CHAPTER 2

Methodology

Geostatistical seismic inversion generates multiple high-frequency models for high-resolution reservoir characterization and uncertainty analysis. This method has allowed the integration of low-frequency seismic information with high-resolution well log data, providing results that consist of a spatial distribution of elastic properties and facies with a vertical resolution intermediate between that of seismic data and well logs (McCrank et al., 2009). The geostatistical inversion methodology provides numerous realizations of elastic properties and facies distribution, all honouring input seismic data, input well log data and geostatistical distribution parameters. Each realization represents a feasible solution with regards to elastic property and facies distributions, and by analysing a complete set of such realizations the range of uncertainty can be captured.

2.1 Theoretical Background

2.1.1 Geostatistics

Geostatistics is a branch of statistics that describes the spatial continuity of natural phenomena (Krige, 1952). Geostatistical techniques rely on statistical models that are based on random-variable theory to model the uncertainty associated with spatial estimation and simulation. It provides reliable estimation of phenomena at locations where no measurement is available. The predictions made by using geostatistical methods are optimal when data are normally distributed and stationary (mean and variance do not vary significantly in space). Conventionally, the most basic tool for describing spatial correlation of any data value is by the use of a variogram (Bohling, 2007).

A variogram is a function describing the degree of spatial dependence of a spatial random field or stochastic process. It is defined as the variance of the difference between field values at two locations across realizations of the field (Cressie, 1993). Conceptually, data points taken close to each other will provide higher correlation than data points taken far apart. The term variogram is also commonly used to refer to a plot of distance (x-axis) and variance (y-axis), constructed from a mathematical calculation or from measured data.

As a function, a variogram is described by the following parameters:

- Type: the type determines how quickly the continuity decreases as a function of distance (h or lag). The variogram types used in this project are Gaussian and Exponential.
- Range: the maximum distance at which two points are still correlated.
- Sill: the sample variance of the property.
- Nugget: amount of variance at zero distance. It determines how closely measured points are to be honored and can be thought of as the sum of measurement uncertainty and microscale variability.

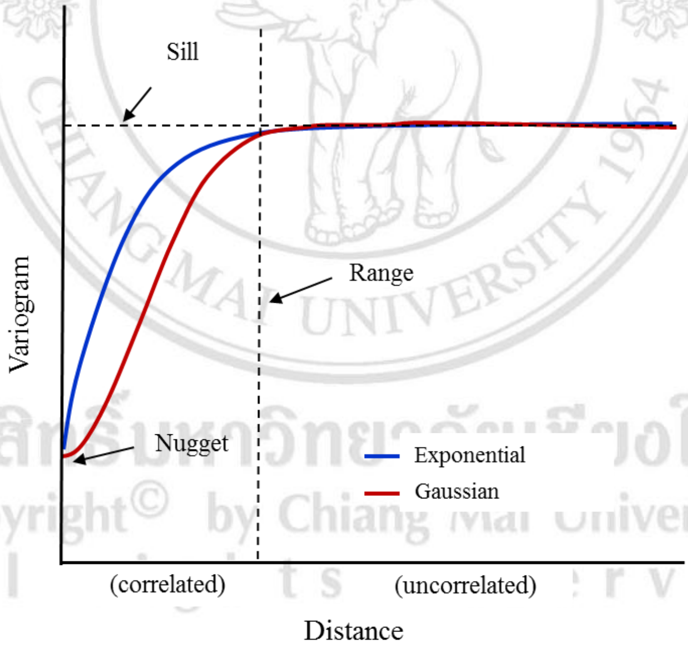


Figure 2.1 Variogram features, including type, range, sill, and nugget (Bohling, 2007).

2.1.2 Bayesian Inference

Bayes' theorem describes the probability of an event, based on conditions that might be related to the event. Bayesian inference is one of the many application of Bayes' theorem which is used to update the probability for a hypothesis as more evidence or information is available. Bayesian inference is used to determine the posterior probability as a consequence of a prior probability and a likelihood function which is derived from a statistical model for the observed data (Gilks et al., 1996). It can be written as:

$$P(X|H,E) = \frac{P(X|H)P(E|X)}{P(E|H)} \quad , \qquad (2.1)$$

where

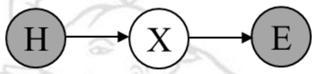
$P(X|H,E)$ is the desired posterior probability.

