Stochastic Modeling and Geostatistics for Reservoirs; Principles and Methods: Basic Concepts Release 5000.8.3

Volume 1

HALLIBURTON | Landmark Software & Services

Part Number 200018F November 2013

Halliburton | Landmark Software & Services
2107 CityWest Blvd, Building 2, Houston, Texas 77042-3051, USA
P.O. Box 42806, Houston, Texas 77242, USA
Phone:713-839-2000
FAX: 713-839-2015
Internet: www.halliburton.com/landmark

Trademark Notice

Note

Third Party Licenses and Attributions

Halliburton acknowledges that certain third party code has been bundled with, or embedded in, its software. The licensors of this third party code, and the terms and conditions of their respective licenses, may be found in a document called Third_Party.pdf. This document is included in the installation directory of each application from Landmark® software. The path name of the file is similar to the following:

ApplicationInstallationDirectory`/docs/Third_Party.pdf`

or

ApplicationInstallationDirectory`\help\Third_Party.pdf`

Disclaimer

The programs and documentation may provide links to external web sites and access to content, products, and services from third parties. Halliburton is not responsible for the availability of, or any content provided on, third party web sites. You bear all risks associated with the use of such content. If you choose to purchase any products or services from a third party, the relationship is directly between you and the third party. Halliburton is not responsible for: (a) the quality of third party products or services; or (b) fulfilling any of the terms of the agreement with the third party, including delivery of products or services and warranty obligations related to purchased products or services. Halliburton is not responsible for any loss or damage of any sort that you may incur from dealing with any third party.

Stochastic Modeling and Geostatistics for Reservoirs; Principles and Methods: Basic Concepts

Chapter 1
Introduction

Landmark's DecisionSpace® Earth Modeling software offers a tool kit for the geoscientist and reservoir engineer to model lithofacies petrophysical properties, or any continuous property, in user-defined stratigraphic intervals within a sealed structural framework or onto 2D or 3D regular grids. The analysis and modeling tools are scale invariant, making the technology a powerful solution for not only reservoir scale modeling, but also for basin scale rock and property modeling. High resolution output models can be upscaled if required using DecisionSpace® utilities for input to dynamic flow simulators such as VIP®, Nexus®, or third party simulators.

DecisionSpace® Earth Modeling software utilities include:

- data analysis,
- variogram computation and modeling tools
- kriging, indicator kriging, and co-kriging estimation algorithms,
- general property modeling algorithms using regular 2D or 3D grids
- lithofacies and petrophysical simulation and co-simulation methods
- static volumetrics with ranking
- uncertainty analysis
- upscaling

The objective of this course is to familiarize you with the tools available in the context of completing 2D or 3D workflows. While this course attempts to provide background information on both classical statistical and geostatistical concepts, it is not a comprehensive guide to these sciences.

To meet these objectives, the course combines a series of lectures and discussions with computer exercises. You will learn how to:

- start the DecisionSpace® software
- create a new session
- load data into the session
- create a workflow
- perform Exploratory Data Analysis
- compute and model variograms
- create a map using Kriging
- integrate data using cokriging or cosimulation

- perform conditional simulation of facies and petrophysical properties
- build a sealed structural framework
- build a stratigraphic model
- calculate proportion curves and proportion maps for facies
- create a calibrated seismic attribute-to-lithology volume for facies modeling using Sequential Indicator Facies Simulation
- post process simulation results for uncertainty analysis
- compute static volumetrics and rank results

Audience

This course is designed for geologists, geophysicists, reservoir engineers, or geostatisticians who need statistical and geostatistical tools to analyze and integrate data derived from diverse sources, such as seismic surveys and well logs, to create high-resolution 3D geological models for well planning or as input to dynamic flow simulators. This manual assumes a basic knowledge of geological and geophysical terminology and a working familiarity of computers.

Organization

The manual and workshops are arranged in a logical order corresponding to a project workflow. However, not all of the software functionality is presented; but the workshops do provide a basic framework upon which you can continue to build. Different data sets are used throughout the course to illustrate various Earth Modeling tools.

At critical points, graphics are provided to illustrate how the screen should appear.

Mouse Button Terminology

In the exercises in this manual, we refer to the left mouse button as MB1, the middle mouse button as MB2, and the right mouse button as MB3.

Why Reservoir Geostatistics?

There remains a continuous need for improved reservoir characterization and new modeling tools—and using a geostatistical approach has its advantages. This section characterizes the problem and summarizes the benefits of using reservoir geostatistics as part of the characterization and modeling solution.

The Need for Improved Reservoir Characterization

- Oversimplified geological reservoir models most often lead to unrealistic production forecasts, resulting in poor reservoir management.
- Erroneous production forecasts are usually attributed to the inaccurate description of reservoir heterogeneities, resulting from too much reservoir connectivity in the geological model.
- Capital intensive projects, such as in deep water or “resource plays”, require more accurate reservoir descriptions earlier in the field development process.
- Improved reservoir characterization is also critical for mature fields where the efficiency of enhanced recovery processes is controlled by our ability to more accurately predict the channeling and breakthrough of injected fluids.

Why Use a Geostatistical Approach?

The reservoir is unique and completely deterministic. However, only sparse and incomplete information is available about the various reservoir properties.

The "hard" data typically comes from well logs, plug or core data, pressure tests, tracer surveys, and so on. Other “soft” data, including geophysical data are often abundant and provide very dense spatial coverage of the reservoir (e.g. 3-D seismic data), but the extracted seismic parameters are affected by many factors, and are only indirectly related to rock and fluid properties; often with limited vertical resolution.

Modeling these data using traditional interpolation methods, including kriging, tend to generate too smooth reservoir models, which underestimate reservoir heterogeneity, thus overestimate the connectivity.

As an alternative, geostatistical stochastic simulation methods attempt to reproduce the spatial variability and connectivity of reservoir properties and provide more realistic input models for fluid flow simulators. These methods effectively integrate both “hard” and “soft” data, capitalizing on the vertical resolution of the “hard” data and the dense spatial sampling of the “soft” data.

Stochastic methods provide alternative reservoir models; each honoring the input data. Post processing the alternative models aids in the quantification of geological uncertainty.

Advantages of Statistical Analysis for Reservoir Characterization

- Geostatistics allows one to analyze and model the spatial continuity and the directional characteristics of reservoir properties.
- Statistical tools are used to gain insight into complex data sets.
- Geostatistics provide the framework to integrate petrophysical, geological, geophysical and engineering data in the reservoir modeling process.
- Geostatistics helps to quantify uncertainty in the geological model, hence aids in risk assessment for more effective reservoir management.

Limitations of a Statistical Analysis

- Not magic or a panacea.
- Not a replacement for good data.
- Not a replacement for a thorough understanding and analysis of the data.

- Models must be interpreted and validated in light of reservoir geology, rock physics and reservoir engineering information and principles.

- Confidence limits calculated from stochastic simulation must be interpreted with caution; uncertainty in the basic data and hence the geostatistical model itself may be large.

Tools

The DecisionSpace® Earth Modeling software offers a number of task panes to help you apply your modeling approach to reservoir characterization:

- **Data Operations** allow you to create a pointset by importing points from a delimited or formatted flat file or by converting log curves that have been loaded into the session from the OpenWorks® database. You can also use the Data Manager to delete properties or subsets of Earth Modeling objects from memory. You can also transform the data distribution, add wells to a 3D grid, and filter data values using a subset of the pointset to create a new point set.

- **Stratigraphic Modeling** creates the stratigraphic layering and 3D geocellular grid, defines lithotypes and assign facies to the lithotypes, and blocks seismic volumes and well log data.

- **Facies Trend Modeling** is used to create proportion curves and proportion volumes, with or without seismic attributes, used in Facies Modeling and Simulation.

- **Variogram Modeling** is used to compute experimental variograms and model the variogram from discrete and continuous properties.

- **Facies Modeling and Simulation** is used to create stochastic facies models using Sequential Indicator, Truncated Gaussian or Plurigaussian Simulation methods.

- **Petrophysical Property Modeling** (typically following Facies Simulation) allows you to populate facies models with petrophysical data, such as porosity and permeability, by interval and facies, using a variety of methods.

- **General Property Modeling** allows you to create a simple 2D or 3D grid and then interpolate and simulate continuous properties on the grid or perform indicator Kriging (deterministic facies interpolation).

- **Post Processing** includes static volumetric computations, realization ranking, and uncertainty analysis.

In addition to the Earth Modeling task panes, you will also use:

- **Data Analysis** allows you to analyze your data using histograms, multi-histograms, cross-plots, Q-Q plots and/or box plots. A Subset Manager allows you to create subsets to be used in data analysis, Filter Point Sets in Data Operations and in some modeling steps.

- **Project Designer** allows you to design a project to guide you through the workflow and its Reporting task pane creates a report of your work on the project as an ODP document or an html file for use in your company intranet.

Input and Output Data

Using the DecisionSpace® Earth Modeling software, you can create pointsets, which are saved to the OpenWorks® database. These pointsets can be used as the input into any Earth Modeling task including creating a regular 2D or a 3D grid. The grid that is created is a .vdb file, which is a geocellular grid using the VDB storage format. VDB is the name of a database created by Landmark Graphics Corporation, mainly used by reservoir simulators and related software. The format is a directory tree structure consisting of multiple sub-directories and files, which hold data of different categories. A .vdb grid can be imported directly into the VIP® or Nexus® software for reservoir simulation. You can use the **Tools > Export 3D Grid** to output a VDB, Rescue, Resqml or an Eclipse grid.

You can load a 3D geocellular grid that was created externally and stored on your system. To load an external grid:

1. Select **Tools > Load 3D Grid** on the DecisionSpace menu bar.

2. In the Select a grid file or directory dialog, browse to and select from the vdb, bin, resqml, grdecl, roff, sg, grid, fgrid, egrid, or fegrid supported formats.

3. Click **OK** to load the grid and close the dialog.

When the grid is loaded, a PM Grid3D is automatically created and appears in your Inventory, the Select Session Data dialog box, and the Data Manager as the name of the file that was loaded. The grid will then appear in all drop-downs where you select an input grid.

Chapter 2
Exploratory Data Analysis

Before using any of the modeling tools, it is necessary for you to learn the fundamentals of Data Analysis tools. The goal of Data Analysis is to provide a general description and analysis of the data. Data analysis is an important precursor to the final modeling objectives, which may be interpolation or simulation, uncertainty assessment, and volumetrics. It is absolutely necessary to have a good understanding of the data, and the time taken in data analysis to quality control, explore, and describe the data set will be rewarded by much improved results.

The Data Analysis component includes univariate (Histogram, Multi-histogram, Box Plot, and Q-Q Plot) and bivariate (Cross-plot) displays to help you understand the data distributions and relationships within a sample data set. The Data Analysis task pane also accesses the Subset Manager allowing you to create subsets to be used in data analysis, the Filter Point Set in Data Operations, and in some subsequent modeling steps.

In this chapter you learn about:

- the fundamentals of classical data analysis
- the concept of the sample population
- univariate data analysis
- bivariate data analysis
- linear regression
- data transformations

Fundamentals and the Sample Population

Quite often the goal of a project is to provide a general description and analysis of a data set using classical statistical tools. This process is commonly referred to as Exploratory Data Analysis or EDA. EDA is an important precursor to the final goal of a reservoir modeling study, which may include interpolation, simulation, assessment of uncertainty, or volumetric computations. Unfortunately, in many studies, including "routine" mapping of attributes, the importance of EDA is often overlooked. It is absolutely necessary to have a good understanding of your data, and the time taken in EDA to quality control, explore and describe the data set will be rewarded by much improved results.

There is no one set of prescribed steps in EDA; you should follow your instincts in explaining anomalies in the data set. By using the various EDA tools, you not only gain a clearer understanding of your data, but also discover possible sources of errors. Errors can be easily overlooked, especially in large data sets and when computers are involved; we simply become detached from our data. Thorough EDA fosters an intimate knowledge of the data that can flag bogus data and lead to improved results.

The Dilemma of Modeling

The classic problems of forward modeling still exist.

- There are often very few direct "hard" observations (well data); they are spatially spare, but typically of high vertical resolution.
- The "Soft" data (seismic) are only indirectly related to the "hard" data; they are spatially dense, but the vertical resolution is limited.
- Few observations lead to much uncertainty.
- You still need to make predictions about the reservoir and its properties.
- Remember this: *"Garbage-In, Garbage-Out"*. Take the time to explore your data!

Sample Population

Each reservoir is unique and completely deterministic, and the reservoir, in statistical terms, is referred to as N, the sample universe. Because your sample data set (n; the sample population) represents a very small subset of reality (N), you will never know the true distribution function of any property N. Not only does your sample population provide you with a sparse and incomplete picture of the reservoir, your sampling program (drilling wells) is highly biased, and rightly so. However, despite these limitations, your task is to infer properties describing the entire reservoir from your sample data set. In other words, you need to infer the properties of N from n. To accomplish this you need to use various statistics tools to understand and summarize the properties of n, the sample population.

The EDA component provides a variety of statistical tools to analyze your data sets. In the following pages basic concepts behind the statistical methods are reviewed; followed by computer workshops to re-enforce these concepts. There are two general categories of tools: univariate and bivariate.

Univariate Data Analysis

There are several ways in which to summarize a univariate (single attribute) data distribution. Quite often one simply computes the mean and the variance, or plots its histogram. However, these statistics are very sensitive to extreme values (outliers) and also do not provide any spatial information about the data, which is the heart of a geostatistical study.

The EDA component uses three types of univariate data analysis tools:

- Histograms
- Box Plots
- Q-Q Plots

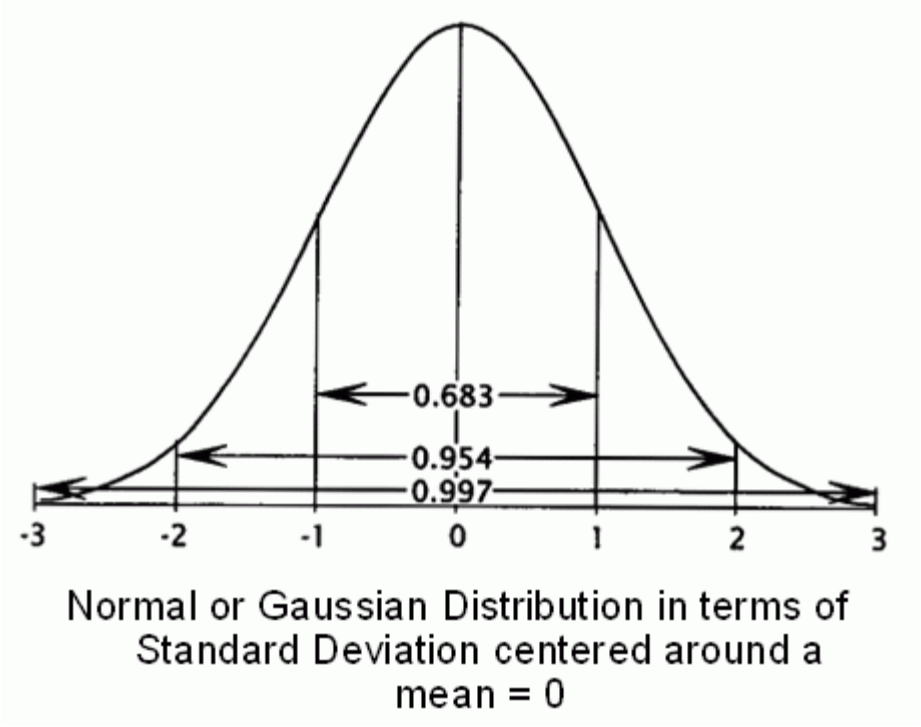
Histograms

A *histogram* is a graphical representation of a frequency table and records how often data values fall within certain intervals or classes. It is common to use a constant class width for a histogram so that the height of each bar is proportional to the number of values within that class. Data is often ranked in ascending order, thus can be represented as a cumulative frequency histogram, where the total number of values below certain cutoffs is shown, rather than the total number of values in each class. The following figure shows a histogram of a normal, or Gaussian, distribution, graphically illustrating the terms defined below.

The Normal distribution is the ideal distribution because if we know the mean and the standard deviation, its shape is perfectly predictable. Most of the statistical tests assume that the data have a normal distribution. If not, transform the data, temporarily, to perform certain tests and procedures.

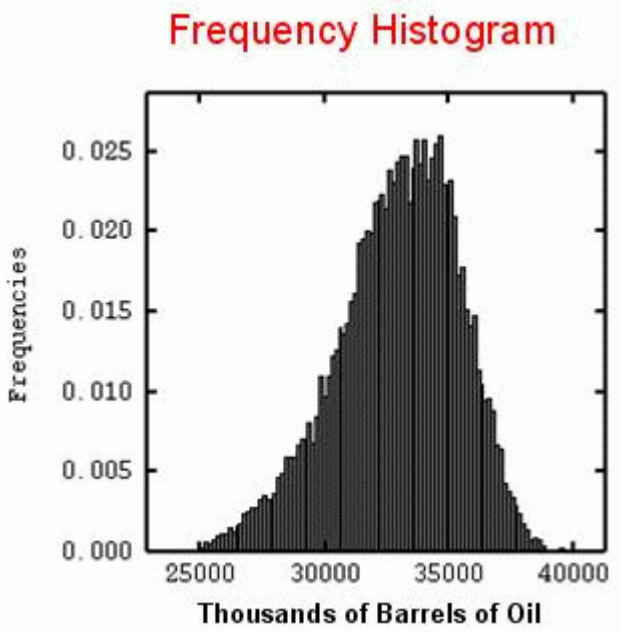
For a Normal distribution, the area under the standard normal curve is:

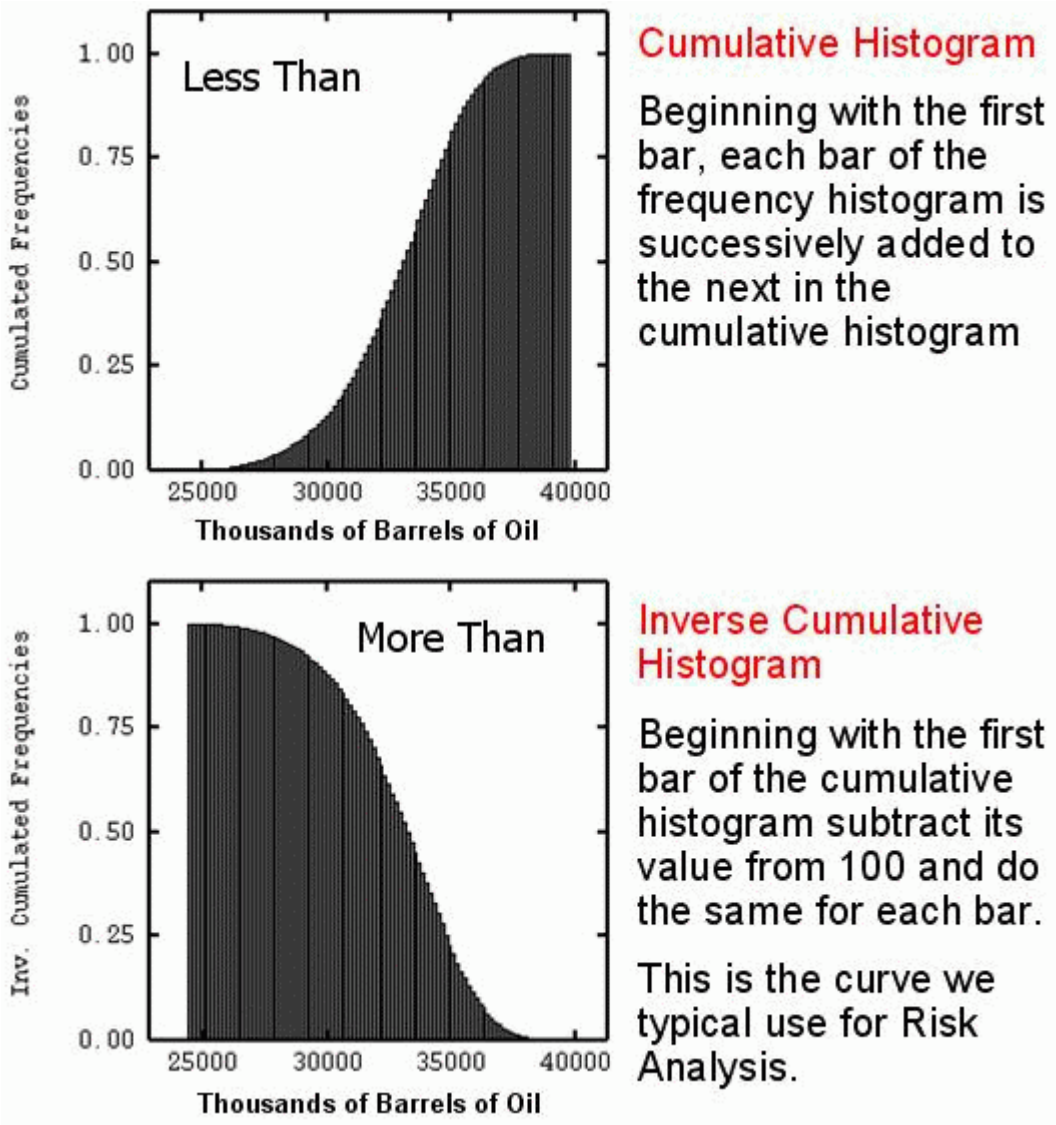
- 68.3% of the data lie within one standard deviation of the mean
- 95.4% of the data lie within two standard deviations of the mean
- 99.7% of the data lie within three standard deviations of the mean



Normal or Gaussian Distribution in terms of Standard Deviation centered around a mean = 0

The histogram can be drawn in three ways, as a frequency distribution, a cumulative, or an inverse cumulative distribution and illustrated below.





The summary statistics represented by the histogram can be grouped into three categories:

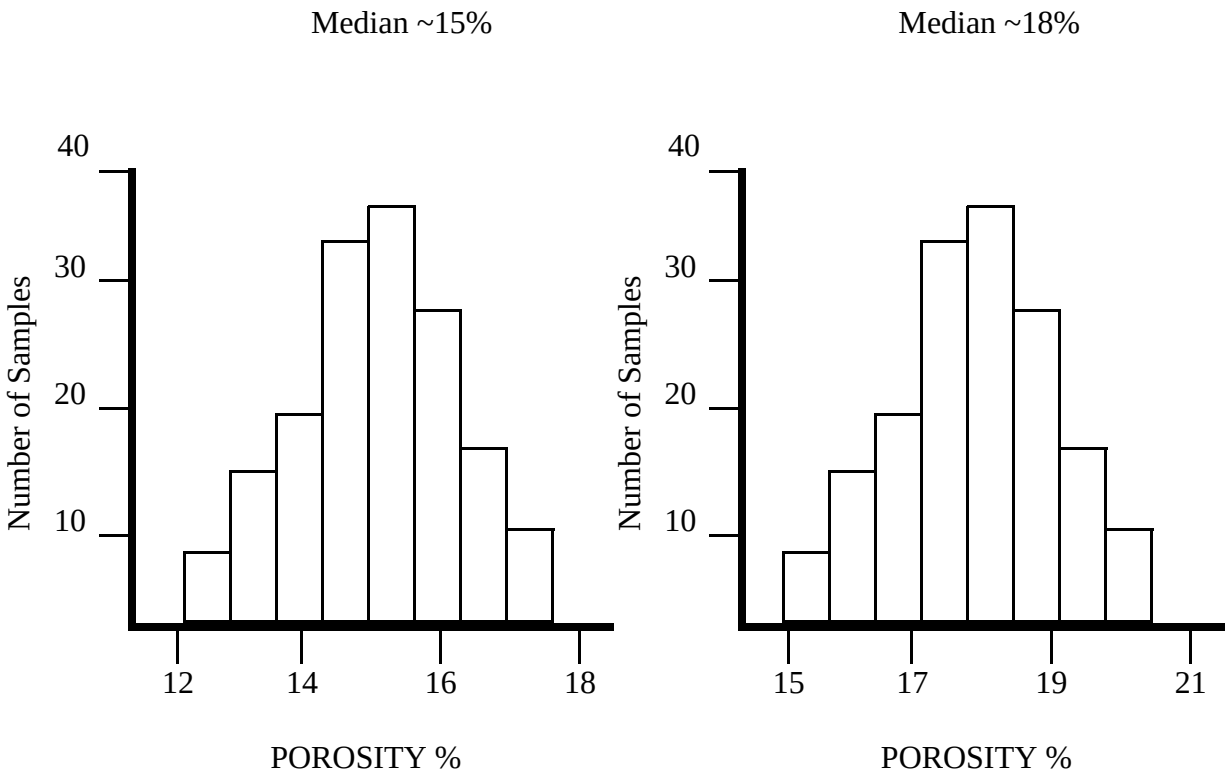
- **measures of location (mean, mode, or median)**
- **measures of spread (variance or standard deviation)**
- **measures of shape (skewness)**

A fourth statistic to consider is the *Z-score*, which can be an indicator of outliers or spurious data.

The following sections discuss these measurements.

Measures of Location

Two data sets may seem to have nearly identical distributions; they may have the same shape and the same spread, and the only thing that is different is the median. The diagram below illustrates this.



Measures of location provide information about where the various parts of the data distribution lie, and are represented by the following:

- **Minimum**— Smallest value in the data set.
- **Maximum**— Largest value in the data set.
- **Mean**— The arithmetic average of all data values. This statistic is quite sensitive to extreme high or low values. A single erratic value can unfairly bias the mean.

