Multivariate Reservoir Property Modeling with Hierarchical Truncated Pluri-Gaussian and Projection Pursuit Multivariate Transformation Techniques

by

Chuks Anthony Sunday

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

Master of Science

in

Mining Engineering

Department of Civil and Environmental Engineering
University of Alberta

© Chuks Anthony Sunday, 2022

Abstract

To characterize a petroleum reservoir and to develop a reliable static and dynamic model for assessing resources and reserves, the rock and fluid properties of the subsurface formation are modeled. In most cases, multiple reservoir properties (categorical and continuous variables) are jointly utilized to represent rock properties. Standard practice is to model and characterize the categorical properties of the reservoir such as facies, before modeling the continuous petrophysical properties.

A framework that applies the Hierarchical Truncated Pluri-Gaussian (HTPG) and Projection Pursuit Multivariate Transformation (PPMT) techniques for multivariate modeling of rock properties in a petroleum reservoir is developed and documented in this thesis. The HTPG technique and PPMT workflow are established methods for modeling categorical and continuous variables, respectively. Unlike other truncated simulation techniques such as Truncated Gaussian Simulation (TGS) and Truncated Pluri-Gaussian Simulation (TPGS), HTPG can be applied to a complex geological domain and does not limit the number of latent Gaussian variables that are used. Although the traditional modeling workflow for multivariate continuous property modeling reproduces important univariate statistics, it does not reproduce the complexity in the multivariate data, and this leads to models that do not reproduce important characteristics of the reservoir. To overcome this limitation and for ease in the geostatistical modeling workflow, the PPMT workflow is implemented. The PPMT decorrelates the multivariate variable and transforms them into univariate Gaussian variables that can be modeled independently.

This proposed integrated modeling method is implemented on the Hekla Reservoir. The HTPG simulation technique is applied to model the five facies in the Hekla reservoir to generate one hundred simulated realizations of the facies, while the PPMT technique is subsequently applied to model the continuous petrophysical properties (porosity and permeability) distributed in the reservoir. This framework is reproducible and applicable to both mining and petroleum projects, which often require multiple categorical and continuous variables of geologic deposits or reservoirs.

Dedication

To the loving memory of my late parents, Mr. Amadi Sunday and Mrs. Alice Mary Sunday.

Acknowledgments

I would like to profoundly thank my supervisor, Professor Clayton V. Deutsch, for his understanding, immense support, and unparalled ability to motivate his students. It was a great life opportunity to learn from Clayton.

I would also like to thank the Centre for Computational Geostatistics (CCG) for the financial support throughout this period.

I am exceedingly grateful to my colleagues at the CCG for their help and willingness to discuss some research and programming challenges.

My sincere gratitude goes to the Graduate Coordinator of the Department of Civil and Environmental Engineering, Arlene Figley, for her tireless support throughout this period.

The completion of my program wouldn't have been possible without the total support, encouragement, and unconditional love of my lovely wife, Angela Ijeoma Sunday, and my beloved siblings and in-laws. Thanks for all your prayers.

Table of Contents

List of Tables

List of Figures

List of Symbols

Symbol	Description
$Cov\{Z_1 Z_2\}$	Covariance between two random variables
$\rho_{Z_1 Z_2}$	Correlation coefficient between two random variables
C_Sand	Coarse sand facies
D_Y	Eigenvalue matrix obtained from covariance matrix
$D(\mathbf{h}) = d_{i,j}(\mathbf{h})$	Dissimilarity matrix
$E\{\}$	Expected value
ϵ_{abs}	Error in reproducing facies proportions
F_Sand	Fine sand facies
$F_Z(z)$	Cumulative distribution function
$f_Z(z)$	Probability density function
$\forall$	For all
$\mathbf{h}$	Lag distance
$H1$	Zone or layer 1 of the Hekla reservoir
$H2$	Zone or layer 2 of the Hekla reservoir
N	Number of categorical variables or facies
L	Total number of realizations
m	Statistical mean
m_{error}	Error in reproducing the mean of the reference data
m_{ref}	Mean of the reference data
m_{real}	Mean of the average of all realizations data
M_IHS	Mud dominated Inclined Heterolithic Stratification
N	Number of Gaussian variables or sampled locations
N	Categorical proportions
σ	Standard deviation
S_IHS	Sand dominated Inclined Heterolithic Stratification
σ^2	Statistical variance
S_Shale	Sandy shale facies
Sh_Sand	Shaly sand facies
$Shale$	Shale facies
σ_{error}	Error in reproducing the standard deviation of the reference data

Symbol	Description
σ_{ref}	Standard deviation of the reference data
σ_{real}	Standard deviation of the average of all realizations data
s_k	Facies type
$\sum(\mathbf{h})$	Covariance matrix
$t_{i,j}(\mathbf{h})$	Downward transition probability matrix, t_d
$t_{i,j}(-\mathbf{h})$	Upward transition probability matrix, t_u
$T(\mathbf{h})$	Transition probability matrix
$\boldsymbol{\theta}$	Unit-length vector
$\mathbf{u}$	Location coordinate vector
V_Y	Eigenvector matrix obtained from covariance matrix
X	Data matrix that is obtained from the rotation of Y by data sphering process
Y	Normal score value of Z
Z	Random variable
z	Outcome of a random variable

List of Abbreviations

Abbreviation	Description
2-D	Two-dimensional
3-D	Three-dimensional
CBM	Cell-Based Modeling
CDF	Cumulative distribution function
EDA	Exploratory Data Analysis
GMM	Gaussian Mixtures Model
GSLIB	Geostatistical Software Library
HTPG	Hierarchical Truncated Pluri-Gaussian
MAF	Minimum/Maximum Autocorrelation Factor
MDS	Multidimensional Scaling
MPS	Multiple Point Statistics
MSNT	Multivariate Standard Normal Transformation
OBM	Object-Based Modeling
PCA	Principal Component Analysis
PDF	Probability density function
PMF	probability mass function
PPDE	Projection Pursuit Density Estimation
PPMT	Projection Pursuit Multivariate Transformation
RV	Random Variables
SCT	Stepwise Conditional Transformation
SESIMIRA	SEdimentological SIMulation in IRAp
SGS	Sequential Gaussian Simulation
SIS	Sequential Indicator Simulation
TGS	Truncated Gaussian Simulation
TPGS	Truncated Pluri-Gaussian Simulation
VB	Variogram-Based

Chapter 1

Introduction

Multiple geological rock properties are jointly used to characterize geological deposits and reservoir formations and for developing reliable geological and engineering models. These models are used for resource and reserves evaluation, to study the economic viability of mining, petroleum, and environmental projects, and to make viable operational and economic decisions. The use of multiple variables improves the performance and reliability of the model. In geostatistical subsurface resource modeling, multiple simulated realizations of the subsurface are often generated to characterize geological heterogeneity at various scales; such models describe transitions and trends in the subsurface formation and assess the degree of uncertainty (Pyrcz & Deutsch, 2014). The motivation of this study is discussed, an overview of some statistical concepts used in geostatistical modeling is given, the research problem statement is highlighted, and the thesis statement is summarized.

1.1 Background and Motivation

Knowledge of the distributions and relationships between the continuous (petrophysical) variables of a reservoir is vital to the effective management of the reservoir and future field development. An integrated reservoir study involves the determination of optimum well placement for field development purposes and assessment of in situ resources and reserves. To assess hydrocarbon resources in-place and reserves, petrophysical properties such as porosity, permeability, volume of shale, and connate water saturation, are measured by several methods including special core analysis, well logging, well test analysis, and reservoir simulation techniques.

The hydrocarbon resources in-place depends on the pore volume (porosity) of the formation and the amount of in situ water (water saturation). Reserves assessment or calculation depends on several factors including the ability of the reservoir to transmit fluid through the inter-connected pores (permeability), flow rates of existing and future producing wells, availability of technology to develop the field, favorable economics or current prices of oil and gas, and government regulations. Also, during field development, wells should be optimally placed to penetrate reservoirs at high porosity and permeability regions, and to avoid interference with adjacent wells. Geological and flow simulation models are often developed to study the rock and fluid properties of the reservoir and to predict the performance of the reservoir under certain operating conditions; these models are continually updated with improved knowledge of the reservoir and when more data comes available in future. This process of consistently updating models helps to effectively manage the

reservoir throughout its life.

Geostatistical modeling provides a timely and effective approach to developing reservoir models, assessing the level of uncertainty, and making predictions with observed data (Caers, 2011); missing data are also accounted for. Most projects in the mining and petroleum industries require the characterization and modeling of multiple continuous variables. In petroleum reservoirs, the values and distribution of continuous variables such as porosity and permeability are controlled by the types and distribution of facies in the reservoir. Facies are categorical variables that constitute either the flow units or the impermeable zones of the reservoir. Production history and research have proven that the flow units or facies distribution has a strong influence on the fluid-flow or dynamic behavior of the reservoir (Hatloy, 1994). Hence, it is essential that a numerical facies model is incorporated into a detailed reservoir description, as a means of emphasizing reservoir heterogeneity, which helps to quantify the inherent uncertainty and possible economic risks in reservoir management decisions.

The standard practice in geostatistical reservoir modeling is such that the categorical variables in the reservoir are first characterized and modeled, and used as an input for modeling the continuous petrophysical properties, and as a basis for developing the flow simulation model (Hatloy, 1994; Murray, 1994; Pyrcz & Deutsch, 2014; Silva & Deutsch, 2018). Various stochastic methods and tools have been successfully established to model categorical variables or facies architecture and distributions in deposits and petroleum reservoirs. These tools are broadly divided into object-based and cell-based methods. The object-based methods reproduce the morphological shapes or architecture of the facies such as meandering channels in fluvial environments, while the cell-based methods are those that assign facies on a cell by cell basis conditioned to the observed data from wells or drillholes within the facies.

The SEdimentological SIMulation in IRAp (SESIMIRA) is one of the established object-based facies modeling tools, which has been successfully used to demonstrate how the stratigraphic sequence or layers of facies in a fluvial channel are modeled with a combination of some geological rules, empirical relations, and probability distributions (Hatloy, 1994); IRAp is a Geomatic mapping system. The SESIMIRA modeling concept involves a two-stage modeling approach; the first stage is the generation of simulated facies realizations and the second stage is the conditional simulation to model the multiple continuous petrophysical properties.

Cell-based methods such as Multiple Point Statistics (MPS), Sequential Indicator Simulation (SIS), Truncated Gaussian Simulation (TGS), Truncated Pluri-Gaussian Simulation (TPGS), Hierarchical Truncated Pluri-Gaussian (HTPG) facies modeling methodologies, have been successfully em-

ployed to model facies or categorical variables in geologic deposits or reservoirs (C. V. Deutsch, 2002; Pyrcz & Deutsch, 2014). The MPS methodology uses training images or reference models as the basis for characterizing and modeling facies types in the reservoir. MPS has a capacity to reproduce complex features in the reservoir; however, this method has the limitations of difficulty in handling non-stationarity with the training images and the relative simple features embedded in the training images (Silva & Deutsch, 2017).

SIS, TGS, and TPGS are variogram-based facies modeling methodologies. Variogram-based modeling methods are useful tools in modeling facies. SIS is commonly used to model reservoirs with diagenetically controlled facies, which are characterized by anisotropy that is modeled by the variogram. The HTPG simulation technique is an established categorical variable modeling technique, which is more suited when the geology of the domain of interest has a complex structure.

Modeling multiple continuous variables in a reservoir enables the construction of not only static reservoir models, but also flow simulation models, which are essential for reservoir performance prediction. Like the simulated facies realizations, realizations of continuous petrophysical properties depict the heterogeneity of the subsurface formation, and enables the modeler to quantify the degree of uncertainty. When modeling multiple variables, the older geostatistical modeling methods transform the variables to univariate Gaussian variables and assume a multivariate Gaussian relations before cosimulating the transformed variables with the Sequential Gaussian Simulation (SGS) technique (C. V. Deutsch, 2002; C. V. Deutsch & Journel, 1998; Kelkar & Perez, 2002). However, there is a potential limitation in this assumption, as geological variables often exhibit complex multivariate relationships such as non-linearity, heteroscedasticity, and constraints. To overcome this potential limitation, the Projection Pursuit Multivariate Transformation (PPMT) and other multivariate transformation techniques are motivated. The HTPG and PPMT techniques will be jointly incorporated in an integrated fashion and applied in a single geostatistical reservoir modeling workflow.

1.2 Basic Statistical Concepts

The branch of statistics which involves the modeling of subsurface phenomena is known as Geostatistics. Simply put, geostatistics is defined as the study of spatially distributed random variables. To perform geostatistical analyses, some statistical concepts must be understood. Statistical concepts such as stationarity, random variables, continuous and categorical variables, cumulative distribution functions, and probability density functions are discussed. These statistical parameters are important modeling input parameters in many geostatistical estimation and simulation processes such as kriging and SGS.

1.2.1 Geological Population and Stationarity Concept

In statistics, an important decision on how to group data into populations is made before any detailed analysis. These pools of data or populations are subsequently used for statistical analysis. The decision to pool the data depends on the objectives of the proposed study, availability of specific data types, and the geological setting of the data, among other factors (C. V. Deutsch, 2002).

A reservoir model is often constructed on the basis of specific facies in each stratigraphic layer of the reservoir and each pool of data is referenced to specific facies. In this way, sample statistics are representative of the defined populations. The decision of which data is classified or pooled together for further use as a representative sample statistics is known as stationarity (Pyrcz & Deutsch, 2014). A decision of stationarity implies that the statistics of the pooled data such as the mean, variance, or covariance are independent of any data location (Hohn, 1999). The choice of the pooled data also depends on the geological context and expertise of the modeler. The decision of stationarity is vital in geological modeling because when all data in the domain of interest are grouped into only one population, important trends might be masked.

1.2.2 Univariate Statistics

Generating univariate statistics of regionalized variables gives the modeler knowledge of the quantity and quality of the available data used for geostatistical modeling. The following are the most relevant concepts of geostatistics.

1.2.2.1 Random Variables (RV)

The major aim of predictive statistics is to characterize any unsampled value z as a RV Z, and the probability distribution which models the uncertainty about z (C. V. Deutsch & Journel, 1998; Journel, 1989). A RV is a variable that takes various outcomes from a random experiment according to some probability distribution, such as uniform or normal distribution. The RV is conventionally denoted as Z, while its outcome values are denoted as z. The use of uppercase letter Z indicates that the value is not known. The outcome of the RV is denoted by the corresponding lowercase letter z. Note that the RV is both location and information dependent, that is, the notation $Z(\mathbf{u})$ is often used, where $\mathbf{u}$ is the location coordinate vector. When the RV is distributed through space, it is called a $regionalized\ variable$. Examples of RVs are categorical or continuous RV.

$Categorical\ Random\ Variables$ are variables that can only take a limited set of values or a small number of integers. This can be 1 for Sand and 0 for Shale. The frequency at which each outcome occurs is described by a probability mass function (PMF), as given in Equation 1.1. The categorical variables are also known as $discrete$ RVs. Examples are geological properties such as facies or rock

types, counts of fossil species, etc.

$$F_{Z(\mathbf{u})}(z) = Prob(Z(\mathbf{u}) = z) \tag{1.1}$$

Where $\mathbf{u}$ is a location coordinate vector. For ease in documentation, $\mathbf{u}$ is dropped from the notation $Z(\mathbf{u})$ beyond this point. $Continuous\ Randon\ Variables$ are variables that take real values. In the categorical variable case, the frequency of possible outcomes can be counted. However, for the continuous random variable case, the number of possible outcomes cannot be counted because there are infinite possibilities. Examples of such varying quantities are petrophysical properties (porosity, permeability, water saturation, acoustic impedance, metal concentrations) and geographical attributes (population densities, topographic elevations). The variations of a continuous RV are commonly described in two equivalent ways: (1) Cumulative distribution function (CDF) and (2) Probability density function (PDF).

1.2.2.2 Cumulative Distribution Function

The CDF is a general way of expressing a state of incomplete knowledge of a continuous RV. It is mathematically expressed as:

$$F_Z(z) = Prob(Z \leq z) \tag{1.2}$$

The CDF is a non-decreasing function between 0 and 1, while the values of z are in the range of the working data.

1.2.2.3 Probability Density Function

The PDF, denoted as $f_Z(z)$, is the derivative of the CDF, $F_Z(z)$ provided $F_Z(z)$ is differentiable, Equation 1.3. Generally, the PDF is defined as the integral or surface area of a positive function and this integral denotes a probability, as given in Equation 1.3, (Caers, 2011; Mallet, 2002).

$$f_Z(z) = \frac{d}{dz} F_Z(z) \tag{1.3}$$

$$Prob(a \leq Z \leq b) = \int_a^b f_Z(z)dz$$

Some important properties to be noted are:

$$\int_{-\infty}^{+\infty} f_Z(z)dz = 1$$

$$f_Z(z) \geq = 0$$

$$Prob(Z = z) = 0$$

Although the CDF is commonly used in summary statistics and other calculations, the PDF and $histogram$ are often used for data visualization and interpretation purposes (C. V. Deutsch, 2002). The histogram could be considered as an approximation to the PDF. For a specific interval, the area

of the histogram bar is equal to the proportion or relative frequency of measurement in that interval (Montgomery & Runger, 2003). The relative frequency is simply an estimate of the probability that the observed value of the continuous RV falls within the interval. Similarly, the area under $f_Z(z)$ over an interval, say $\left[a\ b\right]$, is the actual probability estimate that the observed value of the continuous RV falls in that interval, as expressed in Equation 1.3 (Montgomery & Runger, 2003).

1.2.2.4 Expected Values

The probability-weighted average of a random variable RV is known as the expectation or expected value of the random variable. It is a statistical way to summarize the distribution function of the random variable of interest. The expected value of any RV is known as the $mean$, m or the first moment and it is mathematically expressed in Equation 1.4.

$$\int_{-\infty}^{+\infty} z dF_Z(z) = \int_{-\infty}^{+\infty} z f_Z(z) dz = m \tag{1.4}$$

The expected value could be considered as a statistical or linear operator (Rossi & Deutsch, 2013). The expectation of the squared difference from the mean is called the $variance$, σ^2 of the random variable distribution. This is expressed in Equation 1.5.

$$E\left\{[Z-m]^2\right\} = Var\left\{Z\right\} = \sigma^2 \tag{1.5}$$

The $standard\ deviation$, σ is defined as the square root of the variance of the sample or population and it is expressed in the units of the RV of interest. Traditionally, data is summarized by providing the m and σ.

1.2.3 Multivariate Geostatistics

Multivariate geostatistical modeling involves the modeling of random variables, with consideration to the relationship between multiple variables. Such relationships may be important when predicting the variables relevant for resource and reserves evaluation. The fundamental tool that measures the relationship between two RVs is known as the $covariance$. The covariance is used to characterized bivariate distributions and it's unit is the product of the units of the two variables of interest. Consider two RVs Z_1 and Z_2 spatially distributed within a domain of interest A. If the mean of the random variables Z_1 and Z_2 are m_{Z_1} and m_{Z_2}, respectively, then the covariance between the RVs is expressed in Equation 1.6.

$$Cov\left\{Z_1, Z_2\right\} = E\left\{[Z_1 - m_{Z1}][Z_2 - m_{Z2}]\right\} \tag{1.6}$$

The $correlation\ coefficient$, ρ between the RVs Z_1 and Z_2 is defined as the covariance between Z_1 and Z_2 divided by the standard deviations of the Z_1 and Z_2 variables, as expressed in Equation 1.7.

$$\rho_{Z_1,Z_2} = \frac{Cov\left\{Z_1, Z_2\right\}}{\sigma_{Z_1}.\sigma_{Z_2}} \tag{1.7}$$

The correlation coefficient is a dimensionless value between -1 and +1 that depicts either an inverse or a direct linear relationship. If the covariance, $Cov\left\{Z_1, Z_2\right\}$ between the two random variables is zero, the variables are said to be uncorrelated and there is no dominant direct or inverse relationship, but the variables may have a non-linear relationship (Journel, 1989; Rossi & Deutsch, 2013).

1.3 Thesis Statement

An integrated workflow that utilizes geostatistical techniques to improve modeling of reservoir facies and spatial prediction of continuous petrophysical properties in complex geological terrain will yield high-performing models, which aid resource and reserves assessment, uncertainty quantification, and management decision-making.

Modeling of multivariate geological and petrophysical properties is key to generating reliable geostatistical reservoir models for resource and reserves evaluation. Multivariate modeling is required for the analysis and assessment of most subsurface processes in the mining and petroleum industries. The standard practice in geostatistical reservoir modeling is to model the categorical variables in the formation and consider the simulated realizations of the categorical variables as input to model the continuous properties.

HTPG, among other techniques, has been successfully established for modeling categorical variables (Silva & Deutsch, 2018). HTPG was motivated to overcome the limitation of using only 2 or 3 latent Gaussian variables in other truncated Gaussian Simulation techniques. Unlike other truncated Gaussian simulation techniques, HTPG is more suitable to very complex geological settings and the use of a truncation tree structure makes it more convenient to define the truncation rule in the HTPG technique.

The PPMT is a multivariate transformation technique used for multivariate continuous variable modeling. It was motivated to overcome the limitations of the older geostatistical modeling workflows. The application of PPMT ensures that all the transformed variables are uncorrelated and multi-Gaussian; back transformed values reproduce the complex relationships that have been removed. The models developed by the PPMT modeling workflow are likely to out perform those of a conventional modeling workflow (Barnett, Manchuk, & Deutsch, 2016).

This thesis presents a framework that integrates the application of the HTPG and PPMT techniques to model a petroleum reservoir. This methodology seems to provide a practical approach to model the categorical and continuous variables in the reservoir.

1.4 Thesis Outline

In Chapter 2, a comprehensive review is made on the geostatistical modeling techniques used for modeling of categorical and continuous variables. The linear and non-linear transformation techniques, as applied to multivariate continuous property modeling, are discussed. Chapter 3 presents an overview of the theories of the HTPG and the PPMT techniques and highlights the research workflow that integrates the application of both techniques. Chapter 4 discusses the implememtation of HTPG - PPMT integrated methodology to generate multiple realizations of the Hekla reservoir facies, porosity, and permeability. This chapter also discusses model validation and checking of the Hekla reservoir models. Chapter 5 draws conclusions from the results and highlights future works.

Chapter 2

Literature Review

This chapter provides a review of methodologies and tools relevant to this thesis. The first section discusses the techniques that are employed to model categorical variables, while the second section discusses multivariate continuous property modeling techniques.

2.1 Modeling of Categorical Variables

Modeling of categorical variables such as facies is an important step in most reservoir modeling workflows. A facies is a unique body of sediment with some specified characteristics, formed under certain depositional conditions or settings; the unique characteristics include lithology, diagenetic features, primary and secondary sedimentary structures (Heriot-Watt, 1999). Different facies in the reservoir have different porosity and permeability values. Also, the lateral and vertical changes of a particular facies bring about porosity and permeability changes in the reservoir. This means that facies models capture heterogeneity at different scales, help to quantify the uncertainty in the reservoir, and are a key input in the field development plan for the purposes of well placement and risk mitigation.

All facies of interest should have geological significance and sufficient data to allow reliable inference of the required statistics for reservoir modeling (C. V. Deutsch, 2002). The facies of a subsurface reservoir must be properly identified by well logs from the wells penetrating the reservoir and thus, must have clearly defined petrophysical properties and spatial features. Some stochastic methods have been developed and employed successfully to characterize and model categorical variables. These methods are divided into Object-Based Modeling (OBM) and Cell-Based Modeling (CBM) methods.

2.1.1 Object-Based Modeling

The object-based method is suited for facies modeling when there is a good knowledge of the depositional environment and reservoir geometry. By this method, a reservoir model with a defined geometry is initialized with a background facies, while some parameterized geometries (representing other facies) are sequentially placed into the model by some geological rules, until certain criteria including the expected global facies proportions and data conditioning are met (C. V. Deutsch, 2002; C. V. Deutsch & Tran, 2002; Pyrcz & Deutsch, 2014). OBM methods became famous in the oil and gas industry in the mid-1980s and early 1990s when they were extensively used to characterize

the facies in the Norwegian North Sea reservoirs (C. V. Deutsch & Wang, 1996; Egeland, Georgsen, Knarud, & Omre, 1993; Hatloy, 1994; Holden, Omre, & Tjelmeland, 1992). Reservoirs in fluvial deposits tend to be highly heterogeneous; hence, simulating a 2-D or 3-D model of the geological and petrophysical properties of such reservoirs is essential for planning production strategies and for the prediction of recovery efficiency (Egeland et al., 1993).

Egeland et al. (1993) presented a stochastic method for modeling the facies architecture of fluvial reservoirs by using a computer program known as SISA. Their model, which included four different facies types: a background, channel sand, sheetsplay sand, and barriers, was established with a prior model for each facies. The prior model was then developed into a posterior model with conditioning or observed well data. The prior model was based on general knowledge of fluvial reservoirs and specific information from the reservoir of interest such as the size and geometry of the channels. Two examples were demonstrated to explain their methodology: one synthetic case with a limited number of facies and the other case with a North Sea fluvial field (Egeland et al., 1993). The examples demonstrated the following features of the facies model: (1) location, architecture, and behavior of facies bodies, (2) relationship or spatial correlation between the different facies types, (3) conditioning to observed well data.

Hatloy (1994) developed the SESIMIRA concept, an OBM method for facies characterization. This approach is considered general, as it enables the modeler to combine deterministic and probabilistic parameters, with some geological rules and empirical relations, to model facies bodies in a fluvial channel. The model developed by this approach is used for both uncertainty quantification in the facies architecture and as a conditioning input model for developing models of petrophysical properties in the reservoir. The SESIMIRA concept was utilized on IRAp (a Geomatic mapping system) and was used in the mid 1990s (Hatloy, 1994).

A hierarchical object-based method for modeling complex reservoirs in fluvial depositional environment was developed by C. V. Deutsch and Wang (1996), an approach that involves stochastic modeling of fluvial channel shapes and then filling up those shapes with petrophysical properties such as porosity and permeability. This approach was motivated by the obvious geometries observed at outcrops and those viewed aerially in modern fluvial deposits; and emphasis is placed on geologically valid principles and conditioning statistics including vertical and lateral facies proportions (C. V. Deutsch & Wang, 1996). The modeling approach that is hierarchical, whereby a stratigraphic coordinate transform is established for each layer of the reservoir, a number of channel complexes are positioned in the reservoir layers, and then channels are placed in each of the channel complexes. Codes were written in C and Fortran for this hierarchical OBM method. This was later optimized to the $fluvsim$ program (C. V. Deutsch & Tran, 2002), which is derived from

the Geostatistical Software Library (GSLIB) programs (C. V. Deutsch & Journel, 1998).

One of the objectives of OBM methods for facies modeling is to reproduce the geometry or shape of the geo-objects as geometric shapes and preserve the large and small scales features of the geo-objects. The chosen geometric shapes depend on the modeler's knowledge of the depositional environment and the reservoir architecture. In most OBM methods, geo-objects are represented with a template of cells, which provides a reasonable computational time; however, the preservation and connectivity of the large and small scale features of the geo-objects are highly sensitive to the modeler's chosen grid size (C. V. Deutsch, 2002; Hassanpour & Deutsch, 2010). In recent years, a grid-free OBM method has been adopted, whereby geo-objects are represented with parametric shapes instead of using a template of cells (Hassanpour & Deutsch, 2010; Pyrcz & Deutsch, 2014; Zagayevskiy & Deutsch, 2015). This grid-free modeling approach helps to preserve both the large and small scale features in a particular model.

Although the OBM methods generally have the capability of reproducing the appropriate morphological shape or geometry of the objects in reservoirs, they have three notable limitations: (1) Each geometric shape or object that represents facies are individually parameterized, which implies that each OBM algorithm can handle just one type of reservoir, as in $SISA$ (Egeland et al., 1993) and $fluvsim$ (C. V. Deutsch & Tran, 2002); (2) Difficulty in honoring dense well data when the data spacing is compared to the average size of the object; (3) Difficulty in constraining the reservoir models to spatial trends that are represented by maps of areal proportions or vertical facies proportion curve (Strebelle & Journel, 2001). However, beyond these limitations, the OBM can be used to model the facies of reservoirs in complex geological settings, both in fluvial and non-fluvial depositional environments.

2.1.2 Cell-Based Modeling

The cell-based modeling (CBM) methods for modeling categorical variables are used to generate simulated realizations of the variables within a pre-defined grid that represents the domain of the mineral deposit or reservoir. CBM methods such as Multiple Point Statistics (MPS) and Variogram-Based (VB) facies modeling are data-driven, as they are conditioned to the sampled data from wells (core and/or well log data) and seismic within the domain of interest. A number of CBM techniques have been recently employed to model facies prior to modeling the continuous petrophysical properties in the reservoir.

Most cell-based modeling and simulation algorithms target the reproduction of statistics that honor well data. These well data are often sparse and only give estimates of two-point correlation in the

reservoir or domain of interest (Strebelle & Journel, 2001). Since the variogram considers the spatial continuity between two points at a time, the VB methods for facies modeling may sometimes give poor representation of the facies architecture of the reservoir (Pyrcz & Deutsch, 2014; Strebelle & Journel, 2001). In comparison, the OBM methods give a good representation of the actual geometries of the facies bodies but in most cases, fail to honor dense well data. The MPS technique, which expresses spatial correlation at more than two points at a time, is capable of both honoring reservoir data and adequately reproducing the facies geometries. The MPS method uses training images or reference models as the basis for characterizing and modeling facies bodies in the reservoir and has the capability of reproducing complex features in the reservoir. However, this method has difficulty in handling non-stationarity with the training images and reflects only relatively simple features embedded in the training images (Pyrcz & Deutsch, 2014; Silva & Deutsch, 2017).