$P(X|H)$ is the prior distribution of event X given only our hypothesis H.

$P(X|E)$ is the likelihood of observing evidence E given an event X.

$P(E|H)$ is a normalizing term.

Equation (2.1) can be graphically represented as:



The grey circles represent known data (hypothesis and evidence) and the white circle is a simulated property or event. Using this approach, we hypothesize that a variogram (v) has given rise to a P-impedance (Zp) whose corresponding reflectivity produced a synthetic field (syn) within a given noise level of the measured seismic (s). This can be written as:

$$P(Zp|v,s) \propto P(Zp|v)P(s|syn) \qquad (2.2)$$

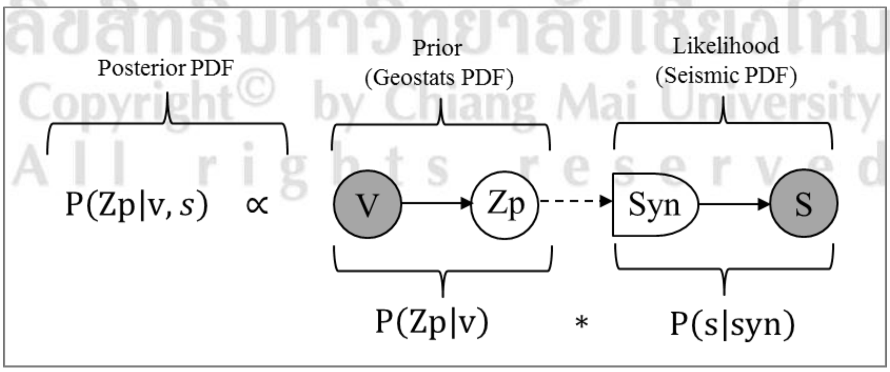


Figure 2.2 Schematic representation of how the posterior PDF is built in Markov Chain Monte Carlo (MCMC) (Jason, 2013).

In this example, the posterior PDF (probability density function) represents the probability of simulated P-impedance (Zp), given that we have a hypothesized property variogram (v), a convolutional model (syn), and measured seismic data (s) (Jason, 2013). Moreover, the target PDF can be generated from various sources of information, i.e. pdfs for other elastic properties, lithotypes, noise in the seismic data, and so on. Figure 2.3 illustrates how various sources of information are combined to get the target posterior PDF.

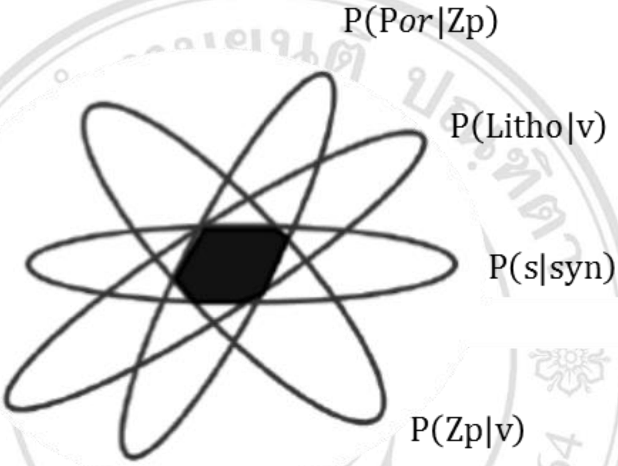


Figure 2.3 The PDF of each input source is shown as a circle. The target posterior pdf we are interested in is shown in black as the intersection of all input PDFs (Jason, 2013).

2.1.3 Geostatistical Inversion Algorithm

The geostatistical inversion algorithm used in this study is a PDF representing P(reservoir | geostatistics, seismic), which is the probability of the produced reservoir model, given the input geostatistics (from well logs) and measured seismic; then, sample the posterior PDF using Markov Chain Monte Carlo (MCMC) algorithm to get desired property volumes (Jason, 2013). MCMC results do not provide a single best solution, but rather drift around the space of solutions which is within the probability distribution (the posterior PDF) (Gilks et al., 1996 and Oh and Kwon, 2001). Therefore, the range of uncertainty can be quantified from various inversion results.