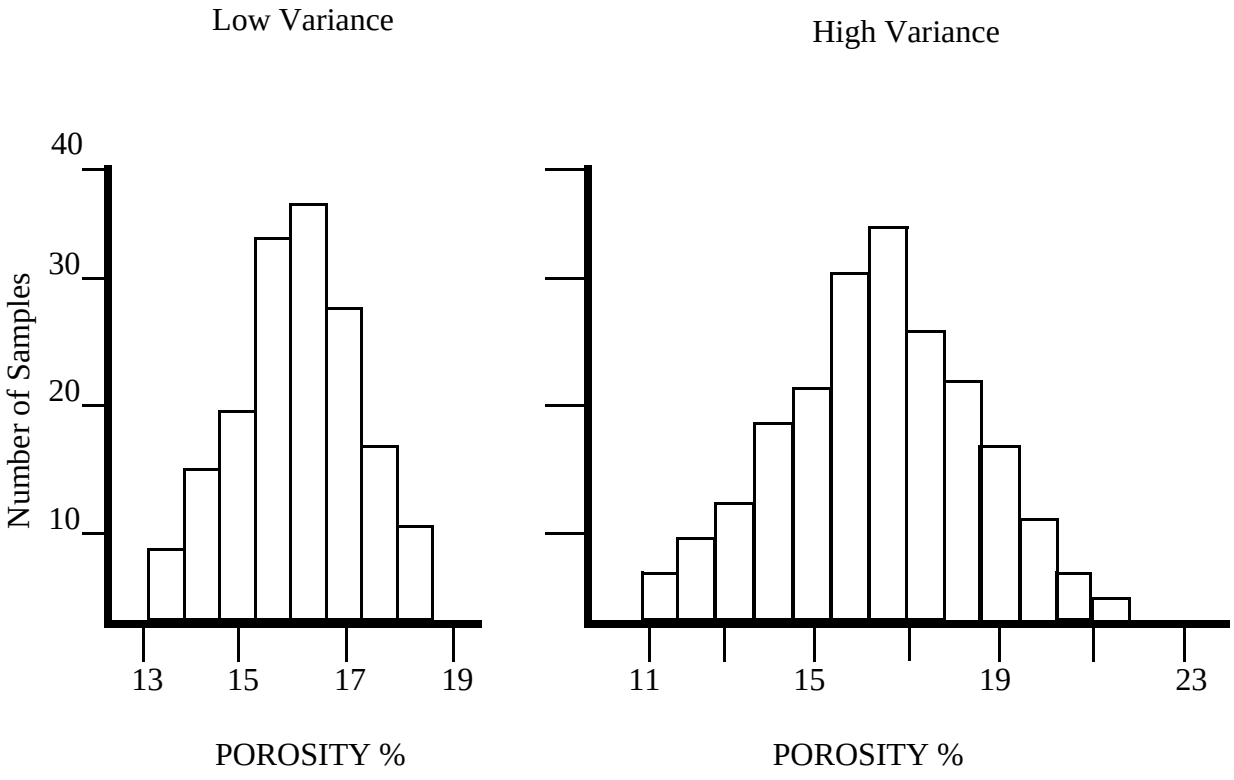
$$Mean = m = (1/n)\sum Z_i$$

- **Median**— Midpoint of all observed data values, when arranged in ascending order. Half the values are above the median, and half are below. This statistic represents the 50th percentile of the cumulative frequency histogram and is not generally affected by an occasional erratic data point.

- **Quartiles**— In the same way that the median splits the data into halves, the quartiles split the data into quarters. Quartiles represent the 25th, 50th and 75th percentiles on the cumulative frequency histogram.

Measures of Spread

Two data sets may seem to have nearly identical distributions; they may have the same shape and the same center, and the only thing that is different is the measure of spread (variance or standard deviation). The diagram below illustrates this.



Measures of spread describe the variability of the data values, and are represented by the following:

- **Variance**— Average squared difference of the observed values from the mean. Because the variance involves squared differences, this statistic is very sensitive to abnormally high/low values.

$$Variance = \sigma^2 = \frac{1}{n}\sum (Z_i - m)^2$$

  In the case of small data sets (<30), it is recommended to divide by n-1 when computing the variance to correct for small sample statistics.

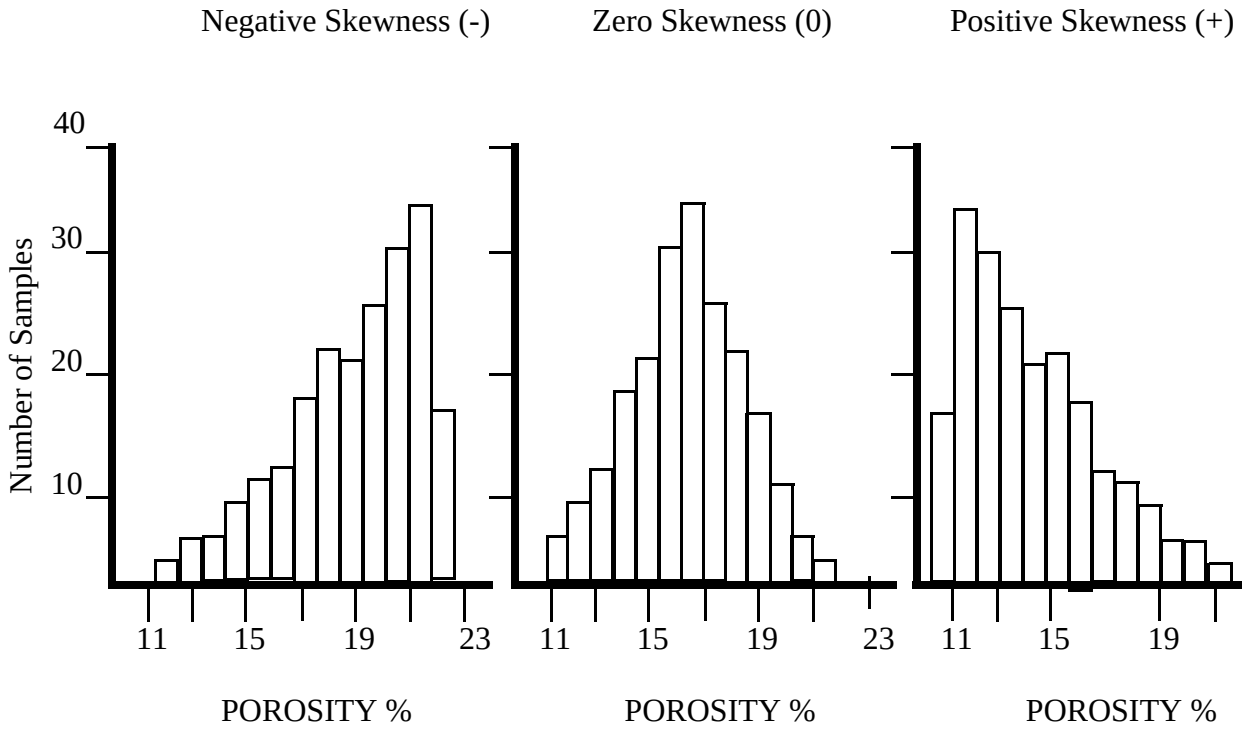
- **Standard Deviation**— Square root of the variance. It is often used instead of the variance because the units are the same as the units of the attribute being described.

$$StandardDeviation = \sigma = \sqrt{\sigma^2}$$

- **Interquartile Range**— Difference between the upper (75th percentile) and the lower (25th percentile) quartile. Because this measure does not use the mean as the center of distribution, it is less sensitive to abnormally high/low values.

Measures of Shape- Skewness

Data sets can have different distributions; but they may have similar centers and similar spreads, and the only thing that is different is the shape of the distribution. The diagram below illustrates this.



Measures of shape describe the appearance of the histogram and are represented by the following:

- **Coefficient of Skewness**—Averaged cubed difference between the data values and the mean, divided by the cubed root of the standard deviation. Very sensitive to abnormally high/low values:

$$CS = \frac{\frac{1}{n}\sum (Z_i - m)^3}{\sigma^3}$$

  - **positive:** long tail of high values (median < mean)
  - **negative:** long tail of low values (median > mean)
  - **zero:** a symmetrical distribution (mean = median)

In the case of small sample data sets (<30), it is recommended to divide by n-1 when computing the skewness to correct for small sample statistics.

Outliers or "Spurious" Data (Standardize)

Another statistic to consider is the Standardized score (Z-score), which is a summary statistic in terms of standard deviation. Data which "appear" to be anomalous based on its Standardized score (values greater than a specified cutoff) are termed outliers. The typical cutoff is 2.5 standard deviations from the mean. The formula is the ratio of the data value minus the sample mean to the sample standard deviation.

$$Z_{score} = \frac{(Z_i - m)}{\sigma}$$

This statistic serves as a caution, or red flag, signifying potentially bad data, or a true local anomaly, which must be taken into account in the final analysis.

Note: Standardization does not change the shape of the input histogram. It only rescales the x-axis into terms of standard deviation.

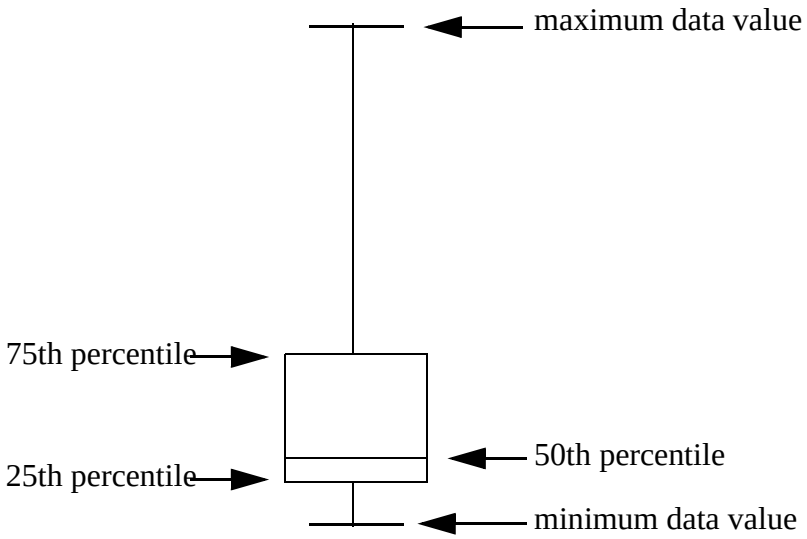
Box Plots

Box Plots (sometimes referred to as box and whisker plots) are useful for displaying a variable's location and spread at a glance; it is kind of a simplified histogram. They also provide some indication of the data's symmetry (skewness). By using a box plot for each data variable side-by-side on the same graph, you can quickly compare data sets. Note, however, that box plots do tend to emphasize the tails of a distribution, which are the least certain points in the data set. They also hide many of the details of the distribution. Therefore, we recommend displaying a histogram in conjunction with a box plot.

A box plot summarizes the following statistical measures:

- Median (second quartile or 50th percentile)
- Upper (75th percentile) and lower (25th percentile) quartiles
- Minimum and maximum data values
- Potential outliers (if any; displayed as red open circles) when the outlier option is applied. The vertical line extends to a maximum of

1.5 times the inter-quartile range. Any points beyond the ends of the vertical line are potential outliers based on statistical analysis.



Q-Q Plot

The Q-Q (quantile-quantile) plot is a graphical technique for determining if a data set follows a known distribution. By a quantile, we mean the fraction (or percentage) of data samples falling below a given quantile level. For example, the 0.3 (or 30%) quantile is the point at which 30% percent of the data fall below and 70% fall above that quantile level as displayed with a cumulative histogram. The quantiles from two distributions are plotted versus one another. A Q-Q plot of two identical distributions plots as a 45-degree straight line.

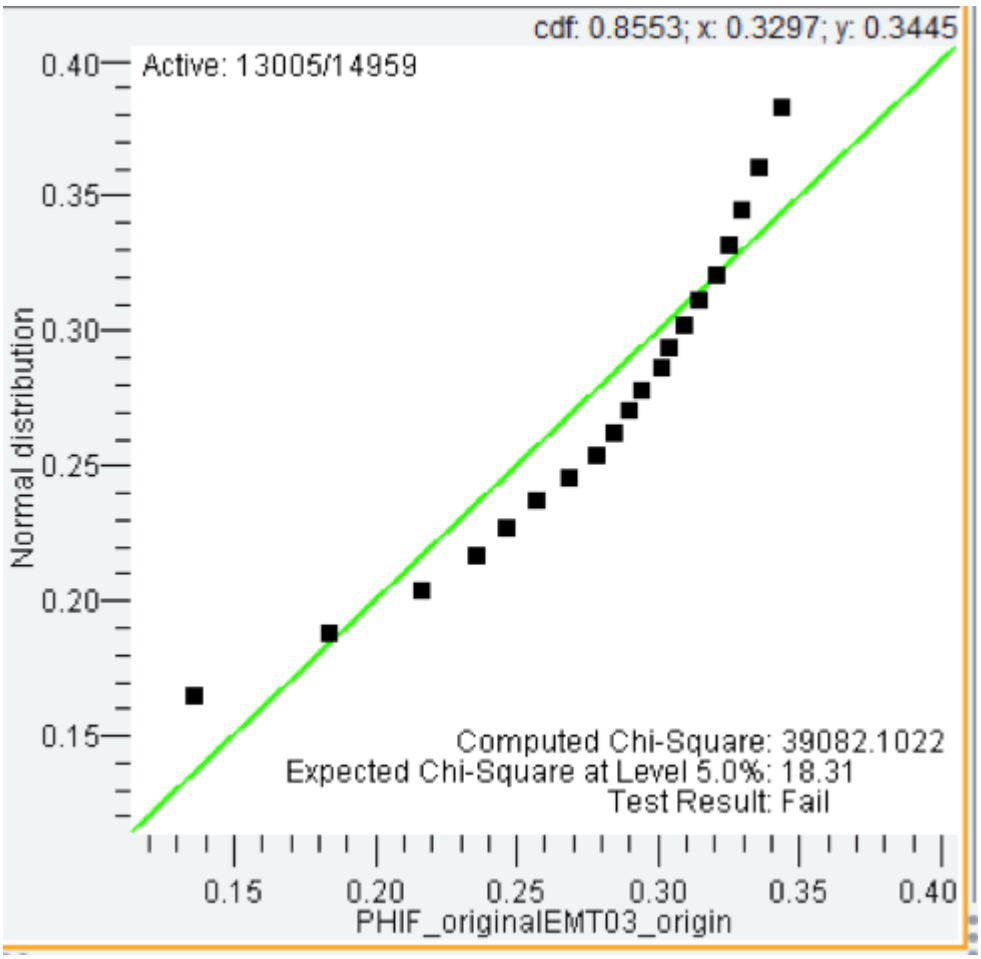
The most common use of the Q-Q plot is to test whether the data follow a Normal (Gaussian) distribution. Because geostatistical simulation algorithms such as Sequential Gaussian and Turning Bands assume that the data are normally distributed, the data must be transformed using a Normal Score Transform, prior to simulation, if the assumption of Normality is not met.

Similarly, Q-Q plots are used to compare data distributions against other known distributions, such as the lognormal distribution, to determine the similarity. Because histograms do not always reveal subtle differences between distributions, the Q-Q plot can be used to obtain a good visual and statistical comparison. The Q-Q plot is essentially a cross-plot of matching quantile values from two distributions.

A Q-Q plot is formed by:

- the X-axis (horizontal), which displays the quantile values from the data set to be tested (the unknown distribution or test data), and
- the Y-axis (vertical), which displays the quantile values from the known distribution (the theoretical Normal distribution, for example).

Both axes are in units of the test data set. That is, the actual quantile level is not plotted, but rather the quantile value. For a given point on the Q-Q plot, we know that the quantile level is the same for both points, but values at that quantile level differ.



Porosity distribution compared to its theoretical normal distribution; it failed the Chi^2 test. The display shows a cross plot of the distribution values at 20 quantiles (e.g., at the a 5th, 10th, 15th, etc.)

A limitation of the Q-Q plot is that it is only a graphical indication of the "Goodness-of-Fit" and only a visual inspection could lead to a wrong

conclusion about whether the distributions are similar or not. Therefore, the Chi-Square test is used to test statistically if the distributions are the same.

Besides the quantile values, a 45-degree reference line is also plotted on the graph. If the unknown data distribution follows the known (theoretical) distribution, the points should fall on this reference line. The greater the departure from this reference line, the greater the evidence for the conclusion that the unknown distribution does not have a Normal distribution, for example.

The following are a few guidelines to remember when interpreting a Q-Q plot:

- A systematic departure above or below the 45-degree line indicates that the center or mean of the distributions is different.
    - A shift above implies that the Y-distribution is higher valued than the X-distribution.
    - A shift below implies that the X-values are higher.
- A slope other than 45 degrees indicates that the spread or variance of the two distributions is different.
    - A slope greater than 1 (> 45 degrees) indicates that the Y-variance is higher than X.
    - A slope less than 1 (< 45 degrees) indicates that the X-variance is greater.
- Curvature indicates that the two distributions have different shapes.
- All but a few points fall on a line indicates outliers in the data.
- Left end of pattern is below the line and right end of pattern is above the line indicates long tails at both ends of the data distribution.
- Left end of pattern is above the line and right end of pattern is below the line indicates short tails at both ends of the distribution.
- Curved pattern with slope increasing from left to right indicates that the data distribution is skewed to the right.

- Curved pattern with slope decreasing from left to right indicates that the data distribution is skewed to the left.

- Staircase pattern (plateaus and gaps) indicates that the data have been rounded or are discrete.

Q-Q plots are used to determine if a data set corresponds to a known distribution type (Normal or Lognormal).

The experimental Chi Square (computed from the data) and Expected Chi Square (critical value of Chi Square determined from the degrees of freedom) for the selected Significance Level (the default is 5%) are shown in the lower right corner of the display. It also indicates whether the Chi Square test passed or failed. In probability theory and statistics, a "Chi Square distribution" is one of the most widely used theoretical probability distributions in inferential statistics, i.e. in statistical significance tests. This is because, under reasonable assumptions, easily calculated quantities can be proven to have distributions that approximate to the Chi Square distribution if the null hypothesis is true, i.e., the data distribution is the same as the distribution tested (Normal or Lognormal). The Chi Square value indicates whether the distributions differ from one another. If the computed Chi Square value is less than (<) the Expected Chi-Square, then there is no reason to reject the Null Hypothesis, and we accept that the distributions are the same and the Test result indicates PASS. If the computed Chi Square value is greater than (>) the Expected Chi Square, then there is a statistical basis to suggest that the distributions are different and the Test result indicates FAIL. At the 5% significance level, there is only a 5% chance that the Null Hypothesis is in fact false.

Summary—Univariate Statistics

Advantages

- are easy to calculate,
- provide information in a very condensed form, and
- can be used as parameters of a distribution model (e.g., normal distribution defined by sample mean and variance).

Limitations

- Summary statistics are often too condensed and do not carry enough information about the shape of the distribution.
- Certain statistics are sensitive to abnormally high/low values that properly belong to the data set (e.g., mean, variances, standard deviation, Z-Score).
- Univariate summary statistics offer only a limited description, especially if your real interest is in a bivariate data set (e.g., attributes are correlated).
- Univariate summary statistics contain no information about the spatial characteristics of the data set.

Bivariate Data Analysis

Methods for bivariate description not only provide a means to describe the relationship between two variables, the *dependent* variable (the one to be estimated) and the *independent* variable (the estimator), but also provide the foundation for the tools used to analyze the spatial content of a random function (to be described in the Variography section). The bivariate method described in this section measures only the linear or non-linear relationship between the two variables, not their spatial features.

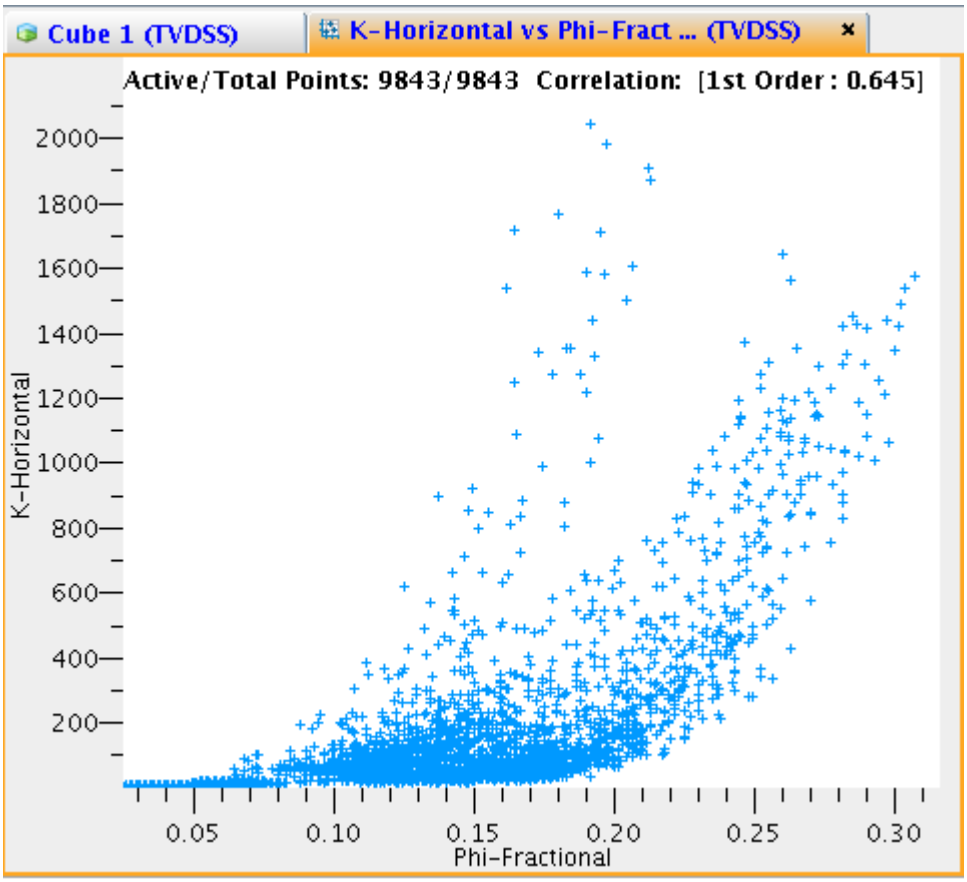
The bivariate data analysis used is the cross-plot.

Cross Plot

The cross plot is the most common bivariate plot. The independent variable (e.g., porosity) is plotted on the X-axis (ordinate) with the dependent variable (e.g., permeability) plotted on the Y-axis (abscissa). This type of plot serves many purposes:

- detect a linear or non-linear relationship,
- positive or inverse relationship
- potential outliers
- overall data quality control check.

A cross plot display should be created before calculating bivariate summary statistics such as the correlation coefficient.



Covariance

Covariance is a statistic measuring the correlation between all points of two variables (e.g., porosity and permeability as shown above). The magnitude of the covariance statistic is dependent upon the magnitude of the two variables.

$$COV_{x,y} = \frac{1}{n}\sum(X_i - m_x)(Y_i - m_y)$$

The covariance is strongly affected by extreme values in data pairs (outliers). This statistic also forms the foundation for the variogram, which measures spatial variability.

In the case of small sample data sets (<30), it is recommended to divide by n-1 when computing the covariance to correct for small sample statistics.

Correlation Coefficient

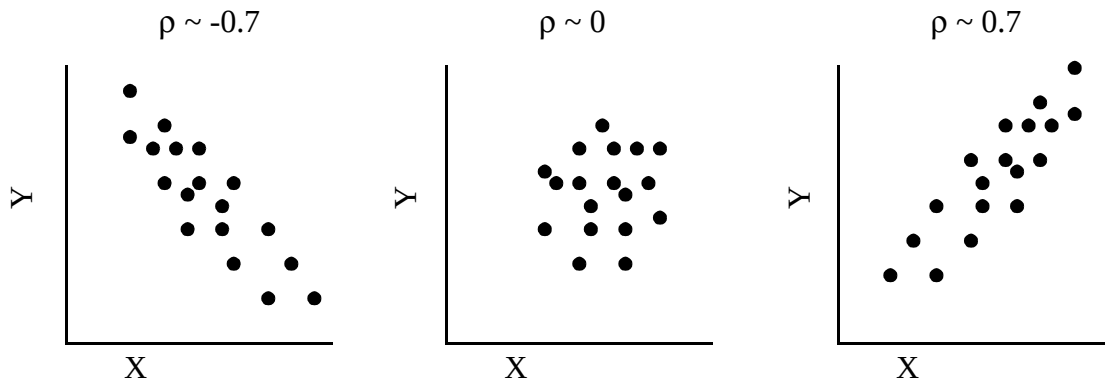
The correlation coefficient is a statistic measuring the linear relation between all points of two variables. For example, in the cross-plot shown above, we plotted porosity on the X axis and permeability on the Y axis. We often use a logarithmic scale for KH (permeability) to linearize its relationship with PHI (porosity)).

Correlation Coefficient values range between +1 (perfect, positive correlation) to -1 (perfect inverse correlation). A value of zero is no correlation or a totally random relation. The numerator for the correlation coefficient is the covariance. This value is divided by the product of the standard deviations for variables X and Y. This normalizes the covariance, thus removing the impact of the magnitude of the data values. Like the covariance, outliers adversely affect the correlation coefficient.

$$CorrCoeff_{x,y} = \frac{\frac{1}{n}\sum (X_i - m_x)(Y_i - m_y)}{\sigma_x \sigma_y} = \rho_{x,y}$$

The Coefficient of Determination is the square of the correlation coefficient, is referred to as R-Squared (R^2), and is a measure of the variance accounted for in the linear relation.

