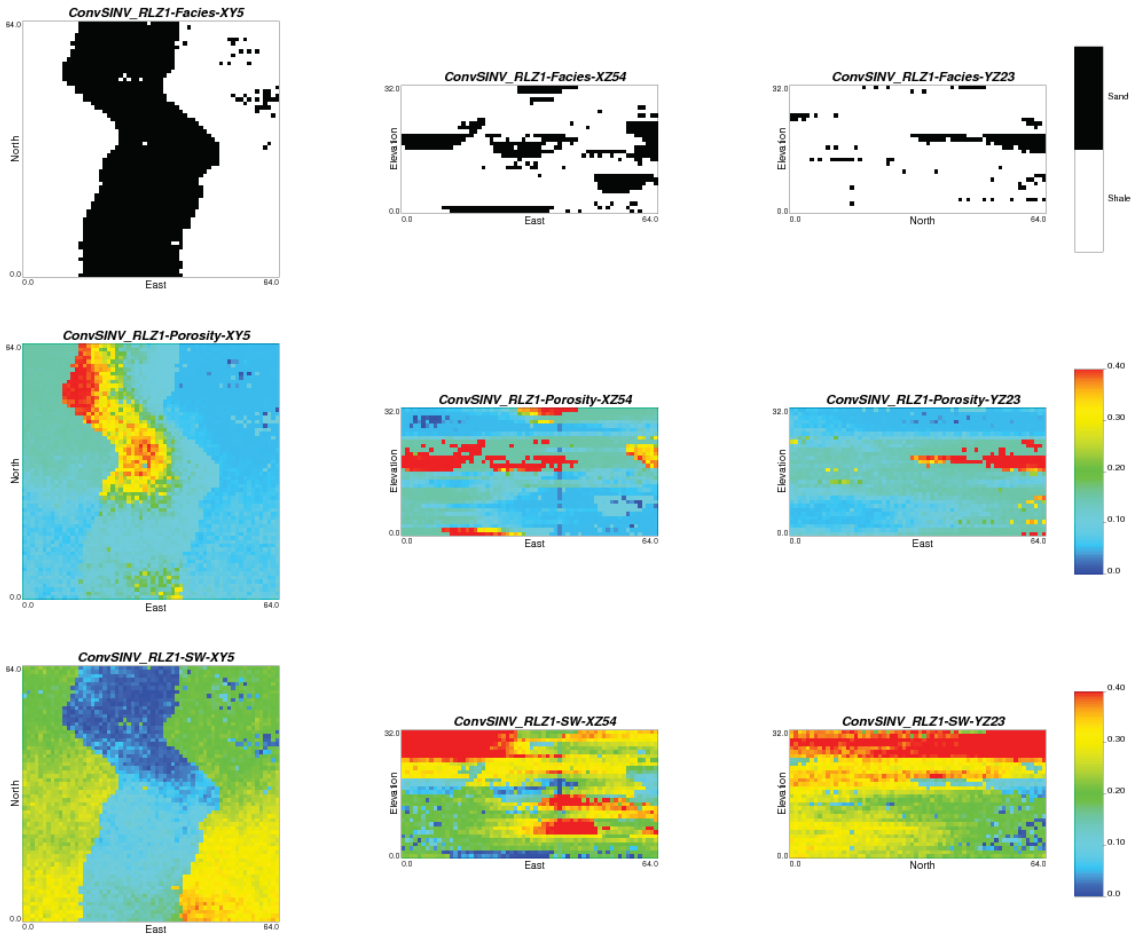

Figure 7.15: Different slices of facies, porosity and water saturation of first realization in different orientations for conventional stochastic inversion approach that are conditioned to the well and seismic data, axes represent grid cells, color bar unit (m^3/m^3).

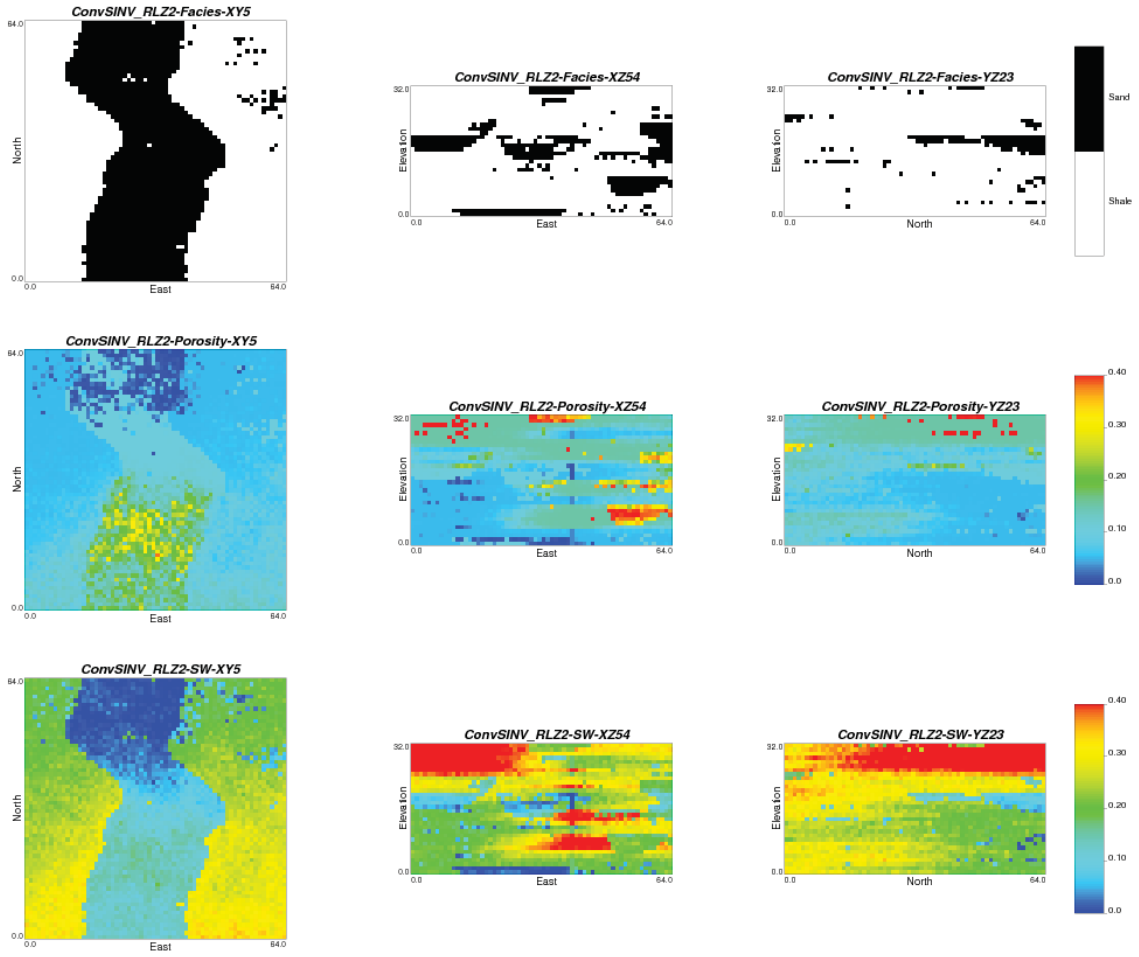

Figure 7.16: Different slices of facies, porosity and water saturation of second realization in different orientations for conventional stochastic inversion approach that are conditioned to the well and seismic data, axes represent grid cells, color bar unit (m^3/m^3).