Brief details of the customized MCMC algorithm used for geostatistical inversion (Jason, 2013) are as described below.

- Individual PDFs for each input data source (i.e., well logs, variogram, seismic data) are computed and combined together to obtain the target posterior distribution T(x) by using Bayesian Inference.
- A proposed distribution π(x) is constructed to approximate T(x), using a Gaussian function. T(x) is too complex to sample fairly and efficiently.
- A candidate prop1 was drawn from $\pi(x)$ and evaluated whether the prop1 is better or worse than prop0 by using MCMC.
- Hastings factor H (Neal, 1993) is computed to make a decision and compared to a number sampled from a uniform distribution U~(0,1). If the sampled number U~(0,1) is smaller than H, implying that prop1 is better than prop0, the current property will be updated by replacing prop0 by prop1. Otherwise, prop1 will be rejected. Both discrete (facies) and continuous properties (AI, Vp/Vs, and density) are inverted simultaneously.
- The algorithm stops after a sufficient number of iterations yielding a distribution approximately that of the target distribution T(x).

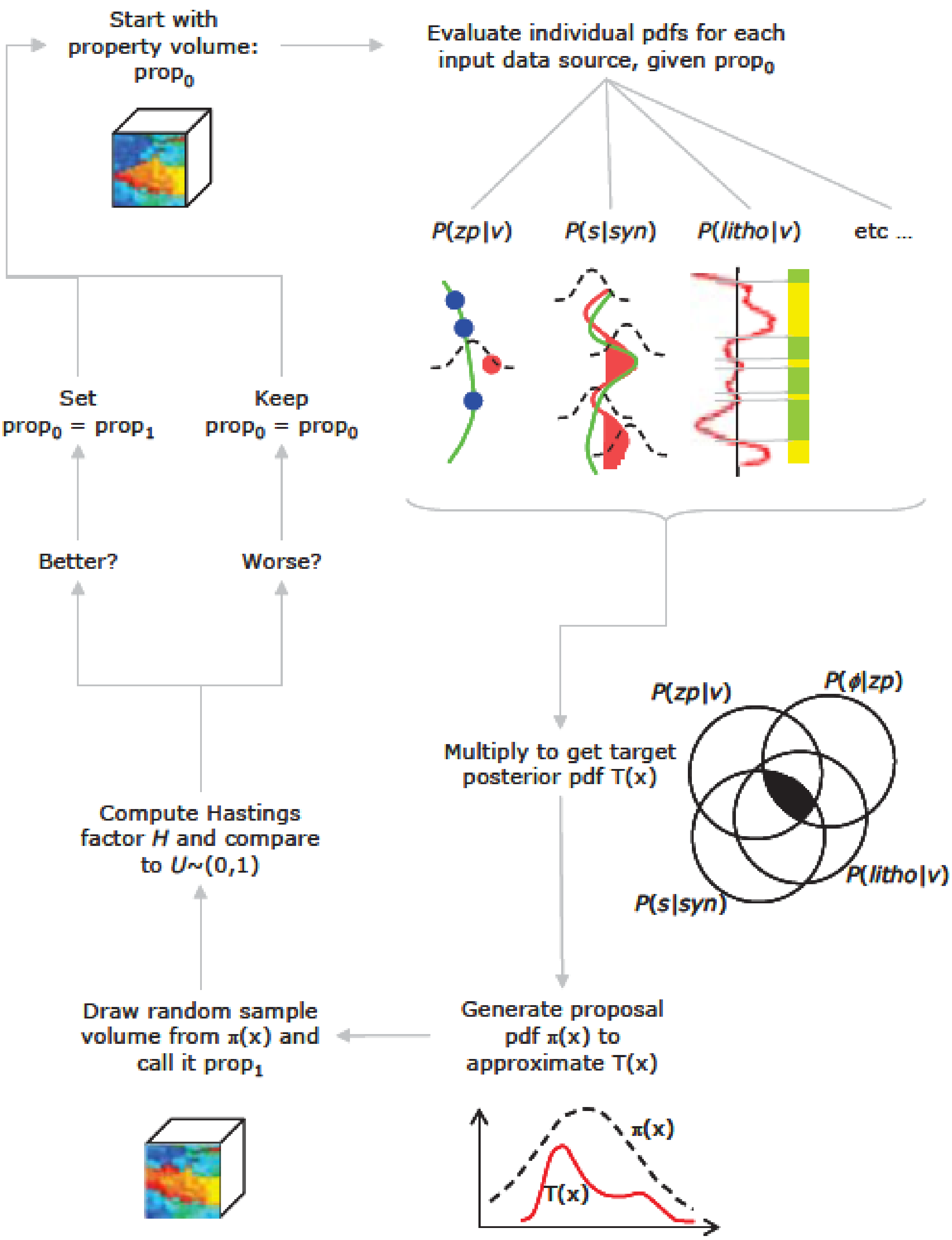


Figure 2.4 Simplified schematic describing the customized MCMC algorithm used for geostatistical inversion (Jason, 2013).

2.2 Geostatistical Inversion Workflow

The workflow used in this pre-stack geostatistical inversion study consisted of six main steps as follows:

2.2.1 Well data conditioning

The quality of the input well data is of great importance to this study. This is ensured by using final logs, after environmental corrections and petrophysical corrections have been applied. All prior PDFs of elastic properties used for geostatistical model fitting are derived from geophysical logs such as P-sonic, S-sonic and density logs. These measured logs are normally affected by borehole wash-out and mud-filtrate invasion due to the shallow depth of investigation of these logging tools. Therefore, appropriate log editing must be carried out before further analysis can be initiated.

The log data conditioning included bad-data removal, depth-shift correction, and other quality improvements. Cross-plots of well log data were used to investigate any abnormal data points diverting from the normal trend. The cros-plots used for well-log data conditioning were as follows:

- compressional velocity (Vp) *versus* density (Rho)
- compressional velocity (Vp) *versus* shear velocity (Vs)