The Sequential Indicator Simulation (SIS) technique is commonly used to model reservoirs with diagenetically controlled facies, which are characterized by anisotropy that is modeled by the variogram (C. V. Deutsch, 2002; Silva & Deutsch, 2018). Consider K mutually exclusive facies types, $s_k, k = 1, \ldots, K$, the indicator transform at a specific location $\mathbf{u}_\alpha \; \forall$ facies s_k is known as the probability of facies s_k present at that location: this probability is 1 if the facies is present and 0 if not present. The SIS algorithm is such that each grid node is visited randomly and at each node a facies code is assigned by: (1) finding nearby and previously simulated grid nodes and then estimating the conditional distribution by kriging, that is, the probability that a particular facies is present at that node, $p_k, k = 1, \ldots, K$, then (2) drawing simulated facies realizations from the set of probabilities. This procedure is repeated with different random seeds to generate multiple realizations of the facies (Pyrcz & Deutsch, 2014).

The Truncated Gaussian Simulation (TGS) technique generates realizations of a continuous Gaussian variable and then truncates the realizations at a series of thresholds to create realizations of facies (Pyrcz & Deutsch, 2014). TGS is commonly applied to heterogeneous facies that have been diagenetically altered, with no well-defined geometric shapes, and are ordered in some predictable way. In practice, the chosen thresholds for truncating the Gaussian variables to facies are ordered based on the facies codes generated from an underlying continuous variable (Pyrcz & Deutsch, 2014; Silva & Deutsch, 2017). One notable limitation of TGS is that it uses only one latent Gaussian variables.

The Truncated Pluri-Gaussian Simulation (TPGS) technique is one of the variants of the truncated Gaussian simulation method for categorical variable modeling. Unlike TGS, it permits the use of 2 or 3 latent Gaussian variables, which makes it more suited for modeling more complex facies relationships. The simulation of categorical variables is also done by multiple truncations of the la-

tent Gaussian variable realizations. The mapping between the categorical and continuous variables could be modeled with one of the established methods for simulating Gaussian random functions. TPGS is a viable method for modeling facies, however, its application is limited to not more than three Gaussian latent variables (Pyrcz & Deutsch, 2014; Silva & Deutsch, 2018). The HTPG facies modeling methodology, which is discussed in Chapter 3, was motivated to overcome the limitations of the TPGS and TGS methods.

2.2 Multivariate Property Modeling

A detailed reservoir description needed for field evaluation requires models of continuous (petrophysical) properties. The variability or distribution of the petrophysical properties is a function of the facies distribution in the reservoir. Greater variability in the petrophysical properties is observed when there is transition from one facies type to another facies type, while the variability is less for transition within the same facies types. Models of continuous variables such as porosity and permeability are not only essential for performing volumetric calculations, but also capture reservoir heterogeneity and are utilized for sensitivity analysis and uncertainty quantification in fluid production from the reservoir under different production regimes.

A good knowledge of the reservoir, expertise of the modeler, and the availability of valuable information from sources including cores, well logs, seismic, well tests, and field production, will aid the construction of a reliable and high-performance reservoir model. Data from these sources are often multivariate in nature; these multiple rock and fluid properties are jointly used to build models that give a detailed description of the reservoir. This section discusses the old (traditional) and recent techniques that are used in multivariate modeling.

2.2.1 Traditional Modeling Method

Most petrophysical rock and fluid properties tend to be correlated. As such, a geostatistical multivariate modeling approach that employs a co-simulation technique is commonly used to generate realizations of the multiple continuous variables. This is often done by transforming the variables into univariate Gaussian by the traditional normal score transformation method (C. V. Deutsch & Journel, 1998; Isaaks & Srivastava, 1989; Journel & Huijbregts, 1978; Pyrcz & Deutsch, 2014) and assuming that the variables exhibit multi-Gaussian relations. The workflow commonly used in this modeling technique is shown in Figure 2.1. The underlying issues with this modeling workflow are (1) geological variables often have complex non-Gaussian relationships, which violates the assumption of multi-Gaussianity. (2) the cosimulation and associated back-transformation do not effectively reproduce the multivariate relationships of the original variables; hence, the validity and performance of models developed by this workflow are questionable (Barnett & Deutsch, 2015; Bar-

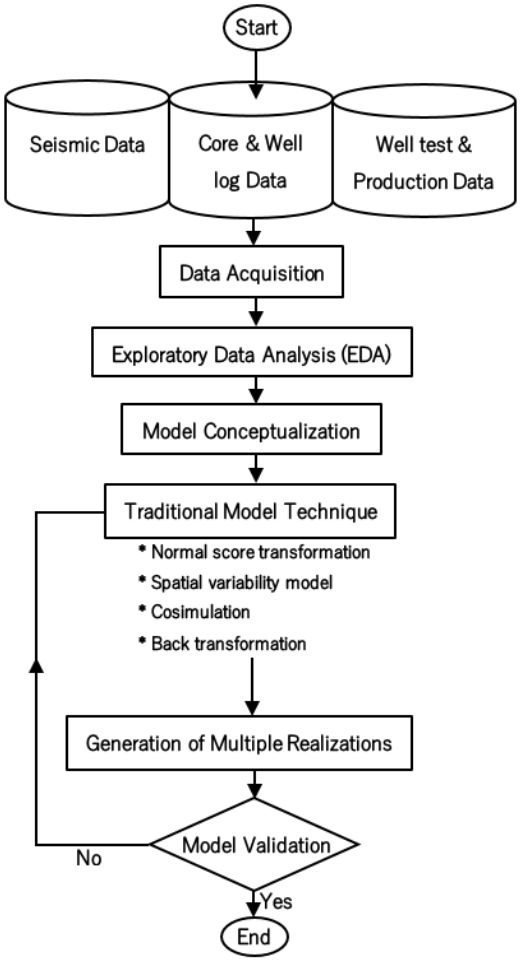


Figure 2.1: Traditional modeling workflow.

nett et al., 2016).

Although the traditional modeling technique has the above-mentioned limitations, it has been successfully applied to evaluate resources and reserves of subsurface mineral deposits and petroleum reservoirs, with some acceptable degree of accuracy (Barnett et al., 2016). The steps involved in the traditional modeling workflow are highlighted below:

1. Normal score transformation of the original multiple variables to Gaussian distribution
2. Variogram modeling with the normal score data
3. Perform conditional Gaussian cosimulation
4. Normal score back-transformation of the generated Gaussian realizations

Another notable downside of this modeling method is the complexity involved in building a linear model of co-regionalization prior to running co-simulations, which becomes more cumbersome

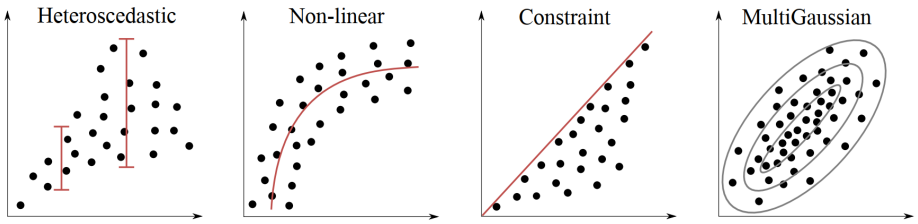


Figure 2.2: Complex relationships of multivariate properties in comparison with a multiGaussian distribution (Barnett & Deutsch, 2015)

with an increase in the number of variables. Recently, this limitation is overcome by removing the correlations between the variables through a multivariate transformation technique and then running independent simulations.

2.2.2 Multivariate Transformation and Independent Simulation Method

The common practice in univariate property modeling is to transform the regionalized variable from its original unit to the Gaussian unit, a process known as normal score transformation (Journel & Huijbregts, 1978; Pyrcz & Deutsch, 2014). Simulation is run to generate multiple realizations of the variable. However, the evaluation of geologic deposits and petroleum reservoirs requires the use of multivariate properties of the subsurface mineral deposits and reservoirs. When multiple variables are modeled, the assumption that univariate Gaussian variables are multiGaussian is not always realistic since geological variables often exhibit complex non-Gaussian features such as non-linearity, heteroscedasticity, and constraints (C. V. Deutsch, 2011). Figure 2.2 shows some of the complexities exhibited by geological variables; note the elliptical contour shape of the multi-Gaussian relationship. A potential issue with the traditional geostatistical modeling workflow, as highlighted above, is that it does not reproduce the multivariate relationships of the multiple variables. To address this issue, the development of multivariate transformation techniques were motivated.

2.2.2.1 Linear Decorrelation Transformation

Principal Component Analysis (PCA) is a multivariate transformation technique that reduces the dimension and decorrelates a correlated multivariate distribution into orthogonal linear combinations of the original variables (Barnett, 2017a). Decorrelation of the multivariate variables facilitates independent geostatistical modeling of the PCA transformed variables, before the PCA back-transformation is used to restore the correlations of the original variables. Consider the data matrix Z: $z_{\alpha,i}$ and the corresponding Gaussian data Y: $y_{\alpha,i}$, $\alpha = 1, \ldots, N, i = 1, \ldots, K$, where N is the number of sampled locations or observations and K is the number of variables. The spatial coordinates of the N observations in Z are denoted by the vectors $\mathbf{u}_\alpha, \alpha = 1, \ldots, N$. The PCA uses the covariance matrix $\sum(\mathbf{h})$ as its input, see Equation 2.1, where $\mathbf{h}$ is the separation or lag distance.

The algorithm performs a pre-processing step, which involves the spectral decomposition of $\mathbf{h} = 0$ covariance matrix, $\sum(0)$ to yield the value in Equation 2.2.

$$\sum(0) : C_{ij} = \frac{1}{N}\sum_{\alpha=1}^{N} y_{\alpha,i}.y_{\alpha,j}, i, j = 1, \ldots, K \tag{2.1}$$

$$\sum(0) = V_Y D_Y V_Y^T \tag{2.2}$$

where V_Y is the eigenvector matrix of the standardized data Y and D_Y is the diagonal eigenvalue matrix. The PCA transform is then performed by multiplying Y with V_Y to yield P in Equation 2.3. This rotates the multivariate data such that the resultant principal components in P are uncorrelated. When P is multiplied by the transpose of V_Y, it is rotated back to Y, which forms the basis of the PCA back-transformation.

$$P = YV_Y \tag{2.3}$$

The Minimum/Maximum Autocorrelation Factor (MAF) transform is an extension of PCA that performs a two-step spectral decomposition of $\sum(\mathbf{h})$ at $\mathbf{h} = 0$ and $\mathbf{h} > 0$ lag distances; this will lead to a better cross-covariance reproduction in simulated realizations (Barnett, 2017b). Since the PPMT, which is discussed in Chapter 3, only removes the correlation at $\mathbf{h} = 0$, the MAF transform is an effective post-processing tool if any $\mathbf{h} > 0$ correlation persists after PPMT.

2.2.2.2 Non-linear Transformation

Gaussian techniques are commonly applied in geostatistical reservoir modeling to develop geological models due to their simplicity; these techniques require the variables of interest to be multivariate Gaussian. The traditional normal score transformation usually generates univariate Gaussian distributions but does not enforce bivariate or multivariate Gaussianity. Some techniques have been developed to transform multiple variables in a non-linear fashion to generate multi-Gaussian distributions of the transformed variables.

The Stepwise Conditional Transformation (SCT) technique was conceptualized and first implemented by Rosenblatt (1952) for multivariate data transformation and has been applied successfully in geostatistical reservoir modeling workflow (Leuangthong, Deutsch, Haas, & Shtuka, 2000). SCT technique ensures that all the transformed variables exhibit multi-Gaussianity with a correlation of zero. Thus, independent simulation is run to generate multiple realizations with no requirement to fit a model of coregionalization. The SCT technique is similar to the conventional normal score transformation when the first variable is transformed. Consider a K - variate problem with N observations, the K variable is conditionally transformed with reference to the $(K-1)$ variables as

described below.

$$y_1 = G^{-1}[F_1(z_1)] \tag{2.4}$$

$$y_2 = G^{-1}[F_{2|1}(z_2|y_1)] \tag{2.5}$$

$$y_3 = G^{-1}[F_{3|2,1}(z_3|y_2, y_1)] \tag{2.6}$$

$$\vdots$$

$$y_K = G^{-1}[F_{K|1,\ldots,(K-1)}(z_K|y_1, \ldots, y_{(K-1)})] \tag{2.7}$$

Where $y_i, i = 1, \ldots, K$ are the multi-Gaussian variables that are independent at lag distance of zero, $\mathbf{h} = 0$, that is, covariance is equal to zero (Leuangthong & Deutsch, 2003).

$$C_{i,j}(\mathbf{h}) = C_{i,j}(0) = C[Y_i(\mathbf{u}), Y_j(\mathbf{u})] = 0, i \neq j, i, j = 1, \ldots, K \tag{2.8}$$

The SCT technique requires the data to be binned; the data are separated into classes based on their conditional probabilities, then each class of data is normal score transformed. It is worthy to note that the multivariate spatial relationship of the original regionalized K-variables is not transformed at lag distances greater than zero ($\mathbf{h} > 0$), which means no changes are made to the bivariate spatial distribution $Y(\mathbf{u})$ and $Y(\mathbf{u} + \mathbf{h})$, or higher multivariate spatial distributions (Leuangthong & Deutsch, 2001). Hence, the multivariate distributions at $\mathbf{h} > 0$ are not Gaussian and the covariance at $\mathbf{h} > 0$ is not equal to zero, unless if such assumption is made for simplicity.

$$C_{i,j}(\mathbf{h} > 0) = C[Y_i(\mathbf{u}), Y_j(\mathbf{u} + \mathbf{h})] \neq 0, i \neq j, i, j = 1, \ldots, K \tag{2.9}$$

For the bivariate case, this covariance is given as

$$C_{1,2}(\mathbf{h} > 0) = C[Y_1(\mathbf{u}), Y_2(\mathbf{u} + \mathbf{h})] \neq 0 \tag{2.10}$$

The primary motivation for employing the SCT technique for multiple variable transformation is because it facilitates independent simulation of the transformed variables instead of the rigorous cosimulation method that requires models of coregionalization for all the cross-covariances at $\mathbf{h} > 0$. Although this data transformation has proven very useful for modeling multiple variables, it has the limitation of under-performing with large amount of data and there are often some minor artifacts due to binning and reduced data in transformation of high dimensional distributions (C. V. Deutsch, 2011).

The Multivariate Standard Normal Transformation (MSNT) concept was motivated to overcome the limitations of the SCT and the multivariate linear transformation tools. Consider N observations of K variables from an arbitrary multivariate distribution, the MSNT technique uses a modified Latin hypercube sampling algorithm to generate N observations from a K-variate multivariate standard normal distribution; then each of the original N data are technically mapped to the standard normal

observation (C. V. Deutsch, 2011). This concept is similar to the traditional normal score transform and the SCT in that each data observation in the original units is mapped onto specific observation in the standard normal distribution. However, while the MSNT technique maps the multivariate observations directly, the normal score and SCT consider only one variable at a time (C. V. Deutsch, 2011).

Like the SCT technique, the MSNT technique yields transformed variables that are treated independently, which would thus facilitate quick independent simulation to generate multiple realizations of the K variables. Although the MSNT concept is promising, some challenges implicit to this technique include (1) the requirement or need for large amount of data to reasonably span the input multivariate space; limited data implies far apart neighbors and interpolation becomes difficult, and (2) the requirement for equal data sampling, that is, having equal number of variables at all sample locations.

To overcome the potential limitations of the SCT and MSNT techniques, Silva and Deutsch (2015) developed an SCT algorithm, gmm_sct, a derivative of the GSLIB that transforms multivariate data with SCT using Gaussian Mixtures Model (GMM) as reference distribution. First, a GMM is fitted to the data and the GMM is referenced in the transformation steps of SCT. This permits the use of SCT in high dimensions and for sparse data, thereby eliminating artifacts due to binning. Note that the transformed multivariate data may not be perfectly decorrelated. Such deviations are anticipated due to the use of limited data and the use of a fitted distribution as reference (Silva & Deutsch, 2015). The associated GMM back-transformation is achieved by using gmm_sctb.

Although these linear and non-linear decorrelation techniques have been successfully employed to transform multiple variables prior to simulating them geostatistically, the PPMT technique has been frequently applied in recent times for multivariate data transformation. PPMT is not constrained by the dimensionality of the multivariate data and perfectly removes the correlation between multiple variables at lag distance $\mathbf{h} = 0$. The concept of the PPMT technique is discussed in the next chapter and it is utilized by the modeling workflow adopted in this study.

Chapter 3

Methodology

This chapter is divided into three sections. The first section discusses the HTPG concept, which is applied in this research for categorical variable modeling. For more insight into the theory and implementation of the HTPG, see the detailed documentation by Silva and Deutsch (2018). The second section presents the PPMT theory; see Barnett and Deutsch (2015) for the complete guide on modeling with the PPMT. The third section discusses the workflow adopted in this study.

3.1 Hierarchical Tuncated Pluri-Gaussian

Categorical variable is often available to geostatistically model and characterize a petroleum reservoir and the categorical variable has a number of possible outcomes or categories. In reservoirs, for example, the distinct rock categories commonly observed are sand, shales, and carbonates. Based on the particle size distribution, sand could be observed and characterized as very fine sand, fine sand, and coarse sand. In fluvial depositional environments with mixed sediments and extremely strong currents, such mixed sediments are often characterized as unique categories or facies, such as shaly-sand and sandy-shale. Modeling of facies architecture of a reservoir helps to discern the regions and layers of the reservoir that will probably be most productive or constitute barriers to fluid flow. The conventional modeling approach is either to combine the categories or to model them independently. When multiple categories are modeled independently, key statistics, such as the joint categorical proportions, are not properly reproduced (Rossi & Deutsch, 2013).

The concept of truncating continuous variables to model categorical variables is very flexible. The TPGS approach for modeling categorical variable does not explore the full capacity of the truncation concept (Pyrcz & Deutsch, 2014; Silva & Deutsch, 2017). Regardless of the number of categories under consideration, for practical purpose, the TPGS is limited to the use of only two or three continuous variables (Silva & Deutsch, 2018); however, there is no theoretical restriction on the number of Gaussian variables that could be used. Also, the concept of using truncation masks in TPGS is limiting, since the geological interpretation degrades when a larger number of categories and/or continuous latent Gaussian variables are available. The increased difficulty in mapping the spatial continuity of the categorical variable to the continuous space is often used to justify the restriction to two latent Gaussian variables in TPGS applications.

The Hierarchical Truncated Pluri-Gaussian (HTPG) technique uses a tree structure to define trunca-

tion rules. This facilitates the definition of the mapping between continuous and categorical space, which makes it possible to effectively use any number of Gaussian variables for the modeling of categorical variables (Silva & Deutsch, 2018). The use of a tree structure gives flexibility to the HTPG approach and makes its application suited for complex geological settings. There are five developed steps for the application of truncated Gaussian methods, which are:

1. Definition of the truncation rules;
2. Mapping of spatial continuity from categorical to continuous space;
3. Imputation of continuous data;
4. Simulation of continuous variables at the modeling nodes; and
5. Truncation of the continuous variable realizations to generate realizations of the categorical variable (Silva & Deutsch, 2018).

Although all these steps are key to the success of the HTPG technique, the truncation rule impacts the entire process and the generated categorical realizations.

3.1.1 Truncation Rule

The definition of a truncation rule is a fundamental step in the HTPG approach to modeling categorical variables. It determines the quality and dependability of the generated models. The HTPG truncation rule is defined by a decision tree, where the parent nodes represent the latent Gaussian variables and the leafs on the tree represent the resultant categories. Thresholds are applied to the nodes and the proportion of categories present in the simulated realizations is controlled by the truncation thresholds.

The main purpose of defining the truncation rule is to introduce geological understanding of the sedimentary depositional setting into numerical models; and this controls the contacts between the categories or facies. When defining the truncation rules for HTPG, geological expertise of the depositional environment should be the main consideration. Geological structures are technically defined from information in the observed data. These vital pieces of information are often obtained from core samples, well logs, and outcrop examination. The geological information that goes into the HTPG truncation rule is mainly qualitative. By carefully analyzing and interpreting the available geological information, structures including folds, faults, intrusions, and unconformities, are properly defined and factored in when defining the truncation rule for the rock layering sequence or strata.

Transition probabilities and Multidimensional Scaling (MDS) mapping are also used to complement the definition of the truncation rule. The calculation of the transition probability matrix is discussed in Chapter 4. It is a quantitative tool that has been used for data-driven definition of truncation rules. Note that the objective of the HTPG technique is to facilitate geological interpretation and allow the geo-modeler to construct the truncation tree based on individual expertise. Also, lack of transition between two categories is often observed, which could be caused by various factors. For instance, the chronological order of a sedimentary sequence may introduce a physical barrier between two geological units in which the transition between one unit to another is not possible without the transition to intermediary units. In this case, there is a sharp well defined barrier between the two geological units and the transition or lack of transition is enforced using the truncation rule. In certain instances, there may be no sharp well defined barrier, and the two categories do not coexist at the same region. In these cases, the lack of transitions could be enforced using locally varying proportions (Silva & Deutsch, 2018).

3.1.2 Global and Local Proportions

The categorical proportion is defined by the probability, p_k, of observing a certain categorical variable at a given location, $\mathbf{u}$. If the proportions of the categories do not change with location, the categorical variables are said to be stationary and the proportions are given by Equation 3.1.

$$p_k(\mathbf{u}) = p_k, k = 1, \ldots, K, \forall \mathbf{u} \in A \tag{3.1}$$

However, categorical variables are often non-stationary, as their proportions change with location in the domain of interest. As such, locally varying proportions are calculated to facilitate the modeling of trends and calculating the indicator residuals in the geological domain, as the spatial variability of the categorical variable is a combination of the continuity of the local proportions and the stochastic residuals (Silva & Deutsch, 2018). Indicator variograms calculated directly from categorical data without consideration of the spatial structure of the trend leads to realizations that are more spatially continuous than the underlying variable. Also, in order to achieve appropriate variogram reproduction while modeling categorical variables with trends, it is important to calculate and utilize the variogram of the indicator residuals (Silva & Deutsch, 2018).

3.1.3 Local Thresholds

Consider the truncation tree in Figure 3.1, defined to interpret the stratification of five facies (coarse sand (C_Sand), fine sand (F_Sand), sand dominated Inclined Heterolithic Stratification (S_IHS), mud dominated Inclined Heterolithic Stratification (M_IHS), and shale ($Shale$)) in a geological domain. Four Gaussian variables and four thresholds are utilized to define the truncation rule,

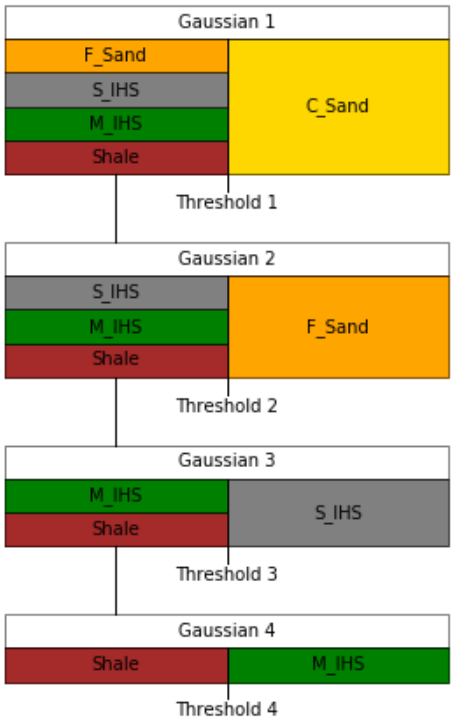


Figure 3.1: Sample truncation tree

where the C_Sand facies is placed above all the facies and the $Shale$ facies underlying all the facies in the geological domain. The F_Sand realizations are mapped from the realizations of Gaussian variable 2, while M_IHS and $Shale$ realizations are controlled by the Gaussian variable 4. The threshold applied on the HTPG truncation tree controls the proportions of the categorical variables in the simulated realizations. Thresholds are often calculated from local varying proportions and are well distributed along the latent Gaussian variables. The Gaussian variable to which thresholds are applied are well specified in the truncation structure. These thresholds are applied when mapping from continuous to categorical space, as discussed in Chapter 4.

3.1.4 Mapping between Continuous and Categorical Variable Spaces

The HTPG algorithm uses the truncation rule and calculated local thresholds to map the spatial continuity from categorical to continuous variable space. The latent variable variograms are calculated and modeled by this mapping process. These variograms are used as a key input parameter in the latent variable data imputation process and in subsequent simulation that generates realizations of the categorical variables.

Imputation of latent Gaussian variables at data locations is done by the HTPG algorithm developed by Silva and Deutsch (2018). This is treated as a missing data problem since in practice, the latent variables are not originally sampled at the data locations but are needed as conditioning input

data during simulation. The imputation process generates multiple realizations of latent variable data that are simulated at a defined grid and from which the categorical variables are mapped. During the simulation process, the HTPG algorithm truncates the latent variable realizations at certain thresholds to generate multiple realizations of the categorical variables.

3.2 Project Pursuit Multivariate Transformation (PPMT)

Realizations generated from simulations are considered representative of the subsurface deposits or reservoir when they reproduce both the univariate and multivariate statistics of the variables of interest. The traditional modeling workflow which is often used in multivariate property modeling uses a cosimulation techniques, with the assumption that the multivariate distribution of the variables is Gaussian. This workflow is capable of reproducing univariate properties but does not reproduce the multivariate relationships between the variables. To resolve this potential limitation, the projection pursuit multivariate transformation PPMT was conceptualized and has been applied successfully in mineral deposit and petroleum reservoir modeling (Barnett & Deutsch, 2015). The PPMT is known to reproduce the multivariate relationships of the variables and thus improve the performance of the reservoir model during evaluation.

The idea of projection pursuit for multivariate data analysis was first implemented successfully by (Friedman, 1987). Friedman (1987) conceptualized the Projection Pursuit Density Estimation (PPDE) technique, which was used for multivariate density estimation. The PPMT is based on the modified components of PPDE. The limitations of linear rotation, PCA and MAF, and the SCT are basically the motivation for the development of the PPMT for geostatistical modeling (Barnett et al., 2016).

Unlike the SCT, the PPMT technique transforms data of any multivariate form. K variables at N sampled locations or observations are transformed to an uncorrelated multi-Gaussian distribution, that is, the PPMT is not constrained by the dimensionality of the multivariate data. The fundamental operation of the projection pursuit algorithm is to find the vector $\boldsymbol{\theta}$ with the most non-Gaussian projection. The multivariate data along this projection are then transformed to make the projection values Gaussian. The PPMT continues the search and transform iteration process until the least non-Gaussian projection approximates to the univariate Gaussian model (Barnett et al., 2016).

The primary goal of the PPMT workflow is to decorrelate the multivariate variables of interest and transform them into a multiivariate Gaussian distribution. This zero correlation could be visually observed from bivariate scatter plots of the PPMT data, the Gaussianity of the PPMT data could also be observed from the multi-Gaussian density contour scatter plot, showing the kernel density. The

PPMT algorithm starts with normal-score transform of the original data Z and removes all marginal complexities in the Gaussian data Y (Barnett & Deutsch, 2015; Barnett et al., 2016; Madani, 2019). Y is then rotated by the data sphering process to yield the data matrix X in Equation 3.2.

$$X = S^{-1/2}Y \tag{3.2}$$

where $S^{-1/2} = V_Y D_Y^{-1/2} V_Y^T$, D_Y and V_Y are the eigenvalue and eigenvector matrices obtained from spectral decomposition of the covariance matrix, $\sum(\mathbf{h}) = V_Y D_Y V_Y^T$. After sphering, the projection pursuit is achieved by the projection of the sphered data X on the vector $\boldsymbol{\theta}$ (Barnett et al., 2016; Madani, 2019). The PPMT modeling steps are shown in Figure 3.2.

1. Normal score transformation of the original multivariate data
2. Apply PPMT algorithm to the normal score transformed data
3. Variogram modeling with normal score data
4. Independent simulation is run with the normal score variogram model as input parameter and the PPMT data as conditioning data.
5. Realizations are PPMT back-transformed to the original space.

Note that the multivariate data used for PPMT must be declustered and homotopic, as the PPMT algorithm is applicable to equally sampled or homotopic data.

3.3 Research Workflow

Five major steps are often taken to develop a reliable model that is used to calculate in situ resources and reserves, and to predict future reservoir performance to a reasonable degree of accuracy. The first step involves building a structural framework or conceptual model of the reservoir. The second step is to construct a model of facies distribution in the reservoir. The third step is to characterize and model the continuous petrophysical properties, using the facies model as conditioning input data. The fourth step involves scaling-up of the geological model to generate a dynamic simulation model. The fifth step involves history matching or calibration of the flow simulation model with observed field data, a process which fine-tunes the final model and makes it fit for future performance prediction. The focus of this research is mainly on the second and third steps, which involves modeling of reservoir facies and petrophysical properties.