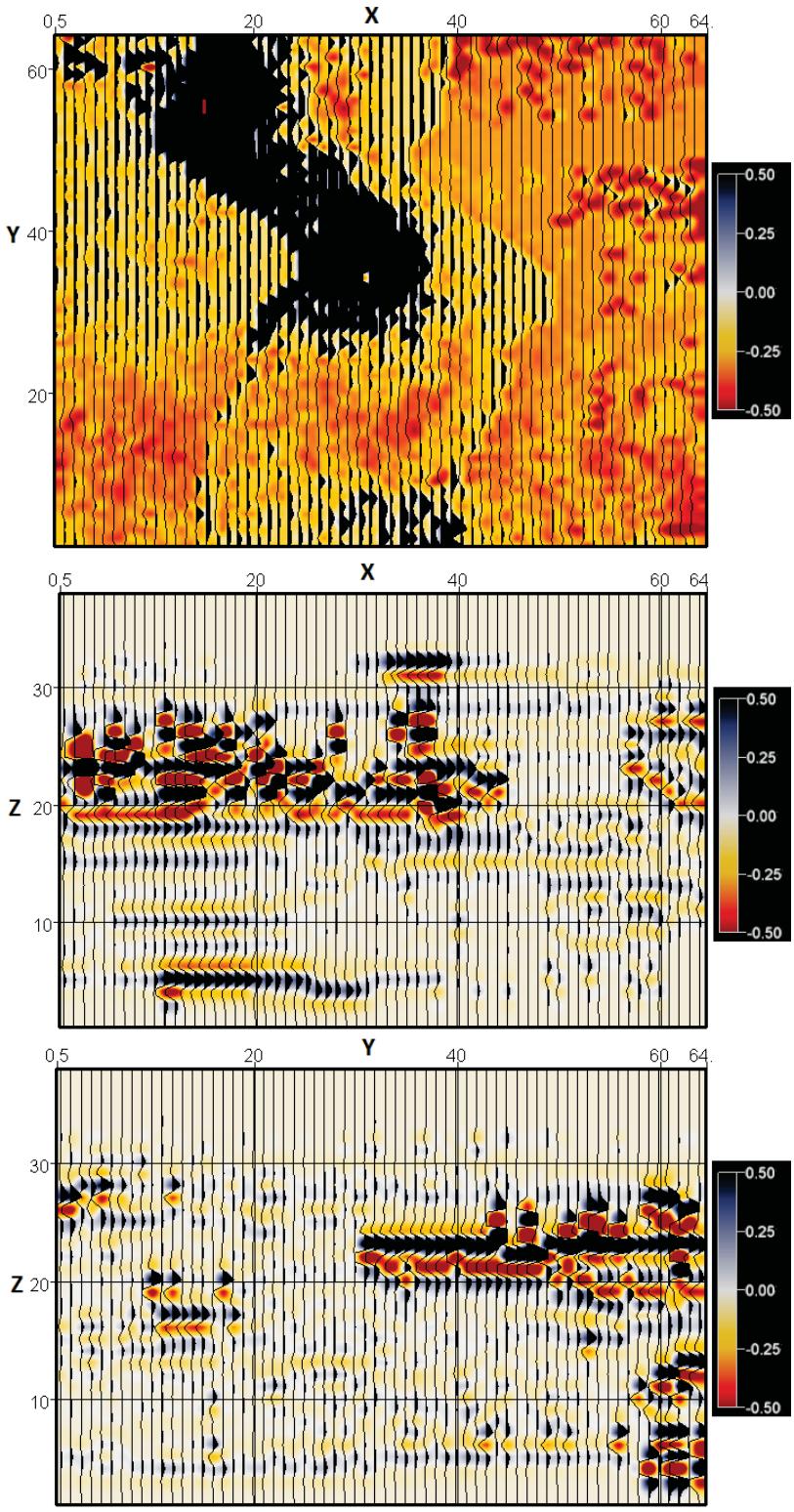

Figure 7.17: Different slices of the synthetic seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom) for the first realization of conventional stochastic inversion approach, axes represent grid cells, color bar - amplitude (m).