This step also involved shear velocity (Vs) prediction in wells without measured shear sonic log data. The prediction methodology was a multi-linear regression (MLR), using compressional velocity (Vp), density (Rho), Gamma ray (GR), Neutron-porosity (NPHIE), and deep resistivity (AT90) as input logs. MLR models the linear relationship between Vs and other log data, and MATLAB software was used for this calculation. The MLR was expressed by equation (2.3)

$$Vs_i = \beta_0 + \beta_1(Vp_i) + \beta_2(Rho_i) + \beta_3(GR_i) + \beta_4(NHPIE_i) + \beta_5(AT90_i)\ , \qquad (2.3)$$

where β_0 = intercept and

$\beta_1 to\ \beta_5$ = partial regression coefficient.

Rock-physics analysis was carried out to review the feasibility of classifying lithology types using elastic properties. Cross-plots of elastic properties such as P-impedance (AI), Vp/Vs, and density were analysed to verify lithology discrimination. This helped define the proper lithology set and stratigraphy interval for geostatistical inversion.

2.2.2 Seismic-to-well tie and wavelet extraction

Wells were tied to seismic data to correlate the seismic data with synthetic seismograms and to obtain time-to-depth relationships. The workflow used for well ties and wavelet extraction is summarized in Figure 2.5.

The check-shot/VSP data provided the initial time:depth relationship. A synthetic seismogram was computed by convolving the reflection coefficient log with a wavelet extracted from the full-offset stack. Synthetic seismograms were computed at each well location, and used to determine a well-to-seismic tie for the full-offset stack using a combination of bulk time shifts and minor stretch/squeeze adjustments. This was followed by a second well-to-seismic tie considering all angle stacks simultaneously. A synthetic seismogram of each angle stack was generated by convolving an angle-dependent wavelet with reflectivity series calculated from Vp, Vs, and density logs. The final time:depth curve was determined using a combination of static time shifts and minor stretch/squeeze to optimize the calibration of the synthetic trace to the seismic trace for all angle stacks. When stretching/squeezing well-log reflection times, the process was carried out with care to avoid unrealistic and abrupt changes in the drift curve. A correlation coefficient was computed to determine the match between seismic and synthetic traces at all well locations. A correlation factor of 1 represented perfect match, while a correlation factor of 0 represented no correlation.