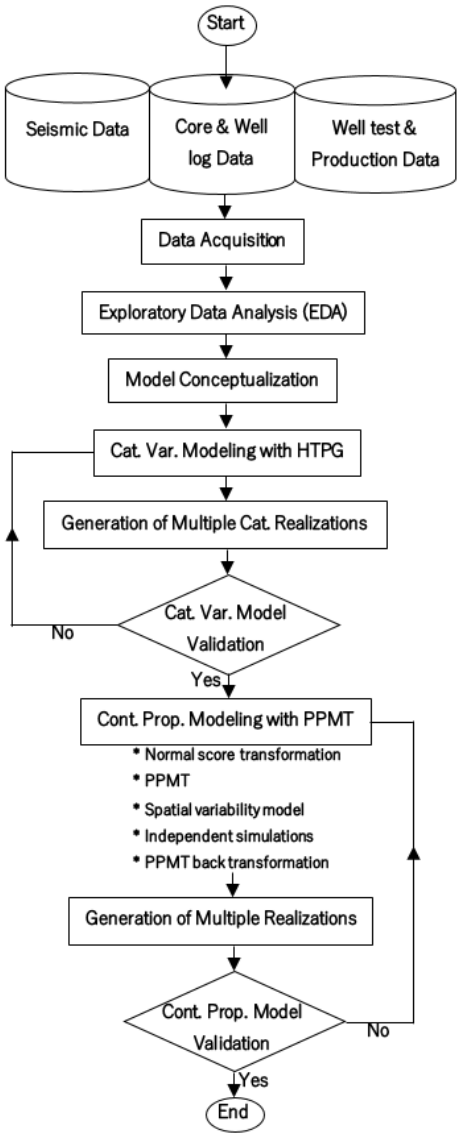


Figure 3.2: Research workflow.

As shown in Figure 3.2, the workflow adopted in this study is an integrated methodology that incorporates the application of the HTPG and PPMT techniques. In the first stage of the research, the HTPG is used for modeling the facies in the reservoir to generate multiple realizations of the reservoir facies. The second stage involves conditioning the facies realizations as input to model the petrophysical properties (porosity and permeability) in the reservoir. The PPMT is employed here as the multivariate transformation tool prior to simulating the properties independently. Note that at different stages in the workflow, the categorical and continuous variable models are validated by data reproduction method. Codes are written in Python in Jupyter notebook and this workflow is implemented with an actual reservoir case study in Chapter 4.

Chapter 4

Implementation and Case Study

This chapter discusses the practical application of the integrated HTPG - PPMT modeling workflow on the Hekla reservoir and it is divided into four sections. The first section is a description of the geometry and location of the Hekla reservoir, the second section provides the multivariate Exploratory Data Analysis (EDA) performed on the Hekla reservoir dataset, the third section discusses the application of HTPG simulation technique to model the facies of the reservoir, while the fourth section discusses the implementation of the PPMT technique to model the continuous petrophysical properties of the Hekla reservoir.

4.1 The Hekla Reservoir

The Hekla reservoir is a portion of a large North Sea reservoir covering about 5000m x 6500m in area. The geology of the reservoir represents a fluvial environment with a channel, marginal depositions, and background lithological elements (Zagayevskiy & Deutsch, 2016). The Hekla reservoir has two distinct layers, $H1$ and $H2$, and the facies in the two layers of the reservoir are uniquely distributed. A close look at the surface maps of the reservoir shows that the reservoir is structurally faulted, with most of the wells strategically positioned between the two noticeable normal faults, and the top-right region of the reservoir is slightly folded, see Figures 4.1 and 4.2. The position of the wells (spherical dots) shows that most of the wells are drilled or concentrated around the 45° line in the x-y plane of the reservoir. Shallow portions of the reservoir are denoted by the blue color, while the deeper regions are coded in red.

4.2 Exploratory Data Analysis

EDA is an important step in all statistical and geostatistical workflows. It is a process that involves cleaning and investigating the data to discover possible anomalies and trends like outliers and make some inferences and decisions from the visualization and statistics that are generated. A detailed EDA is a key to the success of any data-driven research. This process is important as it enables us to eliminate unwanted data (outliers), that is, values that are too large or too small in a range of data. The EDA yields important data visualization and summary statistics, which reveal possible geological trends and anomalies that may probably require further investigation. The approach used to perform a comprehensive EDA is dependent on the objectives of the proposed study and the nature of the dataset. An EDA performed on multiple variables requires more analyses than that performed on a single variable, since the relationships between a variable and other variables are

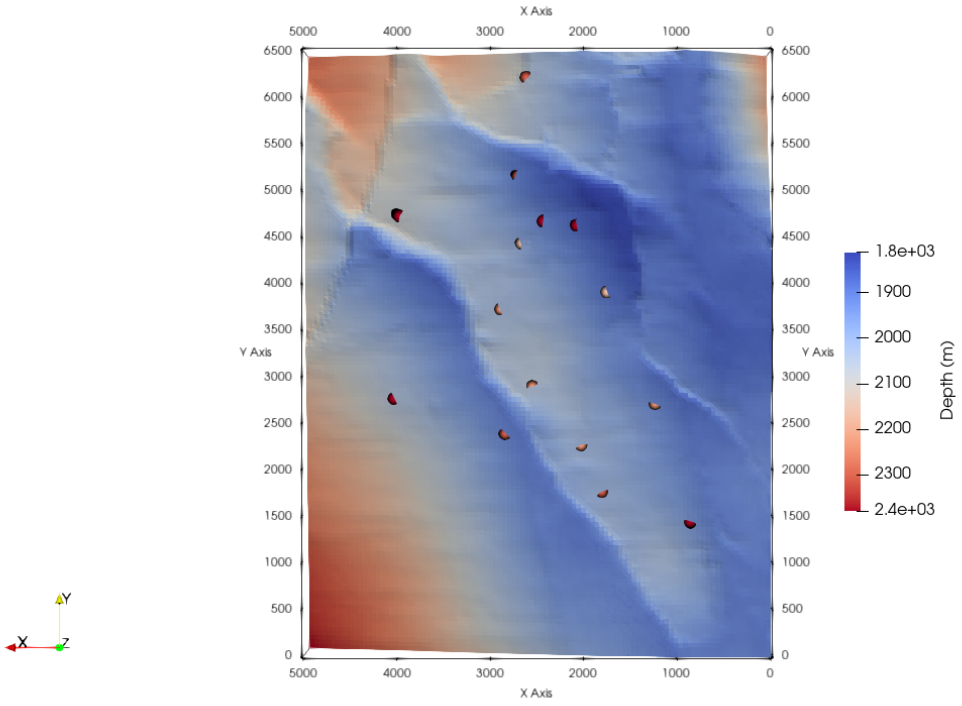


Figure 4.1: Plan view of the Hekla reservoir showing 15 well locations as spherical dots

also carefully studied, whether they are collocated or not. As a major step in the research workflow, EDA is performed on the Hekla reservoir dataset.

4.2.1 Data Inventory

The Hekla dataset comprises two sets of data: (1) 2-Dimensional surface seismic data - at a resolution of 101m x 131m across the entire extent of the reservoir, and (2) 3-Dimensional well data - that contains well locations, facies encountered, petrophysical log porosity, permeability, and acoustic impedance data from fifteen wells in the reservoir. All fifteen wells have well-defined locations, drilled between the two normal faults and across two distinct zones or layers ($H1$ and $H2$) of the Hekla reservoir, as shown in Figures 4.1 and 4.2. The reservoir contains five different facies or rock types that are widely distributed in the area of study of the reservoir. The seismic dataset contains surface depths of each layer ($H1\ Top$, $H2\ Top$, and $H3\ Top$) and acoustic impedance values ($H1\ Impedance$ and $H2\ Impedance$) that are exhaustive of the reservoir areal extent. These regionalized variables of the Hekla reservoir are carefully analyzed in the EDA process.

4.2.2 Well Logs

Well logs are generated from the well data. This is necessary to visualize the variation in petrophysical properties of the reservoir with depth. The well logs shown in Appendix A give a clearer

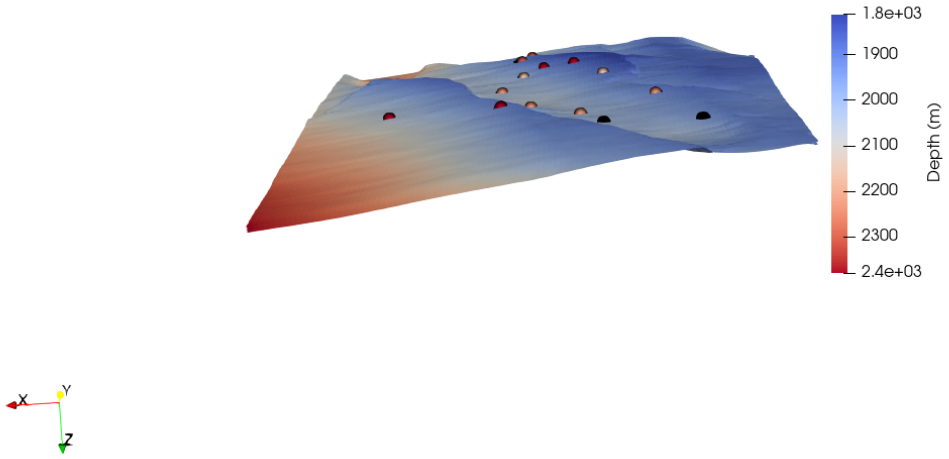


Figure 4.2: 3-D surface map of the Hekla Reservoir, see Figure 4.1 for scale.

Table 4.1: Facies in the Hekla Reservoir

Facies 1	Fine Sand (F_Sand)
Facies 2	Coarse Sand (C_Sand)
Facies 3	Shaly Sand (Sh_Sand)
Facies 4	Sandy Shale (S_Shale)
Facies 5	Shale

picture of the characteristics of each well, in terms of depth, the layers of the reservoirs, the facies encountered by each well, and the associated acoustic impedance, porosity, and permeability, at different reservoir depth. We could infer from all 15 logs that Facies 1, 2, and 3 have relatively good porosity and permeability distribution, which make them easily producible. They could be classified as $Sand$, however, the degree of fineness and cleanliness of the sand is not accurately known. Facies 4 has relatively poor porosity distribution, with little or no permeability; this facies could be a mix or intercalation of mud and sand, with more mud content. The relatively low permeability of Facies 4 encountered by Well 17, as shown in Figure A.13, could result from the presence of interconnected fractures. Facies 5 has very little or no porosity and no permeability, and could be classified as $Shale$. Hence, the codes or names in Table 4.1 are used for identification of the facies in this study.

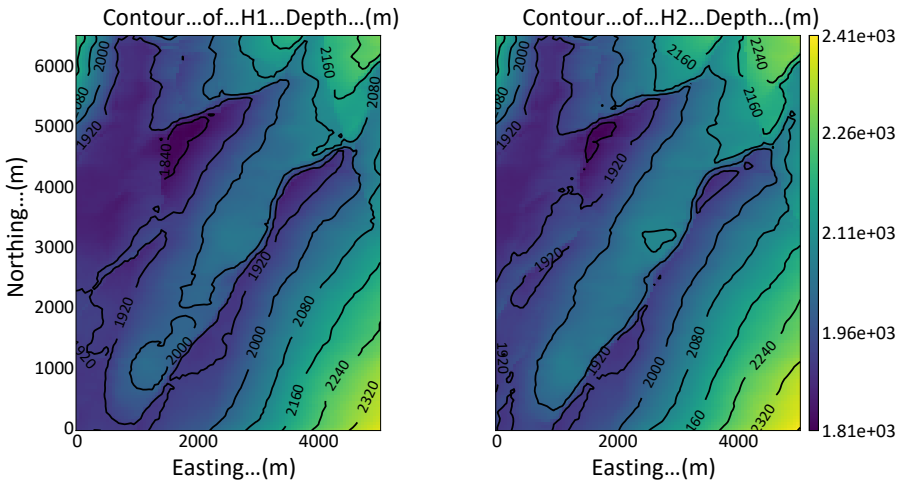


Figure 4.3: Contour map of the $H1$ and $H2$ layers of the Hekla reservoir.

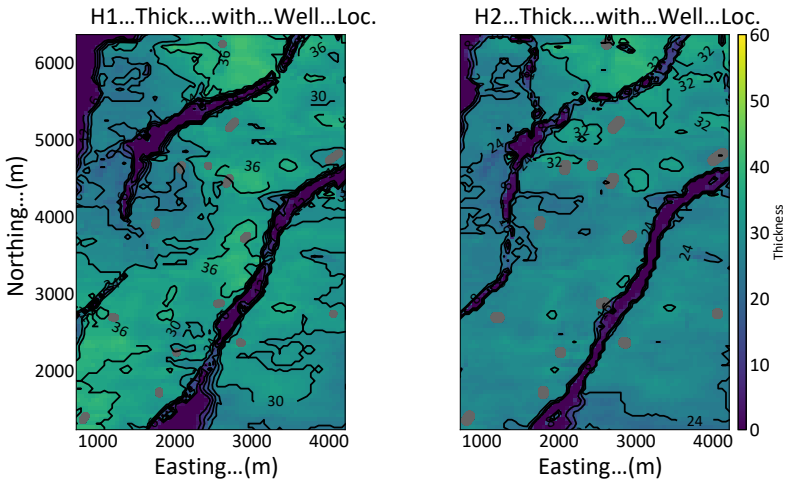


Figure 4.4: Thickness map of the $H1$ and $H2$ layers of the Hekla reservoir.

4.2.3 Contour and Isopach Maps

Lines of constant depth and thickness are plotted to generate the Hekla reservoir contour and isopach maps, respectively. This was done to visually analyze the shallow and deep areas of the reservoir, and the portion with reasonable thickness. The maps in Figure 4.3 show that the deepest areas of the reservoir are on the edge or flank of the two layers of the reservoir, while the maps in Figure 4.4 show the thickness of the two layers of Hekla reservoir. The H1 layer is relatively thicker than the H1 layer, as depicted by the color code of the map, with light-green to yellow color more evident on the $H1$ layer isopach map. The average thickness of the $H1$ and $H2$ layers are approximately $24.8m$ and $23.2m$, respectively.

Table 4.2: Facies count and proportion in $H1$ of the Hekla Reservoir

Facies	F_Sand	C_Sand	Sh_Sand	S_Shale	Shale
Count	103	1355	300	331	2347
Proportion	0.023	0.305	0.068	0.075	0.529
Declustered Proportion	0.023	0.282	0.062	0.095	0.538

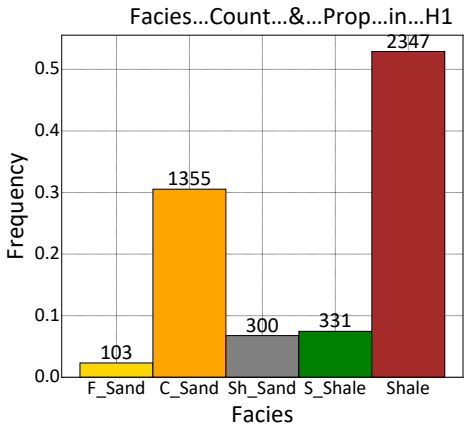


Figure 4.5: Facies count and proportion in $H1$ of the Hekla reservoir.

Table 4.3: Facies count and proportion in $H2$ of the Hekla Reservoir

Facies	F_Sand	C_Sand	Sh_Sand	S_Shale	Shale
Count	36	646	133	397	2695
Proportion	0.009	0.165	0.034	0.102	0.69
Declustered Proportion	0.01	0.152	0.033	0.114	0.691

4.2.4 Facies Analysis

Facies analysis is done by total count and proportions of the facies in $H1$ and $H2$ layers of the reservoir. Tables 4.2 and 4.3 show the summary statistics of facies in each layer of the Hekla reservoir. We can infer from the statistics that layer $H1$ has a higher count of Sand ($Facies$ 1, 2, and 3) than layer $H2$. It implies that $H1$ would probably be more productive than $H2$, with consideration to other parameters such as thickness, average reservoir pressure, fluid viscosity, wellbore flowing pressure, and surface choke size. The plots in Figures 4.5 and 4.6 are visual representations of the summary statistics of facies in the reservoir.

4.2.5 Multivariate Relationship

Since the dataset contains multiple continuous random variables which are distributed in the Hekla reservoir, the bivariate relationship of the variables is studied. The scatter plots of porosity and permeability by facies, in each layer, are plotted to ascertain the correlation between porosity and permeability in the reservoir, as shown in Figures 4.7 through 4.11. The porosity - permeability bivariate relationship in the figures indicates that $Facies$ 2 with high facies count, and high porosity

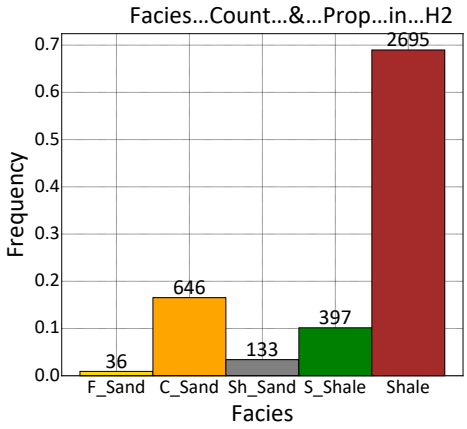


Figure 4.6: Facies count and proportion in $H2$ of the Hekla reservoir.

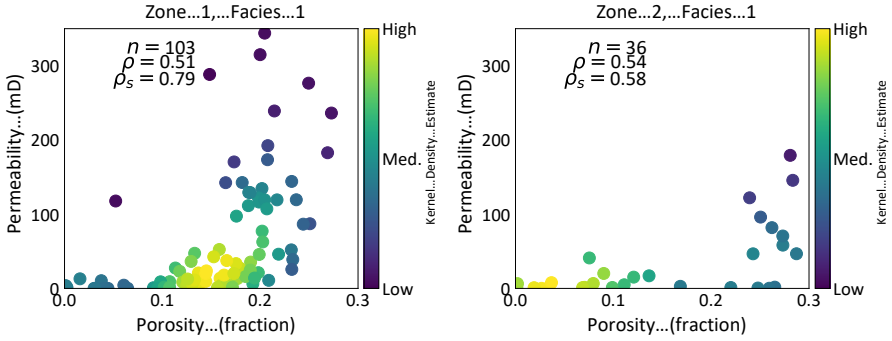


Figure 4.7: Bivariate plot of Porosity vs. Permeability in $Facies$ 1 of $H1$ and $H2$ layers of the Hekla reservoir.

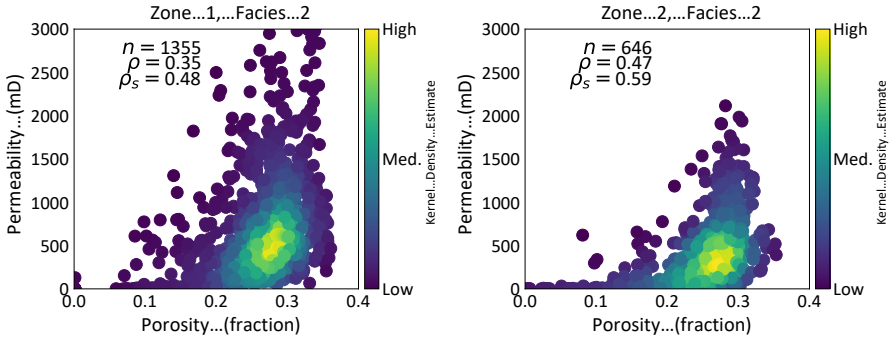


Figure 4.8: Bivariate plot of Porosity vs. Permeability in $Facies$ 2 of $H1$ and $H2$ layers of the Hekla reservoir.

and permeability values, as reflected in the kernel density estimates, is the most relevant of all five facies of the Hekla reservoir and will probably constitute the most productive flow unit in the reservoir. Although $Facies$ 5 has the highest facies count, it is the least relevant because of its low porosity and permeability values.

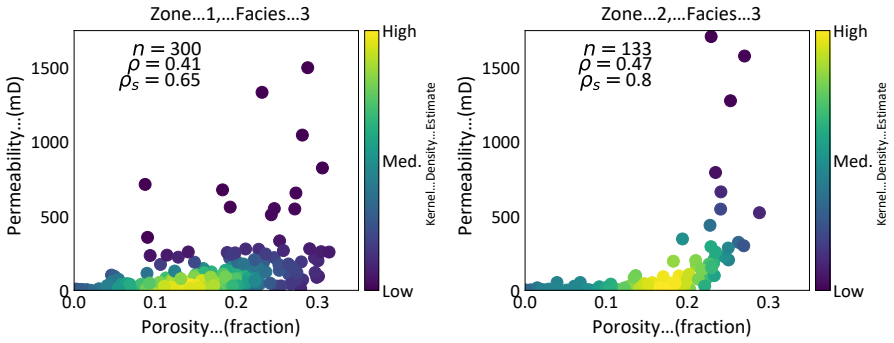


Figure 4.9: Bivariate plot of Porosity vs. Permeability in $Facies$ 3 of $H1$ and $H2$ layers of the Hekla reservoir.

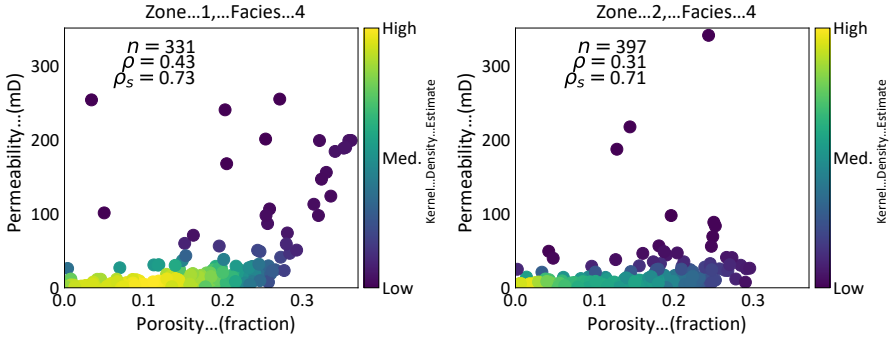


Figure 4.10: Bivariate plot of Porosity vs. Permeability in $Facies$ 4 of $H1$ and $H2$ layers of the Hekla reservoir.

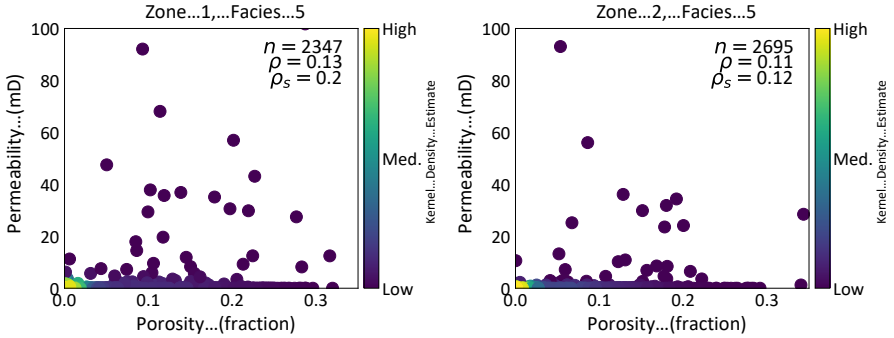


Figure 4.11: Bivariate plot of Porosity vs. Permeability in $Facies$ 5 of $H1$ and $H2$ layers of the Hekla reservoir.

4.2.6 Acoustic Impedance

The Hekla dataset presents acoustic impedance values, which are carefully analyzed to obtain valuable statistics that are representative of the reservoir. The histograms of the gridded (seismic) impedance and those impedance values recorded at well locations in each layer of the reservoir are plotted as shown in Figures 4.12 and 4.13. Although it is expected that the seismic data be exhaustive, the statistics show that the seismic impedance is limited when compared to the well impedance. This is due to large number of missing acoustic impedance values in the Hekla seismic

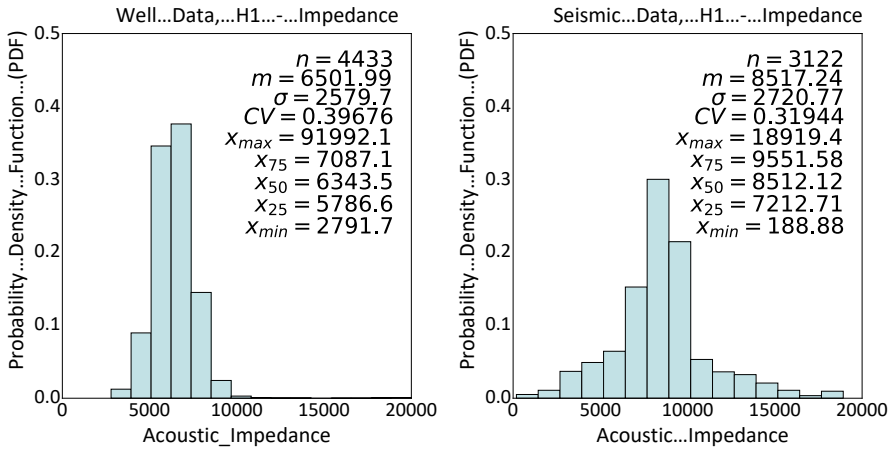


Figure 4.12: Histograms of seismic and well impedance values in $H1$ of the Hekla reservoir.

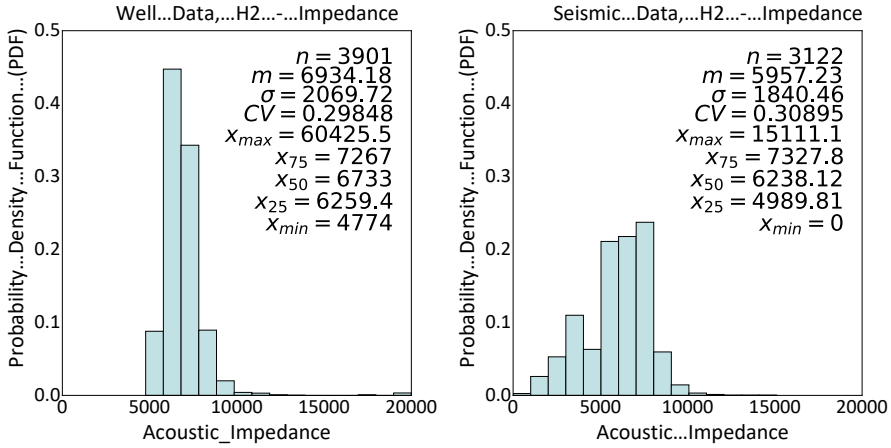


Figure 4.13: Histograms of seismic and well impedance values in $H2$ of the Hekla reservoir.

dataset. Note that the wide difference between the maximum seismic and well impedance values indicates that there could be some discrepancies in the measurement of acoustic impedance either during seismic acquisition or log measurement at well locations.

4.2.7 Calibration of Seismic Property

Integrating seismic data to construct a reliable predictive 3-D reservoir model is very important. Unlike other data sources like core and well logs, seismic provides key information about reservoir properties for the almost or entire areal extent of the reservoir. Seismic attribute such as acoustic impedance is often integrated to construct a 3-D facies model and continuous variable models of a reservoir. However, seismic data are sometimes difficult to use or integrate because of the imprecision in the large-scale seismic data when used for reservoir characterization and modeling.

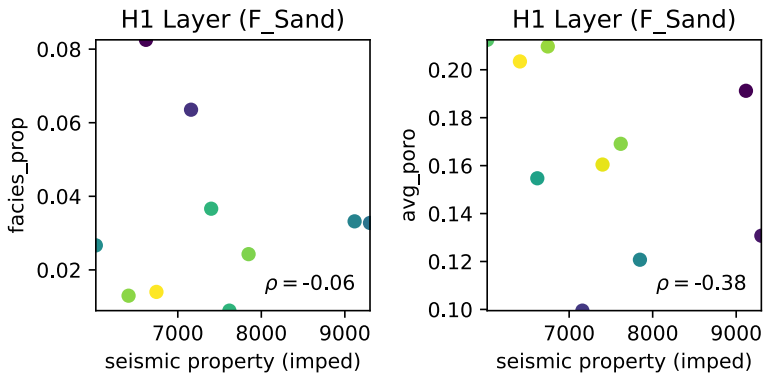


Figure 4.14: Seismic impedance vs. facies proportion and Seismic impedance vs. facies porosity of F_Sand facies in Hekla $H1$ layer

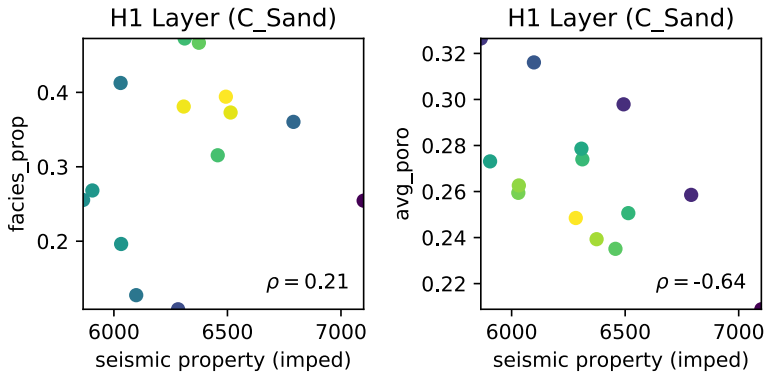


Figure 4.15: Seismic impedance vs. facies proportion and Seismic impedance vs. facies porosity of C_Sand facies in Hekla $H1$ layer

In this study, the Hekla reservoir seismic property is calibrated to determine the quality of the seismic data. Calibration is done by correlating the acoustic impedance at each well location to the proportions of facies and the average porosity of each facies observed in each well cutting across the $H1$ layer of the Hekla reservoir. The acoustic impedance - facies proportion/porosity relationships are shown in Figures 4.14 through 4.18. For the Hekla seismic data to be relevant to this study, it is expected that the acoustic impedance should have a high negative correlation to the sand (F_Sand, C_Sand, and Sh_Sand) facies proportions and porosity, and a high positive correlation to the shale (S_Shale and $Shale$) facies proportions and porosity. Contrary to what is expected, the acoustic impedance is weakly correlated to the facies proportion and porosity of the reservoir. This weak correlation of the seismic acoustic impedance to the facies proportions in both sand and shale makes the seismic impedance data not relevant to this study. The poor correlation between the acoustic impedance and the facies proportions may be due to the small number of wells, among other possible factors.