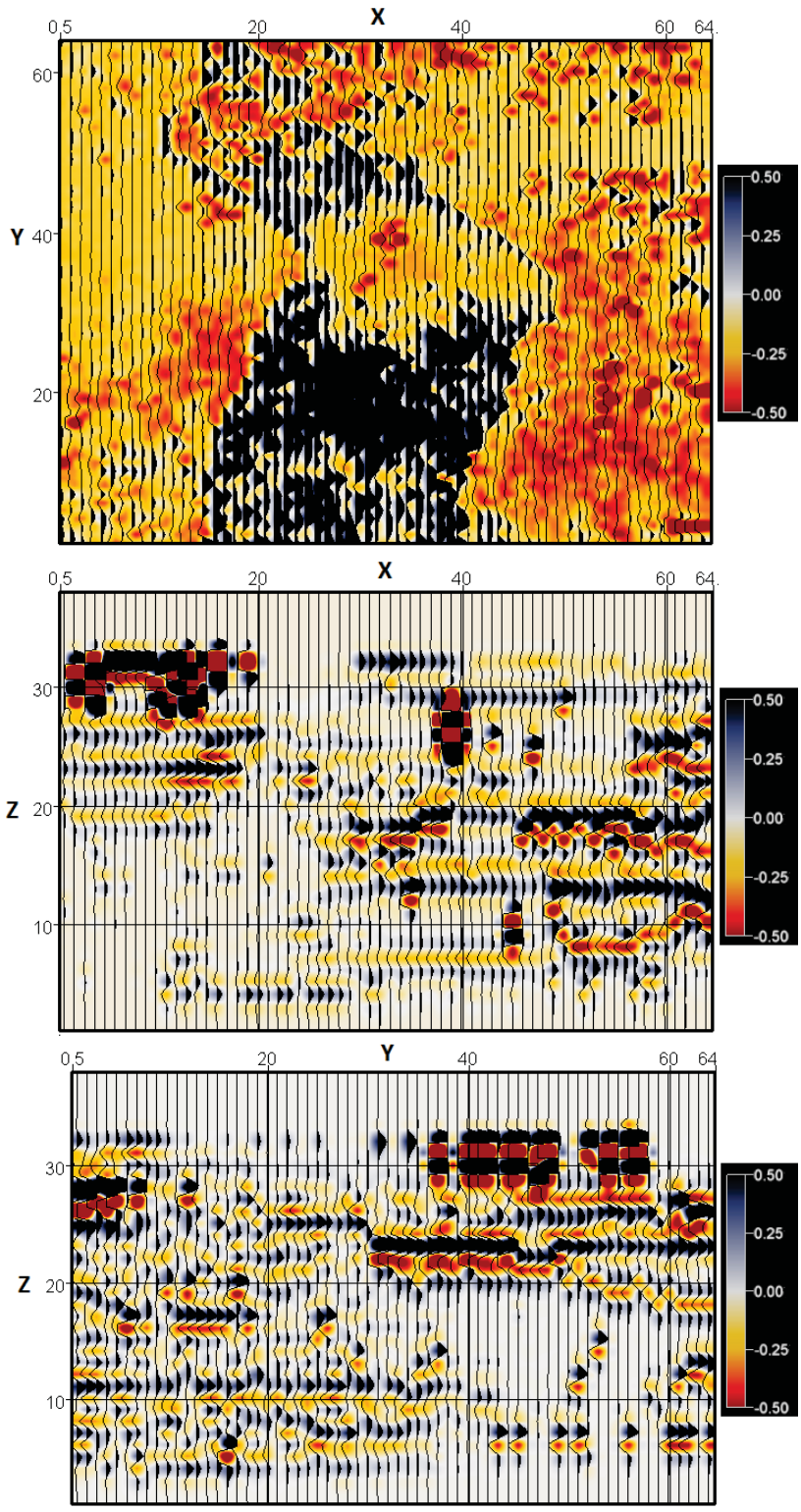

Figure 7.18: Different slices of the synthetic seismic survey, XY-5 (top), XZ-54 (middle) and YZ-23 (bottom) for the second realization of conventional stochastic inversion approach, axes represent grid cells, color bar - amplitude (m).

7.5 Chapter Summary

A case study based on a realistic data set is developed to compare the results of multivariate stochastic inversion with multivariate geostatistical modeling and conventional stochastic inversion. Multivariate geostatistical modeling simulates facies and multiple reservoir properties at the same time that are conditioned only to the well data. Due to limited areal coverage of well data, the reservoir models that are conditioned to the well data do not match the seismic data away from the well locations.

Conventional stochastic inversion, generates a high quality facies model conditioned to the well and seismic data. Then the facies model is linked to the reservoir physical properties including porosity and fluid saturations through the statistical calibrations and sequential connections. Although, the fidelity with original seismic data is lost via the sequential connections. Based on the visual assessment and comparison of MSE, these models do not reproduce the original seismic data within the quality of data.

Multivariate stochastic inversion simulates multiple reservoir properties simultaneously through the multivariate Gaussian simulation technique. This approach applies multivariate geostatistical modeling as part of stochastic inversion technique to condition the reservoir models to the seismic data at the same time. This approach generates high resolution reservoir models that reproduce the original seismic data. The MSE between multivariate stochastic inversion and original seismic data shows this method reproduce the original seismic data better than conventional stochastic inversion approach.

Chapter 8

Summary and Conclusion

This chapter summarizes the contributions and discusses some limitations of the research undertaken in this thesis. It also proposes some future works. A new stochastic inversion approach called ”multivariate stochastic inversion” has been developed and tested on different 2-D and 3-D synthetic and realistic data sets. The main objective of the new approach is to overcomes the limitations of conventional stochastic inversion approach and provide high resolution reservoir property models that reproduce the original seismic data. The conventional stochastic inversion approach simulates acoustic impedance or facies. The high resolution acoustic impedance or facies is linked to the reservoir physical properties like porosity by a set of statistical calibrations and sequential connections. There is no guarantee that the models obtained through the sequential connections reproduce the original seismic data.

8.1 Summary of Contributions

The multivariate stochastic inversion approach applies a multivariate geostatistical modeling technique as part of stochastic inversion. The objective of this close integration is to simulate multiple reservoir properties simultaneously and condition them to the seismic data at the same time. As the seismic surveys are inherently column based, to condition the reservoir properties to the seismic data at the same time all the geostatistical modeling techniques must be implemented in column wise manner. Multiple categorical and continuous reservoir properties are simulated simultaneously by multivariate Gaussian simulation technique. This technique simulates a full column of multiple Gaussian variables to simulate facies and reservoir physical properties. Facies is simulated through the truncated Gaussian simulation and continuous properties are

simulated through sequential Gaussian simulation technique. The number of Gaussian variables simulated by this technique depends on the number of Gaussian variables truncated for the facies categories and the number of continuous variables.

To handle complex and curvilinear geological features in the reservoir models and add more flexibility to the new approach, column based multi points statistics (MPS) simulation is an alternative for facies modeling. To implement the column based MPS algorithm, a new multigrid approach is proposed and developed. The multigrid approach allows the reproduction of large scale features with small templates in facies modeling. In multivariate stochastic inversion for the column based MPS, a $2-D$ multigrid in $X-Y$ direction is applied and at every step of $X-Y$ multigrid a full column of Z is simulated.

A full column of simulated facies and reservoir physical properties are passed to PEM and convolution algorithm to compute synthetic seismogram. The synthetic seismic trace is compared to the collocated actual seismic trace via the adaptive sampling algorithm. The adaptive sampling algorithm defines the stopping criteria when an acceptable match with the original seismic data is attained. The stopping criteria consists of a target MSE and reasonable number of realizations. The target MSE is defined based on cumulative inherent uncertainty of entire modeling process. The size of the space of uncertainty is a good indicator to help define the reasonable number of realizations for the multivariate stochastic inversion approach.

The size of the space of uncertainty for categorical variables is quantified as the product of exponential entropy of the configuration. This is not defined for continuous variables as they vary continuously over a range of values. This size is implausibly large which can not be understood from practical perspective in geostatistical modeling. The size of space of uncertainty is significantly reduced by different factors such as unequal proportions, spatial correlation and conditioning data. The size of the space of uncertainty in presence of these factors helps choose the number of realizations over the simulation domain to reach to the truth in multivariate stochastic inversion. Adaptive sampling in multivariate stochastic inversion means that the simulation process starts with a large number of realizations in the

first columns and gradually reduces the number of realizations going through the simulation path. The previously simulated columns are added to conditioning data and the size of the space of uncertainty reduces through the simulation path.

The multivariate stochastic inversion is applied for a 3-D realistic data set and the reservoir properties and synthetic seismic results are compared with results of multivariate geostatistical modeling and conventional stochastic inversion approach. Multivariate geostatistical modeling simulates multiple reservoir properties including facies , porosity and water saturation through the column based Gaussian simulation technique conditioned only to the well data. This method does not generate reservoir models away from the well data due to limited number of well data.

The conventional stochastic inversion generates facies models conditioned to the seismic data. Then, this model constrains the reservoir physical properties like porosity and water saturation. Although the facies model is high resolution, the quality of porosity and water saturation are reduced through the statistical calibrations and sequential connections due to an element of randomness at each step. Moreover, compared to the multivariate stochastic inversion, the conventional stochastic inversion approach cannot reproduce the original seismic data within the quality of data. As a result, multivariate stochastic inversion provides high resolution facies and reservoir physical properties simultaneously that reproduce the original seismic data within quality of data better than the other approaches.

8.2 Limitations

The integration of surface seismic geophysical data in high resolution geostatistical models is studied in multivariate stochastic inversion. Seismic data are originally in the time domain and well conditioning data are in depth. Although time to depth conversion is an important step, it is not in the scope of this project. To obtain high resolution geostatistical models that honor the seismic data, the actual seismic traces should be reproduced at each column in the geostatistical model. For this purpose the synthetic seismic traces are computed based on a convolution of reflectivity series calculated through a PEM and an extracted wavelet.

Some knowledge of rock above and below the reservoir is required.