Wavelets for all available angle stacks were extracted along each well based on their final cross-correlations. The final wavelet used in the geostatistical inversions were multi-well averaged wavelets which were calculated for each angle stack.

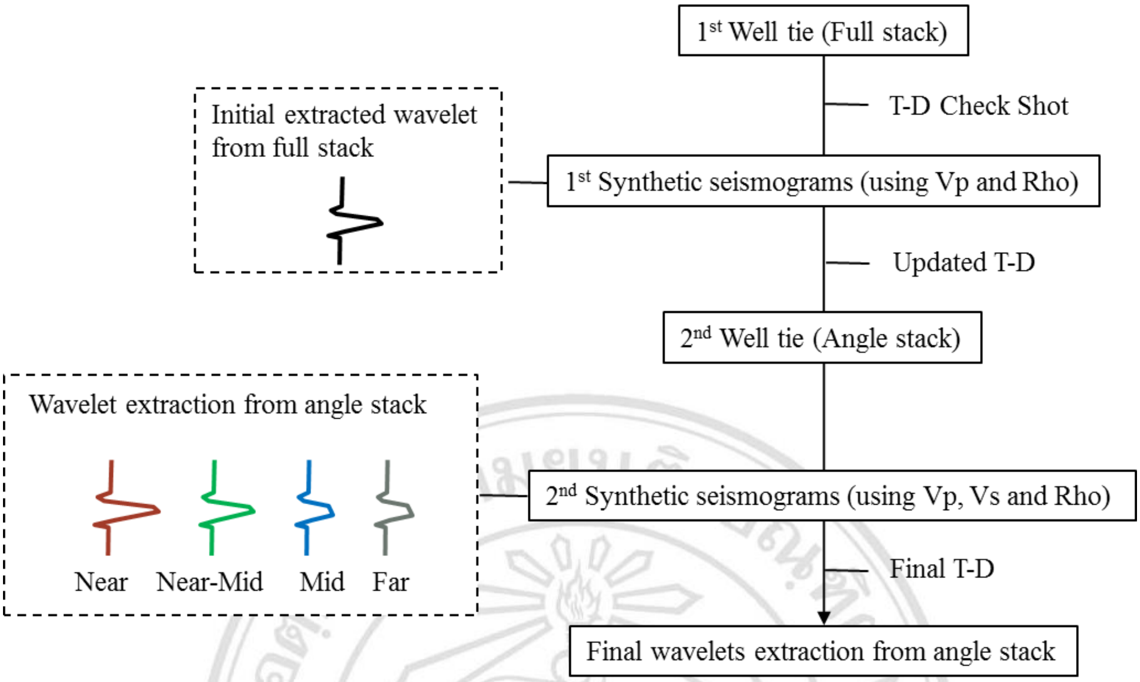


Figure 2.5 Well tie and wavelet extraction workflow.

2.2.3 Building the solid model

The main objective of solid model building is to create a model for the study area. This model was built from interpreted seismic horizons (Jason, 2015). The solid model was built in the time domain, based on an input stratigraphic framework appropriate to resolve the scale of the reservoir layers in the zone of interest (Jason, 2015). The output solid model was a parametric representation of the subsurface covering the study area, and was used to conduct geostatistical modelling and inversion. Descriptions and examples of common stratigraphic layering types are shown in Figure 2.6.