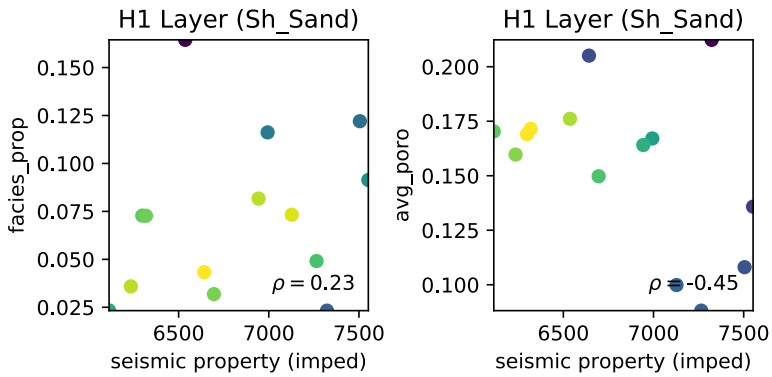


Figure 4.16: Seismic impedance vs. facies proportion and Seismic impedance vs. facies porosity of Sh_Sand facies in Hekla $H1$ layer

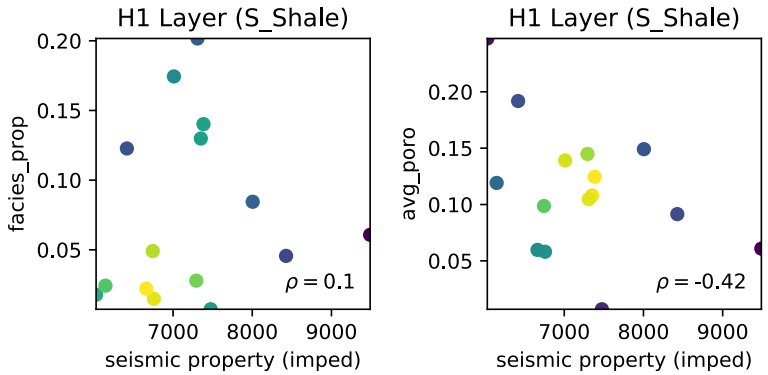


Figure 4.17: Seismic impedance vs. facies proportion and Seismic impedance vs. facies porosity of S_Shale facies in Hekla $H1$ layer

The EDA carried out on the Hekla reservoir is not exhaustive for a multivariate statistical or geostatistical study. EDA is a flexible process, as it depends on the nature and amount of data available and the objectives of the proposed study. It is always advisable that a thorough and exhaustive data processing and multivariate EDA should be performed as a major step in the multivariate geostatistical modeling workflow. The EDA process will provide modelers with sound knowledge of the quality and quantity of data relevant to the proposed study.

4.3 HTPG Application

The Hekla dataset presents five facies. The HTPG simulation technique is applied to model and characterize the facies in the $H1$ layer of the reservoir. This layer seems to have higher proportions of sand when compared to the $H2$ layer, and may constitute the most productive zone of the reservoir.

The HTPG software package developed by Silva and Deutsch (2018) has four programs or executa-

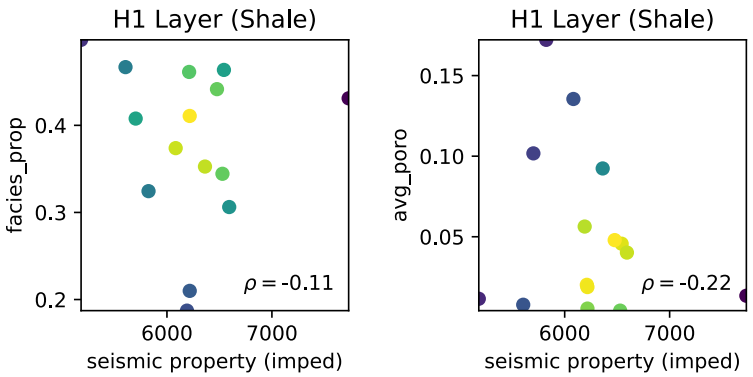


Figure 4.18: Seismic impedance vs. facies proportion and Seismic impedance vs. facies porosity of $Shale$ facies in Hekla $H1$ layer

bles that are utilized to facilitate the HTPG workflow; they are:

1. $htpg\_thresholds$ - which calculates thresholds from categorical (global and local) proportions and defines the truncation parameters

2. $htpg\_gaussvarg$ - which uses the categorical variable indicator variograms to define the spatial structure of the latent Gaussian variables.

3. $htpg\_cat2gauss$ - which performs imputation of latent Gaussian variables based on the observations of the categories.

4. $htpg\_gauss2cat$ - which uses the defined truncation rule to map from continuous space to categorical space, a process that generates simulated realizations of the categories.

The HTPG programs are coded in Fortran and are derived from the GSLIB (C. V. Deutsch & Journel, 1998). Python codes with some Pygeostat functions are written in Jupyter Notebook to run these programs. Beginning with coordinate transformation, the steps taken to model the facies in the $H1$ layer of the Hekla reservoir with HTPG are discussed.

4.3.1 Stratigraphic Coordinate Transformation

Most oil and gas accumulations are found in reservoirs that exist as a series of structurally deformed stratigraphic layers. This deformation mainly results from post-depositional events such as tectonic uplift, differential compaction, and erosion, and leads to the formation of structures and features like folds, faults, and domes. Each layer in the reservoir corresponds to a particular depositional event, which can be correlated with seismic and well log data. The presence of folds and/or faults often constitutes complex reservoir layer architectures that pose a significant challenge in geostatistical modeling. Such architectures make it difficult to capture the original spatial continuity of facies and petrophysical properties.

Before calculating the experimental variogram, some prerequisites must be satisfied. The coordinates of the domain of interest may be transformed to align with the directions of spatial continuity. Hence, folded and faulted reservoirs in petroleum, and tabular deposits in mining, are always flattened or unfolded. This process of coordinate transformation restores the deformed stratigraphic layers to their original depositional state. This facilitates a straightforward calculation and interpretation of two-point statistical variogram calculation and also helps to preserve the original spatial continuity of the facies and petrophysical properties of interest. Calculating an experimental variogram before coordinate transformation may significantly underestimate the extent of horizontal continuity of the facies or reservoir properties of interest.

Each reservoir layer is defined by the existing top and base surface grid, $z_{et}(x, y)$ and $z_{eb}(x, y)$, where x and y are areal coordinates, et refers to the existing or present top, and eb represents the existing or present base (Pyrcz & Deutsch, 2014). The continuity of the facies and the petrophysical properties within a layer does not follow the grids that are based on existing boundary surfaces; additional correlation grids may be required to define stratigraphic continuity within each layer.

Four common correlation styles shown in Figure 4.19 have been successfully used for coordinate transformation to facilitate the restoration of existing deformed layer structures to their original state; they are (1) $Proportional$: where the existing top and base grids coincide with the correlation grids; however, the strata may have varying thickness due to differential compaction and/or sedimentation rate, (2) $Truncation$: where the existing grid bottom coincides with the correlation grid base but the strata top does not conform to the correlation grid top due to surface erosion, (3) $Onlap$: where the existing strata top has not been eroded and thus coincides with the correlation grid top but the existing strata base does not conform to the correlation grid base, (4) $Combination$: here, both the existing top and base of the strata do not conform to the top and base of the correlation grids.

Each stratigraphic layer of the reservoir is independently modeled with a new relative stratigraphic coordinate, z_{rel}, that is defined by the equation proposed by C. V. Deutsch (2002). By this transformation, each layer of the reservoir is modeled in a regular Cartesian x, y, z_{rel} coordinates. Note that the transformation retains the original x and y coordinates of the strata, only the z coordinate is transformed to z_{rel} (M. V. Deutsch & Deutsch, 2014), as expressed in Equation 4.1. The associated back-transformation to the original stratigraphic coordinates is given in Equation 4.2.

$$z_{i,rel} = \frac{z_i - z_{cb}}{z_{ct} - z_{cb}}.T_{avg}, i = 1, \ldots, N \tag{4.1}$$

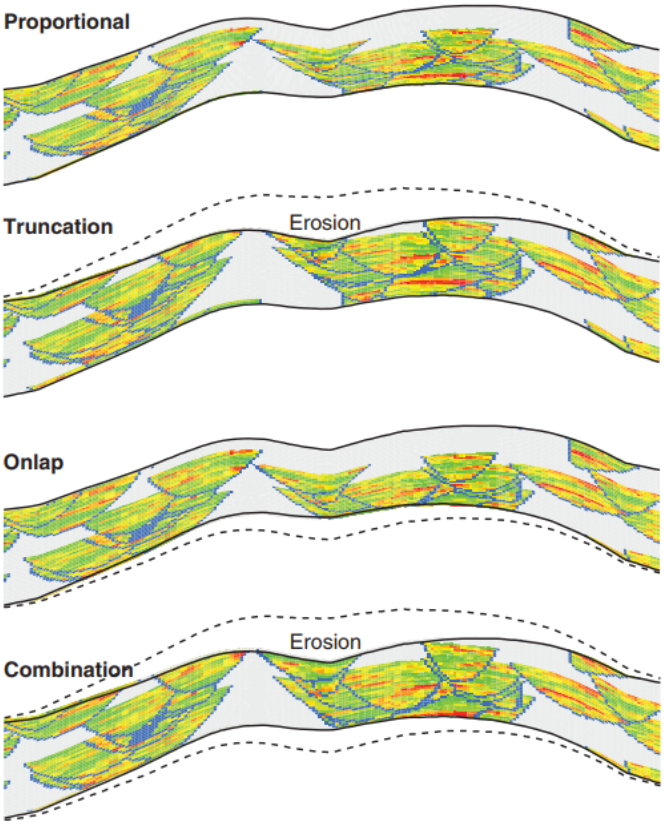


Figure 4.19: Proportion, Truncation, Onlap, and Combination stratigraphic coordinate system transformations, where the original surface is represented by dotted lines and existing surface with solid lines (Source: (Pyrcz & Deutsch, 2014)).

Where:

z_i = depth of the existing strata,

z_{cb} = base of the correlation grids,

z_{ct} = top of the correlation grids,

T_{avg} = average thickness of the existing strata or layer, calculated from z_{eb} and z_{et}, and

N = number of observations or samples

$$z_i = z_{cb} + \frac{z_{i,rel}}{T_{avg}}.(z_{ct} - z_{cb}) \tag{4.2}$$

Since the workflow in this study considers modeling the facies and petrophysical properties of interest independently by layers, the coordinate transformation of $H1$ layer of the Hekla reservoir is considered. $H1$ has an average vertical data spacing $vert_{spacing}= 0.11m$ and an average well spacing in the $x - y$ plane $hor_{spacing}= 887.75m$. The $proportional$ correlation style, defined by the above equation, is applied with the following parameters on per well basis.

z_i = depth of H1 layer in the Hekla well data, z_{cb} = base of the correlation grids of the Hekla seismic data, z_{ct} = top of the correlation grids of the seismic Hekla data, and T_{avg} = average thickness of H1

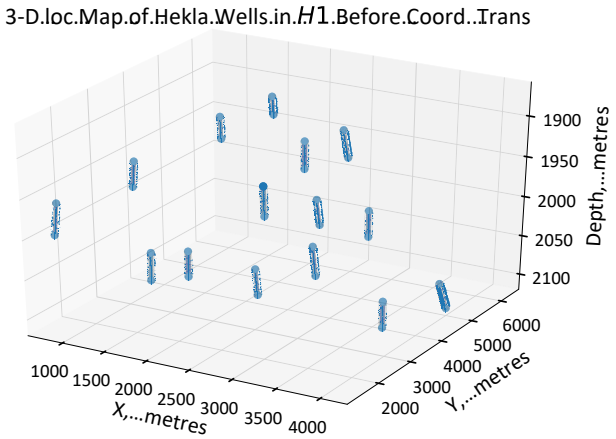


Figure 4.20: 3-D location map of wells in $H1$ of the Hekla reservoir, before stratigraphic coordinate transformation

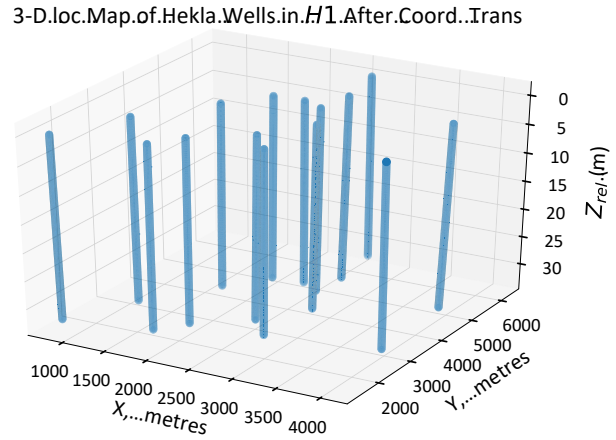


Figure 4.21: 3-D location map of wells in $H1$ of the Hekla reservoir, after stratigraphic coordinate transformation

layer ($228m$), calculated from z_{eb} and z_{et} in the Hekla well data

By this transformation, the z_{rel} Cartesian coordinate of the wells in $H1$ of the Hekla reservoir was calculated and further used for trend modeling, experimental indicator variogram calculation, and subsequent simulations in the workflow. The location maps of the Hekla wells before and after coordinate transformation are shown in Figures 4.20 and 4.21.

4.3.2 Trend Model

The global proportions presented in Table 4.2 show that the facies in $H1$ of the Hekla reservoir are unequally distributed. The declustered proportion varies from a very low value of 0.02 in F_Sand to a high value of 0.54 in $Shale$. With the calculated declustering weights, the local proportions of facies in $H1$ are calculated. It is a common practice to account for trends or locally varying proportions when modeling categorical variables since categorical variables are often non-stationary. The

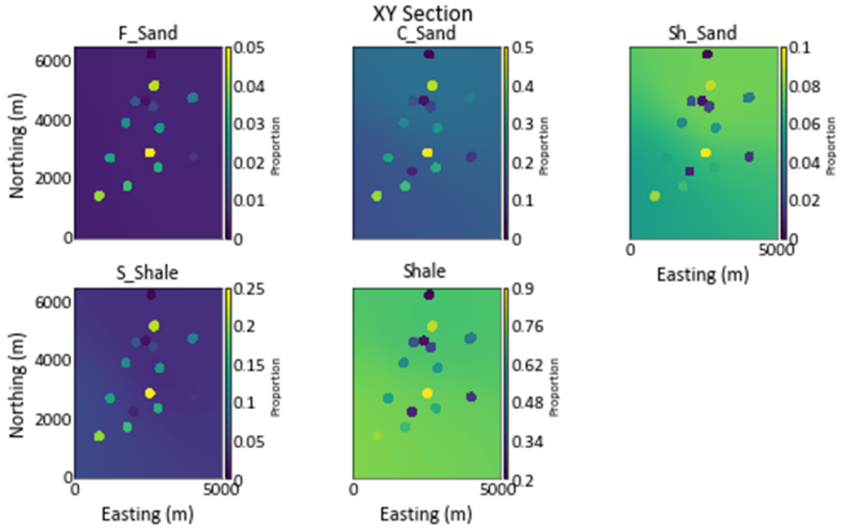


Figure 4.22: Local proportions of the facies in layer $H1$ in $x-y$ orientation, with 15 well locations

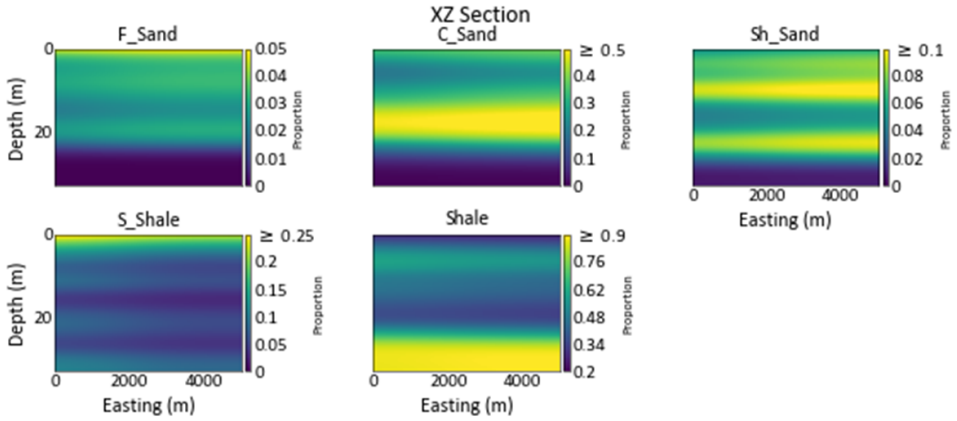


Figure 4.23: Local proportions of the facies in layer $H1$, $x-z$ orientation

local proportions of the five facies in $H1$ layer of the Hekla reservoir are shown in Figures 4.22, 4.23, and 4.24. It is shown that in the lateral ($x-y$) direction of the reservoir, the F_Sand, C_Sand, and Sh_Sand facies trend from low to a high proportion, from the South-West to the North-East region of the reservoir; while the S_Shale and $Shale$ facies trend from high to low a proportion in that direction. In the vertical ($x-z$ and $y-z$) direction, the facies trend from low to a high proportion from the West to East region of the reservoir. The trend also reviews that the reservoir is highly stratified, which is evident in the well logs in Figures A.1 through A.15.

4.3.3 Indicator Residual Variograms

The Indicator (residual) variograms are calculated and modeled directly from categorical variables with trends. Note that the spatial variability structure of the categorical variable is the combination of the continuity of the deterministic trend and the stochastic residuals (Silva & Deutsch, 2017).

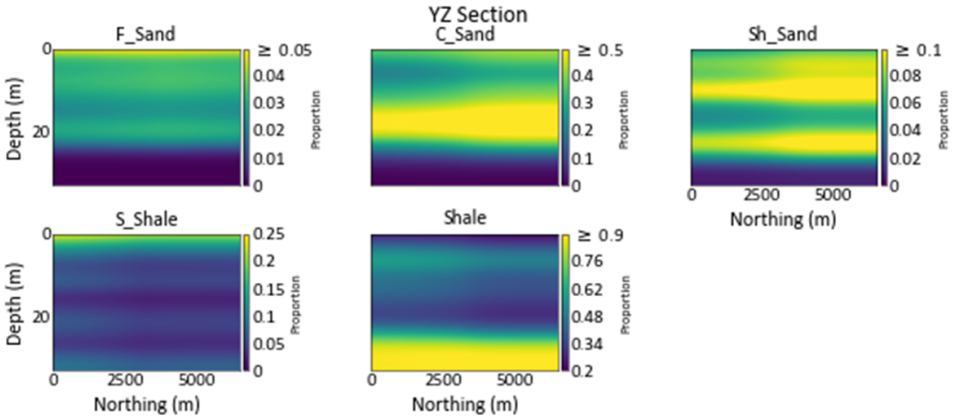


Figure 4.24: Local proportions of the facies in layer $H1$, $y - z$ orientation

To achieve a good variogram reproduction when modeling categorical variables with trends, it is a good practice to calculate indicator residual variograms and use them for subsequent simulations.

Residuals of the Hekla facies and calculated from the trend or local varying proportions and the indicator residual variograms of the five facies in the Hekla reservoir $H1$ layer are calculated and modeled, as shown in Figure 4.25. The variograms are modeled with a spherical model in two directions - vertical direction with a lag distance of $0.5m$ and 10 lags, and horizontal (omnidirectional) with a lag distance of $650m$ and 4 lags. For all five facies, variogram models with good spatial continuity are obtained in the vertical direction than in the lateral direction due to fewer number of wells in the reservoir. Also, the vertical variogram model of C_Sand and $Shale$ facies are more continuous than the vertical variogram model of F_Sand, Sh_Sand, and S_Shale facies, which implies that the C_Sand and $Shale$ facies are widely distributed in the reservoir vertically than the other facies; this is evident in the facies count and proportions in Table 4.2, and the well logs in Figures A.1 and A.15.

4.3.4 Definition of Truncation Rule

The use of truncation (decision) tree makes the HTPG technique very flexible and suited to model complex geological domains. The expertise of the modeler and good knowledge of the sedimentary depositional environment come into play when defining the HTPG truncation rule qualitatively. The transition probability and MDS mapping are important quantitative tools that can be used to define the truncation rule from available data. These quantitative measures compliment the knowledge of the geological terrain to define the truncation rule.

In this study, the truncation rule is defined quantitatively. The transition probabilities are calculated upward and downward along the wells. The downward transition probability matrix, $t_d = t_{i,j}(\mathbf{h})$

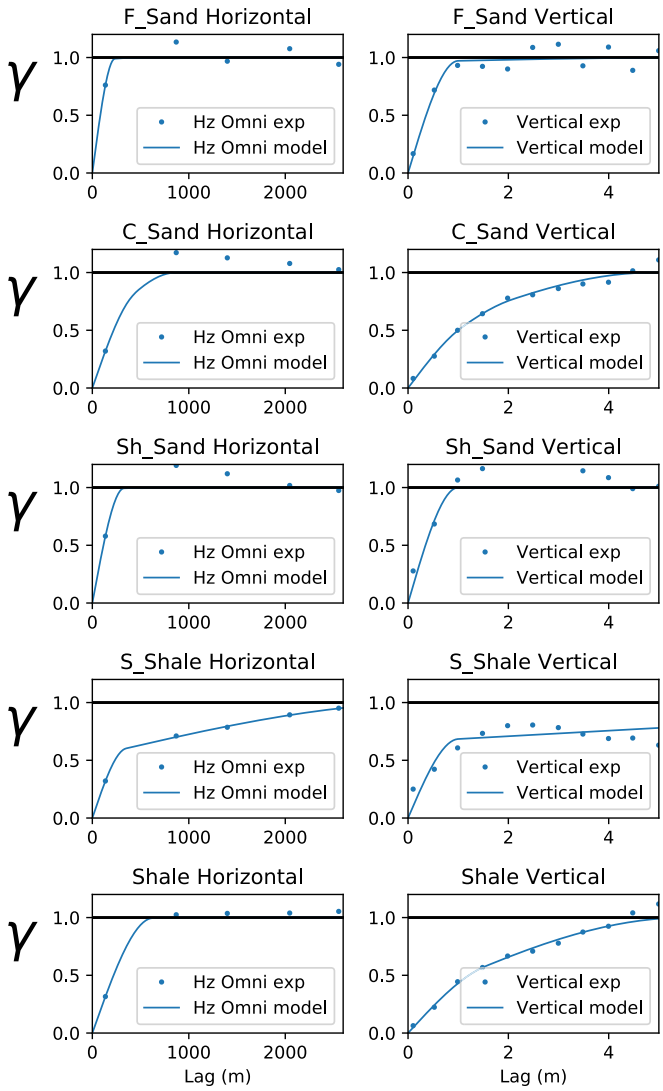


Figure 4.25: Indicator residual variograms of the facies in Hekla reservoir $H1$ layer

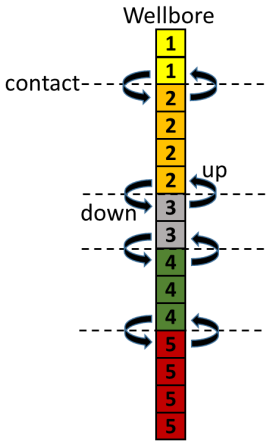
that is calculated along the well is often not equal to the upward transition probability matrix, $t_u = t_{i,j}(-\mathbf{h})$; hence, the transition probability matrix is often asymmetric, that is, $t_{i,j}(\mathbf{h}) \neq t_{i,j}(-\mathbf{h})$ (Silva & Deutsch, 2018). The transition probabilities measure spatial continuity or variability and are calculated for the shortest possible lag distance equivalent to the vertical data spacing derived from well logs; and the high density of data along the wellbore aids the definition and interpretation of geological contacts in the reservoir. For a given lag distance or vertical data spacing $\mathbf{h}$, the transition probability between two facies is defined by Equation 4.3 (Silva & Deutsch, 2018).

$$t_{i,j}(\mathbf{h}) = P\left\{\mathbf{1}_j(\mathbf{u}+\mathbf{h}) = 1 | \mathbf{1}_i(\mathbf{u}) = 1\right\}$$

$$t_{i,j}(\mathbf{h}) = \frac{P\left\{ \mathbf{1}_i(\mathbf{u}) = 1 \text{ and } \mathbf{1}_j(\mathbf{u}+\mathbf{h}) = 1 \right\}}{P\left\{\mathbf{1}_i(\mathbf{u}) = 1\right\}},\ i,j\ \epsilon \left\{1,\ldots,K\right\} \tag{4.3}$$

Where $\mathbf{1}_i(\mathbf{u})$ and $\mathbf{1}_j(\mathbf{u}+\mathbf{h})$ are two different indicator variables separated by lag distance $\mathbf{h}$, i and j are indices that differentiate one facies indicator from another, and K is the total number of categorical variables or facies indicators, which also represents the rank of the transition probability matrix that is derived (Silva & Deutsch, 2018). For illustration purposes, the wellbore in Figure 4.26 is used to demonstrate the transition of five facies along a wellbore and how the resulting transition probability matrix, $T(\mathbf{h}) = \frac{1}{2}(t_d(\mathbf{h}) + t_u(\mathbf{h}))$, is calculated. The diagonal terms of the transition matrix measure the transition between same facies, while the off-diagonal terms measure the transitions from one facies to a different facies. From the matrix T in Figure 4.26, facies 1 has a 75% probability of transitioning between itself along the wellbore and 25% probability of transitioning to facies 2, but however, does not transition or have noticeable geological contacts with facies 3, 4, and 5. Facies 4 does not transition to facies 1 and 2, but it has a 16.5% chance of transitioning into facies 3 and 5, and 67% probability of transitioning between itself. So, the information obtained from the transition probability matrix provides a meaningful guide to defining the HTPG truncation rule.

Tables 4.4 shows the calculated transition probability of the five facies in the Hekla reservoir $H1$ layer. The $Shale$ and C_Sand facies or categories have the highest probabilities of transitioning to same category (expressed by the diagonal terms), which is also an indication of their dominant proportions or wide distribution in the $H1$ layer of the reservoir. The large difference between the diagonal and off-diagonal terms makes it quite difficult to interpret to transition probability matrix. In such a case, the transition matrix is best interpreted visually with the aid of the MDS mapping in two dimensions, as shown in Figure 4.27. The MDS, however, uses a dissimilarity matrix, $\mathrm{D}(\mathbf{h}) = d_{i,j}(\mathbf{h})$, which is obtained by standardizing the off diagonal values of the transition matrix with the transpose; this gives symmetry to the dissimilarity matrix. Also, the diagonal terms of the transition matrix are set as zero since a particular facies type cannot be different from itself, as defined in Equation 4.4. The dissimilarity matrix of the facies in the Hekla reservoir $H1$ layer is given by Table 4.5. For any given pair of facies in the reservoir, the lower their dissimilarity value, the higher the likelihood of transition between the facies, and vice versa. The off-diagonal term of 1.0 indicates a 100% dissimilarity of the $Shale$ from the F_Sand facies, which implies that the F_Sand never transitions into the $Shale$ facies. This is also evident in the well logs in Figures A.1 through A.15, where the F_Sand and $Shale$ are the only facies in the $H1$ layer that do not share the same geological contact.



$$t_d = \begin{bmatrix} 0.50 & 0.50 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.75 & 0.25 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.50 & 0.50 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.67 & 0.33 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \end{bmatrix} \quad t_u = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.25 & 0.75 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.50 & 0.50 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.33 & 0.67 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.25 & 0.75 \end{bmatrix}$$

$$T = \frac{1}{2}(t_d + t_u)$$

$$T = \begin{bmatrix} \mathbf{0.750} & 0.250 & 0.000 & 0.000 & 0.000 \\ 0.125 & \mathbf{0.750} & 0.125 & 0.000 & 0.000 \\ 0.000 & 0.250 & \mathbf{0.500} & 0.250 & 0.000 \\ 0.000 & 0.000 & 0.165 & \mathbf{0.670} & 0.165 \\ 0.000 & 0.000 & 0.000 & 0.125 & \mathbf{0.875} \end{bmatrix}$$

Figure 4.26: Illustration of how transition probabilities of the five Hekla facies are calculated (upward and downward) along the wellbore.

Table 4.4: Transition probability matrix of facies in Hekla reservoir $H1$ layer.

Transition Matrix					
	F_Sand	**C_Sand**	**Sh_Sand**	**S_Shale**	**Shale**
F_Sand	**0.804**	0.104	0.007	0.000	0.000
C_Sand	0.074	**0.956**	0.090	0.004	0.000
Sh_Sand	0.027	0.054	**0.801**	0.068	0.013
S_Shale	0.006	0.019	0.029	**0.787**	0.089
Shale	0.000	0.002	0.011	0.079	**0.976**

Table 4.5: Dissimilarity matrix of facies in Hekla reservoir $H1$ layer.