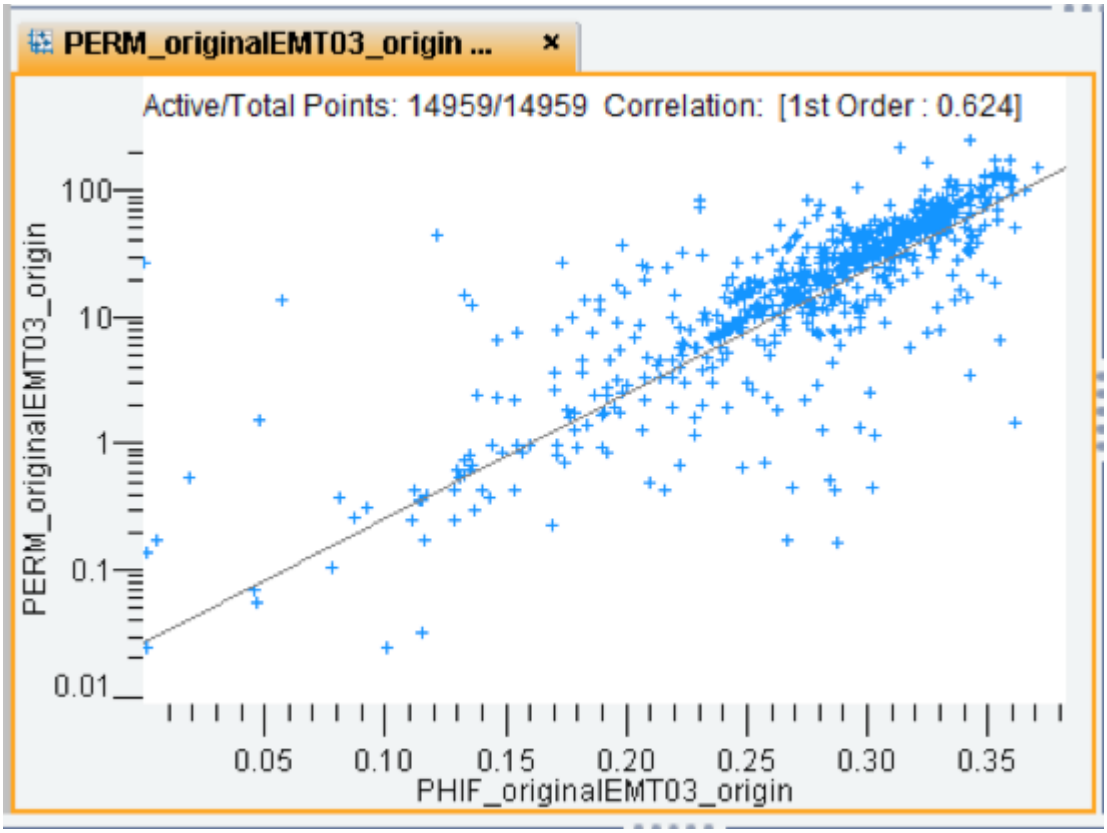
The Correlation Coefficient value of 0.645 between porosity and permeability means that as porosity increases in value, permeability also increases, which has real physical meaning. However, the Coefficient of Determination shows that only about 41.6% of the variability in permeability can be explained by its relation to porosity.



This figure illustrates three cross plots with the correlation coefficients between X and Y. The left-most cross plot shows a strong inverse relationship, the center cross plot shows no correlation, whereas the right-most cross plot shows a strong direct correlation between X and Y.

Linear Regression

Linear regression helps to investigate relationships between two variables and creates an equation to estimate the value of one variable (e.g.permeability) from another variable (e.g. porosity). When two variables have a high covariance (strong correlation), it is possible to predict a linear relationship between the two. The steepness and direction of the regression slope can be represented as a straight line.



Cross plot of PHIF versus PERM (on a log scale) illustrating the first-order linear regression whose equation is

PERM(log10) = -1.566+9.831(PHIF) and a correlation = 0.624

The Mathematical Form of Linear Regression

The regression equation has the following general form:

$Y = a+bX$

Where:

- Y is the dependent variable, or the variable to be estimated (e.g., log permeability);
- X is the independent variable, or the estimator (e.g., porosity);
- *b* is the slope; defined as $b= \rho(\sigma y/\sigma x)$; and
- *a* is a Constant, defined as $a = m_x - bm_y$ *and m = arithmetic mean* and *a* defines the ordinate (Y-axis) intercept.

Linear equations can include polynomials of any degree, and also may include combinations of logarithmic, exponential or any other non-linear variables.

The terms in the equation for which coefficients are computed are independent terms, and can be simple (a single variable) or compound (several variables multiplied together). It is also common to use cross terms (the interaction between X and Y), or use power terms.

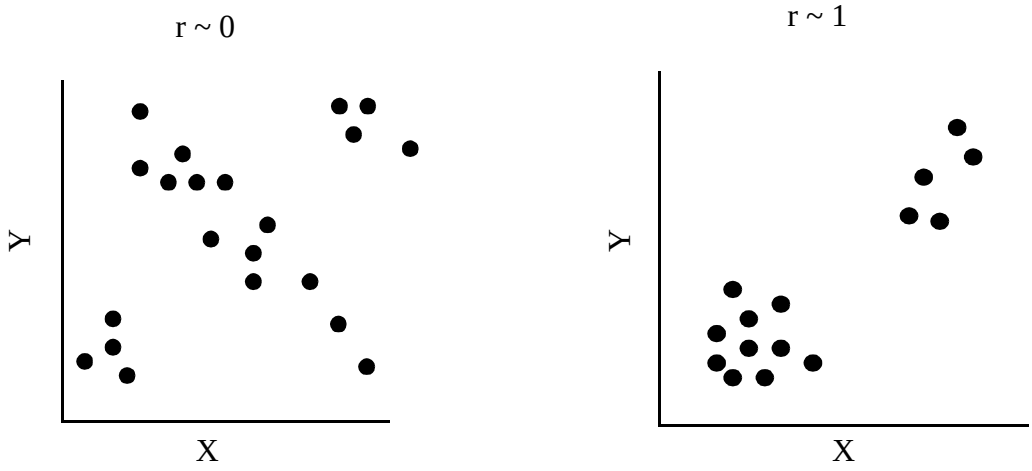
- $Z = a + bX + cY$: uses X and Y as predictors and a constant
- $Z = a + bX + cY + dXY$: adds the *cross* term
- $Z = a + bX + cY + dXY + eX^2 + fY^2$: adds the power terms

Detrending Data with Least Square Polynomial Fit

Geostatistical spatial modeling assumes that there is no trend in the data. If a trend is present, it can be removed using Least Square polynomials. The spatial model is then computed on the residuals after the trend is removed from the raw data. After modeling, the same trend may be re-introduced. Detrending is explained in more detail in the discussion of non-stationarity.

"False" Correlation

Be careful when studying correlation values as you may see a "false" positive, or correlation that occurs by chance. Some correlations may appear good when they are not, some may appear bad when they are very good. The problem will arise when there is more than one population, or in the presence of outliers.



Two examples of misleading correlations: the example on the left could be the result of multiple data populations where the middle grouping might have a strong inverse correlation. The example on the right shows a nearly perfect correlation only because the correlation is between the two data clusters, which could represent two data population or a sampling bias where there are no samples values between the clusters.

Summary—Bivariate Statistics

Advantages

- They are easy to calculate.
- They provide information in a very condensed form.
- Linear regression can be used to estimate one variable from another variable or from multiple variables.

Limitations

- Summary statistics are often too condensed and do not carry enough information about the shape of the distribution.
- Certain statistics are sensitive to abnormally high/low values that properly belong to the data set (e.g., covariance, correlation coefficient). Outliers can highly bias a regression predication equation.
- They provide no spatial information about the data.

Data Transformations

The Data Transform panel on the Data Operations task pane of Earth Modeling allows you to transform the data distribution of the selected input data to another type of distribution; the primary purpose being to approximate a normal distribution. This transformed data is then available in Property Modeling. The coefficients of the transform are saved to allow the simulated results of continuous property data to be back-transformed automatically following the property modeling.

If a method is selected that requires Normal Score data, such as Sequential Gaussian or Turning Banks simulation, the data are transformed automatically if not done so before.

There are numerous transform types but only a few are relevant for reservoir modeling.

Transform Type allows you to select from the following types (note that this list will be filtered so that only those transform types that can be used for selected Property are shown; for example, if the data has been transformed, only the Back_[transformtype] option will be listed):

- **Normal Scor**e transforms the input data distribution into a normal (Gaussian) distribution, which is required when performing simulations. The sill (variance) for normal score transformed data is always one, with the mean value = 0, a standard normal distribution (N0,1).

  Normal Score transformed data should never be used when doing Interpolation in Property Modeling.

- **Logn** (natural log) transforms the input data distribution into a logarithm to the base n, where n is an irrational constant approximately equal to 2.718281818459, or in other words the power to which a number x would have to be raised to equal x.

- **Log10** transforms the input data distribution into a logarithm to the base 10 distribution.

  Log transforms are typically used to look for permeability/porosity relationships when comparing them in cross plots. Because porosity data by nature follows a more normal distribution and permeability data does not, comparing log transformed permeability data against untransformed porosity data in a cross plot should show a more

linear relationship. Log transforms are used in kriging or interpolation, when it is not appropriate to use normal score transformed data.

- **Standardize**, sometimes called Z-score conversion, rescales the X-axis to a mean of zero and a standard deviation of 1. The resulting histogram will look like the histogram of the original data, the only difference being the X-axis values. A cross plot of the standardized data versus the original data will show a perfect correlation. Standardized data are useful in looking for outliers.

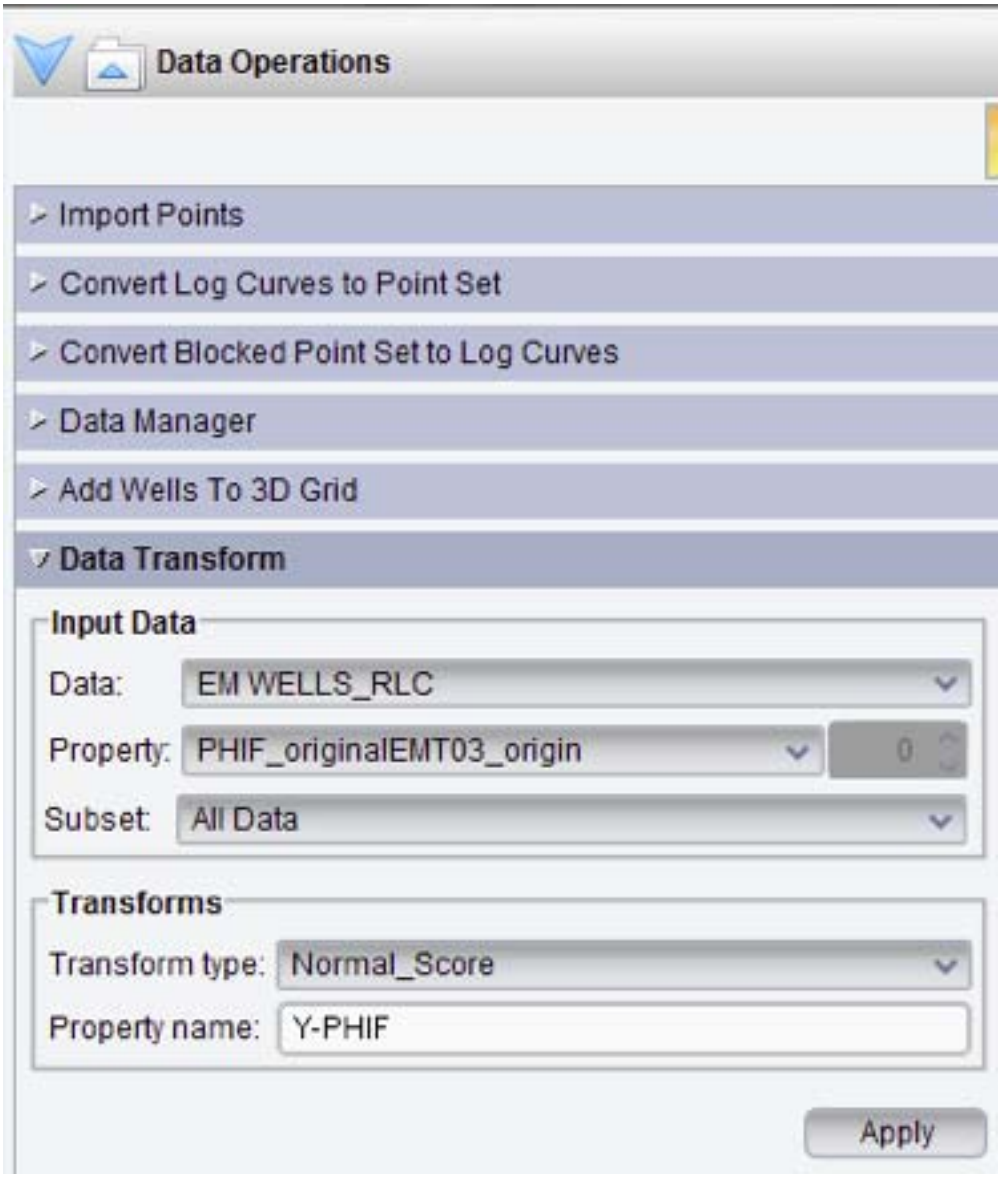
- **Range** rescales the data from 0 to 1. This transform maintains the property relationships and rescales the variances. The output property can be used to grid using interpolation algorithms, particularly in collocated cokriging where the variance ratios could get extreme.

- **Back_Normal_Score** is used to return data that has been transformed to a Normal_Score distribution back to its original distribution.

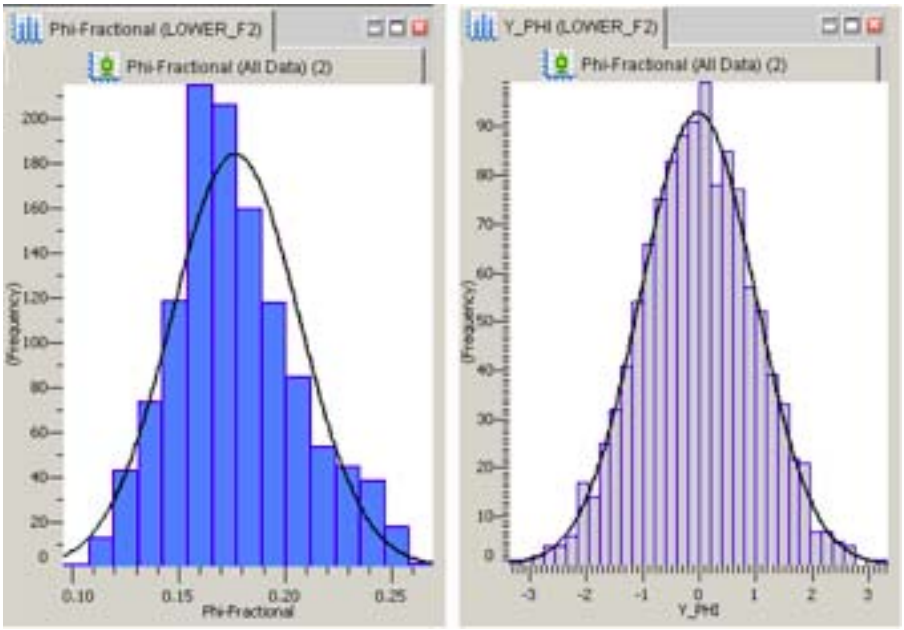
- **Back_Logn** is used to return data that has been transformed to a Logn distribution back to its original distribution.

- **Back_Log10** is used to return data that has been transformed to a Log10 distribution back to its original distribution.

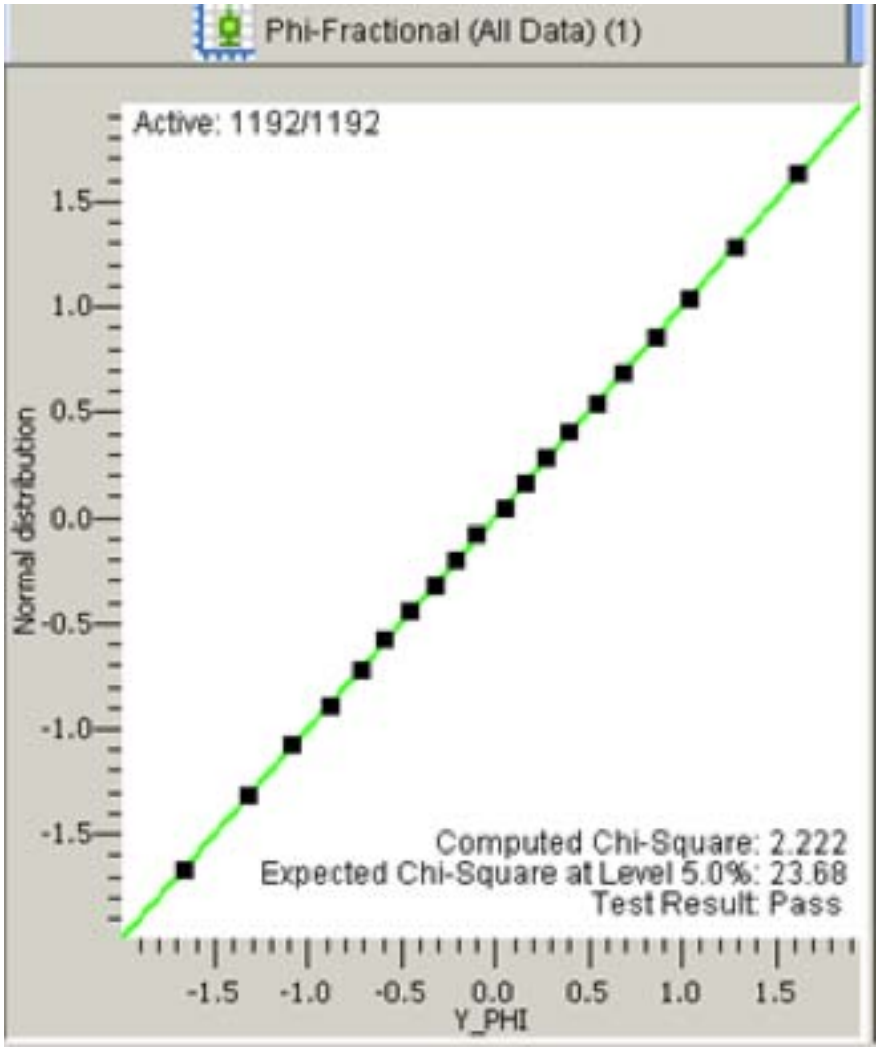
- **Back_Standardize** is used to return data that has been transformed to a Standardize distribution back to its original distribution.



Normal Score Transform example setup



This figure illustrates the use of a Normal Score transform on porosity (left). The frequency histogram has a slight right-skew, when compared to its theoretical normal distribution represented by the black line curve. The histogram on the right shows the porosity following the transformation. Notice that the X-axis is in terms of standard deviation centered about the mean = 0.



We see that the transformed data passes the Chi2 test when using the Q-Q plot to test against a normal distribution.

Chapter 3
Variography

Variography is used to explore, calculate, and model the spatial content of the continuous and discrete properties in a data set. The purpose is to quantify the rate of change of reservoir properties with distance and direction. Variography is often one of the most important and misunderstood steps in a reservoir study because it is used to determine the matrix of weights used in the kriging and conditional (stochastic) simulation algorithms.

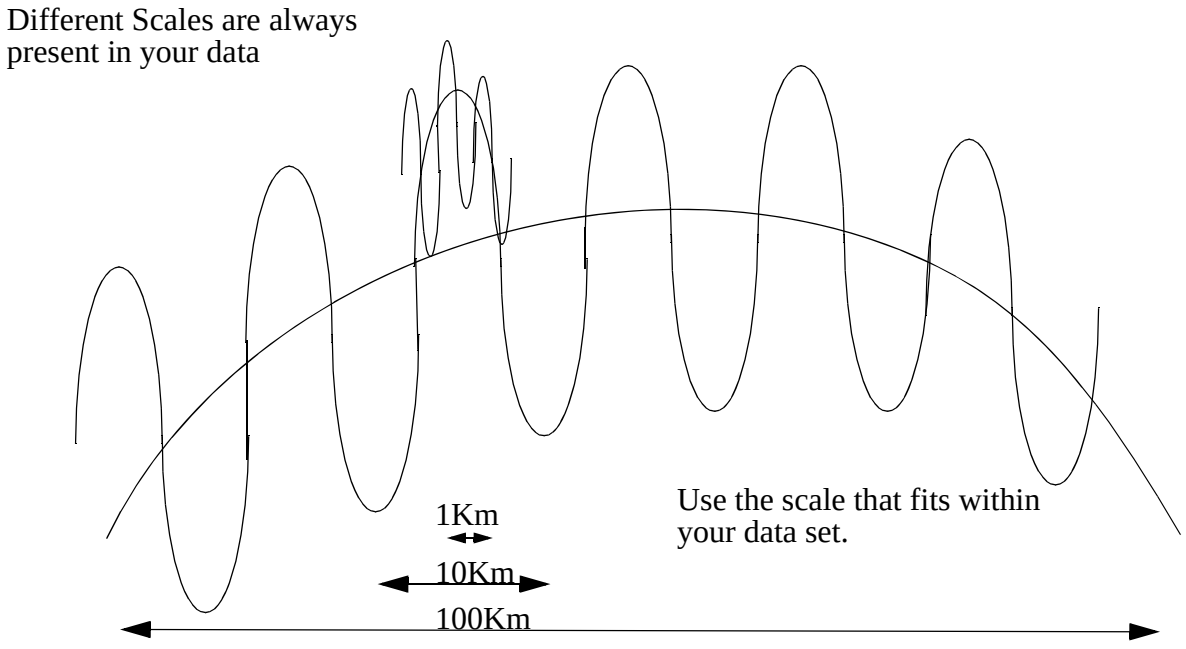
The Variography task panel allows you to compute and model, in 2D or 3D, the spatial continuity and directionality in the data using the traditional semi-variogram (for simplicity called a variogram).

In this chapter you learn about:

- the concepts of spatially correlated data
- variograms
- variogram polar plots, azimuth, lag, and bandwidth of data searches
- types of variogram model functions
- stationarity

Understanding the Spatially Correlated Variable

The variables of interest in the petroleum industry (e.g., porosity, permeability, sand/shale volumes and connectivity, etc.) are the product of a vast number of complex physical and chemical processes. These processes superimpose a spatial pattern on the reservoir rock properties, and it is important to understand the scales and directional aspects of these features for efficient hydrocarbon production.



Because the spatial component makes these variables complex, we are forced to admit that there is uncertainty about the behavior of the attributes between wells. Deterministic models do not adequately handle the uncertainties associated with these variables, so a geostatistical approach is used because it is based not only on deterministic methods, but also on probabilistic algorithms (e.g., variogram models when used with simulation methods) that recognize these inevitable uncertainties.

The spatially correlated variable and its mathematical expression form the foundation of geostatistics. In the previous section, we discussed classical methods for the analysis of single variables or multiple variables. Those methods, however, cannot properly address the spatial continuity and directionality inherent in earth science data. What we require is a model describing the continuity, anisotropy and azimuthal properties within the measured data.

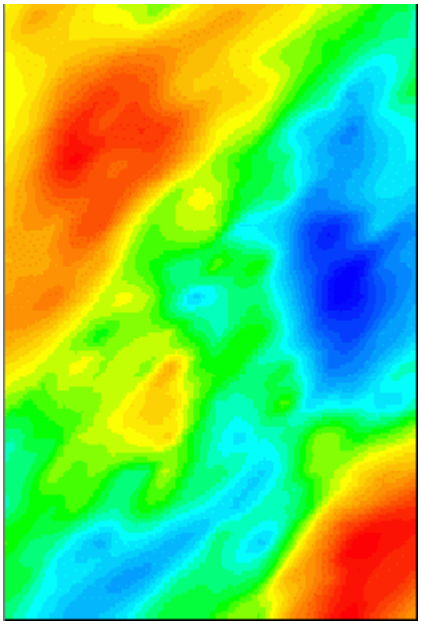
The Basis of Spatial Correlation

The purpose of spatial correlation analysis, or variography, is to quantify the rate of change of reservoir properties with distance and direction. Two wells in close proximity will likely have similar reservoir properties values than from two wells further apart. What we need to determine is how far wells can be separated and still yield similar results. To do this, we need a new statistical measure, because spatial correlation cannot be captured with classical univariate or bivariate statistics

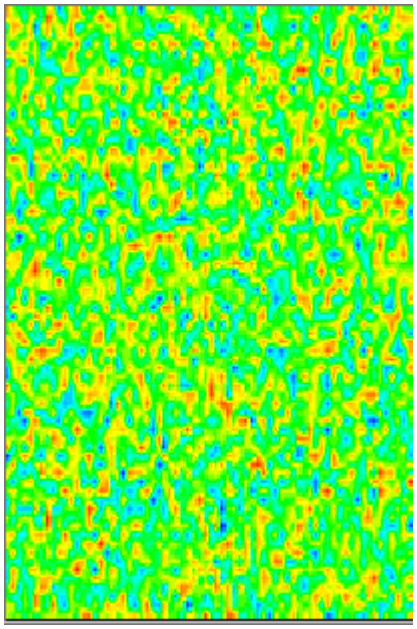
Spatial correlation analysis is often one of the most important steps in a reservoir study because it provides the weights required by the kriging system and conditional simulation.