Wavelet extraction is an important step where the synthetic trace along the well is calibrated with actual collocated seismic trace. This calibration process is necessary to ensure that the actual seismic data matches to convolution of the extracted wavelet with the reflectivity series along the wells. The typical zero phase Ricker wavelet is considered in this study because the specific details of the wavelet are not critical to the developed approach.

Access to real seismic data would make the study more creditable to some practitioners. Although, access to processed 3-D real seismic data is challenging. A realistic data set is generated as a reference model and applied in the case study.

8.3 Future Works

The multivariate stochastic inversion has a great potential to be applied in practice. There are many areas of future work including:

1. Fully developing MPS and other facies modeling techniques
2. Apply local truncation in truncated Gaussian simulation for facies modeling instead of Global truncation. This will improve the quality of facies modeling and accelerate convergence
3. Understanding the sensitivity of grid size for modeling
4. Modeling the rock above and below the reservoir zone of interest
5. Exploring alternative Petro-Elastic models and convolution operators
6. More closely coupling the number of realizations generated to the size of the space of uncertainty
7. Accelerating convergence by considering improved conditional simulation realizations as input to the selection process
8. Practical application for refinement of the technique and adaptation in industrial practice.

Bibliography

Aanonsen, S. I. and Eydinov, D. (2006). A multiscale method for distributed parameter estimation with application to reservoir history matching. *Computational Geosciences*, 10(1):97–117.

Alabert, F. and Journel, A. (1989). Non-gaussian data expansion in the earth sciences. *Terra Nova*, 1(2):123–134.

Alabert, F., Massonnat, G., et al. (1990). Heterogeneity in a complex turbiditic reservoir: stochastic modelling of facies and petrophysical variability. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Alabert, F. G. (1987). *Stochastic imaging of spatial distributions using hard and soft information*. Stanford University Press.

Avseth, P., Mukerji, T., and Mavko, G. (2010). *Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk*. Cambridge university press.

Azevedo, L., Nunes, R., Correia, P., Soares, A., Neto, G. S., Guerreiro, L., et al. (2013). Stochastic direct facies seismic avo inversion. In *2013 SEG Annual Meeting*. Society of Exploration Geophysicists.

Behrens, R. A., MacLeod, M. K., Tran, T. T., Alimi, A., et al. (1998). Incorporating seismic attribute maps in 3d reservoir models. *SPE Reservoir Evaluation & Engineering*, 1(02):122–126.

Birol, F. (2010). World energy outlook 2010. *International Energy Agency*, 1.

Bjørlykke, K. (2010). *Petroleum geoscience: From sedimentary environments to rock physics*. springer.

Blangy, J.-P. D. (1992). *Integrated seismic lithologic interpretation: The petrophysical basis*, volume 51. Department of Geophysics, School of Earth Sciences.

Bortoli, L.-J. (1992). Constraining reservoir models with seismic information. Master's thesis, to the Department of Applied Earth Sciences.Stanford University.

Bortoli, L.-J., Alabert, F., Haas, A., and Journel, A. (1993). Constraining stochastic images to seismic data. In *Geostatistics Tróia92*, pages 325–337. Springer.

Bosch, M., Mukerji, T., and Gonzalez, E. F. (2010). Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review. *Geophysics*, 75(5):75A165–75A176.

Bourennane, H. and King, D. (2003). Using multiple external drifts to estimate a soil variable. *Geoderma*, 114(1):1–18.

Bourennane, H., King, D., and Couturier, A. (2000). Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. *Geoderma*, 97(3):255–271.

Buland, A. and Omre, H. (2003). Bayesian wavelet estimation from seismic and well data. *Geophysics*, 68(6):2000–2009.

Campbell, L. (1966). Exponential entropy as a measure of extent of a distribution. *Probability Theory and Related Fields*, 5(3):217–225.

Carr, J. R., Myers, D. E., and Glass, C. E. (1985). Cokriging—a computer program. *Computers & Geosciences*, 11(2):111–127.

Chiles, J.-P. and Delfiner, P. (2009). *Geostatistics: modeling spatial uncertainty*, volume 497. John Wiley & Sons.

Chilès, J.-P. and Delfiner, P. (2012). Wiley series in probability and statistics. *Geostatistics: Modeling Spatial Uncertainty, Second Edition*, pages 705–714.

Cover, T. M. and Thomas, J. A. (2006). Elements of information theory 2nd edition.

Cressie, N. (2015). *Statistics for spatial data*. John Wiley & Sons.

David, M. (2012). *Geostatistical ore reserve estimation*. Elsevier.

Debeye, H., Sabbah, E., van der Made, P., et al. (1996). Stochastic inversion. In *65th Annual International SEG meeting, Denver*.

Deutsch, C. and Journel, A. (1998). Geostatistical software library and user's guide (gslib).

Deutsch, C. and Silva, D. A. (2014). Guide to mps simulation with snesim algorithm. *Centre for Computational Excellence (CCG), Guidebook Series Vol. 18, University of Alberta, April 2014*.

Deutsch, C. and Tran, T. (2002). Fluvsim: a program for object-based stochastic modeling of fluvial depositional systems. *Computers & Geosciences*, 28(4):525–535.

Deutsch, C. V. (1992). *Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data*. PhD thesis, stanford university.

Deutsch, C. V. (1994). Algorithmically-defined random function models. In *Geostatistics for the Next Century*, pages 422–435. Springer.

Deutsch, C. V. (2001). A short note on: Stochastic inversion for integration of seismic data in geostatistical reservoir modeling. Technical report, CCG Annual Report 3, Edmonton, AB.

Deutsch, C. V. (2005). A sequential indicator program for categorical variables with point and block data:blocksis. Technical report, CCG Annual Report 7, Edmonton, AB.

Deutsch, C. V. and Wang, L. (1996). Hierarchical object-based stochastic modeling of fluvial reservoirs. *Mathematical Geology*, 28(7):857–880.

Deutsch, J. L. and Deutsch, C. V. (2013). Advance in truncated plurigaussian simulation

for reproduction of transition probablities. Technical report, CCG Annual Report 15, Edmonton, AB.

Doyen, P. (2011). Petro-elastic and lithology-fluid inversion from seismic data, a geophysical perspective. In *73rd EAGE Conference & Exhibition-Workshops*.

Doyen, P. and Guidish, T. (1992). Seismic discrimination of lithology and porosity, a monte carlo approach. *Reservoir Geophysics, Investigations in Geophysics*, 7:243–250.

Doyen, P. M. (1988). Porosity from seismic data: A geostatistical approach. *Geophysics*, 53(10):1263–1275.

Eidsvik, J., Avseth, P., Omre, H., Mukerji, T., and Mavko, G. (2004). Stochastic reservoir characterization using prestack seismic data. *Geophysics*, 69(4):978–993.

Escobar, I., Williamson, P., Cherrett, A., Doyen, P., Bornard, R., Moyen, R., Crozat, T., et al. (2006). Fast geostatistical stochastic inversion in a stratigraphic grid. In *SEG 76th Annual Meeting*.

Francis, A. (2005). Limitations of deterministic and advantages of stochastic inversion. *CSEG Recorder*.

Galli, A., Beucher, H., Le Loch, G., Doligez, B., et al. (1994). The pros and cons of the truncated gaussian method. In *Geostatistical simulations*, pages 217–233. Springer.