Dissimilarity Matrix					
	F_Sand	**C_Sand**	**Sh_Sand**	**S_Shale**	**Shale**
F_Sand	**0.000**	0.911	0.983	0.997	1.000
C_Sand	0.911	**0.000**	0.928	0.989	0.999
Sh_Sand	0.983	0.928	**0.000**	0.952	0.988
S_Shale	0.997	0.989	0.952	**0.000**	0.916
Shale	1.000	0.999	0.988	0.916	**0.000**

$$d_{i,j}(\mathbf{h}) = \begin{cases} 1 - \frac{1}{2}\left(t_{i,j} + t_{j,i}\right), & \text{if } i \neq j,\ i,j \,\epsilon\, \{1,\ldots,K\} \\ 0, & \text{otherwise} \end{cases} \tag{4.4}$$

With the information provided by the dissimilarity matrix and the visualization by MDS mapping, the best interpretation of how the facies or categories are ordered from top to bottom is F_Sand, C_Sand, Sh_Sand, S_Shale, and $Shale$, with the $Shale$ underlying the other four facies in the reservoir. However, to define the truncation rule in this study, consideration is also given to sediment fining with depth in the reservoir, that is, from coarse grain at the top, to very fine-grained sediments at the bottom. Hence, the stratification order chosen is C_Sand, F_Sand, Sh_Sand, S_Shale, and $Shale$, with the C_Sand facies placed on the top, while the $Shale$ is placed at the bottom of the

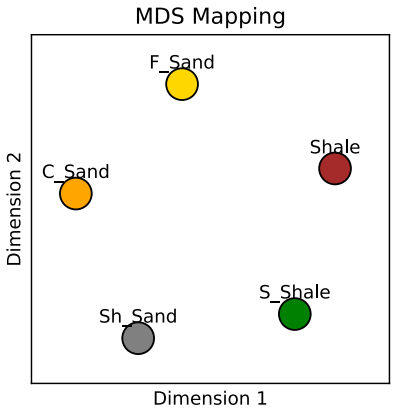


Figure 4.27: Multidimensional Scaling mapping of facies in Hekla reservoir $H1$ layer.

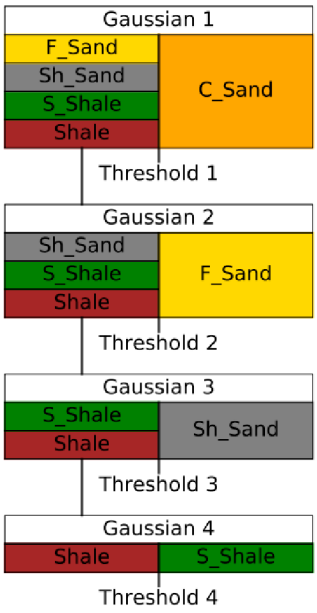


Figure 4.28: HTPG truncation tree for modeling the facies in Hekla reservoir $H1$ layer.

reservoir.

Four latent Gaussian variables and four thresholds are utilized in the truncation tree definition, with the S_Shale and $Shale$ being controlled by the last Gaussian variable, as shown in Figure 4.28. Note that the importance of applying existing knowledge of the reservoir or depositional environment to complement the truncation rule definition cannot be overemphasized. When there is a good understanding of the reservoir architecture, as in the case of redeveloping and producing a brownfield, such understanding should be applied as the dominant tool or measure to define the truncation rule, while being complemented by other quantitative measures. However, if the case study is a greenfield and there is no offset data, quantitative tools such as the transition probability

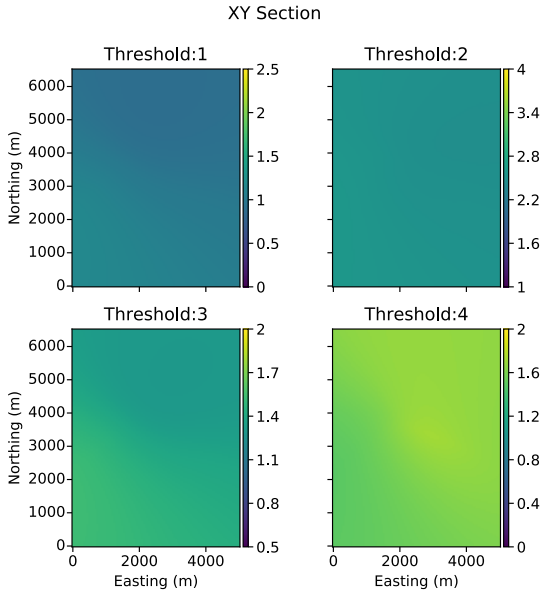


Figure 4.29: Local thresholds for generating realization of facies in the Hekla reservoir $H1$ layer, in the $x - y$ orientation

matrix and MDS, and the modeler's experience will guide the definition of the truncation rule. Generally, the HTPG is a flexible technique for modeling categorical variables because of the flexibility in defining the truncation rule.

4.3.5 Calculation of Thresholds

The $htpg\_thresholds$ software (Silva & Deutsch, 2018) is used to calculate the local thresholds from the local varing proportions of the facies and the defined truncation rule. Figure 4.33 is the defined HTPG truncation structure, while Figures 4.29, 4.30, and 4.31 show the thresholds for mapping the $H1$ layer facies from the latent Gaussian variables. Note that the number of thresholds is often equal to the number of categorical variables minus one, $(K-1)$. The thresholds are spread along the latent variables. So, when defining the truncation tree, the latent Gaussian variable to which the threshold is applied is often specified. Although there are certain cases where more than one threshold is assigned to a latent Gaussian variable, in the Hekla reservoir case study, there is only one threshold to a latent variable. Also, the number of categories or facies is assigned to each threshold; in this study, only one facies is specified per threshold, as defined by the truncation tree in Figure 4.28.

4.3.6 Mapping of Spatial Structure and Imputation of Latent Variables

The $htpg\_gaussvarg$ software (Silva & Deutsch, 2018) is utilized to define four latent Gaussian variable variograms. Spatial continuity is mapped from categorical to continuous space. The experimental points which result from the mapping between the categorical and continuous space are

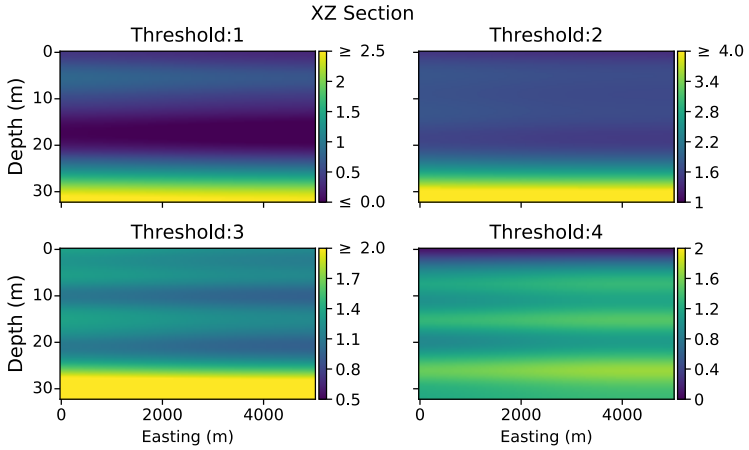


Figure 4.30: Local thresholds for generating realization of facies in the Hekla reservoir $H1$ layer, in the $x - z$ orientation

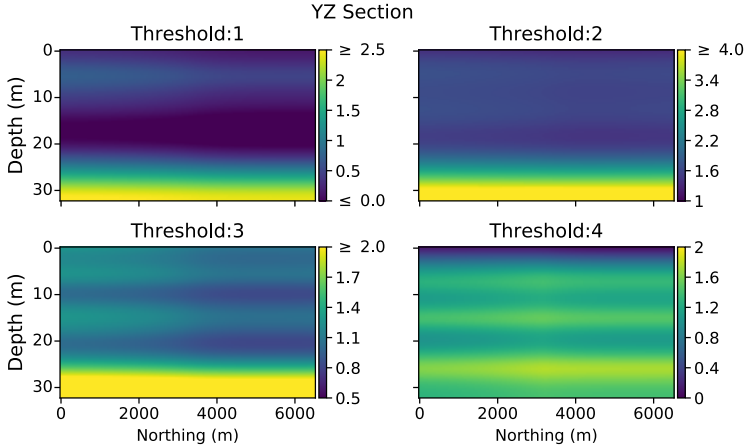


Figure 4.31: Local thresholds for generating realization of facies in the Hekla reservoir $H1$ layer, in the $y - z$ orientation

fitted with the Gaussian variogram model with the $varmodel$ software (C. V. Deutsch & Journel, 1998), as shown in Figure 4.32. The variograms are modeled in two directions - horizontal (omnidirectional) and vertical.

The HTPG algorithm is developed to truncate the realizations of latent Gaussian variables at some thresholds to generate multiple realizations of categorical variables. However, in practice, the latent variables do not exist, but are needed as conditioning data at the sampled categorical data locations. With the calculated local thresholds and defined variograms of the latent variables, the $htpg\_cat2gauss$ software (Silva & Deutsch, 2018) is applied to perform multiple imputation of the latent variables at the sampled data locations. In this case study, multiple ($L = 100$) realizations of the latent Gaussian variable data are generated and used as input data for simulation.

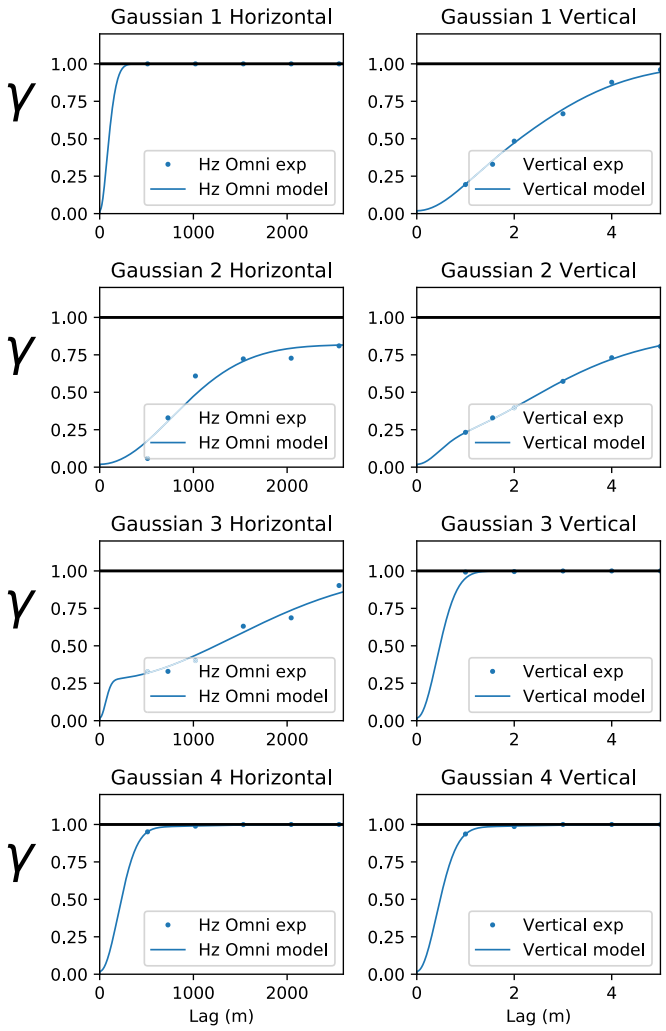


Figure 4.32: Variograms of latent Gaussian variables

4.3.7 Mapping from Continuous Variable to Categorical Variable Space

Gaussian simulation of the latent variable is run with a simulation grid of 101 x 131 x 65 nodes in the x, y, and z directions. The node spacing in the horizontal and vertical directions are $50m$ and $0.5m$, respectively. The truncation system or structure in Figure 4.33 shows that at certain threshold, t_1, the categorical variable 2 or C_Sand facies is mapped from the realizations of latent Gaussian variable Y_1, while at threshold t_4, the S_Shale and $Shale$ facies are mapped from the realizations of latent variable Y_4. The $htpg\_gauss2cat$ software (Silva & Deutsch, 2018) performs this mapping from continuous to categorical variable space by applying the defined truncation rule. Multiple ($L = 100$) realizations of the five facies in the Hekla reservoir $H1$ layer are generated by this truncation process. Four simulated facies realizations in the x, y, and z slices are shown in Figures 4.34, 4.35, and 4.36. The realizations depict how the C_Sand (orange) and $Shale$ (red) facies are widely distributed in the $H1$ layer of the reservoir, which is evident in the high proportions of C_Sand and $Shale$ in Table

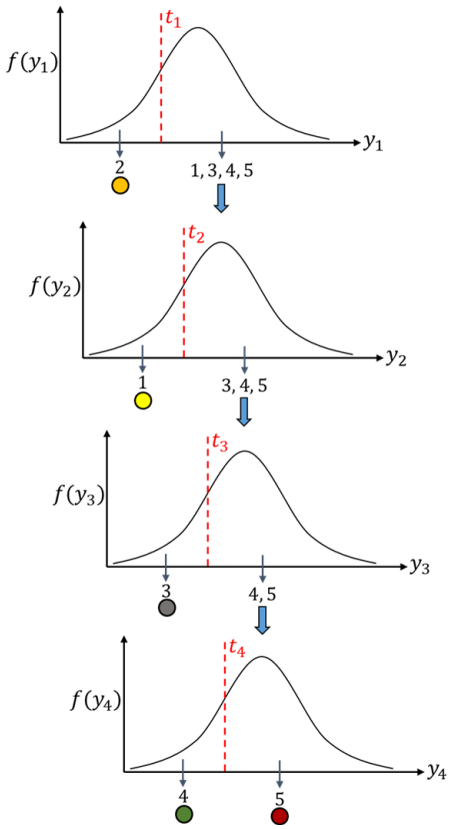


Figure 4.33: HTPG truncation system for generating realization of facies in the Hekla reservoir $H1$ layer

4.2, and the well logs in Figures A.1 through A.15.

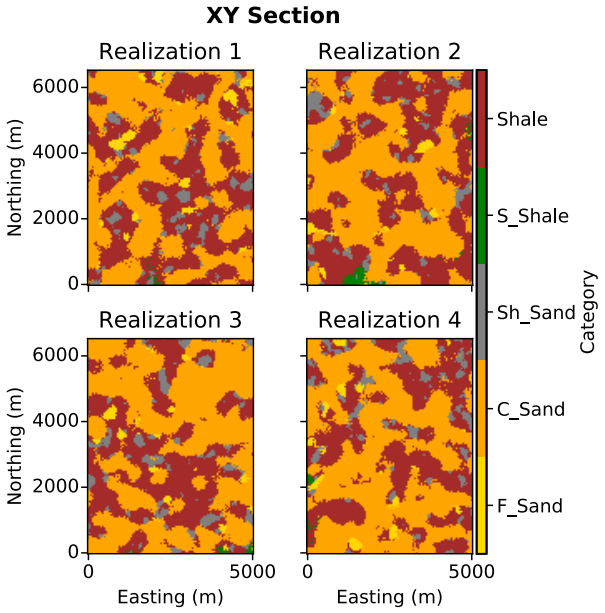


Figure 4.34: Simulated facies realizations in the Hekla reservoir $H1$ layer, in the $x - y$ orientation

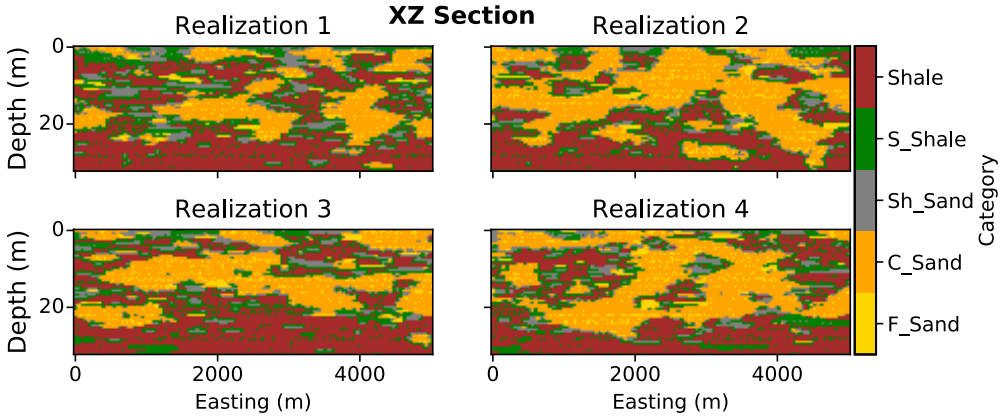


Figure 4.35: Simulated facies realizations in the Hekla reservoir $H1$ layer, in the $x - z$ orientation

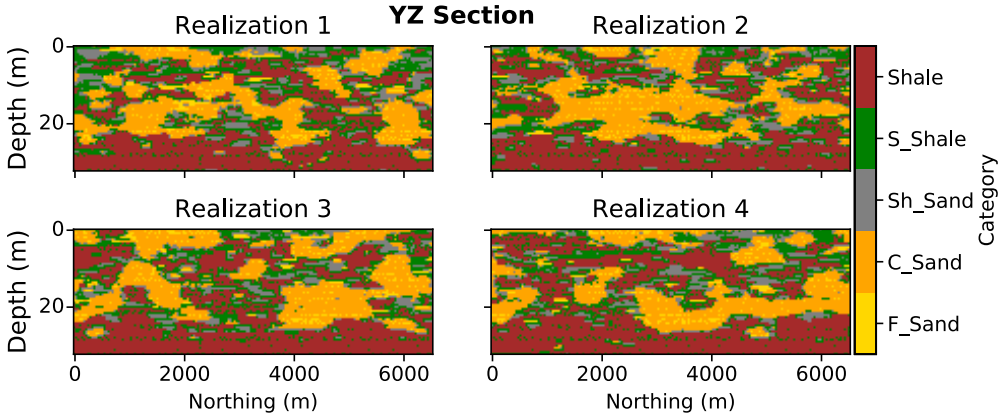


Figure 4.36: Simulated facies realizations in the Hekla reservoir $H1$ layer, in the $y - z$ orientation

4.3.8 Facies Model Validation

All realizations generated from simulations are considered representative of the reservoir when they reproduce the original statistics and input parameters that go into the simulation process. To assess how the constructed facies model will perform, the input univariate statistics and parameters such as facies proportion and indicator residual variogram are checked, as shown in Figures 4.37, and 4.38. As highlighted in Table 4.6, the absolute error, ϵ_{abs}, of the facies realizations reproducing the proportions of F_Sand, C_Sand, and Sh_Sand facies are 8.7% , 7.8%, and 3.2% respectively, while the $Shale$ proportion is reproduced with an error of 2.0%, as calculated from Equation 4.5.

$$\epsilon_{abs} = \frac{|Declustered\ prop. - Avg.\ realization\ proportion|}{Declustered\ prop.} \tag{4.5}$$

The variograms of facies indicator residuals are also fairly reproduced. Some factors affect a better variogram reproduction including the position of the facies or category in the truncation tree and amount of sampled data. By default, the category that overlies other categories in the truncation tree will have the best variogram reproduction, while the category that is placed at the base of the tree will have the least variogram reproduction, putting other factors into consideration as well. The residual variogram of C_Sand facies is well reproduced because C_Sand is topmost in the truncation tree and it has sufficient data. Although the $Shale$ facies underlies other facies in the truncation tree, its residual variogram is also well reproduced due to sufficient data or its high

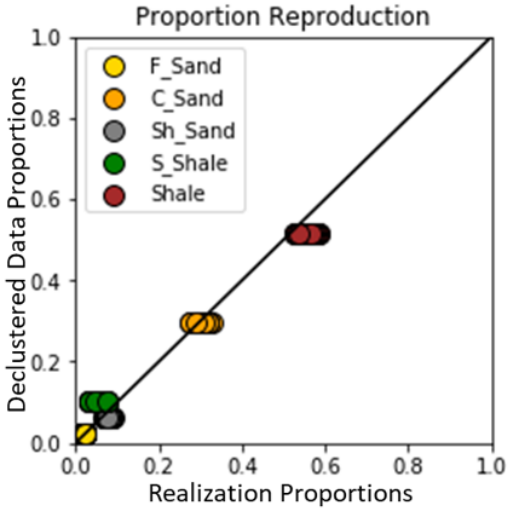


Figure 4.37: Facies proportion reproduction in the Hekla reservoir $H1$ layer

Table 4.6: Facies proportion reproduction in the Hekla Reservoir $H1$ Layer

Facies	**F_Sand**	**C_Sand**	**Sh_Sand**	**S_Shale**	**Shale**
Declustered Proportion	0.023	0.282	0.062	0.095	0.538
Avg. Realization Proportion	0.021	0.304	0.064	0.083	0.527
Error, ϵ_{abs}	**0.087**	**0.078**	**0.032**	**0.126**	**0.020**

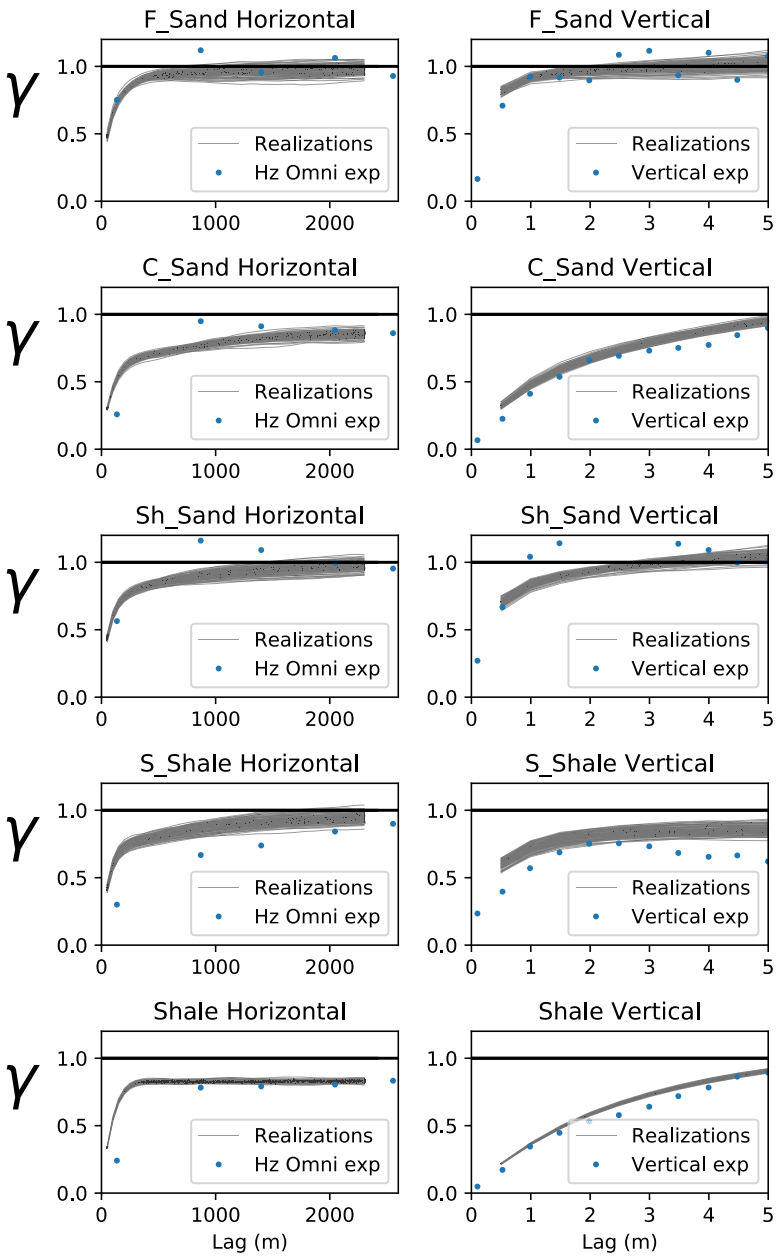


Figure 4.38: Facies indicator variogram reproduction in the Hekla reservoir $H1$ layer

proportion in the $H1$ layer of the reservoir. Note that the horizontal variograms are not fairly reproduced due to fewer number of wells and insufficient data available to model the spatial continuity of the reservoir facies laterally. In general, the fairly reproduced residual variograms and the minimal error in reproducing the facies proportions mean that the facies model will perform well and can be considered as input to develop the model of continuous petrophysical properties in the $H1$ layer of the Hekla reservoir.

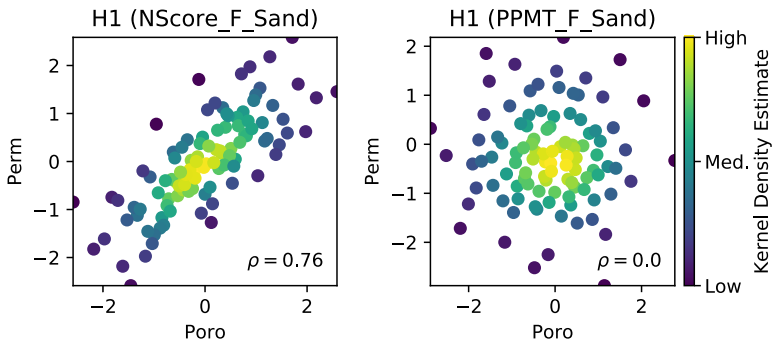


Figure 4.39: Normal score vs. PPMT transformed porosity - permeability bivariate relationship of F_Sand facies in the Hekla reservoir $H1$ layer

4.4 PPMT Application

Having developed the model of facies distribution, the continuous petrophysical properties in the Hekla $H1$ layer are modeled with the PPMT modeling workflow. The continuous properties (porosity and permeability) are controlled by the facies distribution in the reservoir; so the simulated facies realizations are considered as input to simulate the continuous properties. The steps taken to model the Hekla reservoir continuous properties with the PPMT workflow are discussed in this section.

4.4.1 Decorrelation of Variables

To model the continuous petrophysical properties ($Porosity$ and $Permeability$) in the Hekla reservoir $H1$ layer, the $H1$ data was extracted from the entire Hekla dataset and was grouped into five categories or facies data. Each facies data was declustered to have representative distributions of the continuous variables. The PPMT technique was applied to the declustered data to transform the variables to Gaussian and independent variables, as shown in Figures 4.39 through 4.43. As observed in the figures, the kernel density estimate of the PPMT-transformed data approximates the typical multi-Gaussian density contours and the correlation coefficients between the transformed variables are zero, indicating that the variables in each category are uncorrelated and can be simulated independently.

4.4.2 Variograms of Continuous Variables

The variogram is a key input parameter in all geostatistical estimation and simulation exercises. In this study, the variograms of the continuous variables are calculated and modeled by each facies present in the Hekla reservoir. The PPMT algorithm outputs the normal score (NScore) data, which are used to facilitate variogram calculation for the porosity and permeability of each facies in the reservoir. Figures 4.44 through 4.53 show the variogram models of Porosity and Permeability in all five facies in the $H1$ layer of the reservoir. The vertical variograms of C_Sand and $Shale$ are more

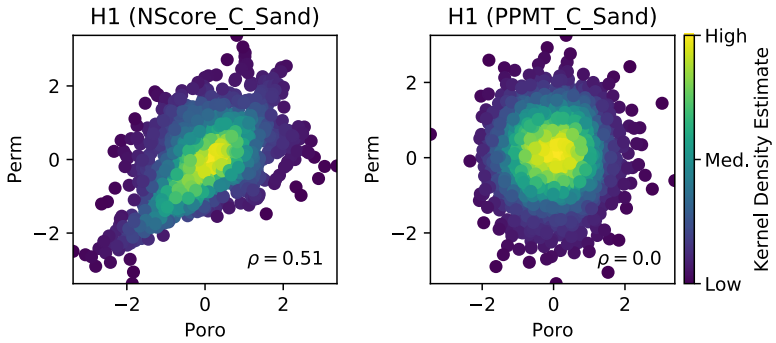


Figure 4.40: Normal score vs. PPMT transformed porosity - permeability bivariate relationship of C_Sand facies in the Hekla reservoir $H1$ layer

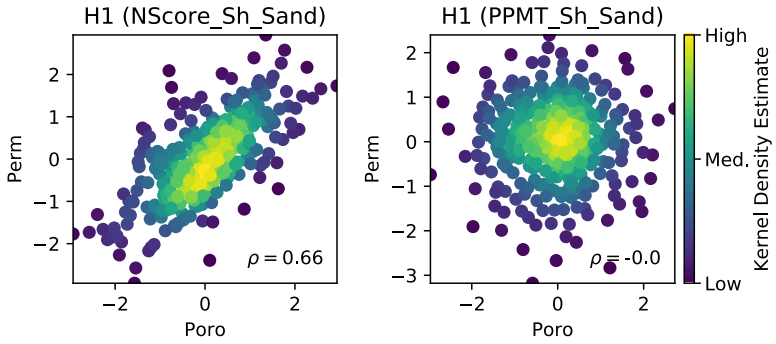


Figure 4.41: Normal score vs. PPMT transformed porosity - permeability bivariate relationship of Sh_Sand facies in the Hekla reservoir $H1$ layer

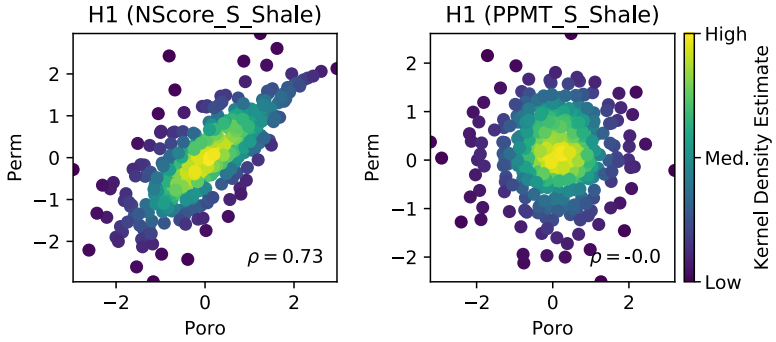


Figure 4.42: Normal score vs. PPMT transformed porosity - permeability bivariate relationship of S_Shale facies in the Hekla reservoir $H1$ layer

continuous than those of other facies due to the high proportion of C_Sand and $Shale$ in the reservoir and consequently, availability of more data location for modeling. The vertical variograms of F_Sand, Sh_Sand, and S_Shale do not look great because of the limited thickness of these facies. Generally, in all the facies, it was difficult to have a reasonable horizontal variogram model since there are few wells in the reservoir.