Regionalized Variable Defined

One data set can have exactly the same univariate statistics as another data set, yet have a very different spatial organization. Consider the following two grids. The property distribution for Porosity A has a distinct orientation, with high continuity in the NE-SW direction and lower continuity NW-SE. Porosity B can be described as a random distribution but both have the same statistical properties, thus identical histograms.



porosity_A
mean = 17.9
std dev = 1.89
min = 11.4
max = 23.5
N = 4900



porosity_B
mean = 17.9
std dev = 1.89
min = 11.4
max = 23.5
N = 4900

Measures of Spatial Correlation

The process of hand contouring data points to create a map is a form of geostatistical modeling. The geoscientist has a geological model in mind before attempting the contouring exercise. Anyone contouring data usually assumes the presence of the spatial component (regionalized variable), but typically ignores the second component (the random variable).

Variography is comprised of two main steps:

1. Compute the experimental measures of spatial continuity, accounting for anisotropy and azimuthal properties (e.g., variogram).

2. Model the experimental variogram for use in the mapping phase.

The Random Function

The complex attributes we deal with in the earth sciences are random functions. The random function has two components:

- a structured component (regionalized variable) exhibiting some degree of spatial auto-correlation

- a local random component (random variable), also referred to as the nugget effect

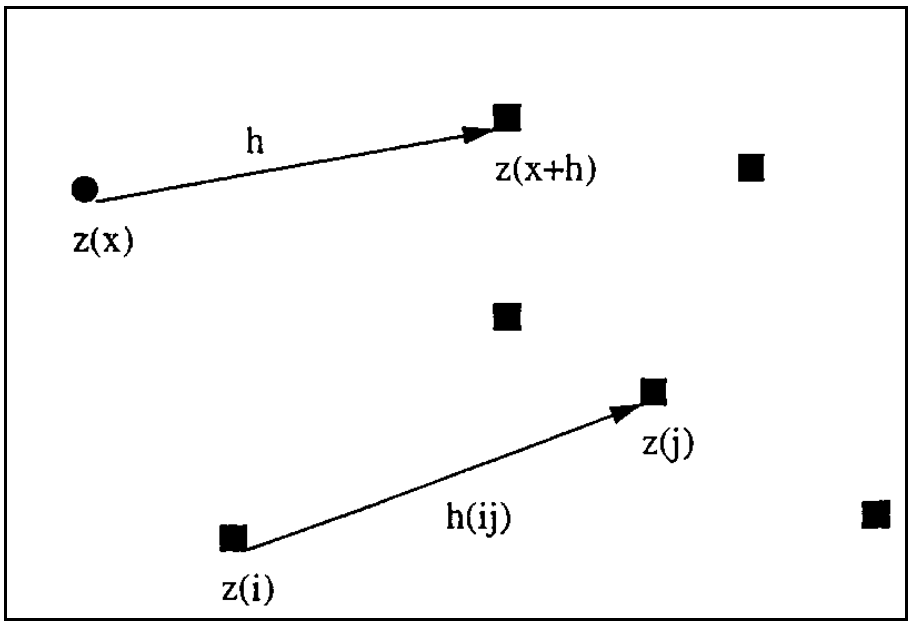
Quantifying the spatial information involves comparing data values measured at one location with values of the same attribute measured at other locations.

The random function model assumes that:

- the single measurement at location $z(x_i)$ is one possible outcome from a random variable located at point $Z(x_i)$.

- the set of collected samples, $z(x_i)$, i = 1, . ..n, are interpreted as a particular realization of dependent random variables, Z(xi), i = 1, . n, known a *random function.*

Studying the spatial dependency between any two measurements of the same attribute sampled a $z(x_i)$ and $z(x_{i+h})$, where h is some measurement of distance along specified vectors, amounts to studying the spatial correlation between two corresponding random functions $Z(x_i)$ and $Z(x_{i+h})$.

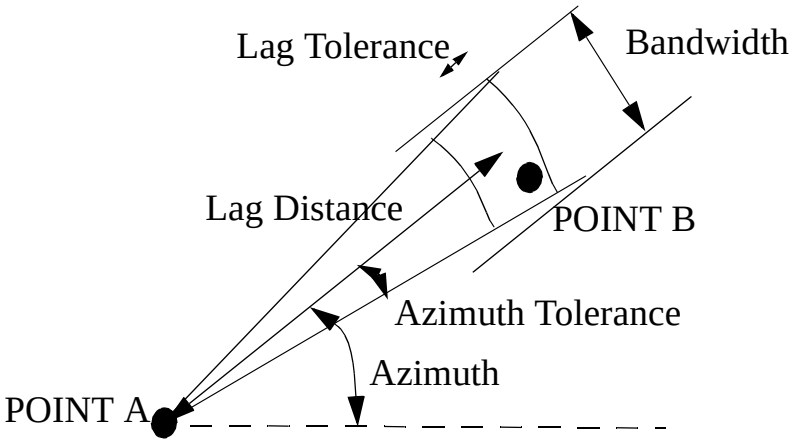


Calculating the Experimental Spatial Model

If the sample data have a regular sampling interval the calculation search strategy is simple. Unfortunately point data (2-D) rarely form a neat regular array, therefore to extract as much information out of the data as possible, rather than searching along a simple vector we search for data within bins.

As shown below, the experimental model is computed from data pairs separated by a specified distance (lag distance) along a given direction (azimuth), which fall within a bin defined by certain tolerances for both the lag and azimuth. The bandwidth limits the width of the azimuth tolerance (i.e., restricts the width for large lags).

Spatial Correlation Azimuth, Lag, Tolerance and Bandwidth

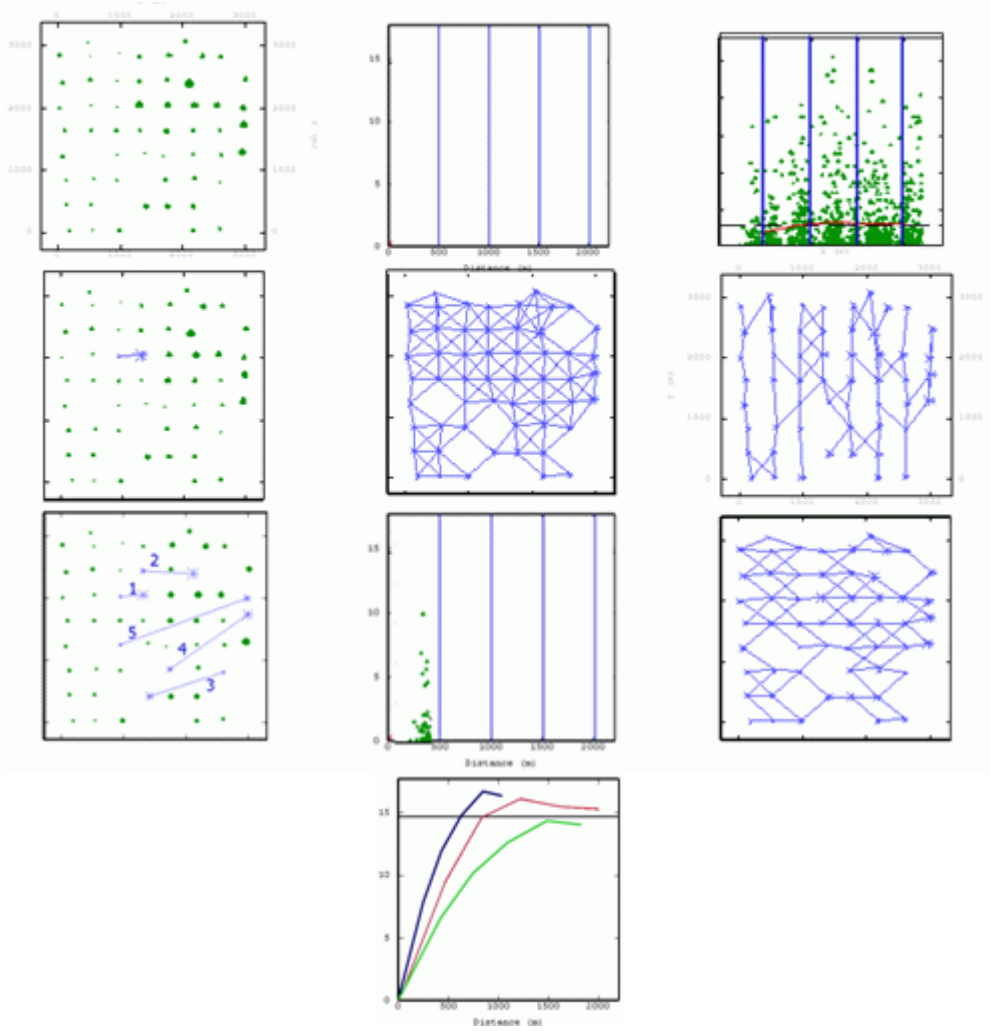


Variogram

For each azimuth and lag distance studied, all measured values can be compared and expressed as a statistical value known as the variogram, and is calculated using the following expression:

$$\gamma(h) = \frac{1}{2n}\sum[z(x_i) - z(x_{i+h})]^2$$

The (semi)-variogram correlation term, gamma (γ), is called the *moment of inertia*, and is a measure of dissimilarity, or increasing variance as a function of distance. Computing and plotting gamma as a function of increasing distance (lag) results in a variogram. The progression of slides below illustrate the concept of developing lags and the associated cloud of points (squared paired differences) for each lag interval. Further, pairs of points can be directionally collected to produce directional variograms.



Why Perform Modeling?

The experimental variogram is calculated only along specific interdistance vectors, corresponding to angular/distance classes.

To use the spatial information, the kriging and conditional simulation applications require a model of spatial dependency, because:

- The kriging system requires knowledge of the variogram for all-possible distances and azimuths.
- The model smooths the experimental statistics and introduces geological information.

Spatial modeling is not curve fitting, in the least squares sense. A least squares model does not satisfy the *positive definiteness* criterion, which prevents the generation of negative covariances during matrix inversion. Only authorized variogram functions ensure this condition.

Consider a random function *Z(x)* with an auto-covariance *C(h)*:

- Define an estimator,

$$Z = \sum \lambda_i Z(x_i)$$

- The variance of Z is given by

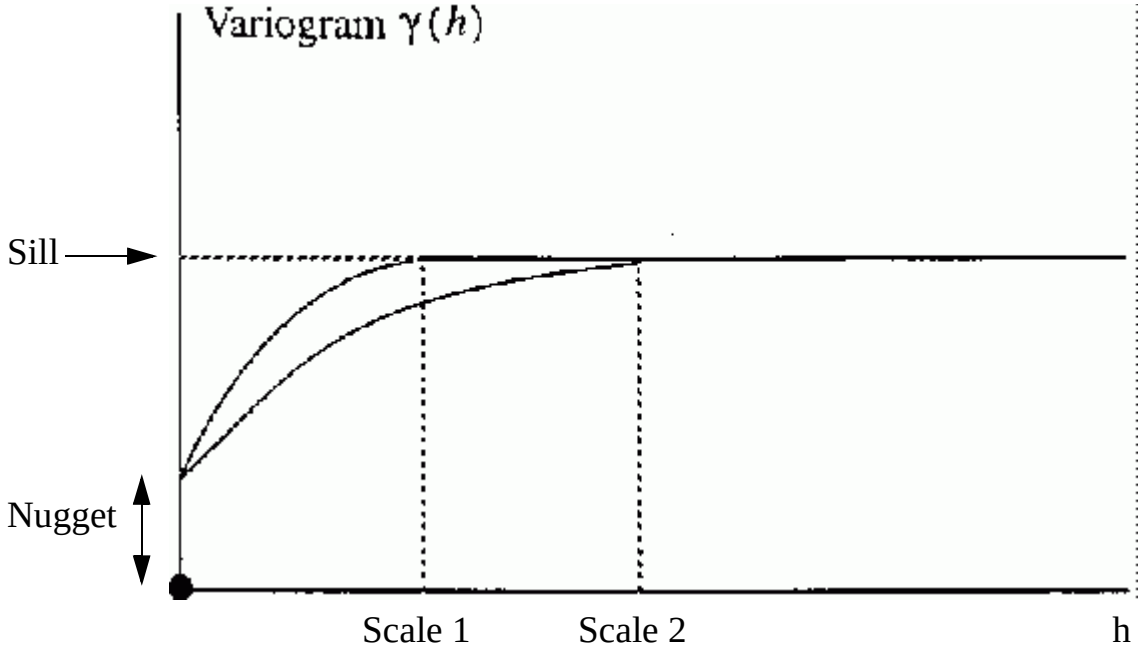
$$\sigma_z^2 = \sum \lambda_i \lambda_j C(x_i - x_j) \geq 0$$

- The variance must be positive for any choice of weights (λ_i and λ_j), and any choice of locations (x_i and x_j).
- To honor the above inequality, fit a positive definite *C(h)* model to the experimental variogram.
- Cannot fit directional variograms independently, but must select a single model from a limited class of acceptable functions.

Anatomy of the variogram model

The variogram model tends to reach a plateau called a **Sill**. This is the variance (σ^2) computed from the measured data values. The distance at which the variogram flattens and reaches a constant variance is called the **Scale or Correlation Range**. The sill and scale are useful properties when comparing directional orientations in the data. Often the variogram shows a discontinuity at the origin, termed the *nugget effect.* The nugget effect is random noise at all scales and may represent short scale variability, measurement error, sample rate, etc., or more likely is

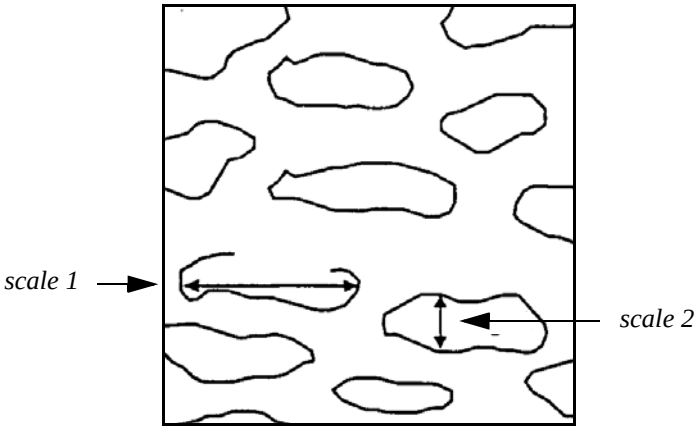
aliasing due to inadequate sampling of features smaller than the sample spacing (well spacing, for example).



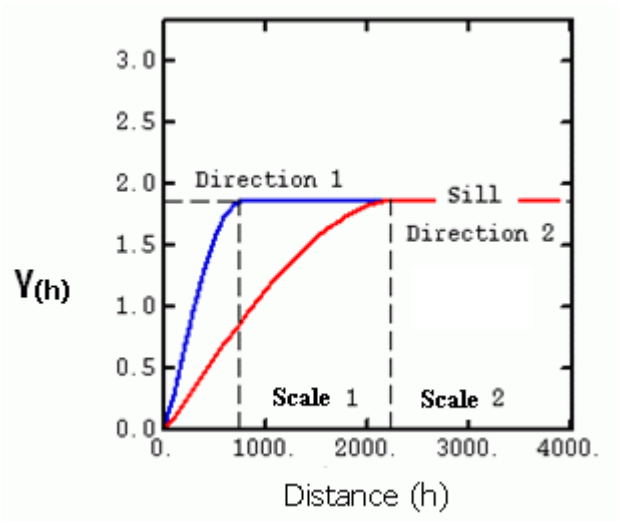
Anisotropy

If the scale computed for all azimuths were constant, then a single, isotropic, variogram would be a sufficiently accurate model over the entire study area. In most cases, however, the scale in one direction is different from another direction, requiring an anisotropic model.

The software models Geometric anisotropy; i.e., variograms that reach the same sill, but have different scales. This is the behavior displayed in the horizontal or XY plane.



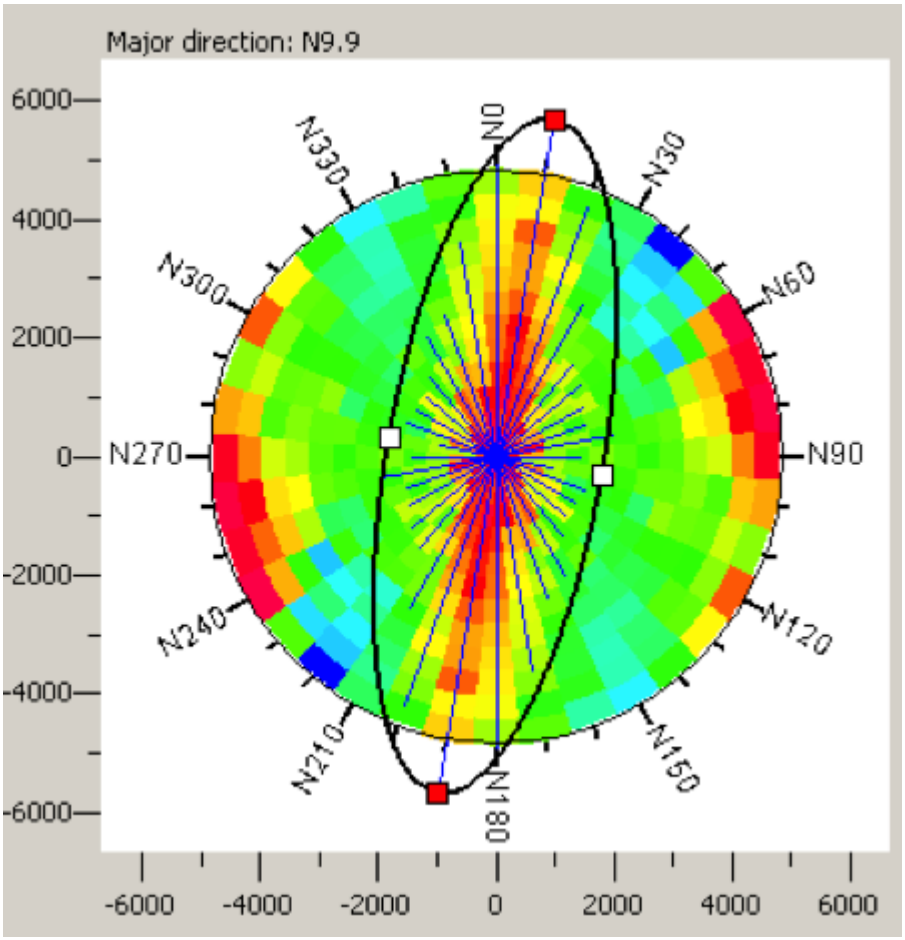
Cross Section of Sand Bodies



Geometric Anisotropy

Anisotropy is more readily apparent in a display known as a variogram map, which is a polar plot of variance and its associated rose diagram. For each direction calculated, the variance is plotted as a function of distance. Fitting an ellipse around the edges of the diagram shows the shape and direction of anisotropy. The major and minor axes represent

the maximum and minimum scales of anisotropy. This model directionally biases the kriging or conditional simulation processes.



The variogram map or more correctly, the variogram polar plot, represents variograms computed for many directions around the compass. In this example, variograms were computed every 5 degrees for twelve 400 meter lags. The center of the polar plot is at zero distance, with each colored “pixel” representing a γ value within a different azimuth and lag bin. The ellipse is fitted automatically based on the length of the rays (scale for each direction). We see the major scale is N9.9E with about a 3:1 anisotropy ratio. Computation of the variogram polar plot requires a lot of densely sampled locations such as from 3D seismic data.

Basic Covariance Functions

The DecisionSpace® Earth Modeling software offers four of the possible fourteen authorized covariance functions for variograms. The other ten functions are typically not used in the petroleum industry.

Qualities of basic covariance functions are:

- Basic models are simple, isotropic functions.
- Independent of direction that are equal to 0 at h = 0 (variograms).
- The functions all reach or approach zero (or the sill) beyond a certain distance which is referred to as the correlation length, a.
- The Cubic, Gaussian, and Exponential functions reach zero (sill) asymptotically. For such functions, the correlation length is arbitrarily defined as the distance at which the value of the function decreases to 5%.

Model Types

DecisionSpace® Earth Modeling offers the following model types:

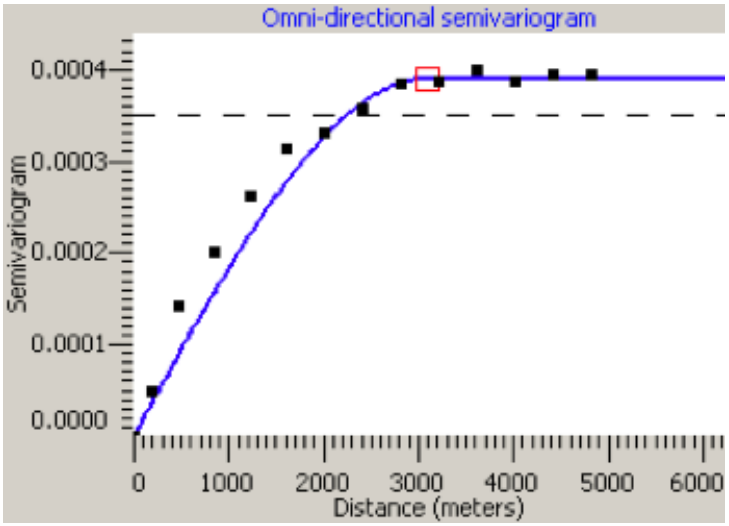
- Spherical
- Exponential
- Gaussian
- vbCubic

In the figures below, the experimental variogram was computed from average interval porosity at 261 well locations with the wells on a 40-acre 5-spot spacing at an average interwell distance of 400-m, which was used as the lag interval during computation. The experimental variogram was modeled with the four available types showing that the Exponential type has the best fit. The dashed line represents the actual value of the data variance. The experimental points rise above the variance indicating a very slight non-stationarity (trend) in the data; however the data do reach a “local” sill.

Spherical Model

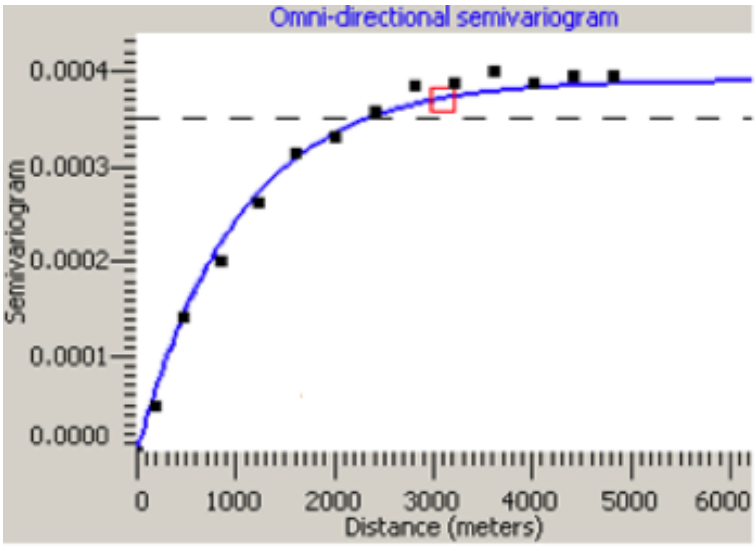
$$\gamma(h) = \frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3 \qquad (h < a)$$

$$\gamma(h) = 1 \qquad (h \geq 1)$$



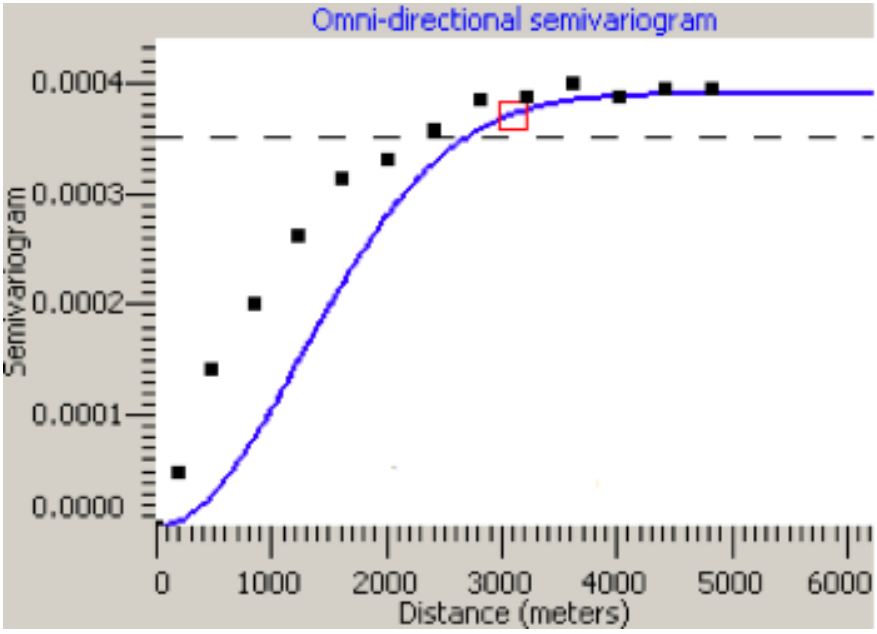
Exponential Model

$$\gamma(h) = 1 - \exp\left(-\frac{h}{3a}\right)$$



Gaussian Model

$$\gamma(h) = 1 - \exp\left(-\frac{1}{3}\left(\frac{h}{a}\right)^2\right)$$



Cubic Model

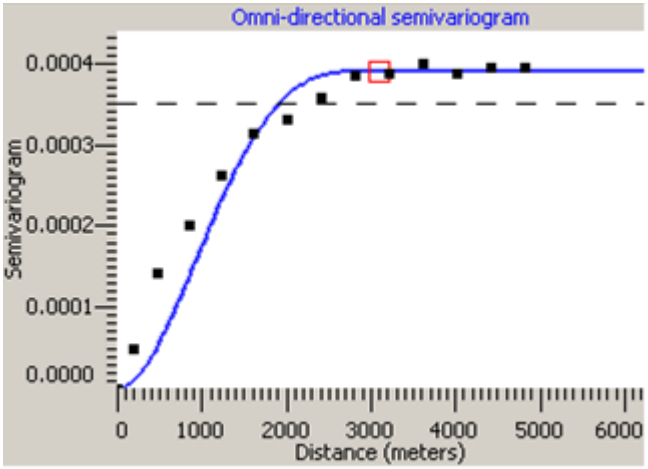
The **Cubic (and Gaussian)** model is a bounded variogram and reaches a sill at some distance "a" along the X-axis, where ‘a’ is the correlation distance or scale of the model. The SF (d) =1. The Cubic (and Gaussian) model belongs to a class of smooth, continuous functions models. Such models can become quite unstable creating matrix inversion problems due to the tail which produces high continuity near the origin. It is recommended when using the Cubic (and Gaussian) model to include very small Nugget, (e.g. 0.0001) as its sill. The default Nugget is C - C/100.