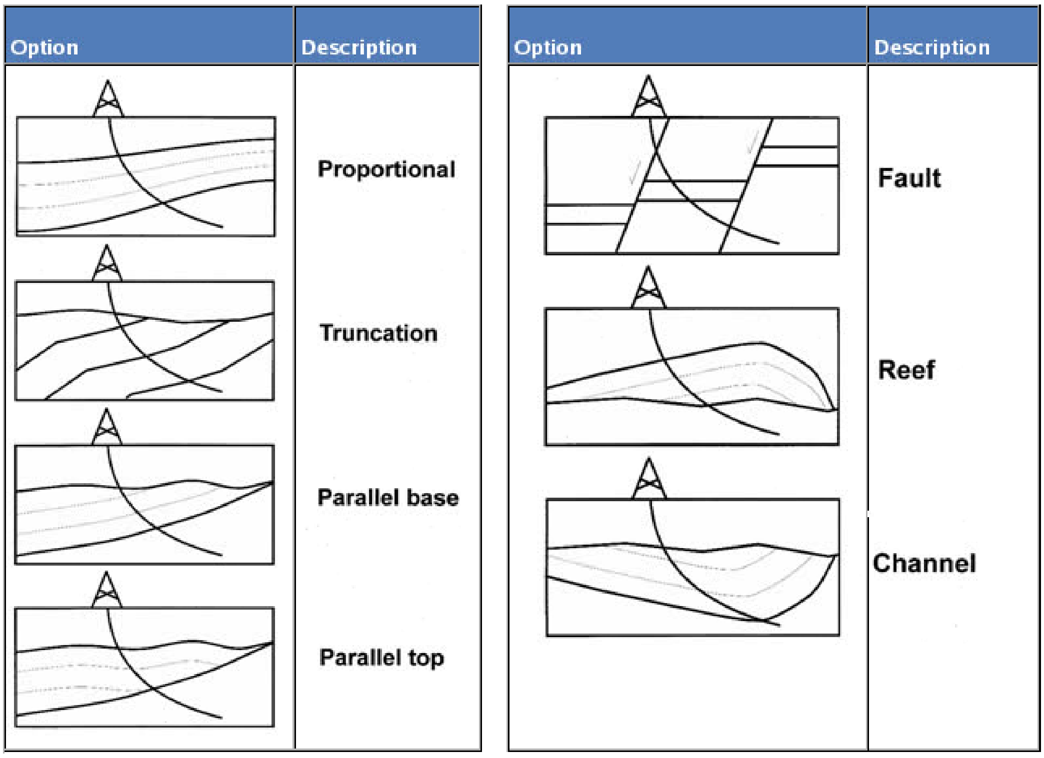


Figure 2.6 Example models for stratigraphic layers (Jason, 2015).

2.2.4 Geostatistical Model Fitting

The objective of geostatistical model fitting is to define the prior PDFs and variogram settings for the discrete facies and continuous elastic properties (Jason, 2015). A probability density function (PDF) is a statistical tool that quantifies and models the spatial continuity of the property. Two types of PDF were used in this project: normal and log-normal PDF. Prior proportions of discrete facies properties were derived from well logs data using a 1D PDF. PDFs of continuous properties were defined using a joint PDF which consisted of the PDF of each property and their linear relationships. When modelling the joint PDF of N properties, the model contained N pdfs and N-1 linear correlation coefficients that expressed the extents of the linear relationships for each pair of properties (Jason, 2015).

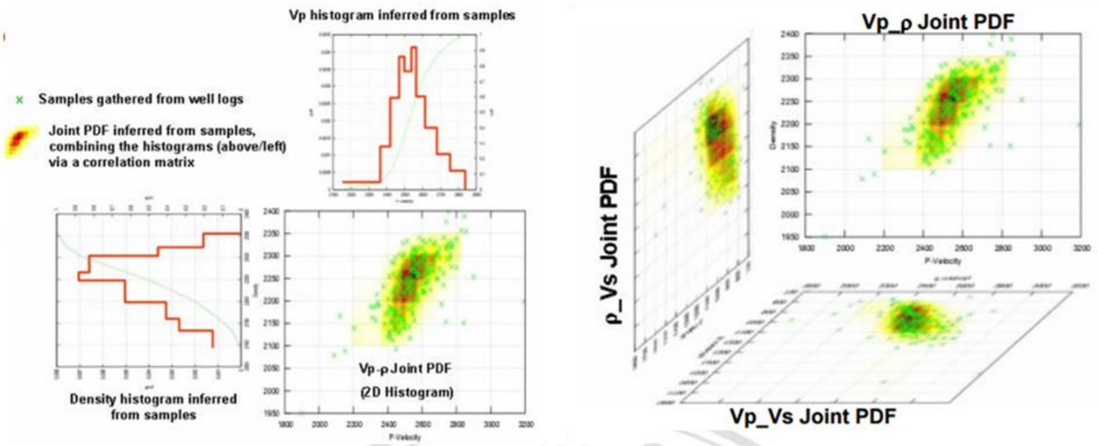


Figure 2.7 Examples of a multidimensional joint distribution generated from well log data: (left) Vp-Density joint PDF (2D histogram) derived from 1D sample histogram, (right) 3D joint PDF of acoustic properties: Vp, Vs, and density (Jason, 2015).