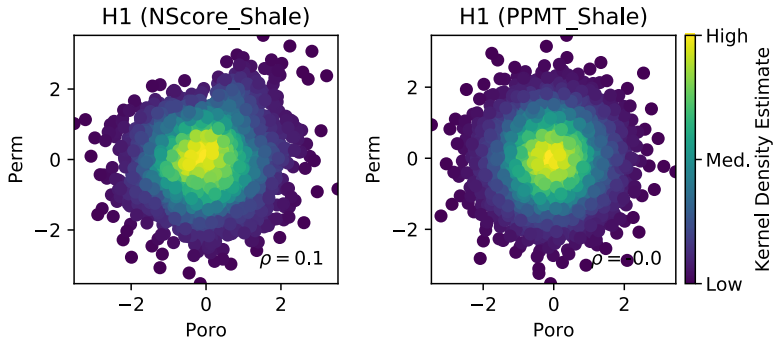


Figure 4.43: Normal score vs. PPMT porosity - permeability bivariate relationship of $Shale$ facies in the Hekla reservoir $H1$ layer

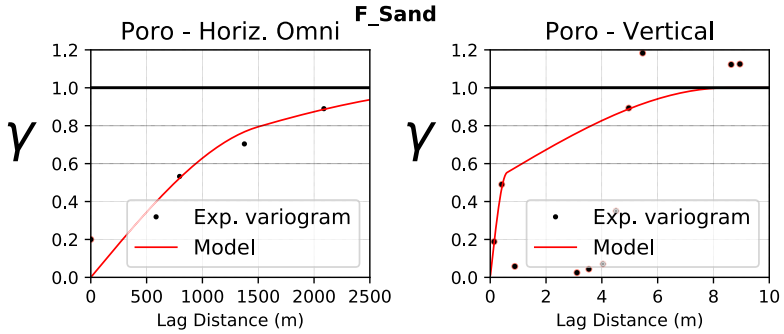


Figure 4.44: Variogram of Porosity in F_Sand facies in the Hekla reservoir $H1$ layer

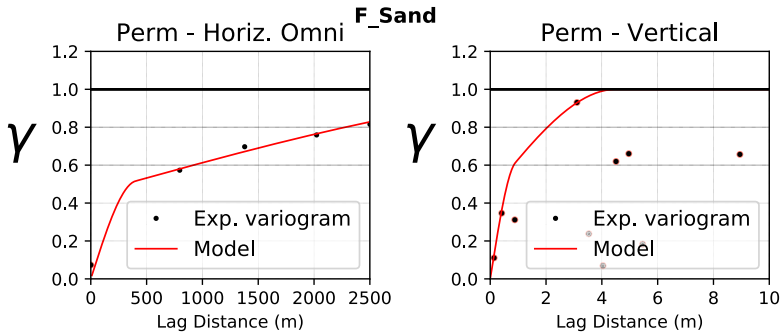


Figure 4.45: Variogram of Permeability in F_Sand facies in the Hekla reservoir $H1$ layer

Since it is difficult to model the variograms of F_Sand, Sh_Sand, and S_Shale facies due to their low count and proportions, and fewer data location, it is reasonable to merge all the sand data and model the variogram of the combined data; a similar approach is used to model the variogram of the merged shale data. Figures 4.54 through 4.57 show the variogram models of the combined sand data (F_Sand, C_Sand, and Sh_Sand) and the combined shale data (S_Shale, and $Shale$). In both cases, the variograms of the merged data show good spatial continuity.

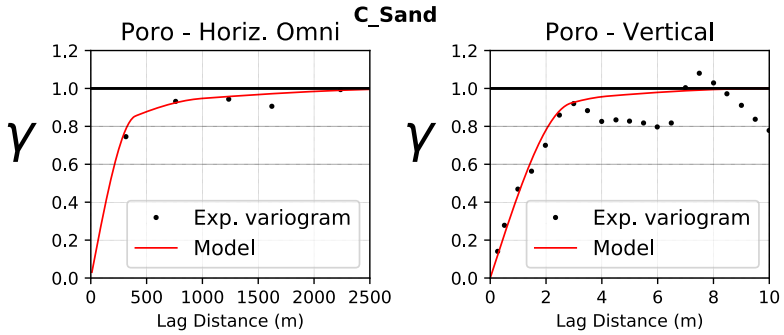


Figure 4.46: Variogram of Porosity in C_Sand facies in the Hekla reservoir $H1$ layer

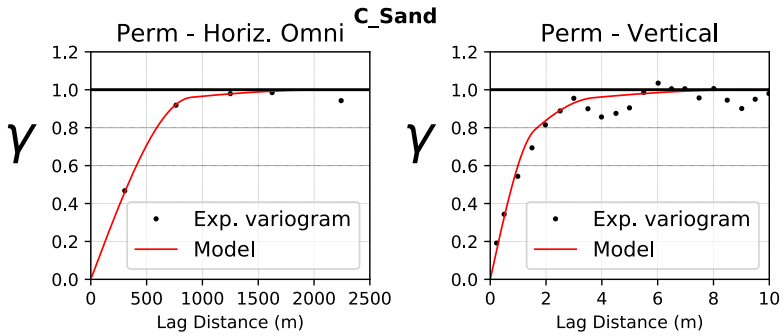


Figure 4.47: Variogram of Permeability in C_Sand facies in the Hekla reservoir $H1$ layer

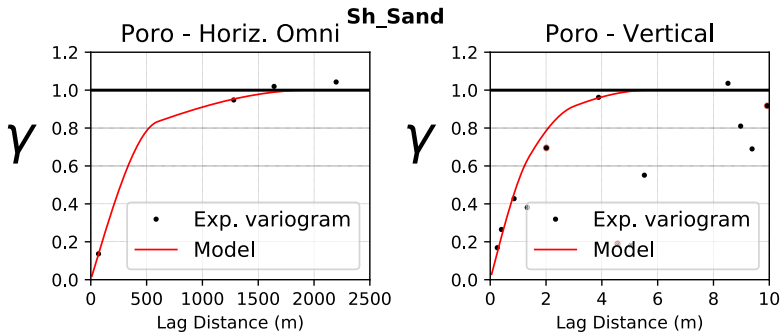


Figure 4.48: Variogram of Porosity in Sh_Sand facies in the Hekla reservoir $H1$ layer

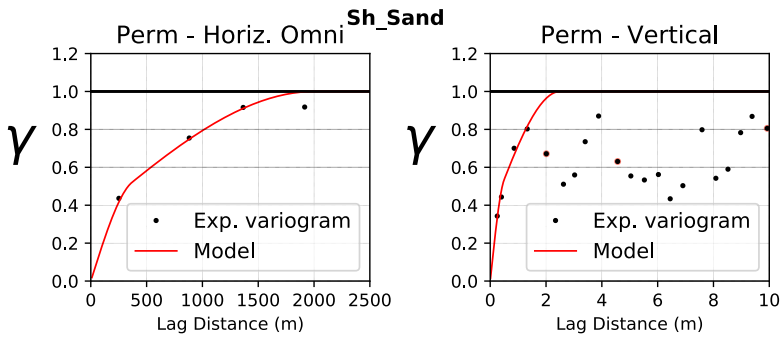


Figure 4.49: Variogram of Permeability in Sh_Sand facies in the Hekla reservoir $H1$ layer

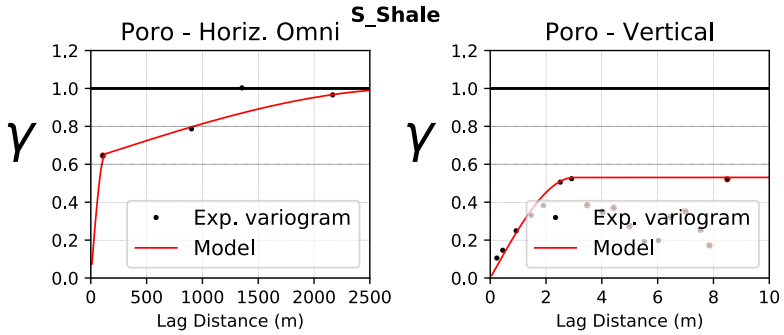


Figure 4.50: Variogram of Porosity in S_Shale facies in the Hekla reservoir $H1$ layer

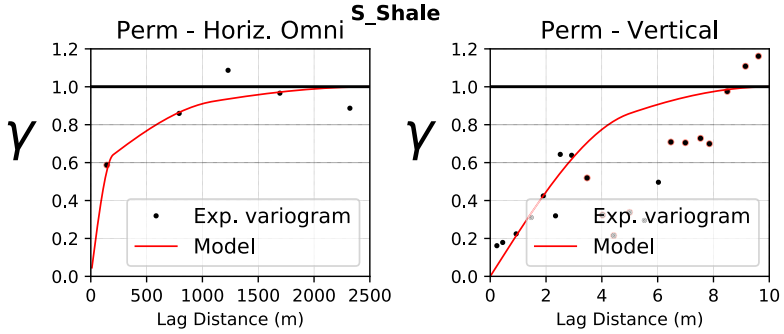


Figure 4.51: Variogram of Permeability in S_Shale facies in the Hekla reservoir $H1$ layer

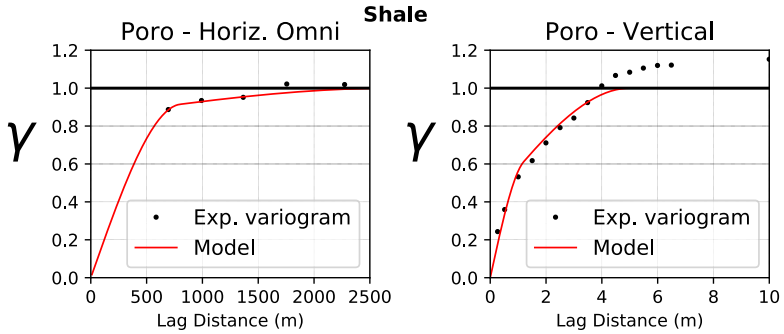


Figure 4.52: Variogram of Porosity in $Shale$ facies in the Hekla reservoir $H1$ layer

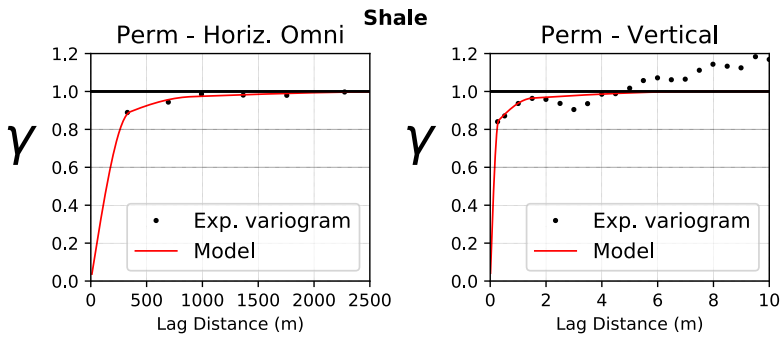


Figure 4.53: Variogram of Permeability in $Shale$ facies in the Hekla reservoir $H1$ layer

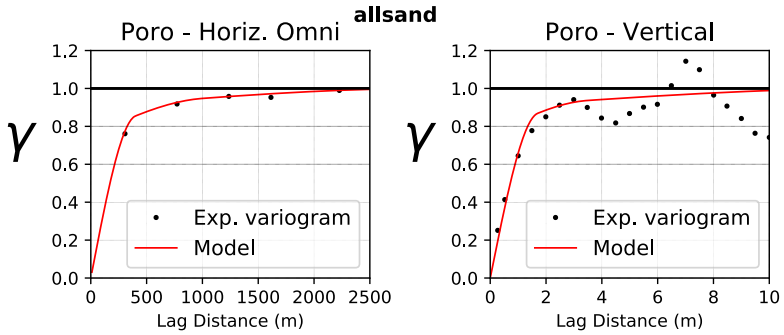


Figure 4.54: Variogram of Porosity of all the Sand in the Hekla reservoir $H1$ layer

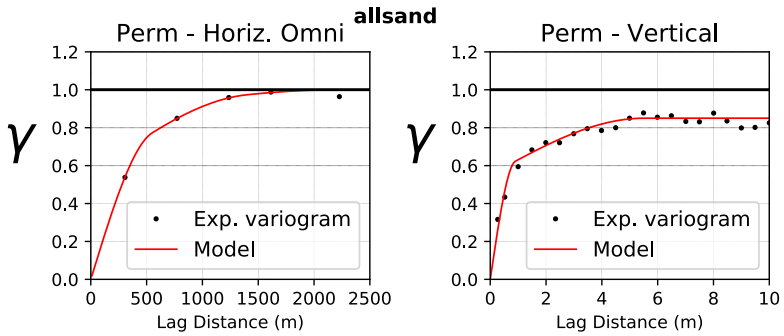


Figure 4.55: Variogram of Permeability of all the Sand in the Hekla reservoir $H1$ layer

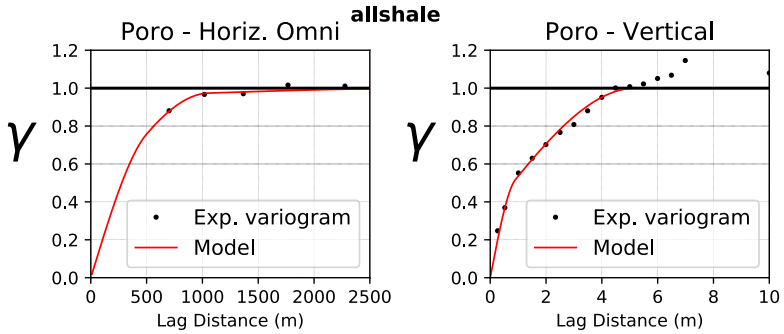


Figure 4.56: Variogram of Porosity of all the Shale in the Hekla reservoir $H1$ layer

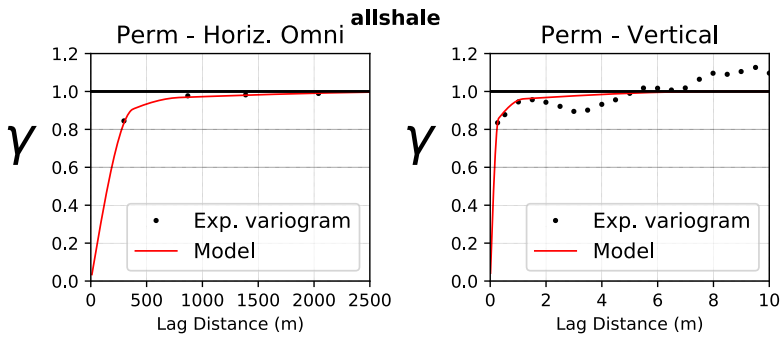


Figure 4.57: Variogram of Permeability of all the Shale in the Hekla reservoir $H1$ layer

4.4.3 Gaussian Simulation

By using the normal score variograms and PPMT transformed data as input parameter and conditioning data, independent Gaussian simulation is run to generate multiple ($L = 100$) realizations of the continuous petrophysical properties in $H1$ layer of the reservoir. The 100 realizations of facies or category variables initially generated are considered as rock type data in the simulation process.

The simulation grid contains 101 x 131 x 65 nodes in the x, y, and z directions. The node spacing in the horizontal and vertical directions are $50m$ and $0.5m$, respectively. The associated PPMT back-transformation is done to restore the original units and dependency of the Porosity and Permeability variables. Four of the 100 realizations of porosity and permeability are shown in Figures 4.58 through 4.63, in different reservoir orientations. As shown in the realizations, the C_Sand (orange) and $Shale$ (brown) facies are widely distributed in the domain of the reservoir; this is expected due to the high count and proportions of C_Sand and $Shale$ facies. The realizations show that the regions of the reservoir with high porosity and permeability distributions are in the C_Sand facies, while the regions with very low to zero porosity and permeability distributions are in the $Shale$ facies. Also, the realizations show that the F_Sand (gold) and Sh_Sand (grey) are poorly distributed in the reservoir but have intermediate porosity and permeability values.

4.4.4 Model Validation

To assess the models developed by the PPMT workflow, the univariate statistics, multivariate relationships, and input parameters such as histogram and variogram, are checked. When performing resource and reserves calculations, and during field development, the sand lithology or regions

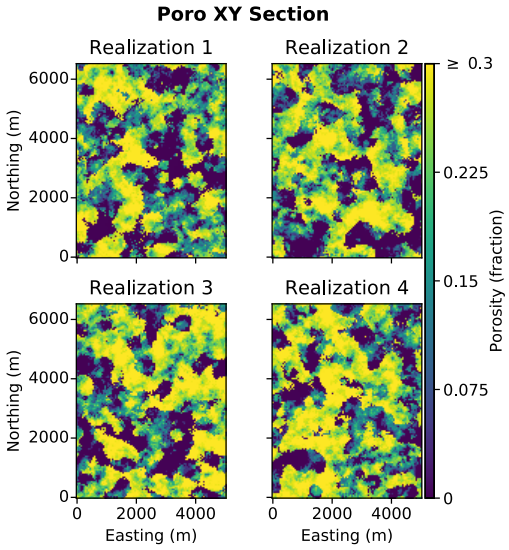


Figure 4.58: Porosity realizations in Hekla reservoir $H1$ layer, $x - y$ orientation

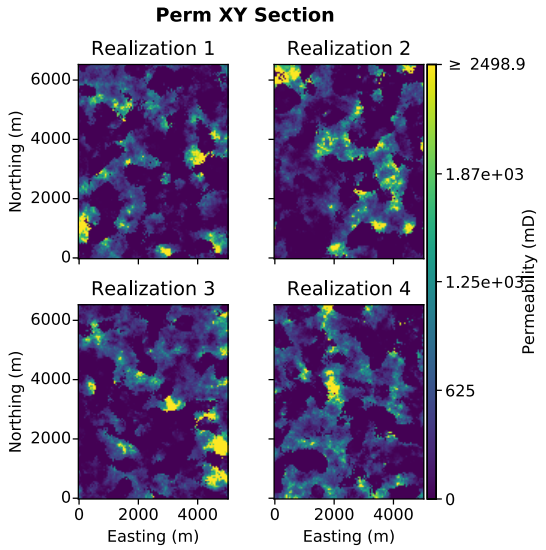


Figure 4.59: Permeability realizations in Hekla reservoir $H1$ layer, $x - y$ orientation

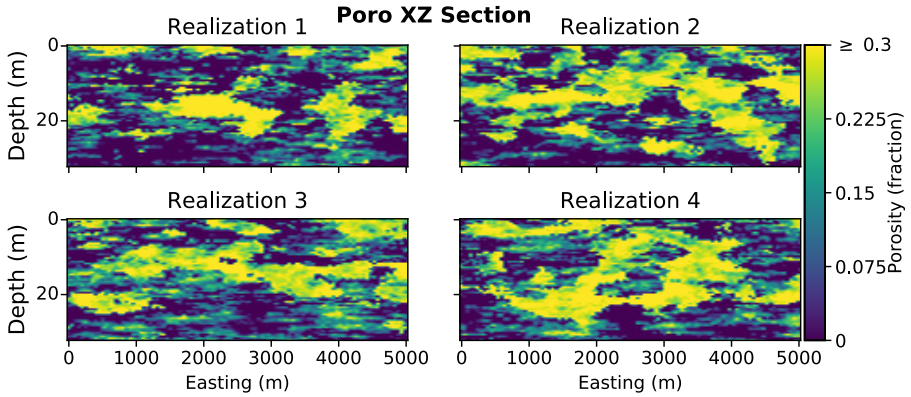


Figure 4.60: Porosity realizations in Hekla reservoir $H1$ layer, $x - z$ orientation

of the reservoir with good sand distribution are often given attention. So, to check the validity of the continuous petrophysical property models, consideration is given to the C_Sand facies and the entire sand in the reservoir. Also, porosity and permeability are continuous variables that take a range of values and cannot be coded as indicators. So, unlike the facies model that is assessed by the error in reproducing the global proportions of the categories, the performance of the porosity and permeability models is checked with the error in reproducing the global mean and standard deviation of the original data.

Figures 4.64, 4.65, and 4.66 show how the simulated realizations of porosity and permeability reproduce the univariate statistics of the original C_Sand, all $H1$ sand, and entire $H1$ data. The error in reproducing the statistics of the reference data is calculated by Equations 4.6 and 4.7. The reproduced univariate statistics in the figures show that the mean of the porosity and permeability data in

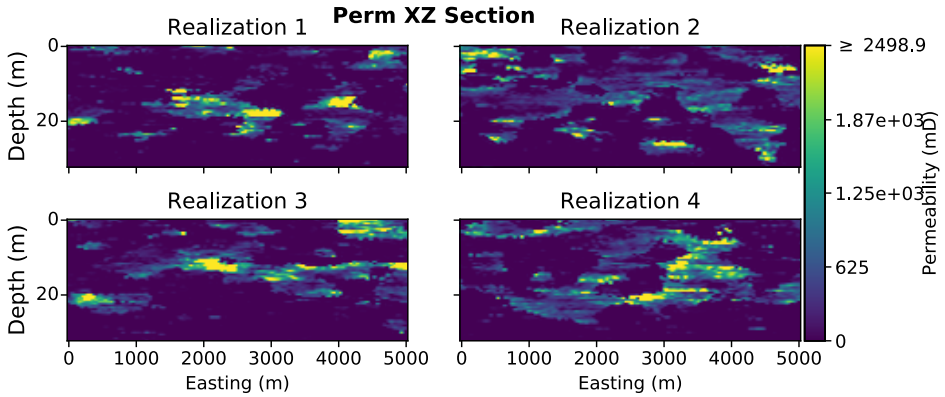


Figure 4.61: Permeability realizations in Hekla reservoir $H1$ layer, $x - z$ orientation

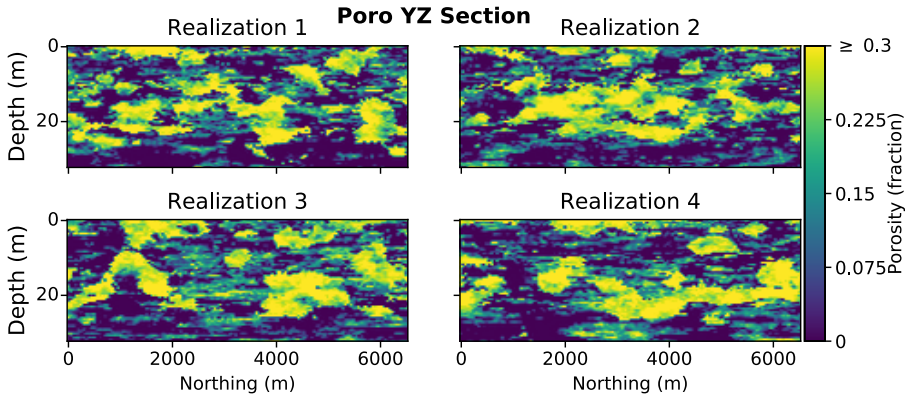


Figure 4.62: Porosity realizations in Hekla reservoir $H1$ layer, $y - z$ orientation

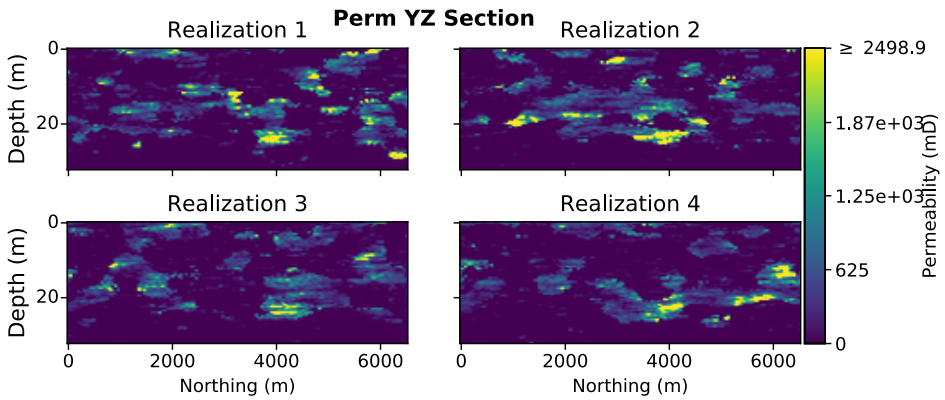


Figure 4.63: Permeability realizations in Hekla reservoir $H1$ layer, $y - z$ orientation

the C_Sand facies are reproduced with absolute error m_{error} of 0.4% and 4.7%, respectively, while the continuous properties in the C_Sand facies are reproduced with an average error of 2.5%. Also, the continuous properties in all the sand facies (F_Sand, C_Sand, and Sh_Sand) and the entire $H1$ layer are reproduced with an average error of approximately 4%, while the average errors in reproducing the standard deviation σ_{error} are 6% and 0.8%, respectively, see Table 4.7. These minimal errors imply that the model of continuous petrophysical properties will perform relatively well. Al-

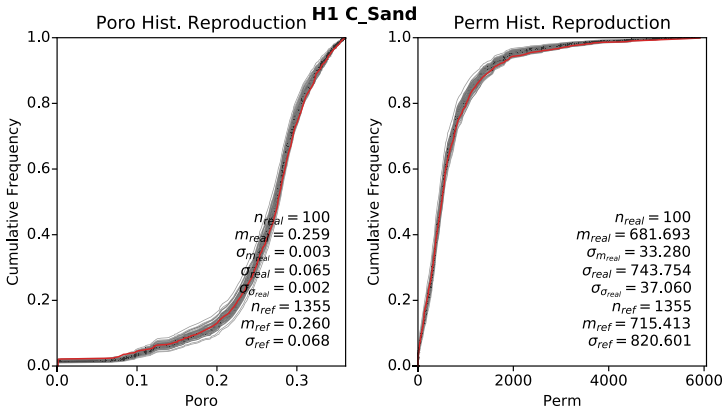


Figure 4.64: Histogram reproduction of C_Sand facies data in Hekla reservoir $H1$ layer

though the mean and standard deviation of the realizations closely match those of the original data, the slight shift of the realizations from original data could be the result of the input parameters such as variogram model. Due to the limited number of wells, it was difficult to model the variograms of porosity and permeability in facies, especially those of F_Sand, Sh_Sand, and S_Shale.

$$m_{error} = \frac{\left| m_{ref} - m_{real} \right|}{m_{ref}} \tag{4.6}$$

$$\sigma_{error} = \frac{\left| \sigma_{ref} - \sigma_{real} \right|}{\sigma_{ref}} \tag{4.7}$$

Where:

m_{error} = error in reproducing the mean of the reference data,

m_{ref} = mean of the reference data,

m_{real} = mean of the average realization,

σ_{error} = error in reproducing the standard deviation of the reference data,

σ_{ref} = standard deviation of the reference data,

σ_{real} = standard deviation of the average realization.

The variograms of continuous properties in the sand facies are also fairly reproduced, as shown in Figure 4.67, while Figures 4.68, 4.69, and 4.70 show the reproduction of the bivariate relationship between porosity and permeability in the data. It is evident that the shape of the bivariate scatter plots and the correlation coefficients between the variables are reproduced well.