The formula for a Cubic model is:

$$\gamma(h) = C\left[7(\frac{\delta h}{a})^2 - \frac{35}{4}(\frac{\delta h}{a})^3 + \frac{7}{2}(\frac{\delta h}{a})^5 - \frac{3}{4}(\frac{\delta h}{a})^7\right]$$

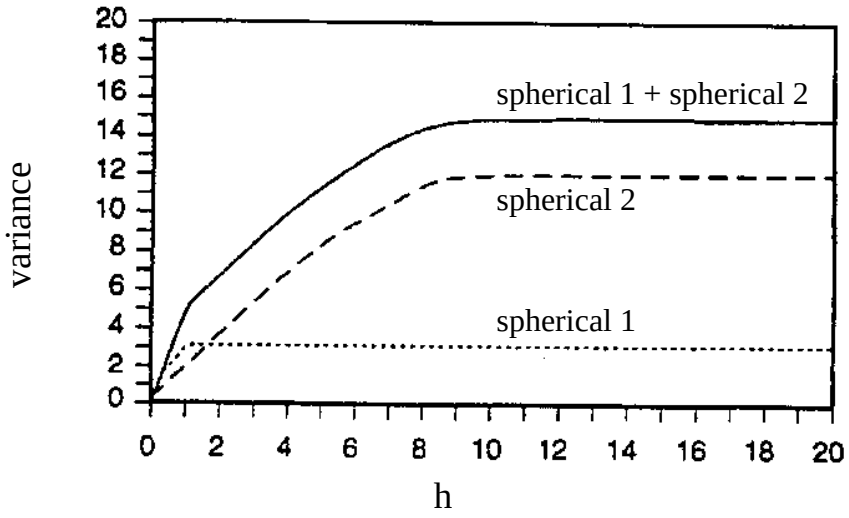
$$\delta = 1$$

where C = Sill value
h = distance
a = correlation distance
δ = scale factor



Nested Model

A nested model consists of the linear combination two or more spatial models. Thus, a nugget model combined with a spherical model is a nested model, although we do not generally think of it in those terms. A nested variogram model (solid line) composed of two different scale spherical models (dashed and dotted lines) is illustrated below.



Practical Considerations and Tips

Practical Considerations

- The omni-directional variogram considers all azimuths simultaneously.
- An omni-directional variogram contains more sample pairs per lag than any directional variogram, and therefore is more likely to show structure.
- If an omni-directional variograms does not show a clear structure, do not expect much success with directional variograms.
- An omni-directional variogram is the average of all directional variograms.
- The Nugget Effect is more easily determined from the omni-directional variogram.
- May need to "clean-up" the data prior to calculating the variogram; variograms are very sensitive to "outliers".
- Do not consider variogram values for distances greater than about one-half the size of the study area.
- Interpret a variogram only if the corresponding number of pairs per lag is sufficient (e.g., 15 to 20 pairs).
- A saw-toothed pattern may indicate a poor choice of the lag increment.
- Consider data clustering.
- The variogram computation involves a decision of stationarity or no trend in the data, which should not be confused with directionality orientation.
- Non-stationary (strong trends in the data) variograms do not reach a sill and are considered unbounded (have a characteristic parabolic upward shape).

Variogram Modeling

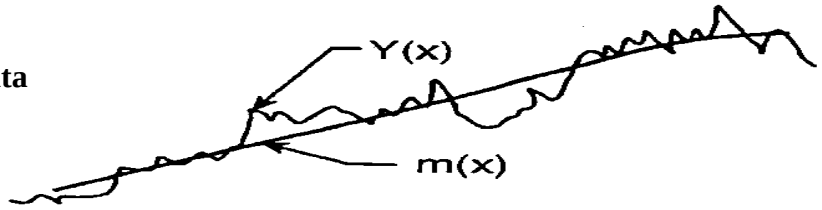
- Do not over fit; use the simplest model (fewest number of structures).
- Do not fit the variogram for distances greater than one-half the study area.
- Pay special attention to the fit for small distances (the first few lags) and the shape of the slope near the origin which determines the variogram model type (e.g. spherical, exponential, etc.).
- The magnitude of the nugget effect
  - acts as a smoothing function during kriging.
  - adds additional variability during conditional simulation.
- Beware of features that may relate to non-stationarity, like a variogram that doesn’t appear to reach its still
- Beware of periodic oscillations
  - "hole" effect related to true structure
  - "hole" effect due to sparse data at given lag intervals

Stationarity

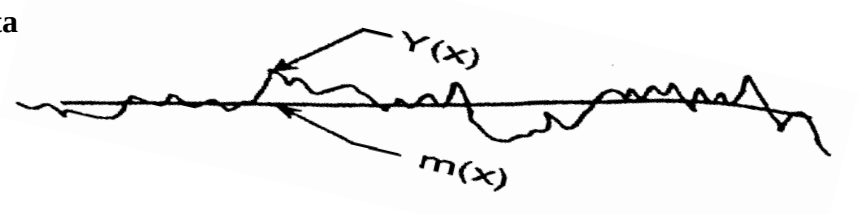
Theoretically, the spatial correlation model should be independent of sample location, however, data often show a regular increase (or decrease) in value over large distances, and the data are said to be non-stationary, or possess a trend.

We define "trend" as a slowly moving function at the scale of the study area (e.g., global behavior), whereas the local behavior is termed "drift." If the regionalized variable is non-stationary, it can be regarded as a composite of two parts, the residual and trend. The regionalized variable can be graphically illustrated as follows:

$Z_{(x)}$ = The observed data (residual data + trend)



$Y_{(x)}$ = the residual data (observed - trend)



$$Z(x) = Y(x) + m(x)$$

Where

- Y(x) has an underlying variogram (residual)
- m(x) can be approximated by a polynomial (trend)

Note

The univariate probability law (the probability of remaining at the same value at location $Z_{(xi)}$ and at $Z_{(xi+h)}$) *does not* depend on the location of x, but only on the separation distance, h.

In practice, it is possible to ignore the trend if the data set contains a "large" number of data points and a “moving” neighborhood search

strategy is used. However, if the data are sparse, the variogram of the residuals should be computed, which means detrending the data.

The steps to handling trends in your data are as follows:

1. "Stationarize" the data (determine the trend) by subtracting the trend from the data.

2. Compute the variogram on the residuals.

3. Perform Kriging or Conditional Simulation of the residuals on the grid.

4. Populate the trend onto the grid.

The final Results are the residuals plus the trend.

Although this is a reasonable approach, the main hurdle is the "correct" determination of the trend to remove from the raw data. In addition, the trend removed by this method is the global trend, but perhaps we should be working at a local scale (e.g., neighborhood scale).

The Dilemma

We should stationarize the data when using sparse data, but how do we determine the trend from a sparse set of data points.

The concept of stationarity is used in every practice and is not merely a set of rules and definitions that make geostatistics work. For example, consider the following:

> *"The top of Formation A occurs at depth of about 980 feet TVDss in Washington County, Oklahoma."*

Such a statement does not preclude the possibility that Formation A varies in depth from well to well. If Z(top) is a stationary random function, and

- at location $Z_{(xi)}$, Formation A occurs at 980 feet TVD, then

- at location $S_{(xi + 1/2 \text{ mile})}$, Formation A should also occur at about 980 feet TVD, plus or minus a few feet.

However, if Formation A is known to be non-stationary, then predicting the depth to the top of Formation A in the new well is more difficult, requiring a more sophisticated model.

Spatial Models

Advantages

With sufficient data, variograms:

- are measures of linear spatial dependence.
- can quantify spatial scales.
- can identify and quantify anisotropy.
- can be used to test multiple geological scenarios.

Limitations

- Variograms are measures of linear spatial interdependence and may not be appropriate for non-linear processes.
- Non-stationary effects often affect variograms.
- Spatial correlation analysis is difficult to perform when data are sparse.
- It is often difficult to select a domain of stationarity (constant mean) for computation.

Chapter 4
Property Modeling Concepts

This chapter introduces basic property modeling concepts illustrated with a series of computer exercises using a 2D data set.

In this chapter you will learn:

- Fundamentals of grids and gridding
- Interpolation Techniques
- Data integration
- Simulation
- Uncertainty analysis

Basic Concepts of Grids and Gridding

One of our many tasks as geoscientists is to create maps of reservoir attributes over a region based on measurements made at the well. For example, your objective may be to estimate (interpolate) the distribution of porosity over the entire study area based on well data only. The DecisionSpace® Earth Modeling software and many other mapping packages require a grid before any interpolation methods can be used. Because it is impractical to compute the variable over an infinite number of points, the DecisionSpace® Earth Modeling software allows you to define a grid that describes locations where the estimates will be calculated.

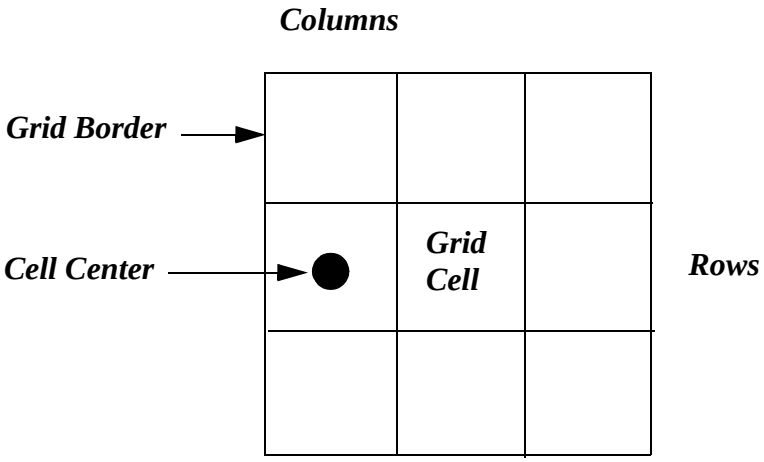
Because these data are located at discrete points, a grid must be designed which reflects the average spacing between the wells, and designed such that the individual data points lie as closely as possible to a grid node. Ideally, the grid is designed such that there are at least four to five grid nodes, on average, between each of the measured data locations. When clustered data are mixed with more widely-spaced control points, it is difficult to achieve such a spacing.

What is a Grid?

A grid is simply a set of numbers representing the X, Y, and Z (if in 3D) coordinates where estimates are made from the control points and are regularly arranged, commonly in a square or rectangle, although other grid forms are also used. The locations of the values represent the geographic locations in the area to be mapped and contoured. For example, you might decide that you want porosity to be calculated every 450 feet in the north-south direction and every 300 feet in the east-west direction, based on well spacing and known geology.

This spacing or increment is used to define a series of rows and columns. The program assumes equal spacing between all rows and columns, but the increment does not need to be equal for both rows and columns.

For DecisionSpace® Earth Modeling, the properties are stored at the corner of the cells for 2D grids (corner-point) and at the center of each cell for 3D grids (cell-centered).

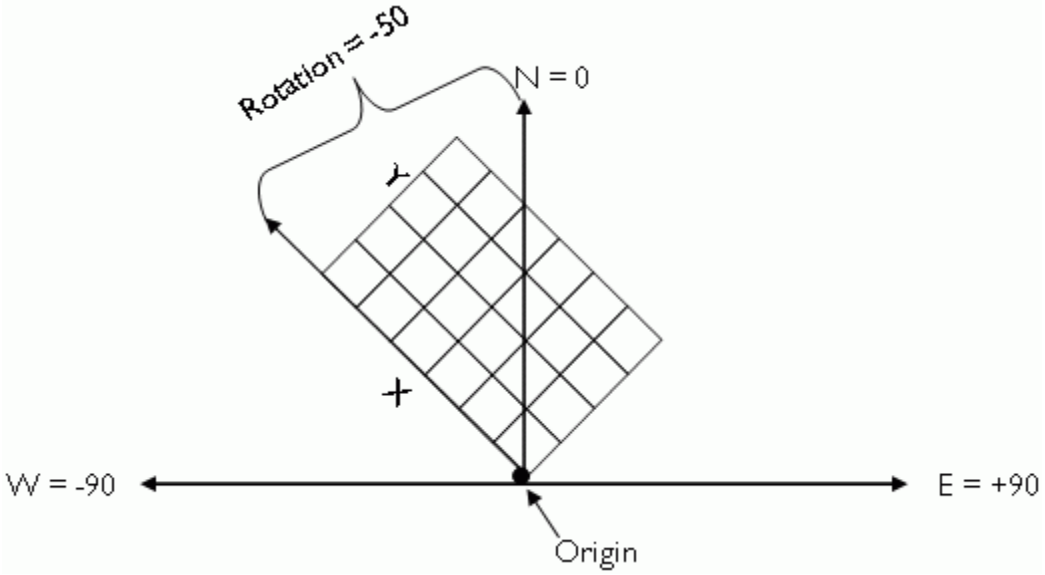


Grid Geometry in the Earth Modeling Software

The word grid may refer to any grid imported from other applications through a grid file or to any set of gridded values defined and produced inside the software. Any variable represented by a grid can be called a 2D variable or a voxel if in 3D. An existing grid is often called an input grid. A grid being defined or calculated is often called the output grid.

Grid Rotation

The grid rotation is based on the geological definition of azimuth where zero degrees equals North. The grid azimuth is defined as 0 degrees plus or minus the rotation value. The absolute value of the grid rotation may not exceed plus or minus 90 degrees. When the grid rotation angle is changed, the lower right-hand corner of the grid serves as the axis of rotation, with the lower X-axis boundary as the reference edge of the grid.



Grid Rotation

What is Gridding?

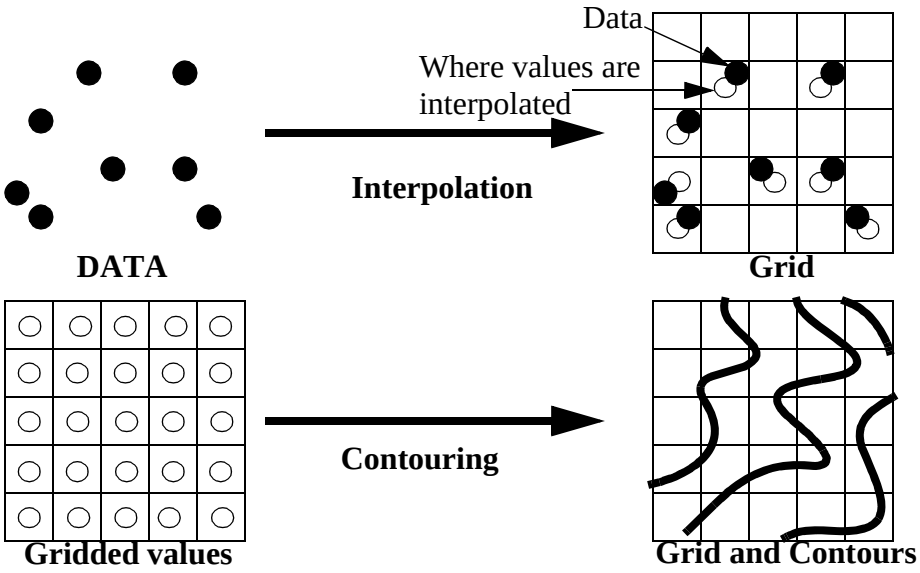
Gridding is simply defined as a process of estimating the value of an attribute from isolated points onto a regularly spaced mesh, called a grid. The mapping (interpolation) and contouring process involves four basic steps:

1. Identification of the area and attribute to be mapped.

2. Designing the grid over the area.

3. Calculation of the values to be assigned to each grid node.

4. Using the estimated grid node values to draw contours.



EarthModeling Grids

The DecisionSpace® Earth Modeling software creates two types of grids, a 2D grid called a GridSet, and a 3D grid called a PM Grid3D. Whenever a reference is made to a grid in this manual, it is referring to these grids as opposed to other types of grids such as OpenWorks surface grids and SeisWorks 3D horizons.

GridSets can be created from a pointset using Create 2D Grid in Property Modeling to create a rectangular 2D grid (these are not .vdb grids). User-defined parameters define the geometry of the basic 2D grid.

PM 3D grids can be created in two ways:

- Using Create 3D Grid in Property Modeling to create a rectangular, non-layered .vdb grid. User-defined parameters define the geometry of the basic, empty .vdb grid; there is no stratigraphic (parallel or proportional) layering involved.

- Using Grid Geometry in Stratigraphic Modeling to create a .vdb grid that includes layer definitions. This utilizes three sets of inputs:

    - inputting a pre-existing structural model, which is an OpenWorks® GeoModel consisting of horizons and faults that are stored in the OpenWorks database. These can be GeoModels

created in the Mapping & GeoModeling Workspace in Dynamic GeoModeling or using the ezModel™ functionality in the GeoProbe® software.

- specifying layering parameters to be used within the stratigraphic grid.

- specifying user-defined parameters such as cell size and rotation of the grid within the model.

Interpolation Techniques

Contouring a map by hand or by computer uses interpolation. There are many algorithms for the interpolation process, all requiring some criterion to be satisfied. Interpolation estimates an attribute value (for example, porosity) at a grid node based on neighboring data values (Input Data). A good estimator

- minimizes the estimation error.
- honors the "hard" data.
- declusters the data.
- does not introduce bias.
- weights control points according to a spatial model.

In the DecisionSpace® Earth Modeling software, you can choose from three "quick" interpolation methods:

- Inverse Distance
- Moving Average
- Moving Median.

Five geostatistical interpolation algorithms are also available:

- Linear Model Kriging
- Simple Kriging
- Ordinary Kriging
- Collocated Cokriging
- Indicator Kriging

For detailed information on Interpolation Techniques, please see Appendix A.

Kriging Techniques

The DecisionSpace® Earth Modeling software offers four Kriging techniques plus a Collocated Cokriging technique. These are described below.

Linear Model Kriging

Linear Model Kriging uses the kriging algorithm but does not require a user-specified variogram model as in regular kriging. It uses a "built-in"

linear variogram model. This is often used as quick screening tool for data QC or where data has mild trends.

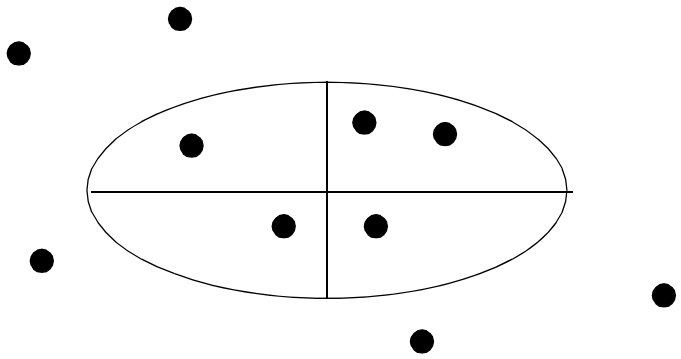
Simple Kriging

Simple Kriging assumes the global mean to be known and constant over the gridded region. Simple kriging uses a univariate model of spatial correlation to guide the interpolation of a primary attribute known only at a number of sample locations.

When this type is selected, you must also specify the **Global mean**, which is the mean of all input property data for the selected subset in the grid. It is assumed to be a global constant. This value is calculated by the software from the input data; however, you can edit this value based on your knowledge of the data or histogram results.

Ordinary Kriging

Ordinary Kriging uses a local mean rather than a global (constant) mean during the interpolation. Ordinary kriging uses a univariate model of spatial correlation to guide the interpolation of a primary attribute known only at a number of well locations. Ordinary kriging differs from Simple Kriging in that the mean is assumed to be locally constant and is estimated using the data within the user-specified Search Neighborhood.



Some practical considerations when considering Simple Kriging versus Ordinary Kriging are:

- Simple Kriging does not adapt to local trends, rather it relies on a constant, global mean.

- Ordinary Kriging uses a local mean, which amounts to re-estimating the mean at each grid node from the data within the search neighborhood.

- When all data points are used, Ordinary Kriging and Simple Kriging yield similar results.

- If only a few data points are available in the local search neighborhood, Ordinary Kriging may produce spurious weights because of the constraint that the weights must sum to 1.

- If the wells are known to provide a biased sampling, it may be better to impose your own value of the mean using Simple Kriging rather than using Ordinary Kriging.

Collocated Cokriging

Collocated Cokriging (COLK) applies kriging with more than one data type. The same basic sets of equations are called for; however, the covariance between the different data types is also required. This requires that the secondary data type is known at the sample locations and at every grid node where a value is interpolated. COLK is a direct extension of Linear Regression and is used when two properties are related but the property of interest is vastly under sampled with respect to the secondary property, but correlated with the primary property. COLK should provide a better solution as more information is used during the interpolation. Collocated means that the secondary data is located on the grid used for interpolation.

Collocated Cokriging:

- Assumes that the secondary variable is known at all nodes of the estimation grid and at the wells.

- Uses the secondary points at the wells and at the target grid node.

- Only requires the knowledge of: $C_{ZZ}(h)$, σ^2_Y, and $\rho_{ZY}(0)$. Assumes that the cross- covariance is a scaled version of the primary variable auto-covariance.

- In general, the system of normal equations is well conditioned.

Indicator Kriging

Indicator Kriging is kriging with a discrete variable (i.e., facies). The input Data must contain an indicator and you must have a variogram model created using the indicator data.

The key idea behind indicator formalism is to code all of the data into a common format as probability values. The reason for this approach is to provide: 1) simplified data integration because of common probability coding, and 2) greater flexibility to account for extreme values. The indicator approach for continuous data variables requires significant additional effort versus more traditional Gaussian approaches. The indicator approach has seen wide spread application in facies modeling. Regardless of the variable type, the indicator approach leads directly to a histogram that is a model of uncertainty at each unsampled location estimated.

For categorical (discrete) data the aim is to directly estimate the probability of occurrence (estimate the distribution of uncertainty) in the categorical facies variable, for example. The probability distribution consists of estimated probabilities for each category. The probability values are estimated by first coding the data as an indicator or probability values, typically 1 if present, 0 if not present, for each sampled depth along the well bore for 3D or at well locations in a 2D approach.

In the current implementation, the user must code data as 1, 2, 3 for a three facies problem for example; continuous variable applications are not available in the current DecisionSpace® Earth Modeling release.

The first step is to select the most common facies code when modeling the variogram, which can be determined by creating a histogram of the facies variable. Once the experimental indicator variogram is computed and modeled it forms the basis for all of the other facies codes. Finally the variogram model is used with the Indicator Kriging algorithm located in the Interpolation panel of General Property Modeling. Indicator Kriging is the only option for categorical data, such as facies.

Advantages and Limitations of Kriging

Advantages of Kriging are that:

- it is an exact, unbiased interpolator
- it does not attempt to reproduce the variability

- it accounts for both distance and directional weighting
- the Kriging variance is a:
  - — Relative index of the reliability of estimation in different regions
  - — Smaller nugget (or sill) gives a smaller kriging variance
- it minimizes the Mean Squared Error (MSE) error
- you can use a spatial model to bias the interpolation process

Limitations of Kriging are that it:

- tends to produce smooth images of reality (like all interpolation techniques).
- poorly reproduces short scale variability.
- underestimates values in the tails of the frequency distribution (histogram).
- requires the specification of a spatial covariance model, which may be difficult to infer from sparse data.

Other Interpolation Techniques

In addition to the Kriging algorithms, the following interpolation techniques are available.

Inverse Distance

Inverse Distance computes the weights based on the Euclidean distance between data points and the target grid node. For Inverse Distance, the resulting weight at each target point is calculated as the weighted average of the surrounding information (neighborhood). The weight is calculated as the inverse of a power (the Inverse distance power) of the distance between the data point and the target point. The weights are normalized so that the sum of all the weights equals 1. The implicit assumption in this estimation technique is that all control data locations/ values are statistically independent; that is, that the underlying model of how these values occur is purely random (non-geologic). The method is

implicitly isotropic. There are no “Rules-of-Thumb” for choosing the Inverse Distance power; it is a trial-and-error procedure.