Galli, A., Le Loch, G., Geffroy, F., and Eschard, R. (2006). An application of the truncated pluri-gaussian method for modeling geology.

Galli, A. and Meunier, G. (1987). Study of a gas reservoir using the external drift method. In *Geostatistical case studies*, pages 105–119. Springer.

Gardner, G., Gardner, L., and Gregory, A. (1974). Formation velocity and density-the diagnostic basics for stratigraphic traps. *Geophysics*, 39(6):770–780.

Goldberger, A. S. (1962). Best linear unbiased prediction in the generalized linear

regression model. *Journal of the American Statistical Association*, 57(298):369–375.

Gomez-Hernandez, J. J. (1991). *A stochastic approach to the simulation of block conductivity values conditioned upon data measured at a smaller support*. PhD thesis, PhD Thesis, Stanford University, CA. Available from Ann Arbor microfilms, Michigan.

González, E. F., Mukerji, T., and Mavko, G. (2007). Seismic inversion combining rock physics and multiple-point geostatistics. *Geophysics*, 73(1):R11–R21.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Oxford university press.

Goovaerts, P. (1999). Impact of the simulation algorithm, magnitude of ergodic fluctuations and number of realizations on the spaces of uncertainty of flow properties. *Stochastic Environmental Research and Risk Assessment*, 13(3):161–182.

Goovaerts, P. (2006). Geostatistical modeling of the spaces of local, spatial, and response uncertainty for continuous petrophysical properties.

Grana, D., Mukerji, T., Dvorkin, J., and Mavko, G. (2012). Stochastic inversion of facies from seismic data based on sequential simulations and probability perturbation method. *Geophysics*, 77(4):M53–M72.

Guardiano, F. B. and Srivastava, R. M. (1993). Multivariate geostatistics: beyond bivariate moments. In *Geostatistics Troia92*, pages 133–144. Springer.

Gunning, J. and Glinsky, M. E. (2006). Wavelet extractor: A bayesian well-tie and wavelet extraction program. *Computers & Geosciences*, 32(5):681–695.

Haas, A. and Dubrule, O. (1994). Geostatical inversion-a sequential method of stochastic reservoir modelling constrined by seismic data. *First break*, 12(11).

Haldorsen, H. H., Lake, L. W., et al. (1984). A new approach to shale management in field-scale models. *Society of Petroleum Engineers Journal*, 24(04):447–457.

Hamilton, J. D. (2015). Demand factors in the collapse of oil prices. *Econbrowser, January*, 11.

Helgesen, J., Magnus, I., Prosser, S., Saigal, G., Aamodt, G., Dolberg, D., and Busman, S. (2000). Comparison of constrained sparse spike and stochastic inversion for porosity prediction at kristin field. *The Leading Edge*, 19(4):400–407.

Honarkhah, M. (2011). Stochastic simulation of patterns using distance-based pattern modeling.

Hong, S., Leuangthong, O., and Deutsch, C. V. (2006). A geostatistical approach to stochastic seismic inversion. Technical report, CCG Annual Report 8, Edmonton, AB.

Jin, L., Xue, Y., and Sen, M. (2013). Application of principal component analysis to simultaneous seismic inversion. In *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.

Johansen, T. A., Jensen, E. H., Mavko, G., and Dvorkin, J. (2013). Inverse rock physics modeling for reservoir quality prediction. *Geophysics*, 78(2):M1–M18.

Journel, A. (1988). Focusing on spatial connectivity of extreme-valued attributes: Stochastic indicator models of reservoir heterogeneities. *SPE paper*, 18324.

Journel, A. (1997). The abuse of principles in model building and the quest for objectivity. *Geostatistics Wollongong*, 96:3–14.

Journel, A. and Zhu, H. (1990). *Integrating soft seismic data: Markov-Bayes updating, an alternative to cokriging and traditional regression*. Stanford Center for Reservoir Forecasting.

Journel, A. G. and Huijbregts, C. J. (1978). *Mining geostatistics*. Academic press.

Journel, A. G. and Xu, W. (1994). Posterior identification of histograms conditional to local data. *Mathematical Geology*, 26(3):323–359.

Kallweit, R. and Wood, L. (1982). The limits of resolution of zero-phase wavelets. *Geophysics*, 47(7):1035–1046.

Kane, J., Rodi, W., elix Herrmann, F., and Toks?oz, M. N. (1999). Geostatistical seismic in v ersionusing well log constraints.

Kjønsberg, H., Hauge, R., Kolbjørnsen, O., and Buland, A. (2010). Bayesian monte carlo method for seismic predrill prospect assessment. *Geophysics*, 75(2):O9–O19.

Lamy, P., Swaby, P., Rowbotham, P., Dubrule, O., and Haas, A. (1998). From seismic to reservoir properties using geostatistical inversion. In *SPE annual technical conference*, pages 535–545.

Larsen, A. L., Ulvmoen, M., Omre, H., and Buland, A. (2006). Bayesian lithology/fluid prediction and simulation on the basis of a markov-chain prior model. *Geophysics*, 71(5):R69–R78.

Latimer, R. B., Davidson, R., and Van Riel, P. (2000). An interpreter's guide to understanding and working with seismic-derived acoustic impedance data. *The leading edge*, 19(3):242–256.

Lee, S. H., Malallah, A., Datta-Gupta, A., Higdon, D., et al. (2000). Multiscale data integration using markov random fields. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Leuangthong, O., McLennan, J. A., and Deutsch, C. V. (2004). Minimum acceptance criteria for geostatistical realizations. *Natural Resources Research*, 13(3):131–141.

Li, Y. and Deutsch, C. V. (2010). Probability estimation with maximum entropy principle. centre of computational geostatistic. Technical report, CCG Annual Report 12, Edmonton, AB.

Lines, L. R. and Newrick, R. T. (2004). *Fundamentals of geophysical interpretation*. Society of Exploration Geophysicists.

Lumley, D. E. (2001). Time-lapse seismic reservoir monitoring. *Geophysics*, 66(1):50–53.

Ma, X.-Q. (2002). Simultaneous inversion of prestack seismic data for rock properties using simulated annealing. *Geophysics*, 67(6):1877–1885.

Manchuk, J. G. and Deutsch, C. V. (2012). A flexible sequential gaussian simulation program: Usgsim. *Computers & Geosciences*, 41:208–216.

Marechal, A. (1984). Kriging seismic data in presence of faults. In *Geostatistics for natural resources characterization*, pages 271–294. Springer.

Margrave, G. F. (2005). Method of seismic data processing, geophysics 557/657. *Geophysics*, 1(1):20–30.

Margrave, G. F. (2009). *Methods of seismic data processing Geophysics 557-course Lecture Notes*. University of Calgary.

Matheron, G. (1962). *Traité de géostatistique appliquée. 1 (1962)*, volume 1. Editions Technip.

Matheron, G., Beucher, H., De Fouquet, C., Galli, A., Guerillot, D., Ravenne, C., et al. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The rock physics handbook: Tools for seismic analysis of porous media*. Cambridge university press.

Moysey, S., Caers, J., Knight, R., and Allen-King, R. (2003). Stochastic estimation of facies using ground penetrating radar data. *Stochastic Environmental Research and Risk Assessment*, 17(5):306–318.