A variogram must be defined for each property in the geostatistical model. The shape of the variogram describes the smoothness of its values (Jason, 2015). The appropriate variogram type is decided based on its respective property distribution. Two types of variogram were considered for this study: exponential and Gaussian (Bohling, 2007). The exponential distribution is often used to model highly fluctuating data, such as elastic properties where correlations between data points decreases rapidly with distance (Bohling, 2007). In contrast, a Gaussian variogram expresses the correlation between two data points with a slower and more gradual reduction in correlation with increasing distance (Bohling, 2007). This variogram type is therefore commonly used to model highly continuous data, i.e., discrete properties.

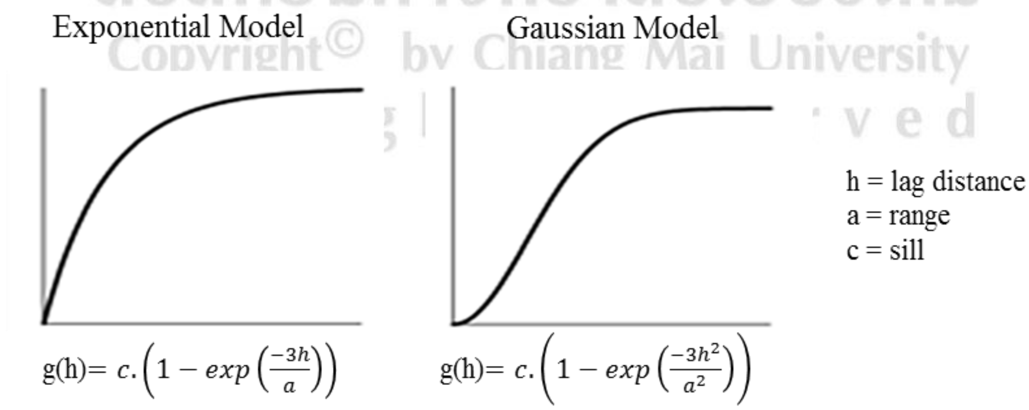


Figure 2.8 Exponential and Gaussian variograms (Bohling, 2007).

For each property, both vertical and lateral variograms are required to control spatial continuity in all dimensions. The vertical variograms are normally derived from well log data, i.e., lithology, P-impedance, Vp/Vs, and density. In this project, mean-value attribute maps derived from deterministically sand probability data were used to derive lateral variograms (Jason, 2015).

2.2.5 Simulation

Simulation is performed to test whether the chosen geostatistical parameters defined by the prior PDFs and variograms can be used to obtain the desired shapes of discrete and continuous properties (Jason, 2015). To capture a range of uncertainty, such geostatistical parameters must also account for minimum and maximum cases, in addition to a most likely case, and be defined as part of the input to the final inversion production (Jason, 2015). In order to QC such simulation parameters, results should be consistent with input statistics, wells and deterministic inverted lithology/property volumes (Jason, 2015).

2.2.6 Inversion

The geostatistical inversion algorithm used in this study was described in topic 2.1.3. The products of geostatistical inversion are multiple realizations of elastic properties, i.e., acoustic impedance (AI), Vp/Vs, and density. After finalizing the selection of geostatistical parameters, inversion is carried out by using various sets of parameters representing different possible scenarios in order to capture a reasonable range of the most probable outcomes through multi-realizations.

Figure 2.9 shows general input and output data involved in geostatistical inversion. Input data mainly include seismic angle stacks, angle-dependent wavelets, and well logs of AI, Vp/Vs, Density, and lithofacies. The geostatistical model included 3D PDF's of elastic properties and variograms derived from input well-log data. The final outputs from the geostatistical inversion process includes multi-realized results of AI, Vp/Vs, density, and lithofacies.