Rather than giving equal weight to all samples, this method of interpolation gives more weight to the closest samples and less to those that are farthest away. This is done by making the weight for each sample inversely proportional to its distance from the point being estimated.

Advantages of the Inverse Distance method are:

- simple to use and understand
- a "quick map"
- excellent QC tool
  - "bulls-eye" effect (lone high or low values)
  - spot erroneous sample locations
  - first indication of trends

Moving Average

Moving Average calculates the resulting value at each target point as the average of the values at each target point of the active data contained in the neighborhood. With this method, you need to use a search neighborhood with an optimum number of samples that is less than the number of input locations (wells). Using a large search neighborhood with All Data results in constant values for each layer. Therefore, when creating the search neighborhood, you should use a subset of the data points by decreasing the Optimum Number of Samples. The fewer data points used, the less smoothing is done.

Moving Median

Moving Median calculates the resulting value at each target point as the median (50th percentile) of the values of the active data contained in the neighborhood.

With the Moving Average or Moving Median method, you need to use a search neighborhood with an optimum number of samples that is less than the number of input locations (wells). Using a large search neighborhood with All Data results in constant values for each layer. Therefore, when creating the search neighborhood, you should use a subset of the data points by decreasing the Optimum Number of Samples. The fewer data points used, the less smoothing is done.

Search Neighborhoods

A search neighborhood allows you to perform a global search to find conditioning data from the entire subset. By default, the algorithm uses the search ranges specified in the variogram model to find these data points. The Create Search Neighborhood dialog in the DecisionSpace® Earth Modeling software allows you to set parameters for a search ellipsoid. Samples are selected in the neighborhood only if they fall inside the specified ellipsoid search criteria.

The search neighborhood is used to identify and select which scattered data points are used by an algorithm to interpolate or simulate values onto a grid node. Designing a search neighborhood is often "trial and error" until satisfactory results are achieved; a somewhat subjective definition of what "looks good." It is difficult to provide precise rules because the criteria depend upon the amount of data (dense or sparse) and how they are distributed across the study area; are the data clustered, evenly distributed, along lines, in combinations. Designing 3-dimensional neighborhoods is more difficult than 2D search designs.

If a directional variogram is selected on the Simulation panel, the major axis azimuth is automatically chosen as the major axis direction for the neighborhood.

The Search ranges and azimuth parameters are used to specify the search ellipsoid. The parameters are:

- **Major** is the maximum search range used for the major axis (i.e., the major axis radius). The default is twice the range of the variogram.

- **Minor** is the maximum search range used for the minor axis (i.e., the minor axis radius). The default is twice the range of the variogram.

- **Vertical** [3D data only] is the maximum search range used for the vertical axis (i.e., the vertical axis radius).

- **Azimuth** is the direction along the major axis direction.

The default ranges are double the variogram ranges. Using ranges less than the variogram ranges is not recommended as you could lose some valuable data. The search area should be large enough to include an

"optimal" number of data points during the interpolation or simulation. We recommend:

- For the major and minor axes use 2-3 times the correlation scale of the variogram model. Larger axes lengths ensure that enough data points are found.

- For the vertical axis it is best to use 1 to 2 times the vertical cell thickness.

If the axes are too short, then few to no data are found; computation time is fast, but can also create local anomalies or artifacts. If too large, too many data are included (limited to 100) and the calculation time is slower and tends to create smoother looking maps.

Sectors are used to divide the search ellipse into many equal angular sections. The Search sectors parameters are:

- **Minimum number of samples** is the minimum number of points required inside a sector in order for a neighborhood search to succeed. If the actual number of points that fall inside the ellipsoid is smaller than this specified number, the neighborhood search fails and the kriged variables at the target point receive an undefined value. The default is 1. Normally, you would increase this value only if you know the minimum number of samples in your data and want to use more than one data point.

- **Number of sectors** is the number of angular sectors. For example, 2 sectors would consist of a sector from 0 to 180 degrees and a sector from 180 to 360 degrees.

  All samples belonging to the neighborhood ellipsoid are sorted against their distance to the target grid node and are grouped in different angular sectors. This parameter is used in combination with the Optimal number per sector.

  You can increase the number of sectors in order to make sure that some samples are selected in different directions of the field. This is particularly useful when the sampling pattern is highly anisotropic.

- **Optimum number of samples per sector** is the ideal number of points per sector. The search of points in the different sectors is a sequential process. The software scans each sector at a time until it has been assigned the specified optimum number if possible within the search ellipsoid.

The number of data points used to interpolate or simulate a value at a grid node affects the final result. For 2D studies, we recommend using 40 samples and for 3D studies 60 samples if possible. When too few data points are used local anomalies are enhanced creating “the “bulls-eye” effect or “figure-eights." We recommend:

- Minimum number of samples: Even thought the default is 1 data value, it is often best to increase the number to ensure that a single value is not populated across many nodes. It is often better to use perhaps 5 to 8 points at a minimum to produce weighted-averaged results over the same sparse data area.

- Number of points per sector: The default is a single sector and the search algorithm will find the nearest number of points specified. Common sector searches use 4 (quadrants) or 8 (octants) sectors with an optimum number of data points per sector. This optimizes a more even distribution of data points around the target node. When using sector, the search algorithm uses a spiral search pattern to collect data points, selecting the closest data point to the target node in the first sector, the closest in the second sector, and so forth until the optimum number of data points are found in each specified sector. The practical limit is to use 100 total data points or less in the search neighborhood to optimize computational time and not create overly smoothed results. For example, if using a 4 sector search, then the maximum allowable points per sector is 25.

A word of caution: Before deciding upon the final search neighborhood parameters it is recommended to test the neighborhood design using an interpolation algorithm, such as kriging or inverse distance. Any neighborhood design problems will show up as artifacts in the results. If a conditional simulation algorithm is used first, such artifacts are most often hidden due to the stochastic nature of the algorithm.

For 3D neighborhoods it is often best to use only 1 or 2 samples vertically, so limit the search to 1 or 2 cells thick to prevent over-sampling along the well. Remember that the search looks for the closest points and including to much vertical information creates local anomalies not based on the geology, but on the search criteria.

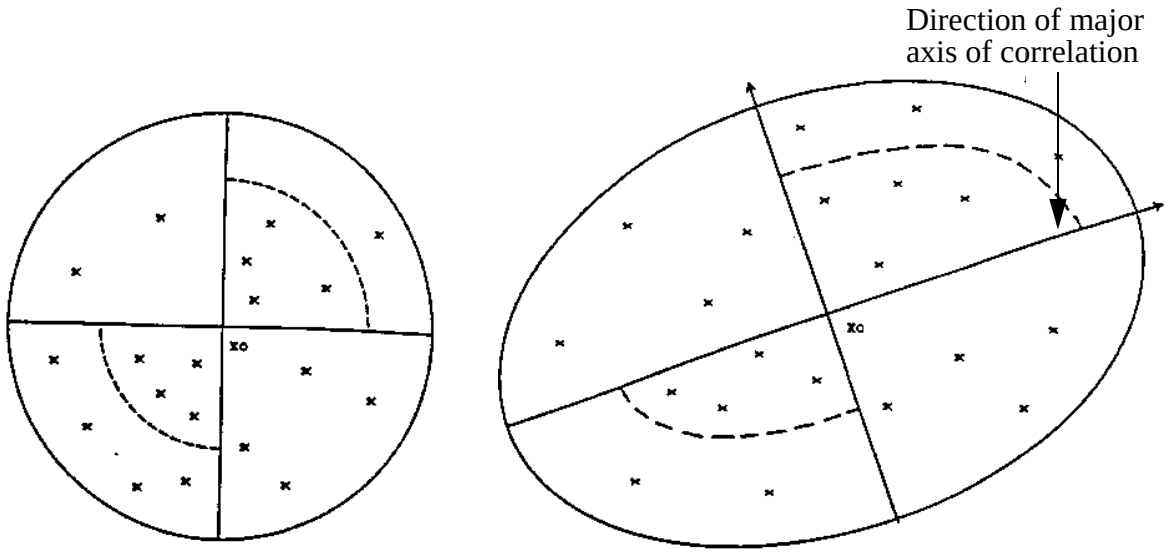
Search Neighborhood Criterion

All interpolation algorithms require good data selection criteria, referred to as the search neighborhood parameters, which include:

- Search radius or radii

- Neighborhood shape (circular or elliptical)
- Number of sectors (4 or 8 are common)
- Number of data points per sector
  - — All Data uses all data points (practical limit is 100)
  - — Search Neighborhoods use a limited number of points per sector
- Azimuth of major axis of correlation (minimum anisotropy)

The figure below illustrates two types of search neighborhoods designed as a quadrant search with a maximum of 8 data points per quadrant.



Circular Neighborhood (Isotropic) Elliptical Neighborhood (Anisotropic)

Practical Considerations in Designing the Search Neighborhood

- Align the search axis with the direction of major direction of continuity.
- Search radii (if anisotropic) should be similar or greater than the correlation scales determined during variography.
- Each quadrant should have enough points (~ 4) to avoid directional sampling bias.

- CPU time and memory requirements grow rapidly as a function of the number of data points in neighborhood.
- In theory, more data in the kriging system reduces the Mean Squared Error (MSE).
- In practice, the covariance is poorly known for distances exceeding about 1/2 to 2/3 the size of the field. Including points that are more distant may actually increase the error.
- The kriging estimator is built from data within the search neighborhood centered at the target grid node location Zo*.

Practical Considerations: Kriging with All Data

- In a search neighborhood, a new Simple Kriging or Ordinary Kriging system of equations is solved at each grid node.
- Using All Data (practical limit of 100 points) tends to prevent artifacts from abrupt changes in the number and values of the data points.
- If sufficient wells are available for Ordinary Kriging, a search neighborhood is preferable to All Data.
- All Data smooths the data more than a search neighborhood.

Effects of Variogram Parameters

In Simple Kriging, Ordinary Kriging, and Collocated Cokriging, a variogram is used to provide the weights based on the modeled distances. Thus the Search ranges (double the modeled scales) and Azimuth come from the user's variogram model. The sector parameters default to a single sector with a minimum and optimum number of data per sector equal to 1 and 16 respectively.

Some practical considerations in setting variogram parameters are:

- Rescaling the variogram (larger or smaller sill)
  - — has no affect on kriging estimate.
  - — changes kriging variance.
- Increasing the nugget component

  - acts as a smoothing term during kriging (weights more similar).
  - adds increased variability during conditional simulation.

- The shape of the variogram near the origin influences the continuity of the interpolation process (e.g., the gentler the slope, the smoother the interpolation).

Increasing the scale tends to increase the influence of more distance data points and leads to smoother maps.

Simulation

In DecisionSpace® Earth Modeling, simulation adds continuous properties to a grid node using the Turning Bands or Sequential Gaussian methods. The Turning Band algorithm was the first large-scale 3D simulation algorithm implemented for continuous properties (e.g. porosity, permeability, etc.). Because the algorithm creates only unconditional simulations (no data required or honored), the conditioning is done by kriging or collocated cokriging when using secondary data. The data (both primary and secondary) require a Normal Score transform and conditioning to the local normal score data also requires a back transformation into the original data units. The algorithm is very fast because the actual simulation is done in 1D along lines (bands) lying on a 2D grid or piercing the 3D volume which partition the grid.

Sequential Gaussian Simulation (SGS), like Turning Bands requires Normal Score transformed data; however, the major difference is that SGS must use some previously simulated values to control the short scale variance seen in the final simulation. The previously simulated points are used to create a local distribution function from which a value is drawn at random for a grid node. Most implementations use a multi-grid approach for the selection of previously simulated points, resulting in a better result. The conditioning is not done through a kriging step like Turning Bands, however the simulations do honor the well data once the data are back transformed.

Collocated Cosimulation using either method produces a simulation with more than one data type. The same basic sets of equations are called for; however, the covariance between the different data types is also required. This method requires that the secondary data type is known at the sample locations and at every grid node where a value is interpolated. This method of simulation is used when two properties are related but the property of interest is vastly under sampled with respect to the secondary, but correlated to, property. It should provide a better solution as more information is used during the simulation. Collocated means that the secondary data are already located at the target grid node and are used during the simulation.

Why Conditional Simulation for Reservoir Characterization?

Conditional simulation is a variation of conventional kriging or cokriging. An important advantage of the geostatistical approach to

mapping is the ability to model the spatial covariance before interpolation. The covariance models make the final estimates sensitive to the directional anisotropies present in the data.

If the objective is to map a structural surface, then the smoothing properties of kriging in the presence of a nugget may be the best approach. However, if the objective is to map directional reservoir heterogeneity (continuity) or estimate reserves and assess model uncertainty, then a method other than interpolation is required. A useful property of conditional simulation is its higher frequency content, which lends a more realistic appearance to maps when compared to kriging.

Conditional simulation is a technique designed to:

- honor measured data values.
- approximately reproduce the data histogram.
- honor the spatial covariance model.
- be consistent with secondary data.
- assess uncertainty in the model.

The first step in Simulation is the random selection of a grid node GN_i, then GN_{i+1}, until all grid nodes contain a simulated value. The order in which grid nodes are randomly simulated influences the cumulative feedback effect on the outcome. The selection process is random, but repeatable:

- For each simulation, shuffle the grid nodes into an order defined by a random seed value.

- Each random seed corresponds to a unique ordering of the grid nodes.

- Different random seed values produce a different path through the grid.

- Even though the total possible number of orderings is very large, each random path is uniquely identified and repeatable.

Stochastic Models and Heterogeneity

Stochastic techniques are capable of producing many plausible outcomes, however, many studies use only a single outcome as the basis of performance prediction. Over the past decade, it has become increasingly apparent that reservoir performance predictions are more accurate when based on models that reflect possible reservoir

heterogeneity. We are painfully aware of the countless examples of failed predictions due to the use of overly simplistic models (Srivastava, 1994). However, use of a single realization could lead to an incorrect assessment of the reservoir. Recall that conditional simulation uses probabilistic sampling and one solution is just as probable as another realization. So, from what part of the distribution is the single realization; is it the P50, P10, or ? That is why it is necessary to generate enough realizations to assess the uncertainty and pick solutions for further evaluation.

Uncertainty Assessment

Anyone who forecasts reservoir performance understands that there is always uncertainty in the reservoir model. Performance forecasts or volumetric predictions are often based on a best case model; the so called P50 solution. The reservoir engineer is also interested in other models, both the pessimistic and optimistic models. These models allow the engineer to assess whether the field development plan, based on the best case scenario, is flexible enough to handle the uncertainty.

When used for this kind of study, stochastic models offer many models consistent with the input data. We could then sift through the many realizations, select one that looks like a downside scenario, and find another that looks like an up-side model.

Uncertainty Analysis

A critical aspect for the use of stochastic modeling is the belief in some space of uncertainty and that the stochastic simulations are outcomes that sample this space fairly and adequately. We cannot possibly look at all outcomes (nor even generate them). However, we believe that we can generate a fair representation of the whole spectrum of possibilities. And, we hope that they do not have any systematic tendencies to show pessimistic or optimistic scenarios.

Uncertainty Analysis involves the idea of a probability distribution, rather than simply sorting through a large set of outcomes and selecting two that seem plausible. In Monte Carlo sampling, we depend on the notation of a complete probability distribution of possible outcomes and that the simulation realizations fairly represent the entire population.

Practical Considerations for Conditional Simulation

- The spatial correlation function is reproduced only for distances within the search radius. Therefore, the search region must extend at least to distances for which the covariance function is to be reproduced.
- Do not expect exact reproduction of the spatial model because of uncertainty in the model parameters.
- In practice, simulated locations are visited according to a random path to avoid artifacts and maximize simulation variability.
- Correct determination of the correlation coefficient between the primary and secondary variable is crucial. Over estimating, the correlation may result in over-constrained simulations and a narrow range of outcomes.
- Sparse data can result in a wide range of outcomes.
- How many simulations should we create and use?
  - This is a difficult question to answer as it depends upon the amount and quality of the input data. We recommend generating at least 50 to 100 simulations to produce reasonable probability maps or confidence margins on global parameters.
  - Only a small number of simulated models, representing minimum, most likely and maximum cases need to be retained for fluid flow simulations.
  - Discard geologically unrealistic simulations.
  - The density and quality of the conditioning data control the amount of variability.

Advantages of Conditional Simulation

- By reproducing the data histogram and the spatial correlation structure, conditional simulations provide more realistic reservoir images.

- Because simulations can reproduce extreme values (tails of the histograms) and their pattern of connectivity, they are useful for simulating hydrocarbon production volumes and production rates.
- They provide alternative models, which are consistent with the data.
- Simulations generate different, but equally probable geological scenarios for use in risk assessment.

Limitations of Conditional Simulation

- CPU and memory intensive.
- Large numbers of simulations may create data management problems.
- Interpret confidence limits calculated from post processing simulations with caution because uncertainty in the conditioning data may be large.
- The simulation error variance is twice the kriging error variance.
- Simulations are very sensitive to covariance model parameters, like the nugget or the correlation coefficient if using collocated cosimulation.
- Sparse conditioning data generally produces a wide range of variability between the simulations.
- Although statistically equally probable, not all images may be geologically realistic.

Simulation Parameters

As in Simple Kriging, Ordinary Kriging, and Collocated Cokriging, a variogram is used to provide the weights based on the modeled distances. And, as with Interpolation, a search neighborhood is used. In addition, the following parameters are used:

- **Number of realizations**, which is the number of simulated results desired for a given run. Each realization is an equiprobable solution to any of the others. Typically, you will want to run many

realizations to get a variety of results in order to understand and quantify the uncertainty or the range of possible results in a model.

- **Seed number**, which is a user-specified value required to start a simulation. The Seed number is used in a random number generator to guide the simulation through the grid. Each Seed number generates a random but repeatable path through the grid. Using the same seed number for a different simulation creates an identical solution. This is a very unique attribute which can be used to test different variograms or neighborhoods using the same seed number for the different parameter changes. If you choose to create 100 simulations, only one Seed number is required as it spawns the additional 99 seed numbers automatically.

For Collocated Cosimulation with Turning Bands or Sequential Gaussian, there are three additional parameters:

- **Collocated input property** is the secondary property in the PM Point set used in the simulation process. The purpose of using secondary data is to improve the simulated results.

- **Correlation coefficient** is the correlation between the primary and the secondary data. It is computed between the Input Data property and the secondary data property (Collocated input property).

> **Note**
>
> Although the software computes the correlation coefficient, this does mean that it is necessarily correct. We recommend that you use the cross-plot to validate the results of the relationship between the primary (well data) and the secondary (usually a seismic attribute) at the well location for the interval of interest. If outliers are present, then mask them using the Data Analysis tools and recompute the correlation.

- **Collocated grid property** is the property created on the grid from the collocated input property.

For Sequential Simulation, specify the **Optimum number of already simulated nodes**. This is the number of previously simulated nodes that will be used in addition to the optimum number of samples per interval specified in the search neighborhood. The default is 10.

Appendix A
Interpolation Techniques

The DecisionSpace® Earth Modeling software allows you to choose from seven interpolation techniques, including three Kriging algorithms and a Collocated Cokriging algorithm. This appendix provides detailed information on the concepts behind the geostatistical interpolation techniques.

In this Appendix:

- Kriging Techniques
- Collocated Cokriging

Kriging Techniques

Kriging is one type of spatial interpolation in which control points are weighted according to the degree of spatial correlation between them using a variogram model. There are two types of kriging available and they are distinguished by how the mean value is determined.

Kriging takes the following weighting factors into consideration:

- closeness of known values to the target grid node
- redundancy between the data values
- direction of continuity
- magnitude of continuity

Kriging accounts for redundancy between the data and a measure of distance specific to the data considered through the variogram model. The disadvantage of kriging is that it smooths, and thus reduces, the heterogeneity in the model. This is acceptable for attributes that are already smooth, but, in many cases, kriged outputs are not acceptable for mapping because the true heterogeneity is removed. Another issue that arises from smoothing is the failure of kriging to honor the histogram; the interpolated values are a weighted mean, thus under representing values in the tails of the histogram.

Kriging is for the interpolation of one property at a time.

How Kriging Works

Kriging uses a search ellipse, the parameters of which depend on:

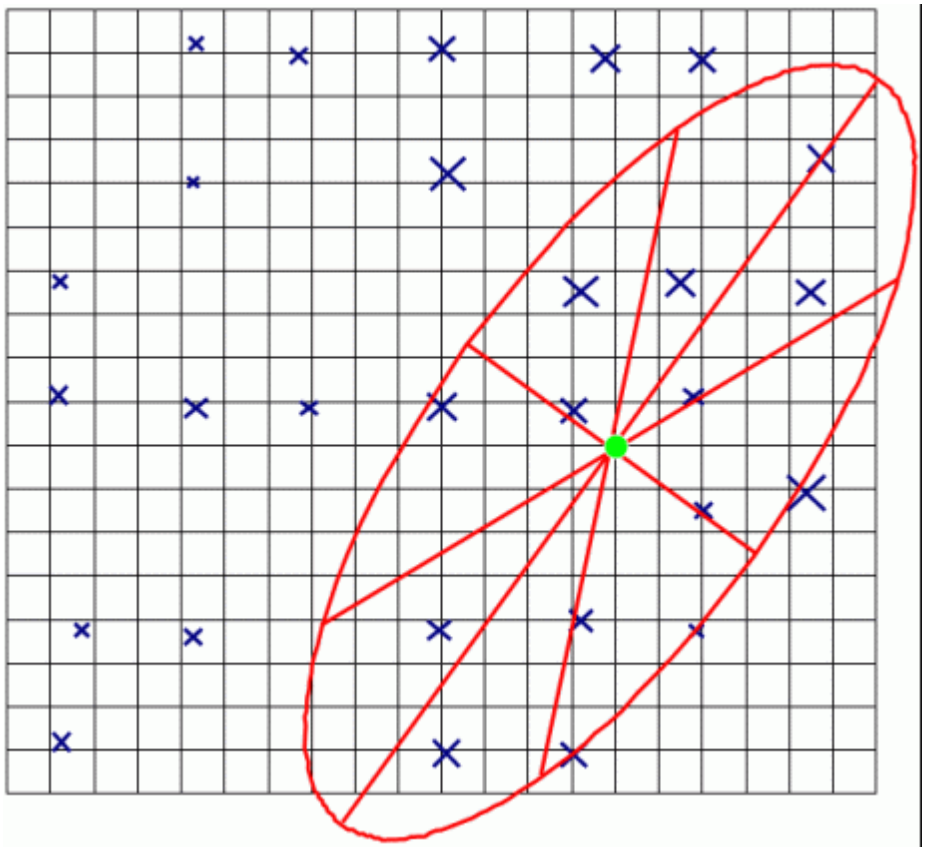
- data density
- variogram spatial scales
- variogram model type
- variogram azimuth

It is a linear estimator expressed as:

$$Z_0 = \lambda_1 X_1 + \lambda_2 X_2 + \lambda_n X_n$$

where Data Points are: $X_1, X_2,...,X_n$
Weights are $\lambda_1, \lambda_2, ..., \lambda_n$
and Z_o is the value estimated at the target grid node (cell center)

Example: How weights are determined



Control point (**x**), an eight sector search ellipse, and target location (●).

General Properties of Kriging

Here are some general properties of Kriging:

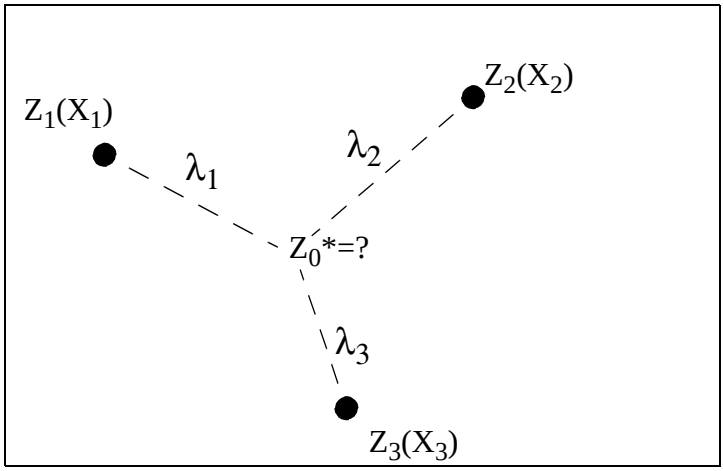
- Univariate estimator; only requires one covariance model.
- Produces a “smoothed” result, like all interpolators.
- Does not reproduce the interwell variability.
- Tends to the mean value when control data are sparse.
- Assigns negative or null weights to control points outside the correlation range.

- Root Mean Square error is a by-product of kriging, which indicates the relative reliability of the estimate. This error value should not be used as a "true" standard deviation, as in classical statistics.
- Kriging matrix is general and easily reformulated, making it a very flexible technique.
- Declusters data.

The Kriging Estimator

Problem

Given samples located at (x_α), find the most likely value of the variable Z at the target point (e.g., at the grid node).