Mukerji, T., Avseth, P., Mavko, G., Takahashi, I., and González, E. F. (2001). Statistical rock physics: Combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization. *The Leading Edge*, 20(3):313–319.

Myers, D. (1994). Geostatistical simulation: thoughts and questions. *Geostatistics*, 7:4–7.

Myers, D. E. (1996). Choosing and using simulation algorithms. *UNITED STATES DEPARTMENT OF AGRICULTURE FOREST SERVICE GENERAL TECHNICAL REPORT RM*, pages 23–29.

Oliver, D. S., Reynolds, A. C., and Liu, N. (2008). *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press.

Park, N.-W., Chi, K.-H., and Kwon, B.-D. (2003). Geostatistical integration of spectral and spatial information for land-cover mapping using remote sensing data. *Geosciences Journal*, 7(4):335–341.

Parker, R. L. (1994). *Geophysical inverse theory*. Princeton university press.

Pereira, P., Azevedo, L., Nunes, R., and Soares, A. (2016). Integration of initial guess models into geostatistical seismic inversion methodologies. In *Proceedings of the tenth International Geostatistics Congress*.

Pyrcz, M. J. and Deutsch, C. V. (2014). *Geostatistical reservoir modeling*. Oxford university press.

Pyrcz, M. J. and White, C. D. (2015). Uncertainty in reservoir modeling. *Interpretation*, 3(2):SQ7–SQ19.

Ripley, B. (1981). Spatial statistics. 1981. *Hayward Wiley, New York*.

Rossi, M. (1994). Of tool makers and tool users. *Geostatistics*, 7(2):7–9.

Rowbotham, P. S., Lamey, P., Swaby, P. A., and Dubrule, O. (1998). Geostatistical inversion for reservoir characterization. In *68th SEG Annual Meeting Expanded Abstracts.*

Russell, B. and Hampson, D. (1991). A comparison of poststack seismic inversion methods. In *61th SEG Annual Meeting Expanded Abstracts. Tulsa: SEG*.

Russell, B. H. (1988). *Introduction to seismic inversion methods*, volume 2. Society of Exploration Geophysicists.

Ryan, H. (1994). Ricker, ormsby, klauder. *Butterworth?a choice of wavelets, Can. Soc. Explor. Geophys. Record*, 19(7):8–9.

Sams, M., Atkins, D., Said, N., Parwito, E., van Riel, P., et al. (1999). Stochastic inversion for high resolution reservoir characterisation in the central sumatra basin. *SPE Asia Pacific Improved Oil Recovery, Kuala Lumpur, SPE*, 57620.

Shuey, R. (1985). A simplification of the zoeppritz equations. *Geophysics*, 50(4):609–614.

Skauvold, J., Eidsvik, J., and Theune, U. (2015). Seismic wavelet estimation and uncertainty quantification using a parametric model. In *Petroleum Geostatistics 2015*.

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

Srivastava, R. (1994). Thoughts and comments on conditional simulation algorithms. *Geostatistics*, 7(2):9–10.

Srivastava, R. (1996). An overview of stochastic spatial simulation. *Mowrer, HT, Ž.*

Srivastava, R. and Sen, M. (2010). Stochastic inversion of prestack seismic data using fractal-based initial models. *Geophysics*, 75(3):R47–R59.

Strebelle, S. (2000). Sequential simulation drawing structures from training images.

Telford, W. M., Geldart, L. P., and Sheriff, R. E. (1990). *Applied geophysics*, volume 1. Cambridge university press.

Thadani, S. G., Alabert, F., Journel, A. G., et al. (1987). An integrated geostatistical/pattern recognition technique for characterization of reservoir spatial variability. In *1987 SEG Annual Meeting*. Society of Exploration Geophysicists.

Tran, T. T. (1994). Improving variogram reproduction on dense simulation grids. *Computers & Geosciences*, 20(7-8):1161–1168.

Ulrych, T. J., Sacchi, M. D., and Woodbury, A. (2001). A bayes tour of inversion: A tutorial. *Geophysics*, 66(1):55–69.

Walden, A. T. and White, R. E. (1998). Seismic wavelet estimation: a frequency domain solution to a geophysical noisy input-output problem. *IEEE Transactions on Geoscience and Remote Sensing*, 36(1):287–297.

Widess, M. (1973). How thin is a thin bed? *Geophysics*, 38(6):1176–1180.

Xu, W., Tran, T., Srivastava, R., Intl, F., and Jownel, A. (1992). hltegrating seismic data in reservoir modeling: The collocated cokriging alternative.

Yao, T. and Chopra, A. (2000). Integration of seismic attribute map into 3d facies modeling. *Journal of petroleum science and engineering*, 27(1):69–84.

Yilmaz, Ö. (2001). *Seismic data analysis*, volume 1. Society of exploration geophysicists Tulsa.

Zhang, R., Zhang, K., and Alekhue, J. E. (2016). Depth-domain seismic reflectivity inversion with compressed sensing technique. *Interpretation*, 5(1):T1–T9.

Zhu, H. (1991). *Modeling mixture of spatial distributions with integration of soft data*. PhD thesis, Stanford University.

Zhu, H. and Journel, A. G. (1993). Formatting and integrating soft data: Stochastic imaging via the markov-bayes algorithm. In *Geostatistics Tróia92*, pages 1–12. Springer.

Zunino, A., Mosegaard, K., Lange, K., Melnikova, Y., and Mejer Hansen, T. (2014). Monte carlo reservoir analysis combining seismic reflection data and informed priors. *Geophysics*, 80(1):R31–R41.

Appendix A

Multivariate Stochastic Inversion

MSInv (Multivariate Stochastic Inversion) is a new stochastic inversion program that applies a multivariate geostatistical modeling technique as part of stochastic inversion algorithm to condition the reservoir models to the seismic data at the same time. The main purpose of this close integration is to overcome the limitations of conventional stochastic inversion approaches and provide high resolution reservoir models that reproduce the original seismic data within the quality of the data. The new stochastic inversion approach simulates facies and multiple reservoir physical properties simultaneously through the multivariate Gaussian simulation technique. As the seismic surveys are inherently in the column based fashion, all the geostatistical modeling techniques in multivariate stochastic inversion approach are implemented in the column wise manner.

In multivariate stochastic inversion approach, a full column of facies and multiple reservoir properties are simulated via the multivariate Gaussian simulation technique. The facies is simulated by TGS and continuous variables are simulated by SGS. To condition these properties to the seismic data at the same time, a full column of the multiple simulated reservoir properties are passed to the Petro Elastic Model (PEM) and convolution algorithm to compute reservoir elastic properties and synthetic seimogram. The computed synthetic seismogram is compared with collocated actual seismic trace to choose the acceptable match. To implement multivariate Gaussian simulation technique and integrate it with the stochastic inversion algorithm, the Ultimate Sequential Gaussian Simulation (USGSIM) that performs multivariate geostatistical modeling is modified to: (1) perform column based simulation, (2) generate multiple realizations per column, (3) apply multivariate simulation for both categorical and continuous variables simultaneously, (4) couple with Petro Elastic Model (PEM) to generate multiple reservoir elastic properties, (5) convolve the elastic properties with extracted wavelet, and (6) apply

a selection criteria to retain one realization. The multivariate stochastic inversion parameter file is described below.