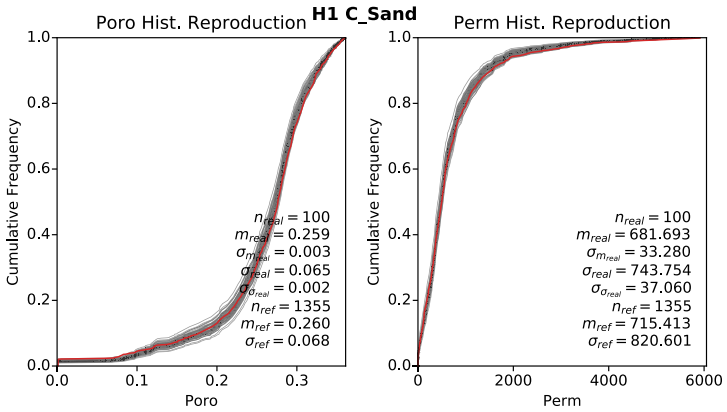


Figure 4.65: Histogram reproduction of all sand data in Hekla reservoir $H1$ layer

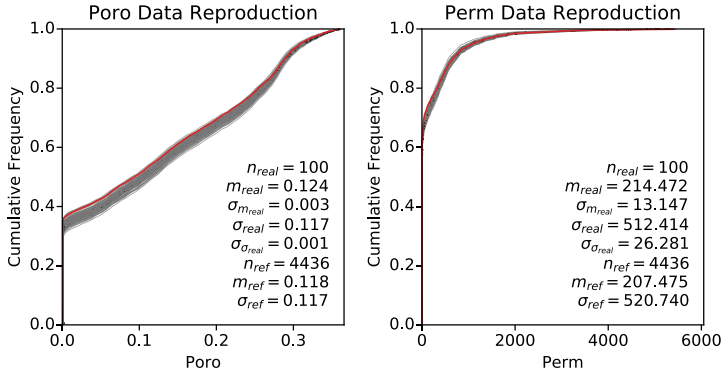


Figure 4.66: Histogram reproduction of data in the entire Hekla reservoir $H1$ layer

Table 4.7: PPMT workflow error assessment for the Hekla Reservoir $H1$ Layer

Facies	**C_Sand**		**All_Sand**		**Entire Layer**	
Data Reproduction Error	m_{error}	σ_{error}	m_{error}	σ_{error}	m_{error}	σ_{error}
Porosity	0.004	0.044	0.013	0.024	0.051	0.000
Permeability	0.047	0.094	0.072	0.096	0.034	0.016
Average	**0.025**	**0.069**	**0.043**	**0.060**	**0.042**	**0.008**

4.5 Discussion

The application of HTPG simulation technique to the Hekla reservoir yields a facies model that is consistent with original data. HTPG has the capability of generating facies or categorical variable model for any number of categories and reservoir complexity. As evident in the final facies model and validation results, the developed facies model of the $H1$ layer of the Hekla reservoir will perform well with minimal error. Although parameterization in the HTPG technique is a bit challenging, the technique seems to generate facies model that describes the facies architecture of the reservoir. Considering the facies realizations as rock type in the PPMT modeling workflow also generates porosity and permeability models that are consistent with the facies model. The

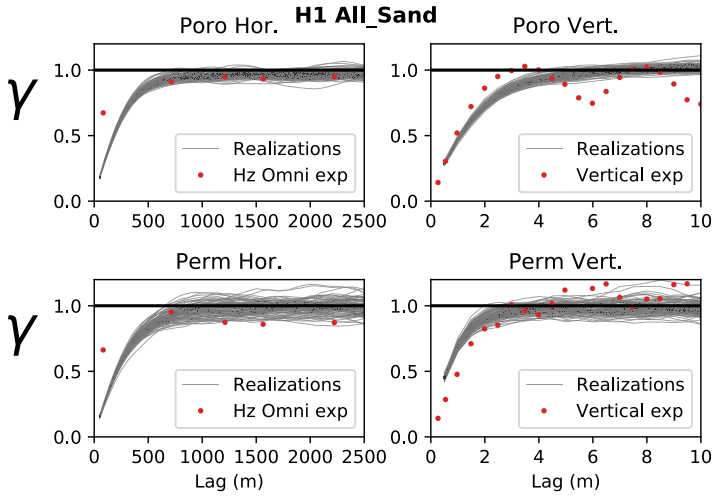


Figure 4.67: Variogram reproduction of all sand data in Hekla reservoir $H1$ layer

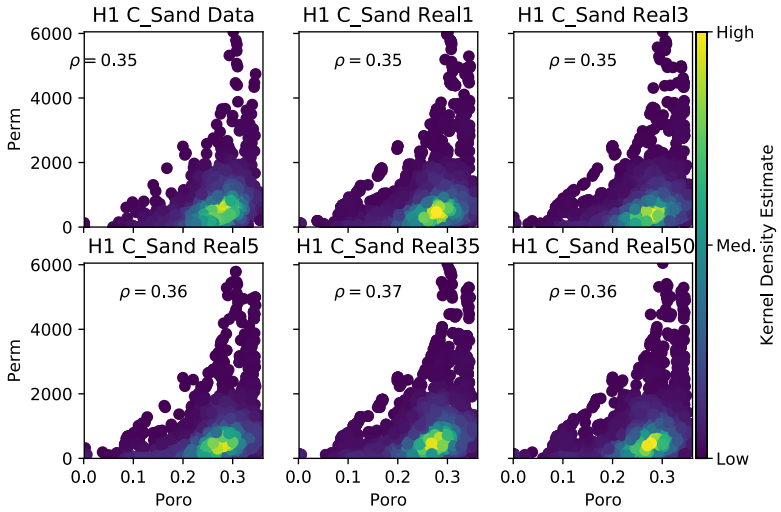


Figure 4.68: Porosity - permeability bivariate relationship reproduction in C_Sand facies in Hekla reservoir $H1$ layer

HTPG-PPMT modeling workflow, as adopted in this research, emphasizes the applicability of both techniques for modeling categorical and continuous variables in a petroleum reservoir or geologic deposit.

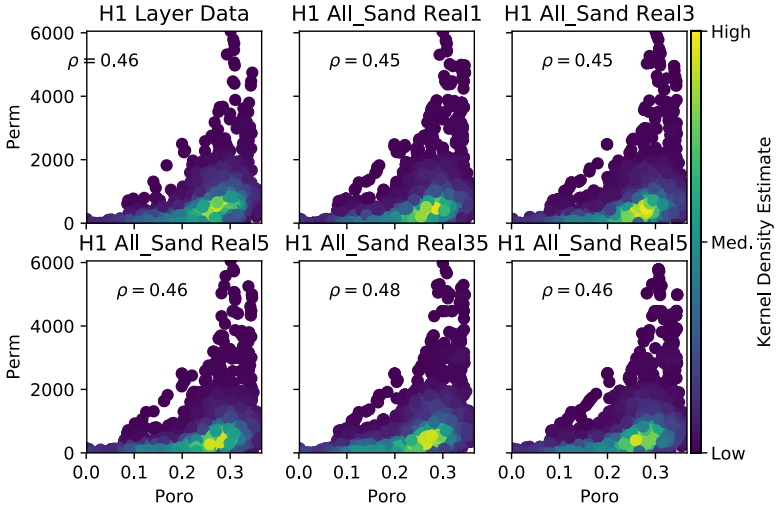


Figure 4.69: Porosity - permeability bivariate relationship reproduction in $allSand$ of Hekla reservoir $H1$ layer

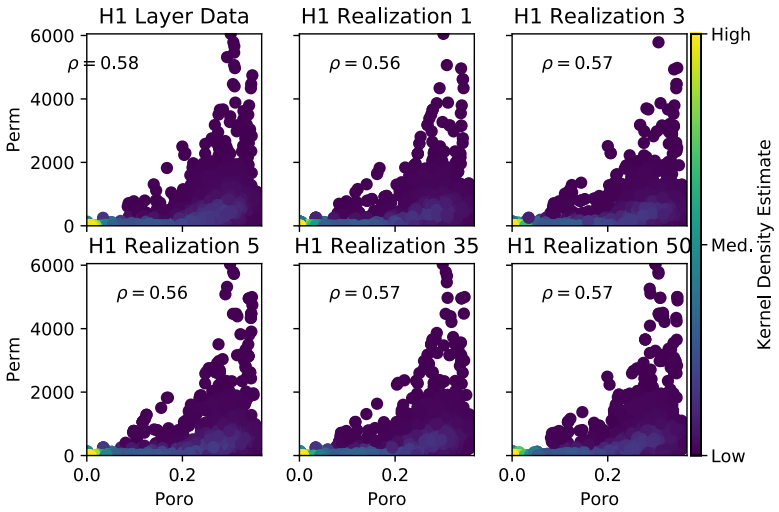


Figure 4.70: Porosity - permeability bivariate relationship reproduction in the entire Hekla reservoir $H1$ layer

Chapter 5

Conclusions and Future Work

Although the workflow adopted in this research is applicable to projects in mining and petroleum industries, this thesis focuses on categorical and continuous property modeling in the petroleum context. These are key steps in the process of developing a reliable a reservoir model that is utilized for reservoir performance prediction and field development planing purposes, and also used to quantify and manage the inherent uncertainty in the reservoir.

5.1 Covered Topics and Contribution

The HTPG and PPMT techniques are established geostatistical methods for modeling categorical and continuous variables, respectively. Chapter 4 discusses how the HTPG is implemented to model the facies in the upper ($H1$) layer of the Hekla reservoir. The definition of the HTPG truncation rule is one of the key steps in the HTPG workflow, as this determines how the categories or facies are mapped from the latent Gaussian variables. This makes the HTPG a flexible technique and suitable for modeling reservoirs or mineral deposits in complex geological settings.

The developed facies realizations of the Hekla reservoir are considered as rock types to develop models of continuous properties (porosity and permeability). The continuous petrophysical properties are controlled by the distribution of facies in the reservoir. To model the porosity and permeability in the Hekla reservoir $H1$ layer, the PPMT modeling workflow is used. PPMT transforms the continuous variables to uncorrelated variables and thereby facilitates independent simulation of the variables. Multiple ($L = 100$) realizations of porosity and permeability are generated for the Hekla reservoir. The porosity and permeability realizations are consistent with the facies realizations, as high porosity and permeability distributions are observed in the regions of C_Sand facies, while low porosity and permeability are distributed in regions of $Shale$ facies. Such facies and continuous petrophysical property models can be utilized to study reservoir heterogeneity, which helps to assess and manage the uncertainty in the reservoir. These models are also used in field development and redevelopment to discern the productive regions of the reservoir; this will facilitate infill and outstep well placement in the reservoir.

The main contribution to this thesis is the integration of the application of HTPG and PPMT techniques into a single geostatistical reservoir modeling workflow. To achieve this, codes are written in Python in Jupyter Notebook, with appropriate use of some Pygeostat functions to run the HTPG

and PPMT software. Note that the models developed in this study are specific to the Hekla reservoir $H1$ layer. The adopted workflow is reproducible and applicable in mining and petroleum to model categorical and continuous variables.

5.2 Limitations and Future Work

One of the challenges encountered in this study is the limitation of the Hekla reservoir dataset. First, the limited number of wells (only 15 wells available) makes it challenging to have a good model of spatial variability or continuity in the lateral direction of the reservoir. This increases the uncertainty in discerning the possible lateral continuity of each facies in the final model. This can be overcome by utilizing a dataset or case study with more wells, as this would help improve the horizontal variogram of each facies indicator residual. Second, the inconsistency in acoustic impedance data. As discussed in Chapter 4, after calibrating the seismic data, the weak correlation of the acoustic impedance values to facies proportions observed in both sand and shale makes the impedance data not helpful in this study. Seismic data are exhaustive and often provide good information about the reservoir lithology and architecture. This means a reservoir model should consider seismic data. A case study or dataset with reliable seismic attributes could be considered in the future.

Note that the models of facies and continuous petrophysical properties developed in this study are static models. Therefore, the scope of this research does not cover the process of upscaling such static models to a flow or dynamic model that is utilized to predict future performance of the reservoir under different production or operating conditions.

References

Barnett, R. M. (2017a). Principal component analysis. *In J. L. Deutsch (Ed.), Geostatistics Lessons*.

Barnett, R. M. (2017b). Sphereing and min/max autocorrelation factors. *In J. L. Deutsch (Ed.), Geostatistics Lessons*. Retrieved from `http://geostatisticslessons.com/lessons/sphereingmaf.html`

Barnett, R. M., & Deutsch, C. V. (2015). Guide to multivariate modeling with the ppmt. *Centre for Computational Geostatistics (CCG) Guidebook Series*, *20*.

Barnett, R. M., Manchuk, J. G., & Deutsch, C. V. (2016). The projection pursuit multivariate transform for continuous variable modeling. *Society of Petroleum Engineers Journal of Petroleum Technology, Paper 184388*.

Caers, J. (2011). *Modeling uncertainty in the earth sciences* (1st Edition ed.). Wiley-Blackwell: A John Wiley & Sons Ltd. Publication.

Deutsch, C. V. (2002). *Geostatistical reservoir modeling* (1st Edition ed.). Oxford University Press, U.S.A.

Deutsch, C. V. (2011). Multivariate standard normal transformation. *Centre for Computational Geostatistics Annual Report*, *13*.

Deutsch, C. V., & Journel, A. G. (1998). *Gslib: Geostastical software library and user's guide* (2nd Edition ed.). Oxford University Press.

Deutsch, C. V., & Tran, T. T. (2002). Fluvsim: A program for object-based stochastic modeling of fluvial depositional systems. *Computers & Geosciences*, *28*. Retrieved from `https://ccg-server.engineering.ualberta.ca/CCG%20Publications/Other/CVD%20Papers/01-Peer%20Reviewed/2002/FLUVSIMprgm-stoch-model02.pdf`

Deutsch, C. V., & Wang, L. (1996). Hierarchical object-based geostatistical modeling of fluvial reservoirs. *Society of Petroleum Engineers Annual Technical Conference and Exhibition (ATCE), Paper SPE-36514-MS*.

Deutsch, M. V., & Deutsch, C. V. (2014). Developments in Lerchs Grossmann Optimization - LG3D. *Centre for Computational Geostatistics Annual Report*, *16*.

Egeland, T., Georgsen, F., Knarud, R., & Omre, H. (1993). Multifacies modelling of fluvial reservoirs. *Society of Petroleum Engineers Annual Technical Conference and Exhibition (ATCE), Paper SPE-26502-MS*.

Friedman, J. H. (1987). Exploratory projection pursuit. *Journal of the American Statistical Association*.

Hassanpour, R. M., & Deutsch, C. V. (2010). An introduction to grid-free object-based facies modeling. *Centre for Computational Geostatistics (CCG) Annual Report*, *12*.

Hatloy, A. S. (1994). Chapter 10 - numerical facies modeling combining deterministic and stochastic methods. *Stochastic Modeling and Geostatistics - Principles, Methods, and Case Study*, *3*.

Heriot-Watt, U. (1999). *Petroleum geoscience* (1st Edition ed., Vol. 9). United Kingdom: Heriot-Watt University, Edinburgh.

Hohn, M. E. (1999). *Geostatistics and petroleum geology* (2nd Edition ed.). Springer Science and Business Media, Dordrecht, Netherlands.

Holden, L., Omre, H., & Tjelmeland, H. (1992). Integrated reservoir description. *Society of Petroleum Engineers European Petroleum Computer Conference, Paper SPE-24261-MS*.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics* (1st Edition ed.). Oxford University Press, U.S.A.

Journel, A. G. (1989). *Fundamentals of geostatistics in five lessons* (Vol. 8). Washington, D.C.: American Geophysical Union.

Journel, A. G., & Huijbregts, C. J. (1978). *Mining geostatistics*. New York: Academic Press.

Kelkar, M., & Perez, G. (2002). *Applied geostatistics for reservoir characterization* (1st Edition ed.). Society of Petroleum Engineers, Richardson, Texas, U.S.A.

Leuangthong, O., & Deutsch, C. V. (2001). Multivariate stepwise transformation for stochastic reservoir modeling. *Centre for Computational Geostatistics Annual Report*.

Leuangthong, O., & Deutsch, C. V. (2003). Stepwise conditional transformation for simulation of multiple variables. *Mathematical Geology*, *35*(5).

Leuangthong, O., Deutsch, C. V., Haas, A., & Shtuka, A. (2000). Multivariate stepwise transformation for stochastic reservoir modeling. *Society of Petroleum Engineers Journal of Petroleum Technology, Paper 63069*.

Madani, N. (2019). Application of projection pursuit multivariate transform to alleviate the smoothing effect in cokriging approach for spatial estimation of cross-correlated variables. , *60*, 583–598.

Mallet, J. L. (2002). *Geomodeling* (1st Edition ed.). Oxford University Press, U.S.A.

Montgomery, D. C., & Runger, G. C. (2003). *Applied statistics and probability for engineers* (Third Edition ed.). John Wiley & Sons, Inc., U.S.A.

Murray, C. J. (1994). Chapter 23 - identification and 3-d modeling of petrohysical rock types. *Stochastic Modeling and Geostatistics - Principles, Methods, and Case Study*, *3*.

Pyrcz, M. J., & Deutsch, C. V. (2014). *Geostatistical reservoir modeling* (2nd Edition ed.). Oxford University Press, U.S.A.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. , *23*(3), 470–472. Retrieved from `https://doi.org/10.1214/aoms/1177729394`

Rossi, M. E., & Deutsch, C. V. (2013). *Mineral resource estimation*. Springer Netherlands.

Silva, D. F., & Deutsch, C. V. (2015). Transformation for multivariate modeling using gaussian mixtures with exhaustive secondary data. *Centre for Computational Geostatistics (CCG) Annual Report*, *17*.

Silva, D. F., & Deutsch, C. V. (2017). Multivariate categorical modeling using hierarchical tpg. *Centre*

for Computational Geostatistics (CCG) Annual Report, *19*.

Silva, D. F., & Deutsch, C. V. (2018). Guide to categorical modeling with htpg. *Centre for Computational Geostatistics (CCG) Guidebook Series*, *23*.

Strebelle, S. B., & Journel, A. G. (2001). Reservoir modeling using multiple-point statistics. *Society of Petroleum Engineers Annual Technical Conference and Exhibition (ATCE), Paper SPE-71324*.

Zagayevskiy, Y., & Deutsch, C. V. (2015). Grid-free geostatistical modeling of categorical variables. *Centre for Computational Geostatistics (CCG) Annual Report*, *17*.

Zagayevskiy, Y., & Deutsch, C. V. (2016). Grid-free petroleum reservoir characterization with truncated pluri-gaussian simulation: Hekla case study. *Petroleum Geoscience*, *22*(3), 241–256. Retrieved from `https://pubs.geoscienceworld.org/pg/article-abstract/22/3/241/323253/Grid-free-petroleum-reservoir-characterization`

Appendix A

Appendices

A.1 Hekla Reservoir Well Logs

The logs shown in Figures A.1 through A.15 are the well logs of all fifteen wells in the Hekla reservoir. The well logs are records of the variation of facies and continuous petrophysical properties (porosity and permeability) with depth, in the two distinct layers ($H1$ and $H2$) of the reservoir. The first log clearly shows the two layers ($H1$ and $H2$) of the Hekla reservoir through which the wells are drilled. The second log shows the vertical distribution of the five facies at the well locations in the reservoir. The third and fourth logs are the record of porosity and permeability, respectively, with depth in the reservoir.

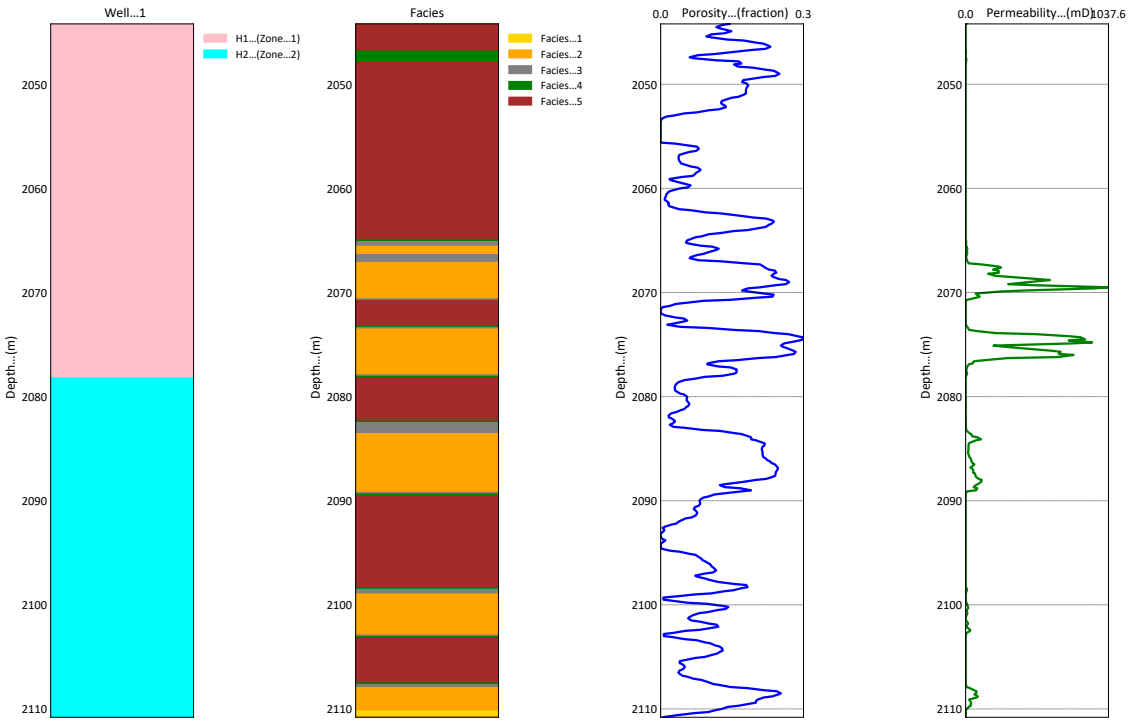


Figure A.1: Well log of Hekla Well 1 showing the distinctive reservoir zones, facies, porosity and, permeability records

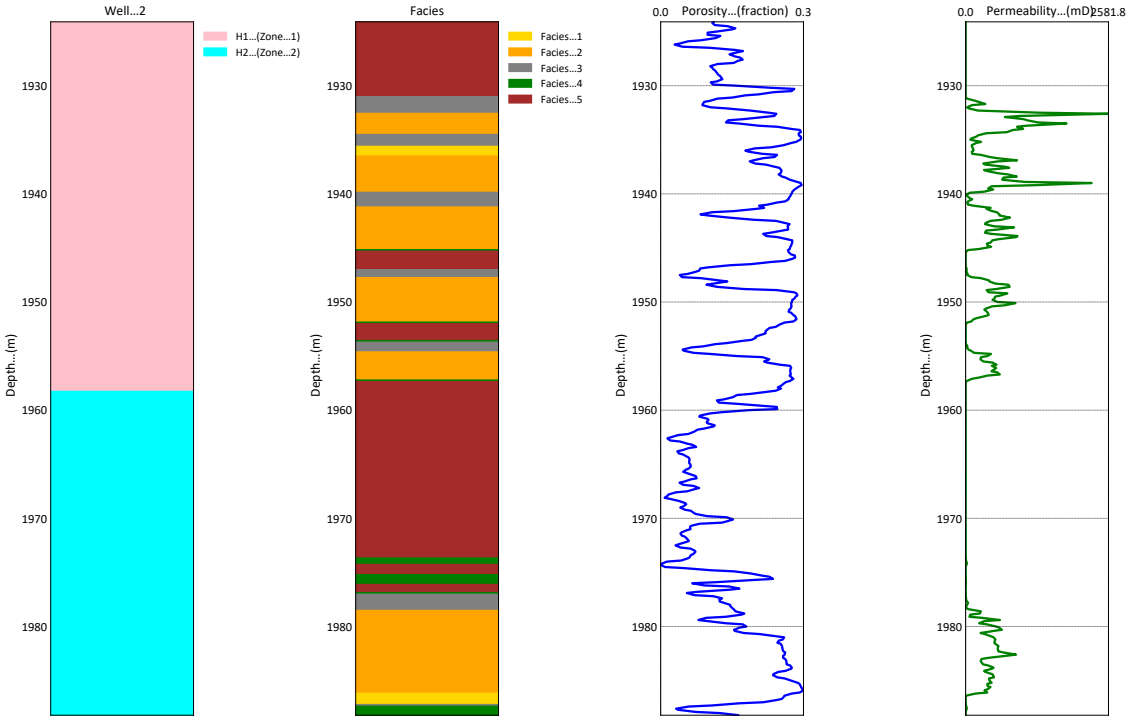


Figure A.2: Well log of Hekla Well 2 showing the distinctive reservoir zones, facies, porosity and, permeability records

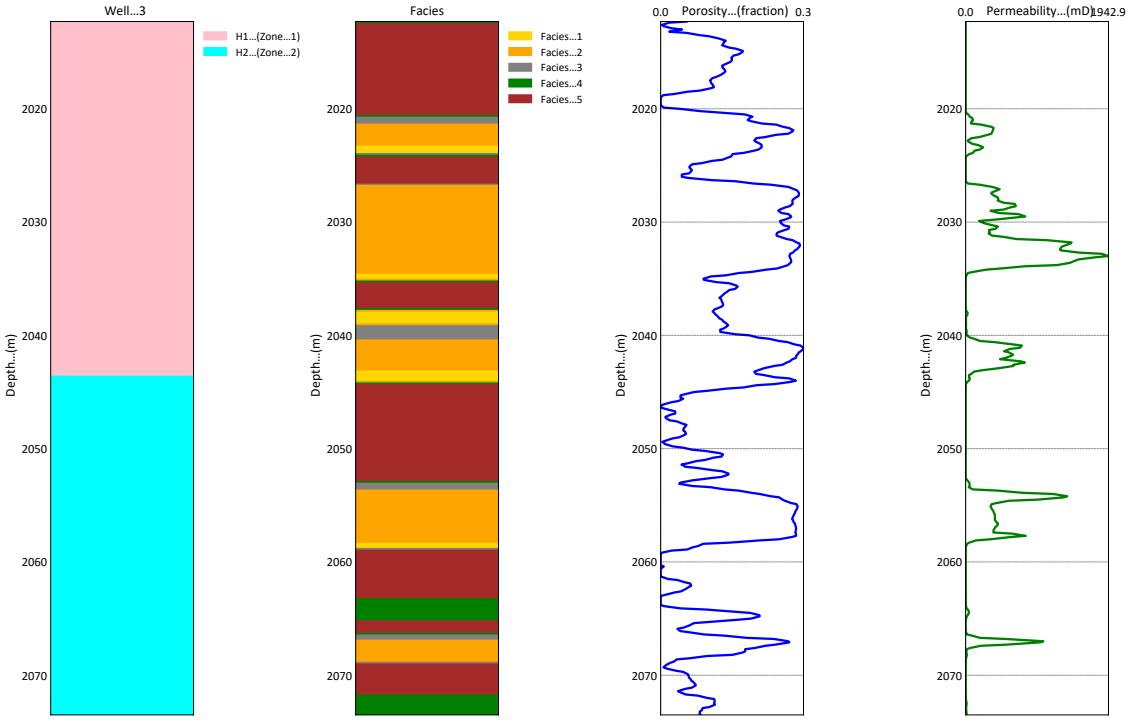


Figure A.3: Well log of Hekla Well 3 showing the distinctive reservoir zones, facies, porosity and, permeability records

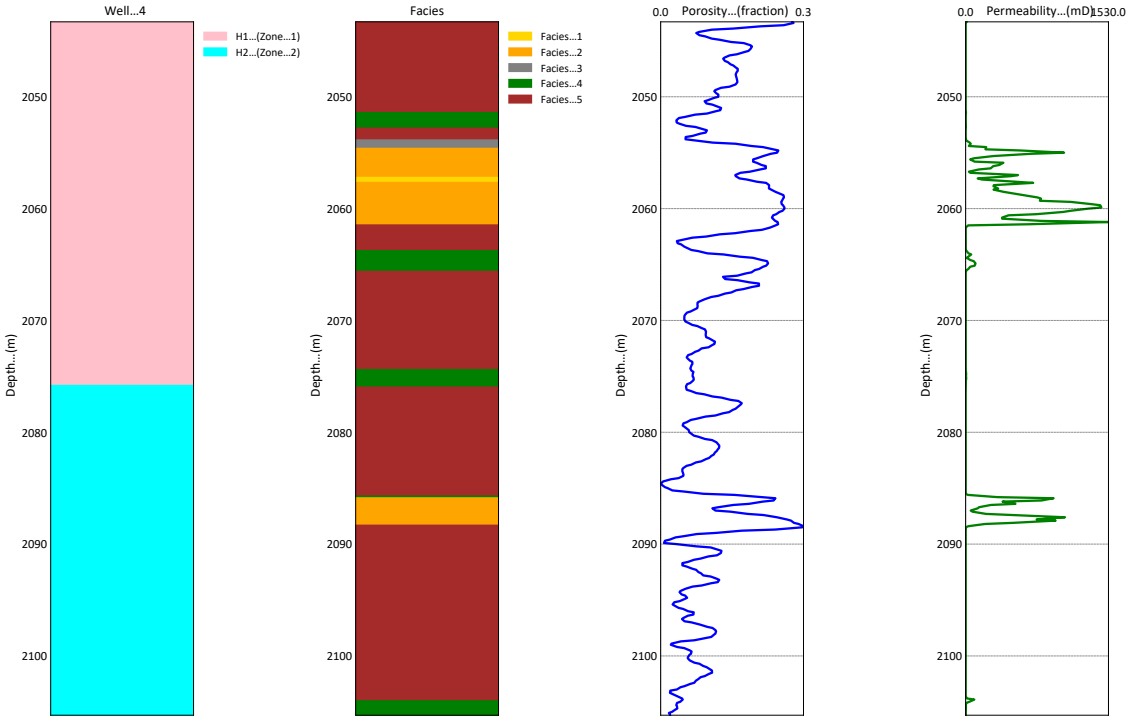


Figure A.4: Well log of Hekla Well 4 showing the distinctive reservoir zones, facies, porosity and, permeability records