Notation

- The data Z_α
- The estimation value at the target point Z_o*

Equation

- Consider Z_0^* as a linear combination of the data Z_α

$$Z_0^* = \sum \lambda_\alpha Z_\alpha$$

Where: $\Sigma\lambda_\alpha = 1$

- Determine λ_α so that:

  — Z_0^* is unbiased: $E[Z_0^* - Z_\alpha] = 0$

  — Z_0^* has minimum mean square error (MSE)
  $E[Z_0^* - Z_\alpha]^2$ minimum

General Kriging System of Equations

Recall that the unknown value Z_0^* is estimated by a linear combination of n data points plus a shift parameter λ_0:

$$Z_0^* = \sum \lambda_\alpha Z_\alpha$$

Transforming the above equation into a set of linear normal equations, we solve the following to obtain the weights λ_α. The set of linear equations takes the following form:

$$\sum \lambda_j \bar{C}(x_\alpha, x_j) - \mu = \bar{c}(x_\alpha, x_0) \text{ for all i = 1,n}$$

or in matrix shorthand notation:

$$\overline{C\Lambda} = \bar{c}$$

where the terms in the equation are represented by:

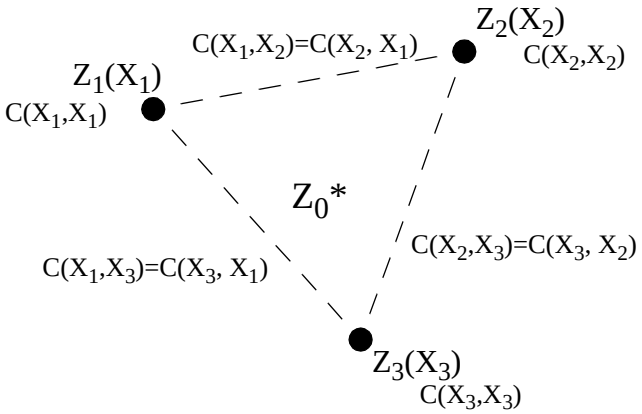
$$\bar{C} = \begin{bmatrix} C(x_1, x_1) & \cdots & C(x_1, x_n) & 1 \\ \vdots & \cdots & \vdots & \vdots \\ C(x_n, x_1) & \cdots & C(x_n, x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}, \bar{\Lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix}, \bar{c} = \begin{bmatrix} C(x_1, x_0) \\ \vdots \\ C(x_n, x_0) \\ 1 \end{bmatrix}$$

All three terms are matrices where:

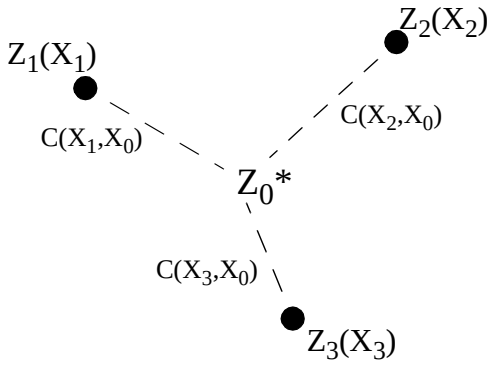
- $C(x_\alpha, x_j)$ represents a covariance between sample points at locations x_α and x_j
- $c(x_\alpha, x_0)$ represents a covariance between a sample located at x_α and the target point, a grid node, x_0; the estimated point.
- λj are the unknown weights
- μ is a Lagrange multiplier used to convert an unconstrained minimization problem into a constrained minimization.

Determine the Matrix of Unknown Weights

- To solve the matrix equation we simply need to substitute in values of covariance as follows:



Covariance between each data point



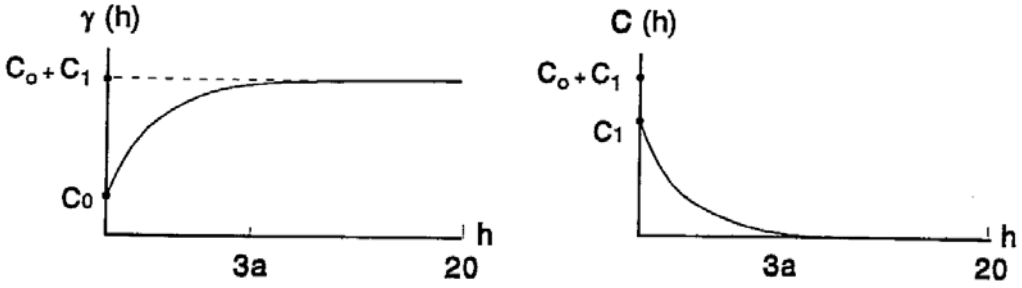
Covariance between each data point and target grid node

- The above diagram describes which covariances between pairs of points are required. The result of spatial correlation modeling was to produce a model which measures covariance as a function of direction and distance. By substituting the directions and distances observed above into the spatial correlation model, the covariances can be determined and substituted into the matrix equation.

- Note that the above equation is written in terms of covariance values; however we modeled a variogram, not the covariance.

  — Computationally more efficient to use covariance matrix
  — Relationship to the variogram

$$Covariance = C(h) = \sigma^2(sill) - y(h)$$



Variogram Covariance

- Once the values of covariance are known, solve the matrix equation for $\underline{\Lambda}$

$$\bar{C} \cdot \bar{\Lambda} = \bar{c}$$

$$\text{where } \bar{\Lambda} = \bar{C}^{-1} \cdot \bar{c}$$

- Once the weights are known, they can be inserted into the original equation and finally the estimate can be determined.

Kriging Variance

In addition to estimating the variable value at an unsampled location, the kriging technique also allows an estimation of the error variance. This is only an estimation of uncertainty with respect to the estimation, therefore, only provides a local estimate of reliability. Because the

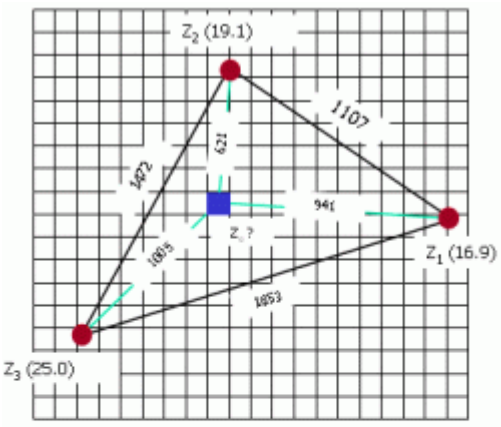
kriging errors are not derived independently from grid node-to-grid node, do not attempt to use it like a measure of standard deviation.

$$\sigma_E^2 = C(x_0, x_0) - \sum_{i=1}^{n} \lambda_j(x_\alpha, x_0) + \mu$$

Do not use Krig σ as a true measure of error. We will get the error term following post-processing of the conditional simulations.

Kriging: Example Using 3 Data Points

Data Configuration

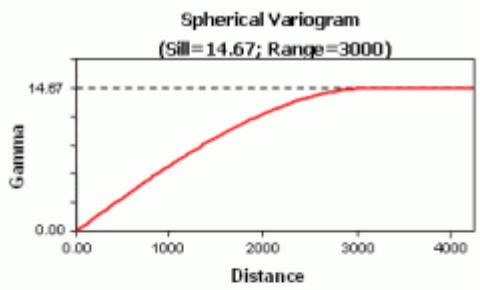


$$\gamma(|h|) = \left[C_1 \left(\frac{3}{2} * \frac{|h|}{a} \right) - \frac{1}{2} * \left(\frac{|h|}{a} \right)^3 \right]$$

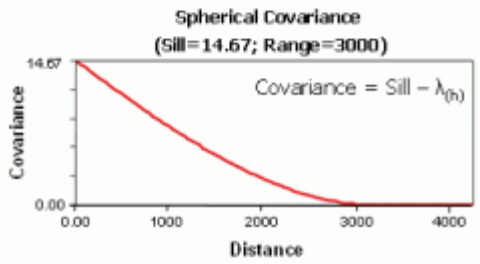
C_1 = Sill of model + any Nugget term

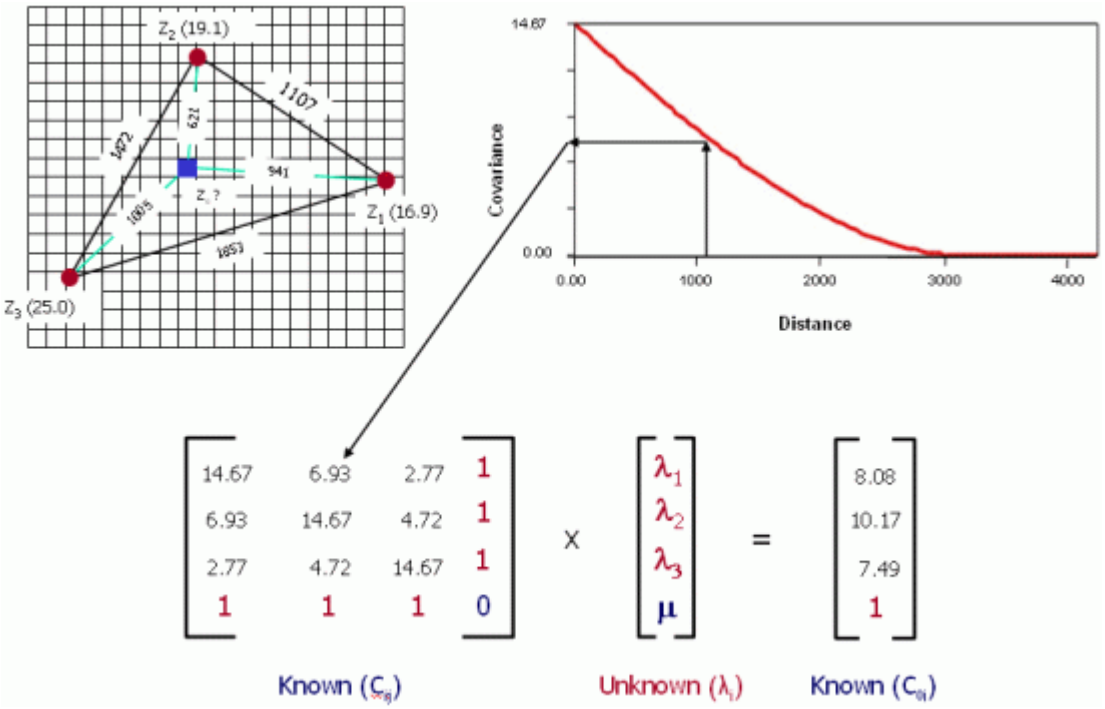
h = Distance between data points and node

a = Range of model



What is used during Kriging





14.67 6.93 2.77 1
6.93 14.67 4.72 1
2.77 4.72 14.67 1
1 1 1 0
X
λ1
λ2
λ3
μ
=
8.08
10.17
7.49
1
Known (Cij)
Unknown (λi)
Known (C0i)

Set up Linear System

$$14.67\lambda_1 + 6.93\lambda_2 + 2.77\lambda_3 + \mu = 8.08$$
$$6.93\lambda_1 + 14.67\lambda_2 + 4.72\lambda_3 + \mu = 10.17$$
$$2.77\lambda_1 + 4.72\lambda_2 + 14.67\lambda_3 + \mu = 7.49$$
$$1\lambda_1 + 1\lambda_2 + 1\lambda_3 + 0 = 1$$

Using the examples above, you would solve for the weights:

$$C_{ij} * \lambda_i = C_{0i}$$

$$\lambda_i = C_{0i} * C_{ij}^{-1}$$

Therefore, the estimated value is:

$$Z_o = \lambda_1 Z_1 + \lambda_2 Z_2 + \lambda_3 Z_3$$

$$Z_o = .2546(16.9) + .4544(19.1) + .2910(25.0)$$

$$Z_o = 20.3$$

And the Kriging variance is:

$$Krig_{\sigma}^2 = C_1 - \Sigma \lambda_i C_{0i} - \mu$$

$$Krig_{\sigma}^2 = 14.67 - (.2546)(8.06) - (.4544)(10.17) - (.291)(7.49) - 0.37$$

$$Krig_{\sigma}^2 = 5.45$$

$$Krig_{\sigma} = 2.33$$

Collocated Cokriging

Collocated Cokriging (COLK) applies kriging with more than one data type. The same basic sets of equations are called for; however, the covariance between the different data types is also required. This method also requires that the secondary data type is known at the sample locations and at every grid node where a value is interpolated. COLK is a direct extension of Linear Regression and is used when two properties are related but the property of interest (e.g., data at the well) is vastly under sampled with respect to the secondary (e.g., seismic data), but correlated, property. COLK should provide a better solution as more information is used during the interpolation. Collocated means that the secondary data is located on the grid and at the well locations when used for interpolation.

In addition to accounting for the proximity of data and the location of the data, cokriging also must account for:

- the correlation between the primary and the secondary data
- the spatial continuity of the primary data to the secondary data
- the spatial configuration of the primary and secondary data, and
- the sampling density of the primary and secondary data.

COLK and kriging estimates are likely to be similar when variables are weakly correlated, when the data are equally sampled or when the cross variogram is the same shape as the direct variograms (the variables have similar spatial correlation). COLK and kriging estimates are equivalent when there is no correlation between the primary and secondary because the variogram goes to zero and the COLK portion of the system of equations vanishes. The COLK estimate is the same as the kriging estimate when the correlation between the primary and secondary data is proportional. Note that when the location of the estimate is beyond the correlation range of both the primary and secondary data, then, as with kriging, the estimate becomes the mean.

The requirement that the secondary variable be known at all locations is not as limiting as one might think. Typically secondary data is seismic and seismic data is often available for all locations required and, if not, it is easy enough to estimate/simulate the attribute at all locations.

Collocated Cokriging is a modification of the general Cokriging case in that it:

- requires only the simple covariance model of the secondary attribute.
- uses all primary data according to search criterion.
- uses secondary data attribute located at the target grid node and each data location during estimation.
- the secondary attribute covariance model is assumed proportional to the primary attribute covariance model.
  - — Product-Moment Correlation Coefficient is the constant of proportionality.
  - — Transforms a univariate covariance model into a multivariate covariance model.
  - — This assumption is termed Markov-Bayes.

Markov-Bayes Assumption

- Cannot guarantee a positive-definite covariance matrix which yields a unique solution
- Modifies the correlation coefficient automatically if a non-positive-definite condition occurs

Estimator

The general Cokriging estimator is expressed as a weighted linear combination of the well data $Z_1, . , Z_n$ and the seismic data $T_1, . , T_m$.

$$Z_0^* = \sum \lambda_\alpha Z_a + \sum \beta_j T_j + \beta_o T_o$$

Where, to ensure unbiasness:

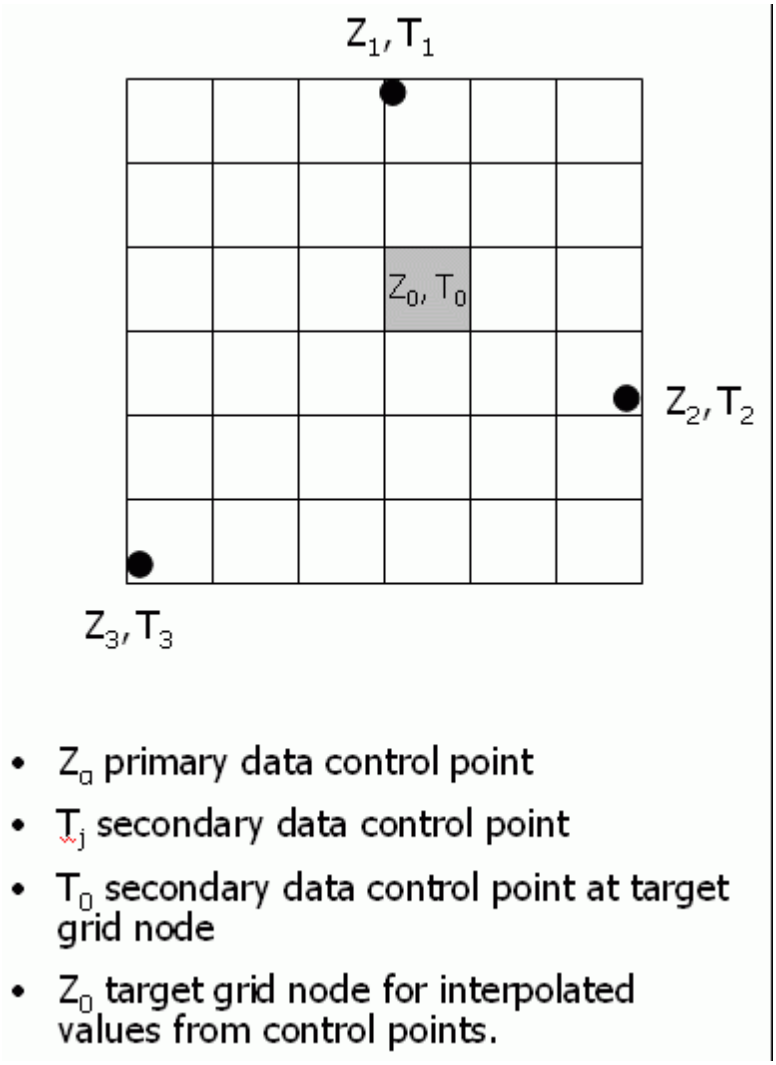
$$\sum \lambda_\alpha = 1$$

$$\sum \beta_j = 0$$

β_o = slope term from a regression between the primary and secondary variable

- Consider the following data configuration. The objective is to estimate the primary attribute at location Z_0* using the secondary

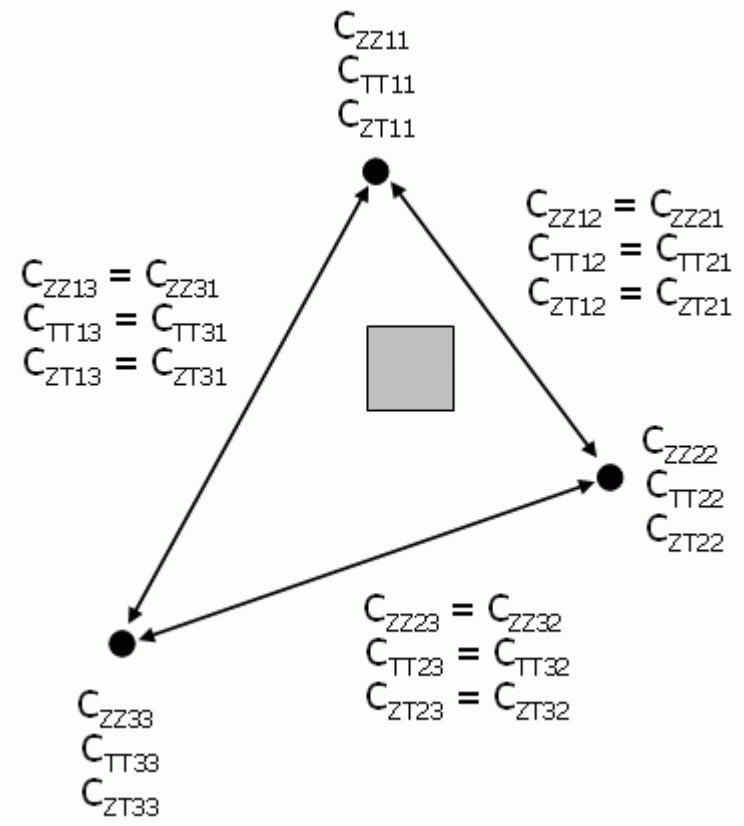
attribute.



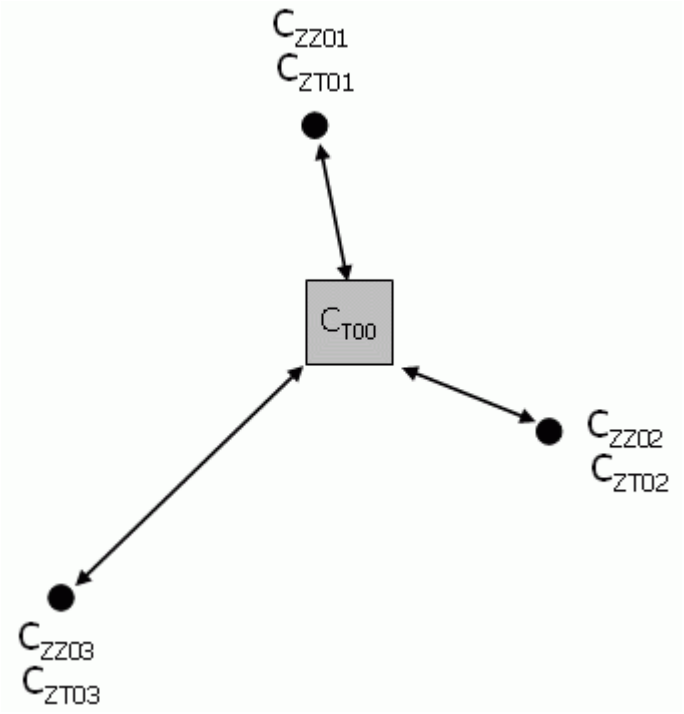
Requirements

- $C_{ZZ}(h)$, spatial covariance model of the primary attribute (well data).
- $C_{TT}(h)$, spatial covariance model of the secondary attribute (seismic data).
- $C_{ZT}(h)$, spatial cross-covariance model of well and seismic data.

Collocated CoKriging: Covariance Requirements Between Control Points



Collocated CoKriging: Covariance Requirements Between Control Points and Target Grid Node



- To solve the matrix equation we simply need to substitute in values of covariance. The result of spatial cross correlation modeling was to produce a model which measures covariance and cross covariance as a function of direction and distance. Use the directions and distances observed above with the spatial correlation model, to determine the covariances.

- Once the values of covariance are known, solve the equation:

$$\begin{bmatrix} C_{z11} & C_{z12} & C_{z13} & C_{zf11} & C_{zf12} & C_{zf13} & C_{zf01} & 1 & 0 \\ C_{z21} & C_{z22} & C_{z23} & C_{zf21} & C_{zf22} & C_{zf23} & C_{zf02} & 1 & 0 \\ C_{z31} & C_{z32} & C_{z33} & C_{zf31} & C_{zf32} & C_{zf33} & C_{zf03} & 1 & 0 \\ C_{zf11} & C_{zf12} & C_{zf13} & C_{f11} & C_{f12} & C_{f13} & C_{f01} & 0 & 1 \\ C_{zf21} & C_{zf22} & C_{zf23} & C_{f21} & C_{f22} & C_{f23} & C_{f02} & 0 & 1 \\ C_{zf31} & C_{zf32} & C_{zf33} & C_{f31} & C_{f32} & C_{f33} & C_{f03} & 0 & 1 \\ C_{zf01} & C_{zf02} & C_{zf03} & C_{f01} & C_{f02} & C_{f03} & C_{f00} & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_{z1} \\ \lambda_{z2} \\ \lambda_{z3} \\ \beta_{f1} \\ \beta_{f2} \\ \beta_{f3} \\ \beta T_o \\ \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} C_{z01} \\ C_{z02} \\ C_{z03} \\ C_{zf01} \\ C_{zf02} \\ C_{zf03} \\ C_{f00} \\ 1 \\ 0 \end{bmatrix}$$

- Once the weights are known, they can be inserted into the original equation and finally the estimate can be determined:

$$Z_0^* = \sum \lambda_\alpha Z_a + \sum \beta_j T_j + \beta_o T_o$$

Advantages and Limitations of Collocated Cokriging

Advantages:

- Allows incorporation of correlated, secondary data into the mapping process.
- Can calibrate and control the influence of the secondary data through the correlation coefficient.
- When compared to traditional least- squares regression, the cokriging technique honors the primary data and accounts for spatial correlation in the variations of the primary data.
- Yields more accurate estimates than simple single variable kriging.

Limitations:

- Cokriging tends to produce a smoothed image, but not as smooth as kriging.
- Inferring a correct linear correlation model is difficult for sparse well data.

Appendix B
Facies Simulation

DecisionSpace® Earth Modeling offers the traditional Sequential Indicator Simulation facies modeling method and a rules based method known as Truncated Gaussian Simulation; both of which are based on sequence stratigraphic principles for the simulation of categorical variables such as geological facies or lithotypes.

Sequential Indicator Simulation

Sequential Indicator Simulation (SIS) is the most frequently used of the facies simulation methods because of its availability in all commercial earth modeling software. The general perception is that because it is "easier" to use than other methods, it is the better choice. However, it has several major limitations and incorrect implementation in most commercial offerings has led to misconceptions and poor modeling practices (Emery, 2004). The major limitations are:

- The mathematical constrains of the method require that only one indicator variogram should be used for all facies within a reservoir interval, even though most commercial offerings allow one indicator variogram for each facies. The result is that the facies proportions don't always add to 100% so a post processing correction is made to ensure a correct summation. In DecisionSpace® Earth Modeling only one variogram is model for each interval. Then the variogram for the other facies are created automatically such that their sills are scales to the facies proportions, ensuring that the summation is 100% without a post processing step.

- Like Sequential Gaussian, SIS requires the use of previously simulated points.

- Boundary conditions between the facies transitions are not always honored resulting invalid contact between some facies. This is corrected partially with the use of a 3D proportion matrix.

Truncated and Plurigaussian Simulation

For these methods, the indicator variogram is not modeled directly, instead indicator data are transformed into the proportions of each indicator variable, thus creating a continuous Gaussian variable. The continuous Gaussian variable is modeled and then simulated using the Turning Bands simulation algorithm and finally the facies are obtained by applying thresholds (cutoffs) to the Gaussian simulated values.

When first implemented this method was known as Truncated Gaussian Simulation (TGS) and the method implies a strict sequence to the facies ordering. Plurigaussian Simulation (PGS) is a direct extension of TGS allowing for more flexibility in the facies relationships with the use of two variogram models; each controlling the orientation and spatial continuity of two lithotype sets. The fundamental concept of non-stationary (trends) proportion curves is central in TGS and PGS, where the so-called rock type rule plays an essential role in producing realistic models that represent the transitions between the different facies. The key point is that the Gaussian variables and the indicators are linked by means of the thresholds but, even if the indicators are not stationary, they can be obtained by truncation of stationary Gaussian variables. TGS and PGS are the recommended methods for facies simulation. A detailed review is given in Armstrong et al. (2003).