A.1 Parameter File Sections

Parameter file of MSInv is similar to the USGSIM parameter file. The parameter file consists of a number of required and optional blocks. All the blocks will be described based on their order in the parameter file. The first block is the **MAIN block** that includes number of realization, number of simulated Gaussian variables, random seed number, grid size and out put file. The number of realization in the MAIN block is to generated a 2-D or 3-D multiple reservoir properties with different seed numbers. The number of Gaussian variables are required by multivariate Gaussian simulation technique which depends on the number of Gaussian variables are needed for TGS, number of facies categories and number of continuous variables to be simulated must be defined in this block.

```
1  Parameters for Multivariate Stochastic Inversion
2  ************************************************
3  START OF MAIN:
4  **************
5  1                       -number of realizations to generate
6  5                       -number of simulated Gaussian variables
7  69069                   -random number seed
8  64    0.5    1.0        -nx,xmn,xsiz
9  64    0.5    1.0        -ny,ymn,ysiz
10 32     0.5    1.0       -nz,zmn,zsiz
11 Props-BP.out            -file for simulation output
12 1                       -output format:(0=reg, 1=coord, 2=bin)
13 1                       -debugging level: 0,1,2,3
14 sgsim.dbg               -file for debugging output
```

The next block is **Facies-Modeling block**. As mentioned before, the primary option for

facies modeling in multivariate stochastic inversion is truncated Gaussian simulation approach. In this block number of facies categories to be simulated, categories and global proportion are defined.

```
Parameters for Multivariate Stochastic Inversion
************************************************
START OF Facies-Modeling:
*************************
2                              -number of categories
0 1                            -categories
0.3 0.7                        -global proportion
```

The **SRCH block** consists number of data, multigrid option, search radius and anisotropy and also the other search parameters. The data are assigned to the closest grid node. In case, they fall outside the grid they are not used. The size of covariance look up table is automatically calculated and allocated based on the search parameter.

```
Parameters for Multivariate Stochastic Inversion
************************************************
START OF SRCH:
**************
32                           -number of data to use per variable
1                            -multiple grid search (0=no, 1=yes)
0     3                      -octant search (0=no, 1=yes), number
50000 50000  1.0             -maximum search radii (hmax,hmin,vert)
0.0    0.0    0.0            -angles for search ellipsoid
```

The variogram definition for facies and reservoir physical properties is defined in the **VARG block**. The specification of variogram is as the other GSLIB programs. The variograms for facies modeling by TGS must be defined first. Then, the variogram for continuous variables are defined. The order of variogram for continuous variables does not matter. The multivariate Gaussian simulation technique simulated continuous variables per facies code. Later, they merged based on facies code of each cell in the

facies model. Therefore, it is necessary to define the facies code, Gaussian variable's number related to this facies code, continuous variable's number related to this facies code and Gaussian variable and variogram parameters for this set. These parameters must be defined in VARG block for all continuous variables per facies code. For example for a facies with shale (code 0) and sand (code 1) categories that needs one Gaussian variable for TGS and two continuous variable such as porosity and water saturation, five Gaussian variables are required to be simulated. The Gaussian variable 1 is for facies modeling. The category 0, Gaussian variable 2, continuous variable 1 is for shale porosity with related variogram parameters. The category 1, Gaussian variable 3, continuous variable 1 is for sand porosity with related variogram parameters and so on. Later, Gaussian variables number 2 and 3 are merged to obtain the final porosity model.

```
1  Parameters for Multivariate Stochastic Inversion
2  ************************************************
3  START OF VARG:
4  **************
5  5                          -number of variograms
6
7  1                          -variable 1
8  1    0.0                   -number of structures, nugget effect
9  2    1.0  0.0   0.0   0.0  -type,variance,ang1,ang2,ang3
10 10.0  100.0  20.0          -a_hmax, a_hmin, a_vert
11
12 1  2  1                    -category, Gaussian var, continuous var
13 1    0.1                   -number of structures, nugget effect
14 2    0.9  0.0   0.0   0.0  -type,variance,ang1,ang2,ang3
15 40.0  40.0  20.0           -a_hmax, a_hmin, a_vert
16
17 0  3   1                   -category, Gaussian var, continuous var
18 1    0.1                   -number of structures, nugget effect
19 2    0.9  0.0   0.0   0.0  -type,variance,ang1,ang2,ang3
20 40.0  40.0  20.0           -a_hmax, a_hmin, a_vert
```

```
21
22 1  4   2                    -category, Gaussian var, continuous var
23 1    0.1                    -number of structures, nugget effect
24 2    0.9  0.0   0.0   0.0   -type,variance,ang1,ang2,ang3
25 40.0  40.0  20.0            -a_hmax, a_hmin, a_vert
26
27 0  5   2                    -category, Gaussian var, continuous var
28 1    0.1                    -number of structures, nugget effect
29 2    0.9  0.0   0.0   0.0   -type,variance,ang1,ang2,ang3
30 40.0  40.0  20.0            -a_hmax, a_hmin, a_vert
```

The **DATA block** is an optional block that is required for conditional simulation. If this block is not in the parameter file or there is no data file for conditioning data, the code applies an unconditional simulation. The conditioning data must be defined per facies categories as the continuous variables are simulated per facies categories. The data point have coordinates and declustering weights.

```
1 Parameters for Multivariate Stochastic Inversion
2 ************************************************
3 START OF DATA:
4 **************
5 Cond-3D.out                    -file with primary data
6 1  2  3  0  0                  -columns for X,Y,Z,wt
7 4 5  6  7  8                   -columns for variables
8 -1.0e21       1.0e21           -trimming limits
```

The **TRAN block** transfers the conditioning data to the Gaussian units and back transforms the simulated values from Gaussian unit to original units. If the option is no transform, then the conditional data must be in Gaussian units and the out-put file for simulation will be in Gaussian units.

```
1 Parameters for Multivariate Stochastic Inversion
2 ************************************************
```

```
START OF TRAN:
**************
1                       -transform the data (0=no, 1=yes)
200                     -number of quantiles to keep (default=200)
4                       -number of min/max values
1  2   0.1   0.4       -category, Gaussian variable, min, max
0  3   0.05  0.15      -category, Gaussian variable, min, max
1  4   0.0   0.25      -category, Gaussian variable, min, max
0  5   0.2   0.5       -category, Gaussian variable, min, max

4                       -number of reference distributions

1  2                    -category, primary variable number
phig.dat                -file with ref. distribution
1  0                    -columns for value and weight

0  3                    -category, Gaussian variable number
phib.dat                -file with ref. distribution
1  0                    -columns for value and weight

1  4                    -category, Gaussian variable number
SWg.dat                 -file with ref. distribution
1  0                    -columns for value and weight

0  5                    -category, Gaussian variable number
SWb.dat                 -file with ref. distribution
1  0                    -columns for value and weight
```