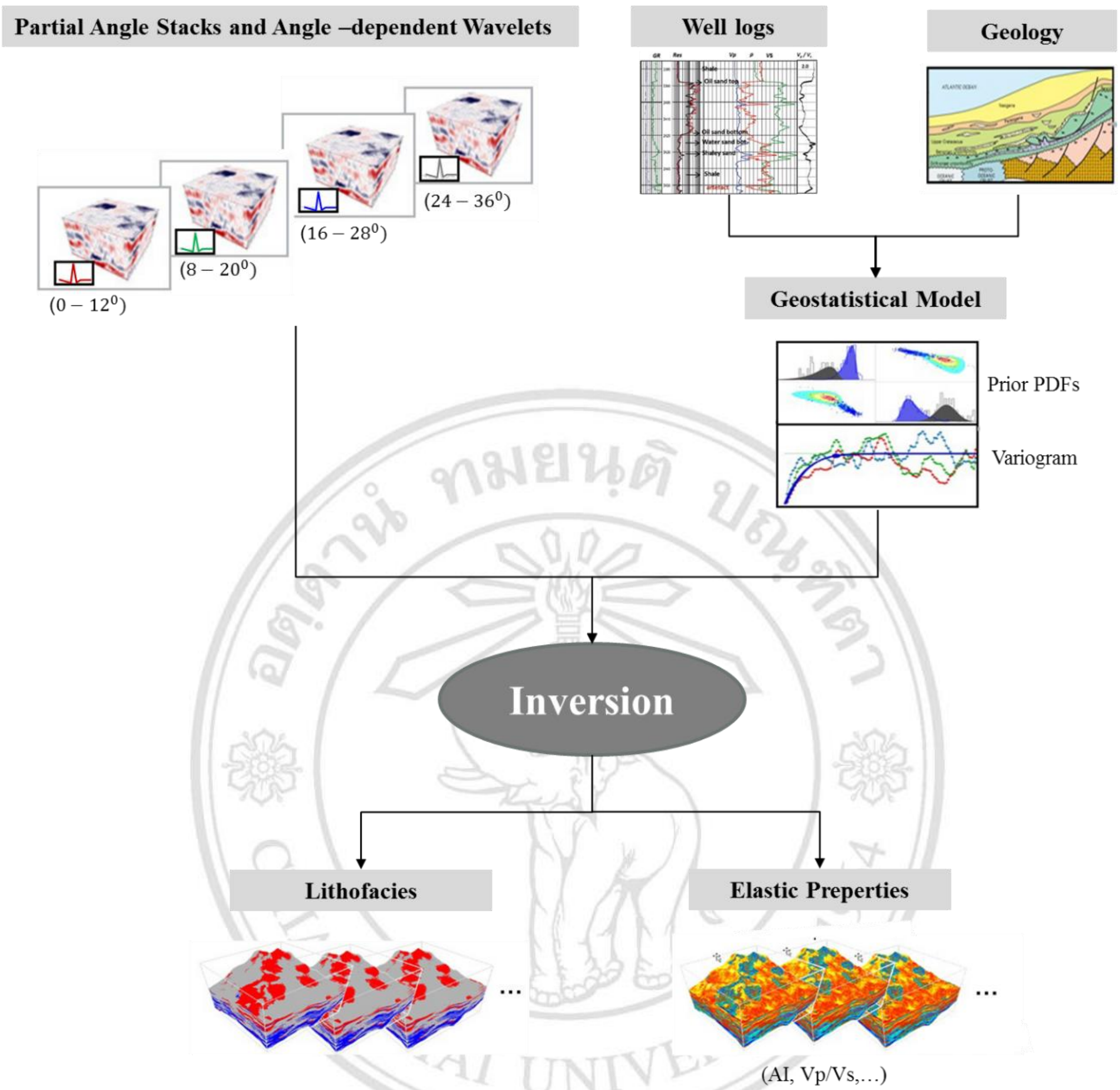


Figure 2.9 Schematic representation of the input and output data related to the geostatistical inversion process (Jason, 2015).