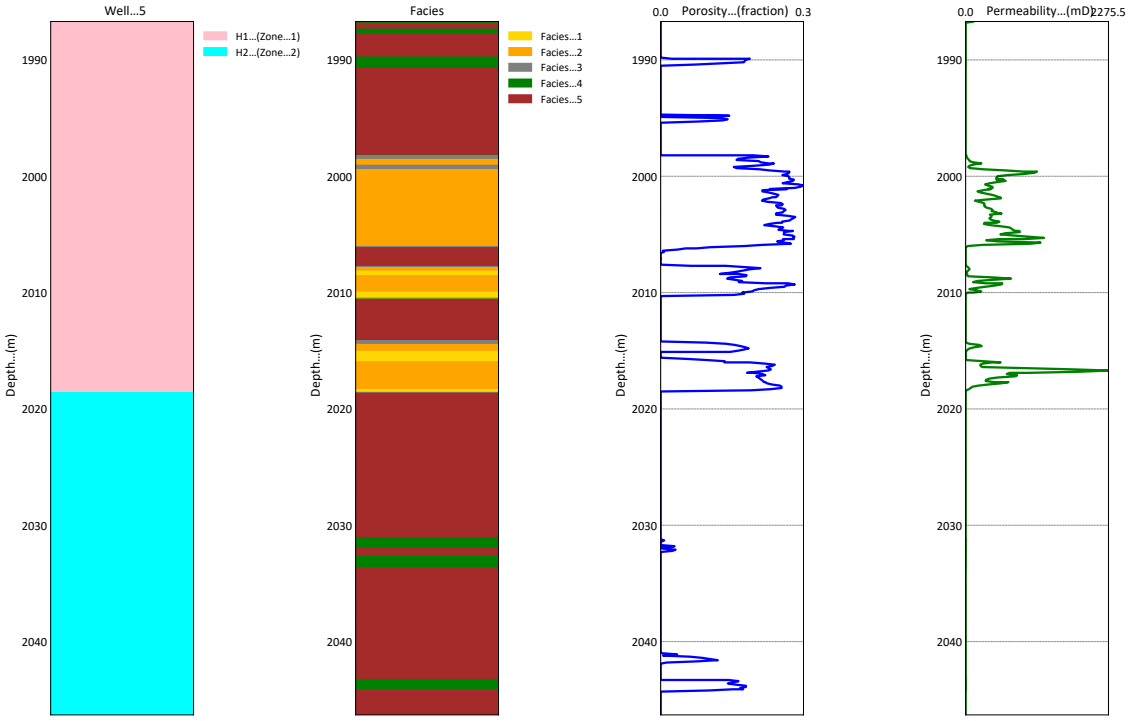


Figure A.5: Well log of Hekla Well 5 showing the distinctive reservoir zones, facies, porosity and, permeability records

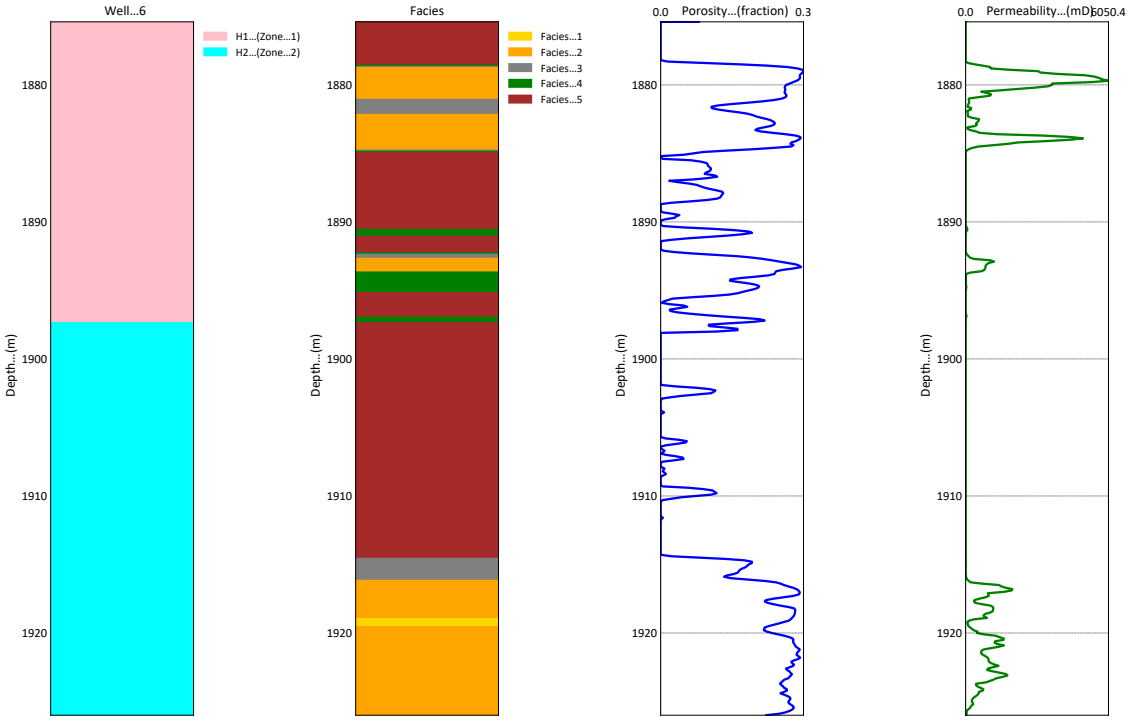


Figure A.6: Well log of Hekla Well 6 showing the distinctive reservoir zones, facies, porosity and, permeability records

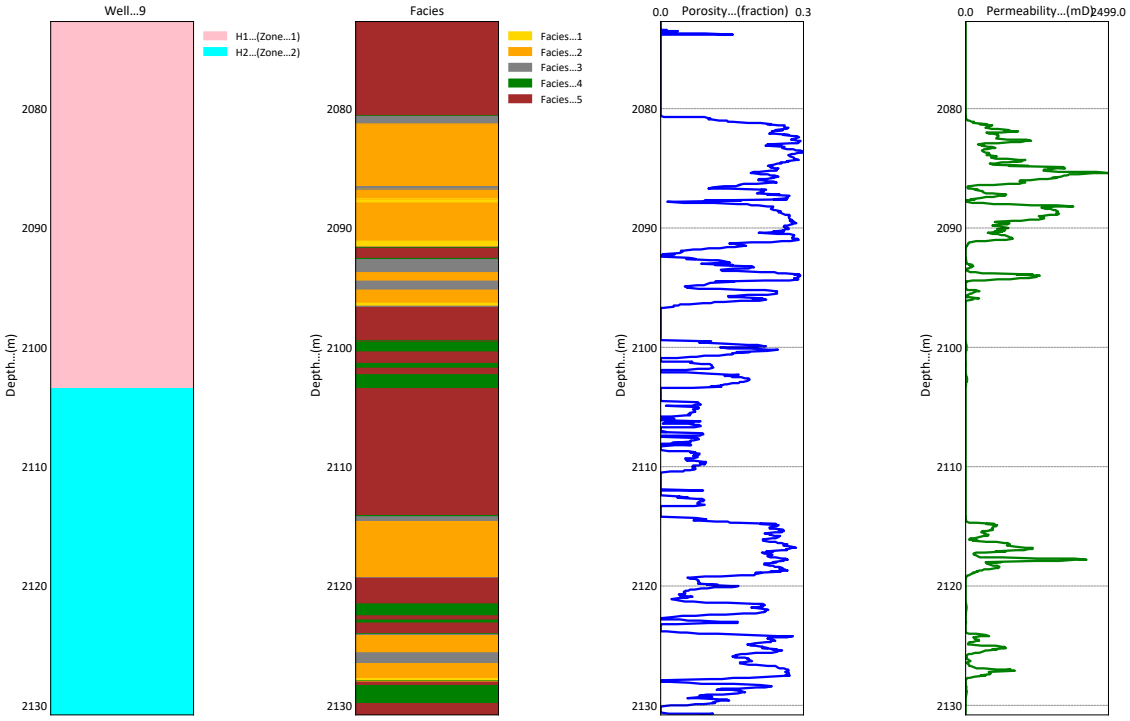


Figure A.7: Well log of Hekla Well 9 showing the distinctive reservoir zones, facies, porosity and, permeability records

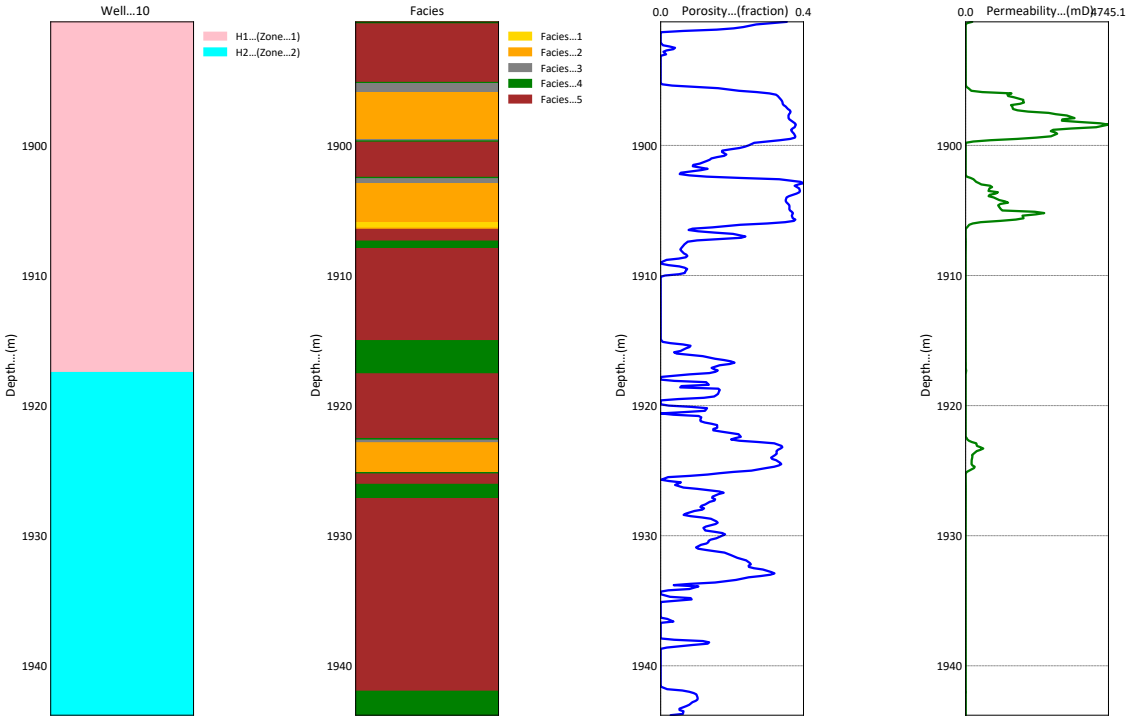


Figure A.8: Well log of Hekla Well 10 showing the distinctive reservoir zones, facies, porosity and, permeability records

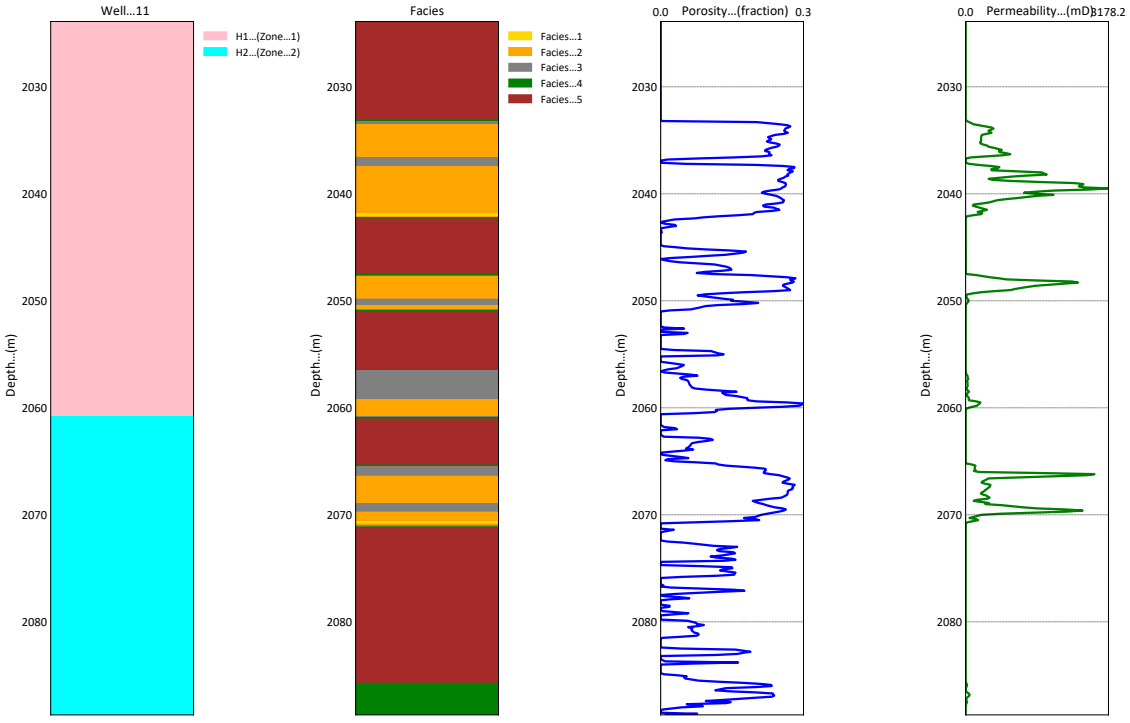


Figure A.9: Well log of Hekla Well 11 showing the distinctive reservoir zones, facies, porosity and, permeability records

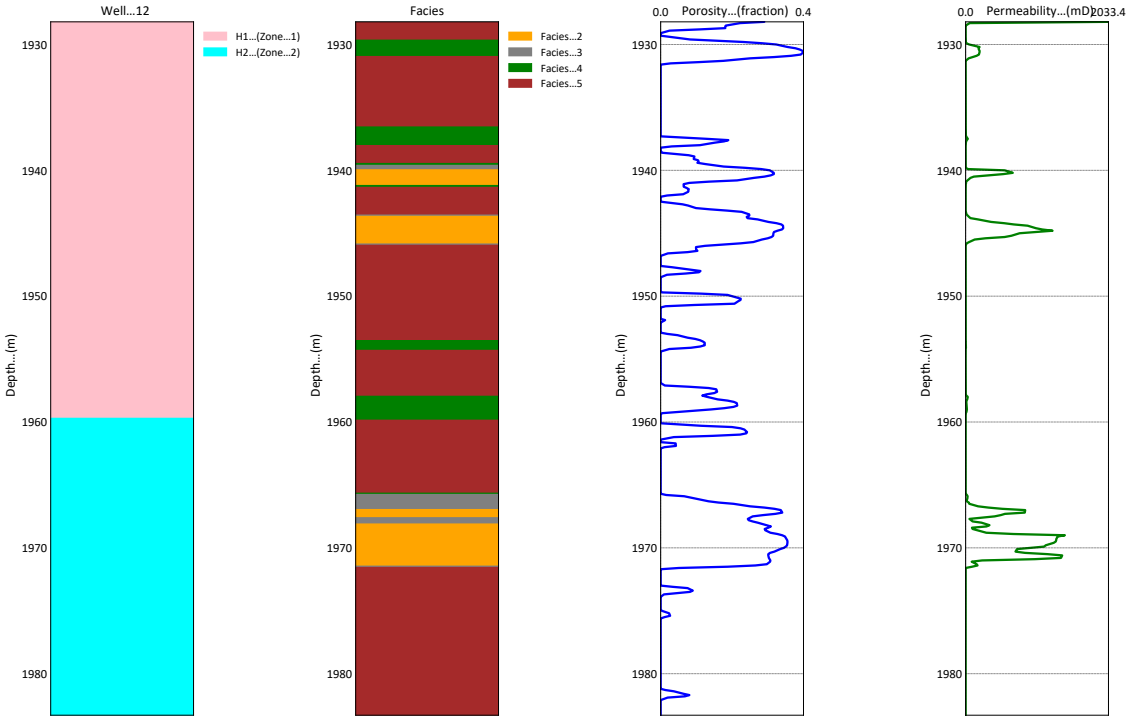


Figure A.10: Well log of Hekla Well 12 showing the distinctive reservoir zones, facies, porosity and, permeability records

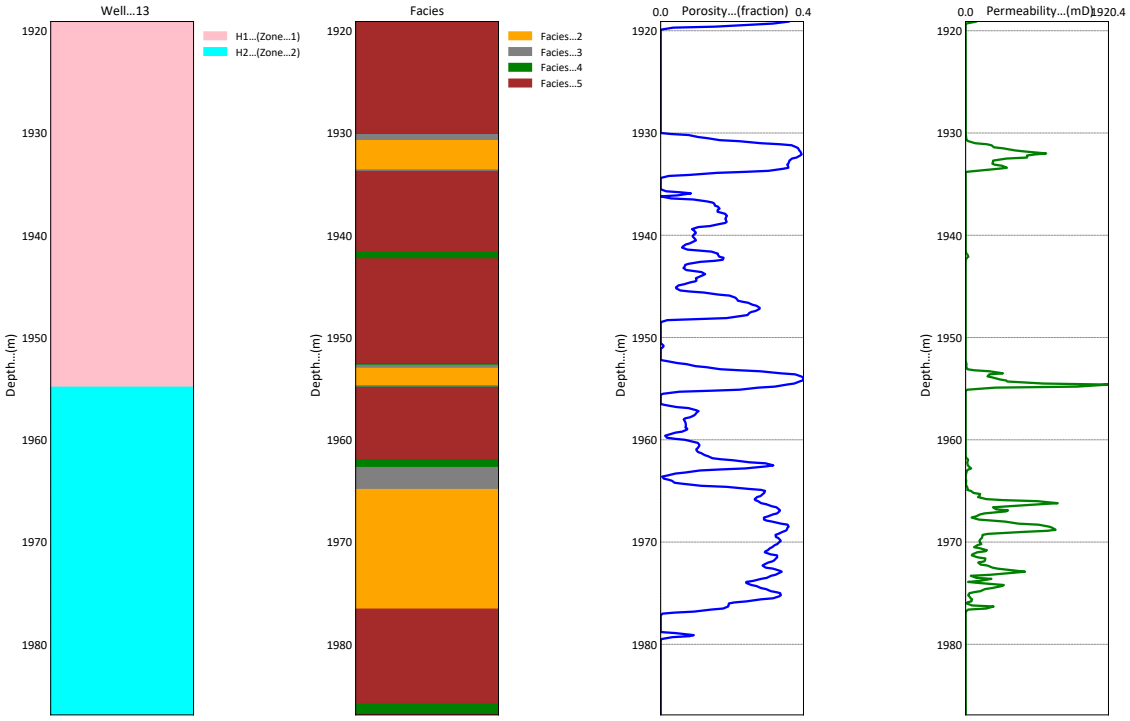


Figure A.11: Well log of Hekla Well 13 showing the distinctive reservoir zones, facies, porosity and, permeability records

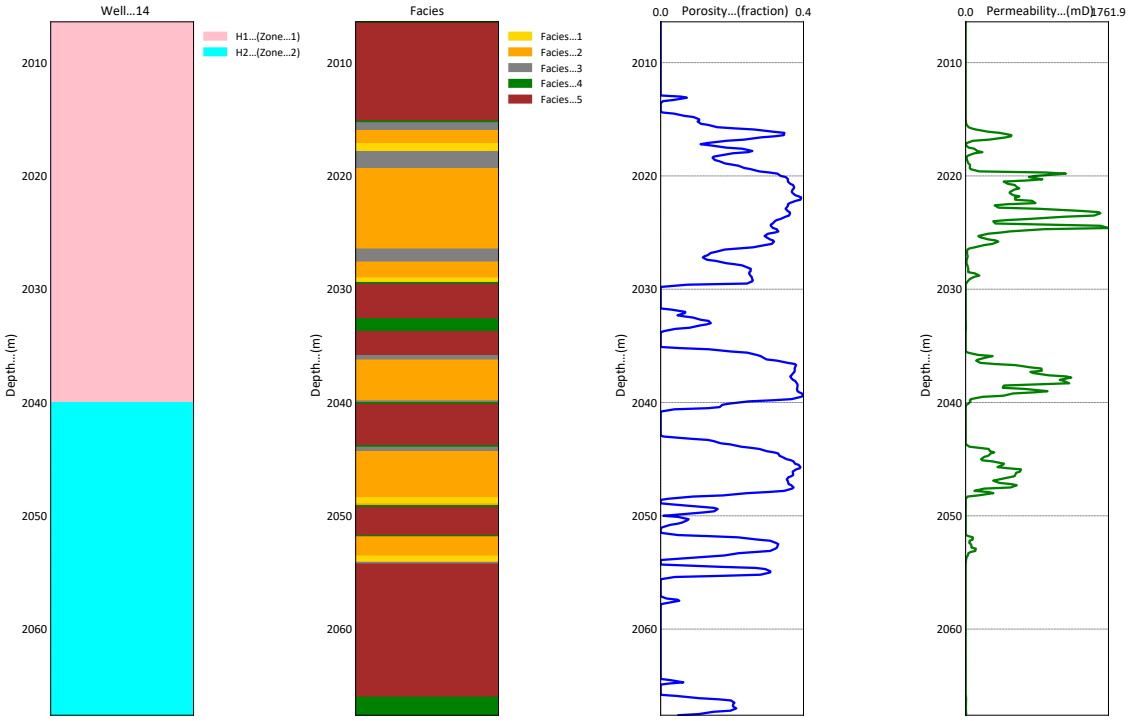


Figure A.12: Well log of Hekla Well 14 showing the distinctive reservoir zones, facies, porosity and, permeability records

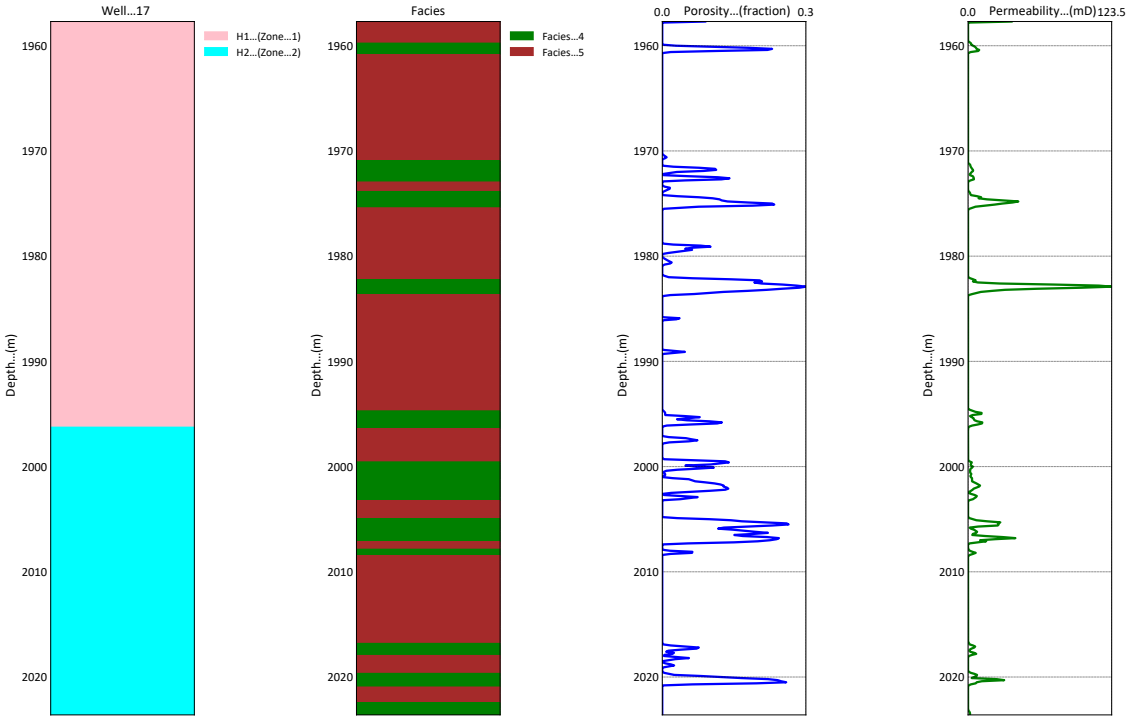


Figure A.13: Well log of Hekla Well 17 showing the distinctive reservoir zones, facies, porosity and, permeability records

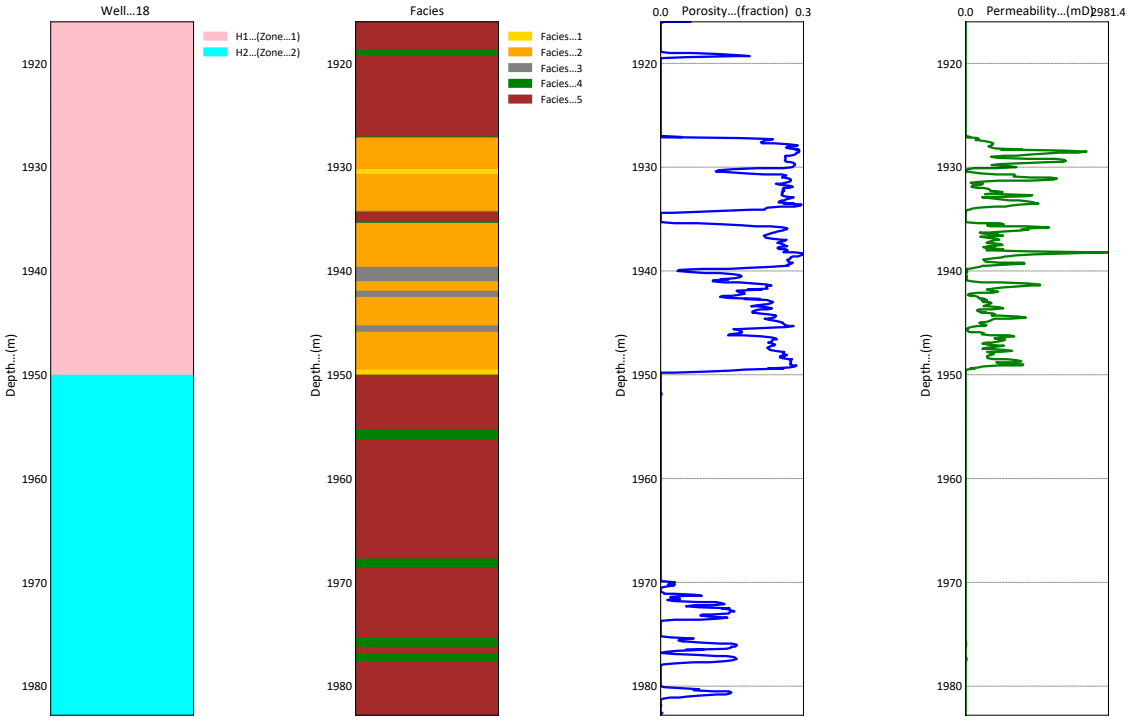


Figure A.14: Well log of Hekla Well 18 showing the distinctive reservoir zones, facies, porosity and, permeability records

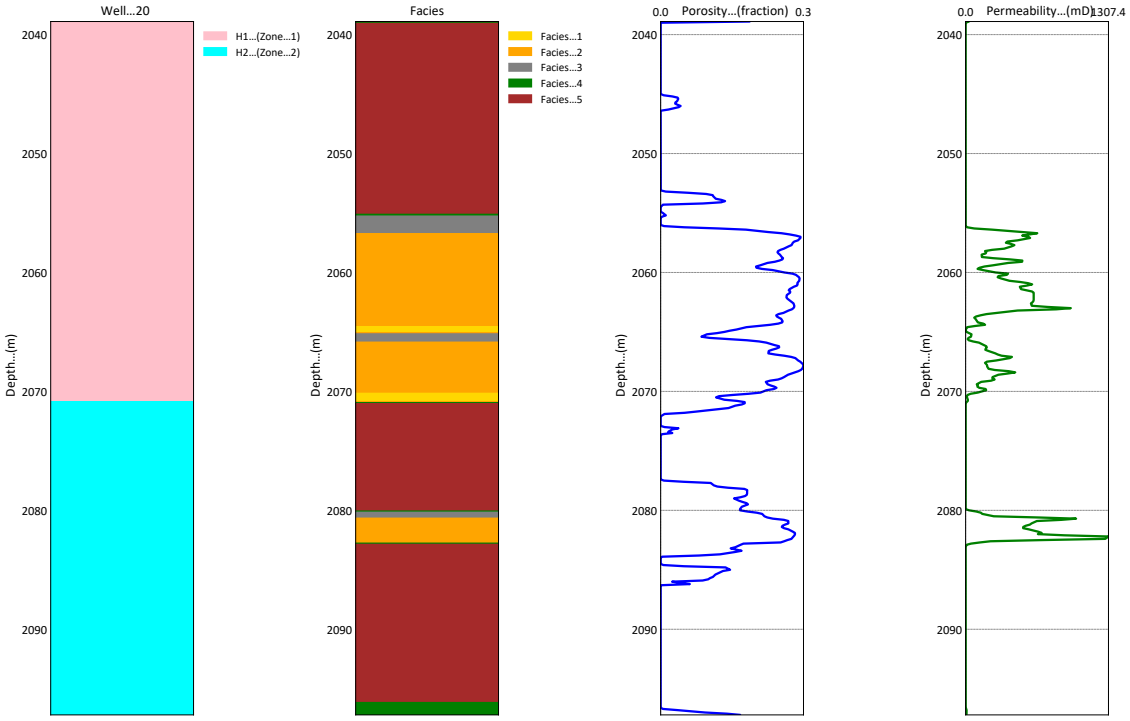


Figure A.15: Well log of Hekla Well 20 showing the distinctive reservoir zones, facies, porosity and, permeability records