The details about multivariate geostatistical modeling and joint simulation of multiple variables are specified in **MULT block**. Multiple variables may be simulated independently (code 1) with no correlation, with collocated cokriging.

```
Parameters for Multivariate Stochastic Inversion
************************************************
START OF MULT:
**************
2                       -Multivariate handling, see comment below
0                       -number of classes for stepwise
nodata                  -file with transformation data
0  0  0 0  0            -columns for variables
0                       -column  for weight
1                       -rock type for correlation matrix
0.0 0.0 0.0  0.0        -correlation matrix: 1-2, 1-3, 1-4, 1-5
    0.0 -0.7  0.0                                 2-3, 2-4, 2-5
         0.0 -0.7                                      3-4, 3-5
              0.0                                           4-5
-----------------------------------------------
Multivariate options:
1 - independent
2 - collocated
3 - stepwise
4 - full cokriging
5 - intrinsic collocated
6 - Bayesian updating
```

The parameters for stochastic inversion algorithm are defined in the **Stochastic-Inv block**. The original seismic 2-D or 3-D is specified in this block. If there is no input file for the seismic data, the code applies multivariate geostatistical modeling, passes the simulated values to the PEM and convolution algorithm to compute synthetic seismic cube through the forward modeling. The PEM parameters, including: bulk densities, fluid densities and Gardner's parameters are defined in this block. The multivariate stochastic inversion approach considers the zero phase Ricker wavelet with a central frequency for the convolution algorithm. The frequency of Ricker wavelet is defined in this part. The

stopping criteria for adaptive sampling algorithm such as target MSE and reasonable number of realization per multigrid level are defined in this block.

```
Parameters for Multivariate Stochastic Inversion
************************************************
START OF Stochastic-Inv.
**********************
seis.out              -file with actual seismic data
seisbp.out            -out put file for synthetic seismic
1 2.62                -category, density
0 2.65                            -category, density
2                                 -number of fluid phase
1 1                               -phase, density
2 0.95                            -phase, density
0.31 0.25             -Gardner's parameter;a m
30                    -Ricker wavelet frequency
5                     -number of multigrid level
1000 500 200 100 10   -number of realization per columns per multigrid
1.e-2                 -target MSE
```

All the input and output files are in GSLIB conventional format. The out-put of multivariate stochastic inversion is two files. One file includes the facies and reservoir physical properties such as porosity and fluid saturations. The other out put file is the synthetic seismic cubes based on stopping criteria in adaptive sampling algorithm.

Appendix B

Column Based SNESIM

The primary option for facies modeling in multivariate stochastic inversion is truncated Gaussian simulation (TGS). Although, to handle more complex geological features, the column based multi point statistics (MPS) simulation is an alternative for facies modeling.The seismic data is inherently column based; therefore, the SNESIM code is modified to perform column based simulation. The concept of original multi grid approach does not work in column based SNESIM because a full column must be simulated at once. For this purpose, a new multigrid approach called directed multigrid is developed and applied in column based SNESIM. There are two different versions of directed multigrid: 1) $X-Y$ multigrid, full column of Z and 2) $X-Y$ multigrid plus Z multigrid.

B.1 Column Based SNESIM Parameter File

The parameter file for column based SNESIM is quite similar to the original SNESIM code. More details about the original SNESIM code can be found in different sources including the guide book "Guide to MPS Simulation" by Deutsch and Silva (2014). Three lines added to the end of the original SNESIM parameter file. First line specifies to apply column based MPS with or without multigrid approach. The next line defines the number of multigrid levels. At the last two line the directed multigrid version can be defined.

```
1 Parameters for SNESIM
2 ********************
3 START OF PARAMETERS:
4 nodata                - file with original data
5 1  2  3  4            - columns for x, y, z, variable
```

```
3                       - number of categories
0   1   2               - category codes
0.7  0.2   0.1          - (target) global pdf
0                       - use (target)vertical proportions(0=no,1=yes)
vertprop.dat            - file with target vertical proportions
0    0.5                - target pdf repro. (0=no, 1=yes), parameter
1                       - debugging level: 0,1,2,3
snesim.dbg              - debugging file
snesim.out              - file for simulation output
1                       - number of realizations to generate
128   0.5    1.0        - nx,xmn,xsiz
128   0.5    1.0        - ny,ymn,ysiz
64   0.5    1.0         - nz,zmn,zsiz
548452                  - random number seed
20                      - max number of conditioning primary data
ti.dat                  - file for training image
128  128  64            - training image dimensions: nxtr, nytr, nztr
1                       - column for training variable
150.0  120.0  64        - maximum search radii (hmax,hmin,vert)
30.0   0.0   10.0       - angles for search ellipsoid
1                       -multiple grid search (0=no, 1=yes)
5                       -number of multigrid level
1                       -multiple grid option,XYMGRD-fullZ(0=no,1=yes)
0                       -multiple grid option,XYMGRD-ZMGRD(0=no,1=yes)
```

Appendix C

Size of Space of Uncertainty

The size of the space of uncertainty in presence of different parameters such as: unequal proportions, spatial correlation and conditioning data is a good indicator to define the dynamic number of realizations per column in multivariate stochastic inversion. The size of the space of uncertainty also quantified as the product of exponential entropy of the system. The probability of each category at location must be estimated to compute size of the space of uncertainty. The BLOCKSIS (conventional GSLIB Code) is modified to: 1) compute p_k for $k = 1, \ldots, K$ sequentially through a random path at each location, 2) calculate entropy H and exponential entropy e^H, 3) repeat this procedure over the entire simulation domain and 4) calculate size of the space of uncertainty by Equation 4.10. The out put file for size of space of uncertainty is in the GSLIB format that represents p_k, H, e^H and size of space of uncertainty.

```
1  Parameters for BlockSIS
2  ***********************
3  START OF PARAMETERS:
4  0                  -0=SK,1=OK,2=L1,3=L2,4=CC,5=BU,6=PR,7=BK,8=BC
5  0                  -Clean: 0=none, 1=light, 2=heavy, 3=super
6  2                  -number of categories
7  1 2                -categories
8  0.5 0.5            -global proportions
9  0.50  0.50         -correlation coefficients for soft data
10 nodata             -file with local data
11 0 0 0 0            -columns for X,Y,Z, and category
12 nodata.dat         -file with gridded prior mean values
13 0   0   0          -columns for each category
14 2                  -2-D areal map (2) or 3-D cube (3)
```

```
nodata            -file with keyout array
1                 -column for keyout indicator
0                 -debugging level: 0,1,2,3,4
BlockSIS.dbg      -file for debugging output
BlockSIS.out      -file for simulation output
sou.out           -file for size of space of uncertainty output
100               -number of realizations
50  0.5  1.00     -nx,xmn,xsiz
1   0.5  1.00     -ny,ymn,ysiz
1   0.5  1.00     -nz,zmn,zsiz
69000             -random number seed
24                -maximum original data  for each kriging
24                -maximum previous nodes for each kriging
1                 -assign data to nodes? (0=no,1=yes)
0                 -maximum per octant    (0=not used)
50.0  50.0  50.0  -maximum search radii
0.0   0.0  0.0    -angles for search ellipsoid
100.0 100.0 100.0 -size of covariance lookup table

1    0.1           -nst, nugget effect
2  0.9 0.0 0.0 0.0 -it,cc,ang1,ang2,ang3
50.0  50.0  50.0   -a_hmax,a_hmin,a_vert

1   0.1             -nst, nugget effect
2  0.9 0.0 0.0 0.0  -it,cc,ang1,ang2,ang3
50.0  50.0  50.0    -a_hmax,a_hmin,a_vert
```