1. Block the wells

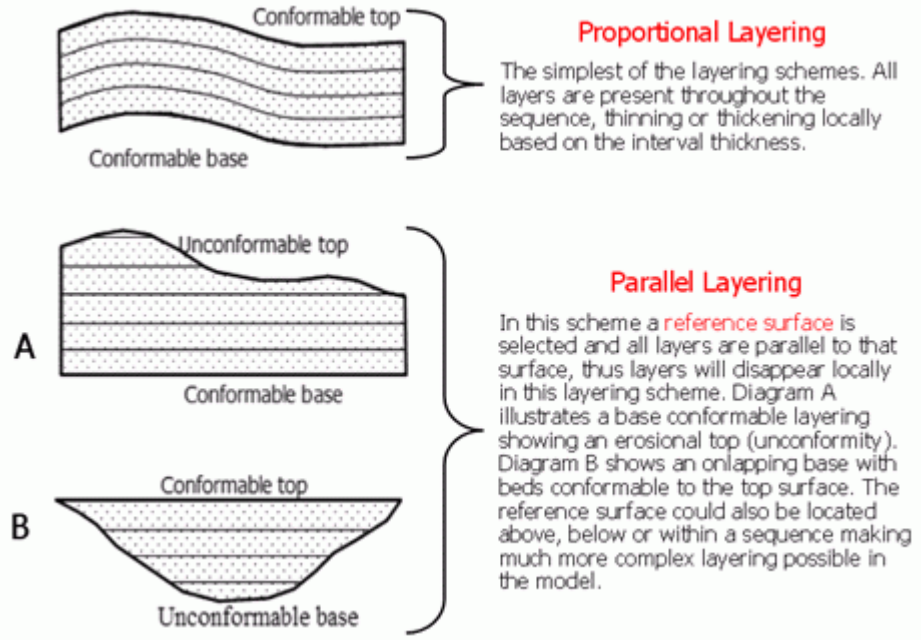
2. Create the Lithotype Proportion Map using regular or custom polygons within the Lithotype Proportion Map Creation task panel. These statistics are highly dependent on the choice of the layering scheme (parallel or proportional bedding) within each reservoir interval; with the layering choice dependent upon the depositional system.

3. Choice of a model describing the relationships between the different facies. This includes the definition of the lithotype rules and the variogram model for each lithotype set.

4. Simulation of the two Gaussian variables followed by truncation to obtain the facies indicators.

Choice of a Reference Surface and Layering Style

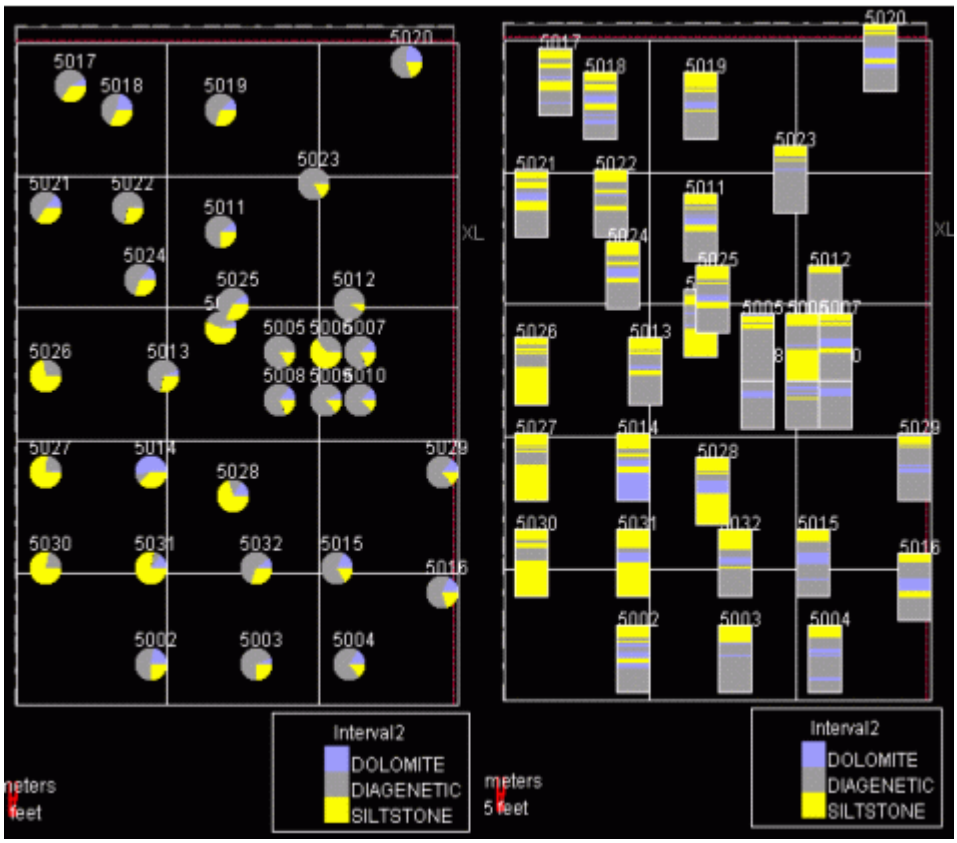
The choice of the layering type with respect to the bounding surfaces of each reservoir interval has consequences on all stages of the process, for example in data analysis and the facies simulation and petrophysical attribution results. In a sedimentary context, the reference surface represents the direction perpendicular to the deposition of the different facies. When comparing facies parallel to that surface, more similarity is expected and consequently higher correlation between the well control. As a result, the simulated results will force the facies to be stacked parallel to the reference surface. In the current release there are three layering style choices: 1) Proportional, 2) Parallel to Top, and 3) Parallel to Bottom.



Vertical Proportion Curves and Pie Charts

The well logs are blocked into the Stratigraphic geocellular grid based on the layering style and vertical grid resolution. The new blocked pointset values are placed at the cell's gravity center.

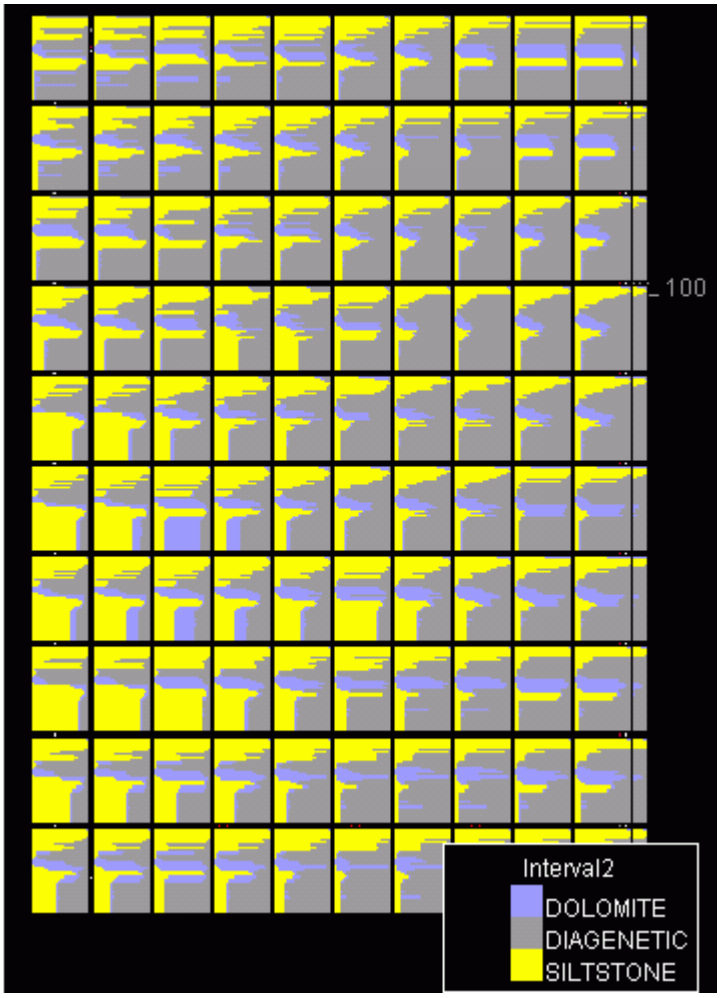
For each interval, regular polygons or custom polygons are drawn around groups of wells with similar lithotype proportions based on the use of pie charts or vertical proportion curves (VPC) displayed at each well location. Once the polygons are defined, an average VPC is calculated within each polygon and displayed at the polygons gravity center either as a pie chart or as a curve. The use of multiple VPCs is to account for lateral and vertical changes in the lithotype proportions across the area-of-interest (AOI).



Pie Charts (left) and Vertical Proportion Curves for each well (right)

3D Proportion Map

Once the average VPCs are created for each polygon, the proportions are interpolated onto each cell of the 3D geocellular grid using an inverse distance interpolation algorithm. This step provides a 3D matrix of the proportions for each interval and is used to calculate the local thresholds on the continuous Gaussian variables. The proportion map serves as a template used in the facies simulation to help control the vertical and lateral transitions, thus accounting for trends in the lithotypes.



Vertical Proportion Map created from Averaged Vertical Proportion Curves

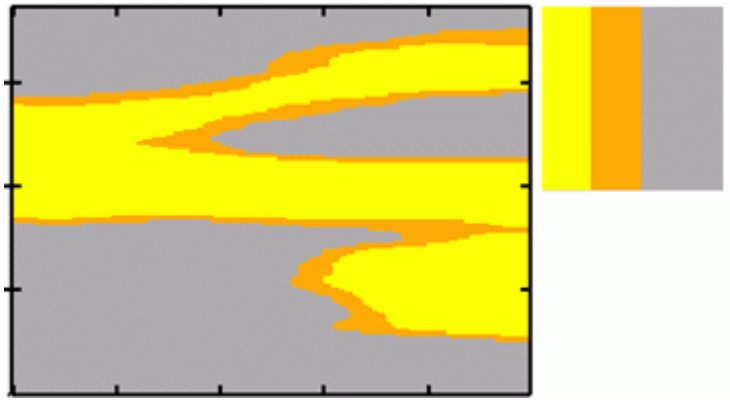
Lithotype Rule

Knowledge of the lithotype proportions is not sufficient to derive the values of the thresholds to be applied on the simulated Gaussian values. Additional information on the relationships between the lithotypes is required. Depending on the number of lithotypes, there are a finite number of rules that may represent the relationships between the Gaussian random functions and the lithotypes. From these we must design a rule set that makes the most sense from a geological point of view, taking into consideration the probable transitions between the lithotypes.

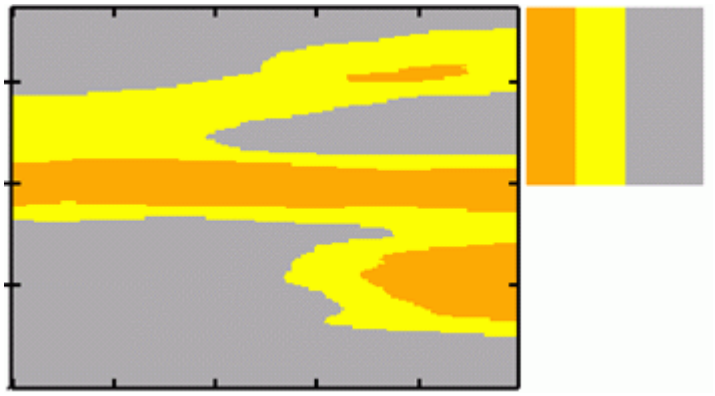
For example, let us consider that we are modeling a depositional non-marine migrating channel environment and we defined only 3 lithotypes. With 3 lithotypes there are 9 possible ways in which to define the facies rules, some of which are mirror images and create the same facies patterns.

Let's start with simplest case where there is a strict transition from one facies to another, requiring only one variogram model. There are 3 possible solutions.

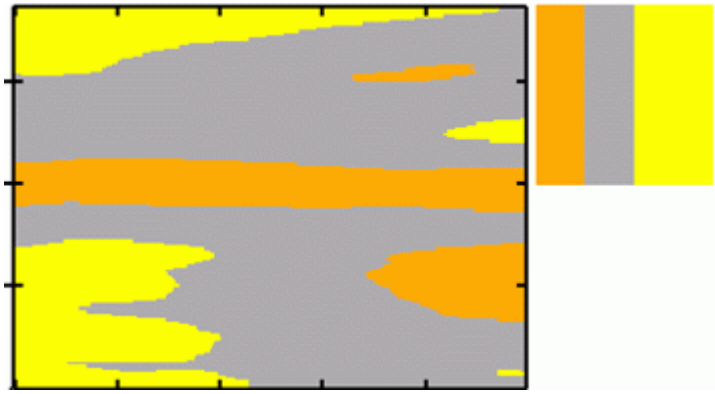
Case 1: Yellow is channel, orange is levee and gray is overbank shale. Notice the facies transition from channel to levee to shale, with shale never contacting the channel lithotype.



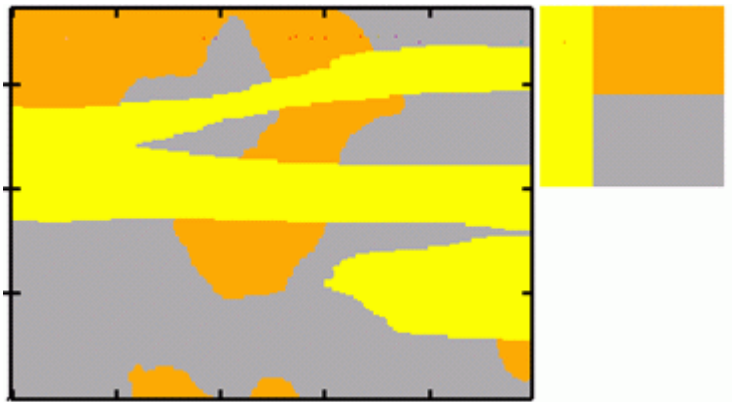
Case 2: The yellow and orange are switched creating a different lithotype relationship, yellow is still channel, but orange is now considered as a channel lag.



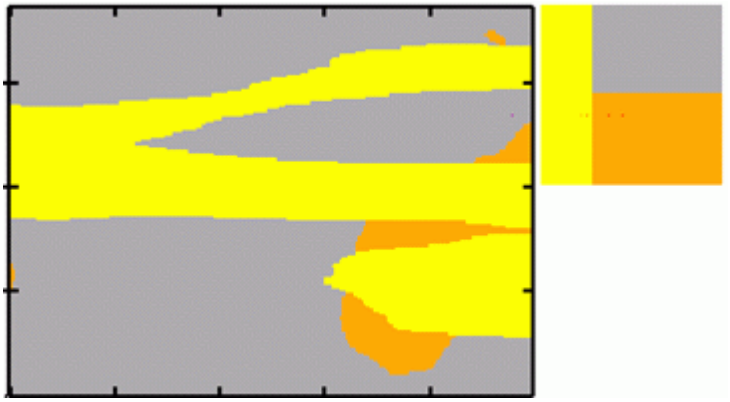
Case 3: Here the rules probably don't make too much sense as the shale contacts both the sand facies (yellow and orange).



Now let's use the same 3 lithotypes, but add a cross-cutting or an attachment based rule set, requiring two variograms.



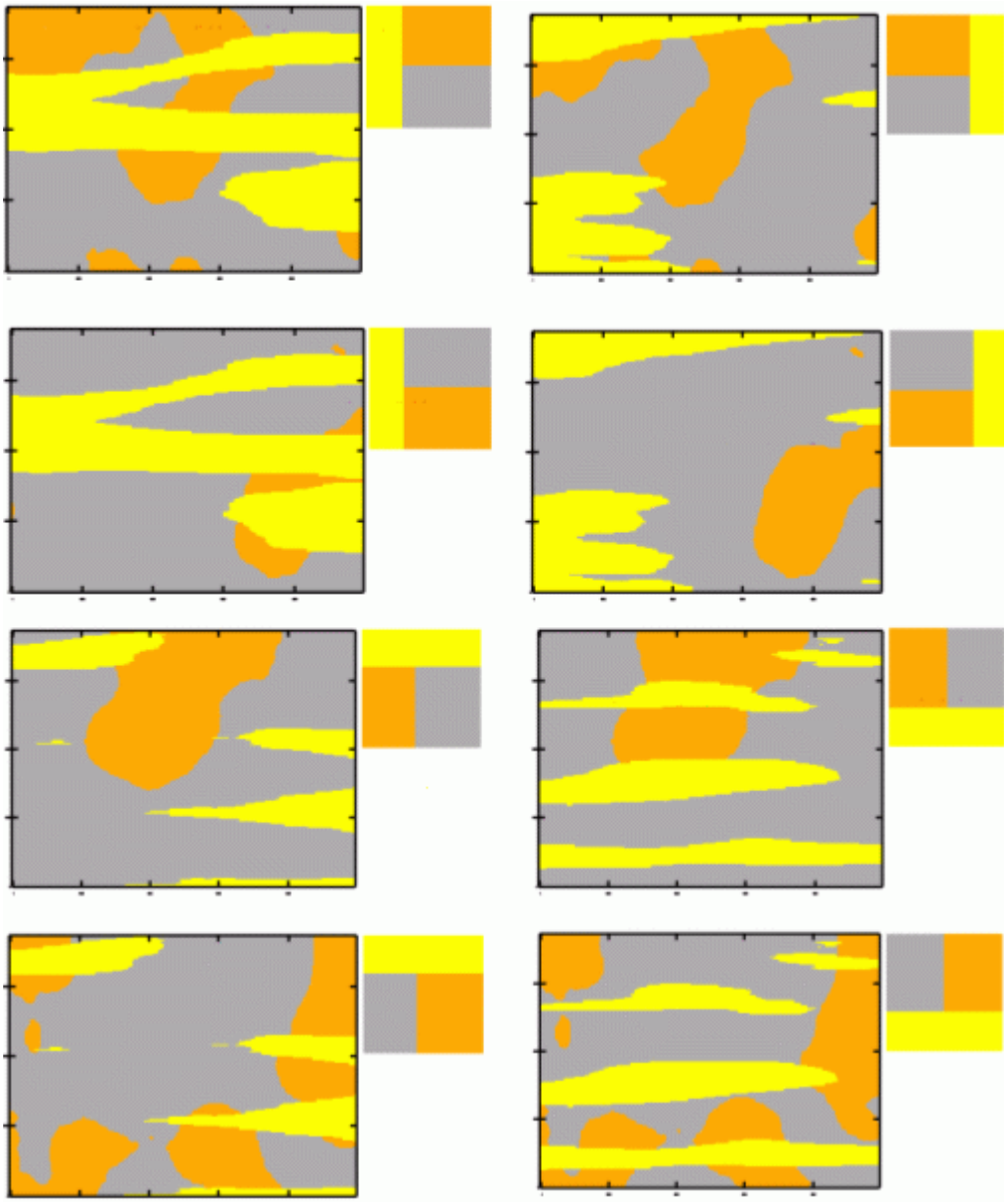
In the example below, we see a channel (yellow) controlled by a variogram with a lone E-W correlation scale comprising Lithotype Set 1. Lithotype Set 2, splay (orange) and shale (gray), is controlled by the second variogram model with the different orientation and anisotropy ratio.



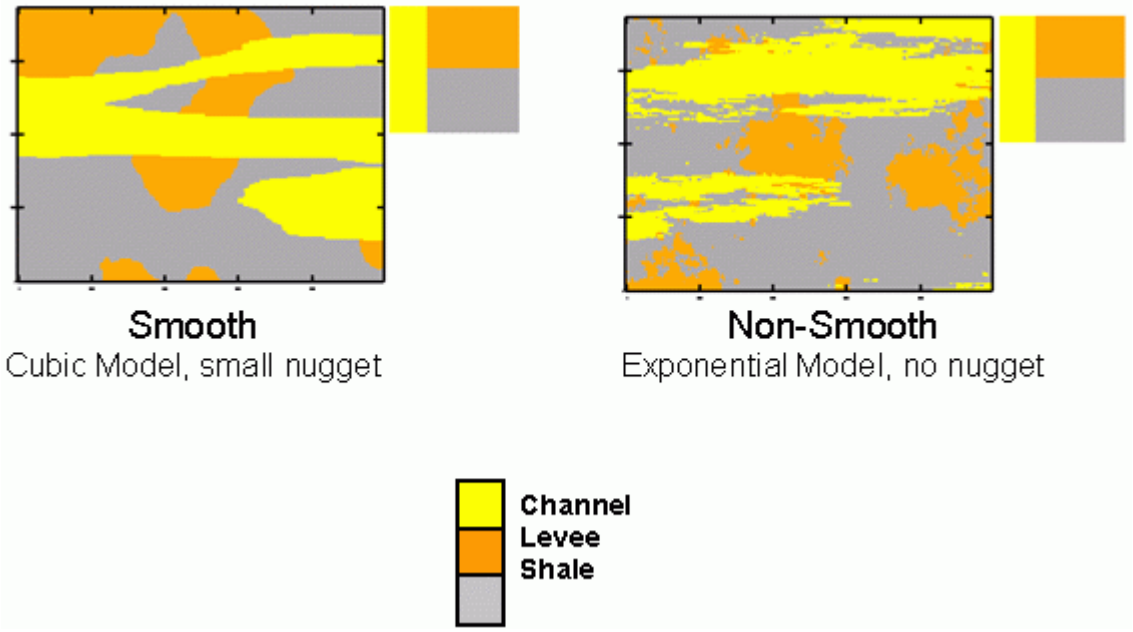
The scenario is the same as in the prior examples except that the splay (orange) and shale (gray) lithotype were switched in position; however, the rules are still honored. The difference between these images is only when the thresholds are determined at the end of the simulation to create the final facies (lithotype) model. For the upper situation the truncation occurs yellow, gray and then orange, but for the lower case the truncation occurs yellow, orange and then gray. So it doesn't matter which Rule is used provided the Lithotype Sets are correct. Note that the same variogram models are used to create both images.

Let's look at some other possibilities using the same 3 lithotypes and rule relationships to see the differences in the simulation

All of the images shown below have the same lithotype relationship but the rule box is rotated and the orange and gray lithotype are switched. In all cases where the Yellow lithotype is a vertical bar it is Lithotype Set 1 and controlled by its own variogram model, when is it a horizontal bar it is known as Lithotype Set 2, but it uses the same variogram model. When the orange and gray lithotypes are horizontal they are Lithotype Set 2 and are controlled by their own variogram model, and when vertical, they are Lithotype Set 1.

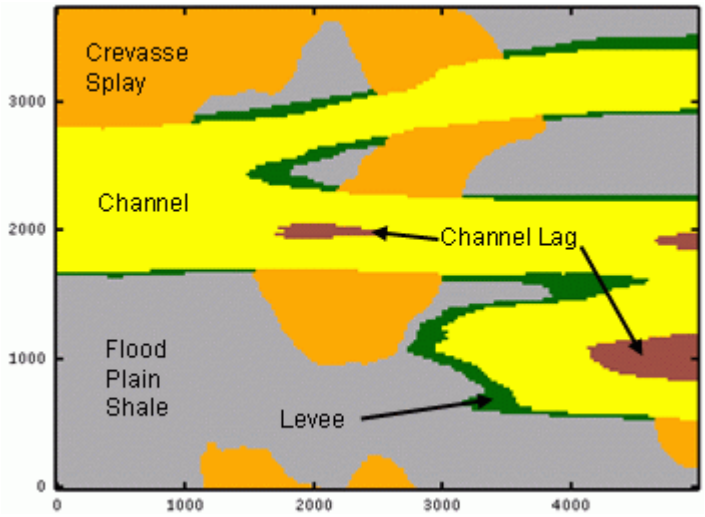


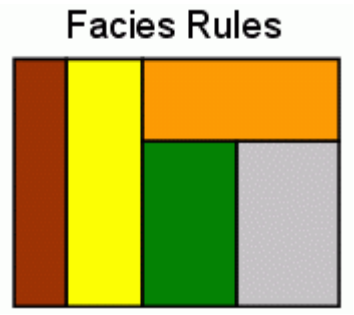
Within DecisionSpace® Earth Modeling there are two variogram model types: Smooth (Cubic model with a small Nugget) and Non-Smooth (Exponential model). The results of these models are shown below.



Lithotype Sets

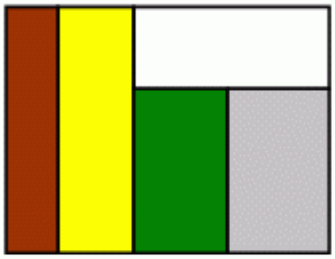
Now let's look at the definition of the Lithotype Sets and which variogram model controls each Set. The example below shows 5 lithotypes in a fluvial environment. The image is created by unconditional simulation using the Facies Rules shown below the display.





The Lithotype Sets and associated variograms are:

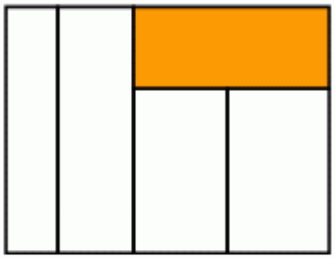
Litho Set 1 (Vertical)



Variogram Model

- Major Scale: 8000
- Minor Scale: 1000
- Major Axis: 90 degrees
- Model Type: Smooth

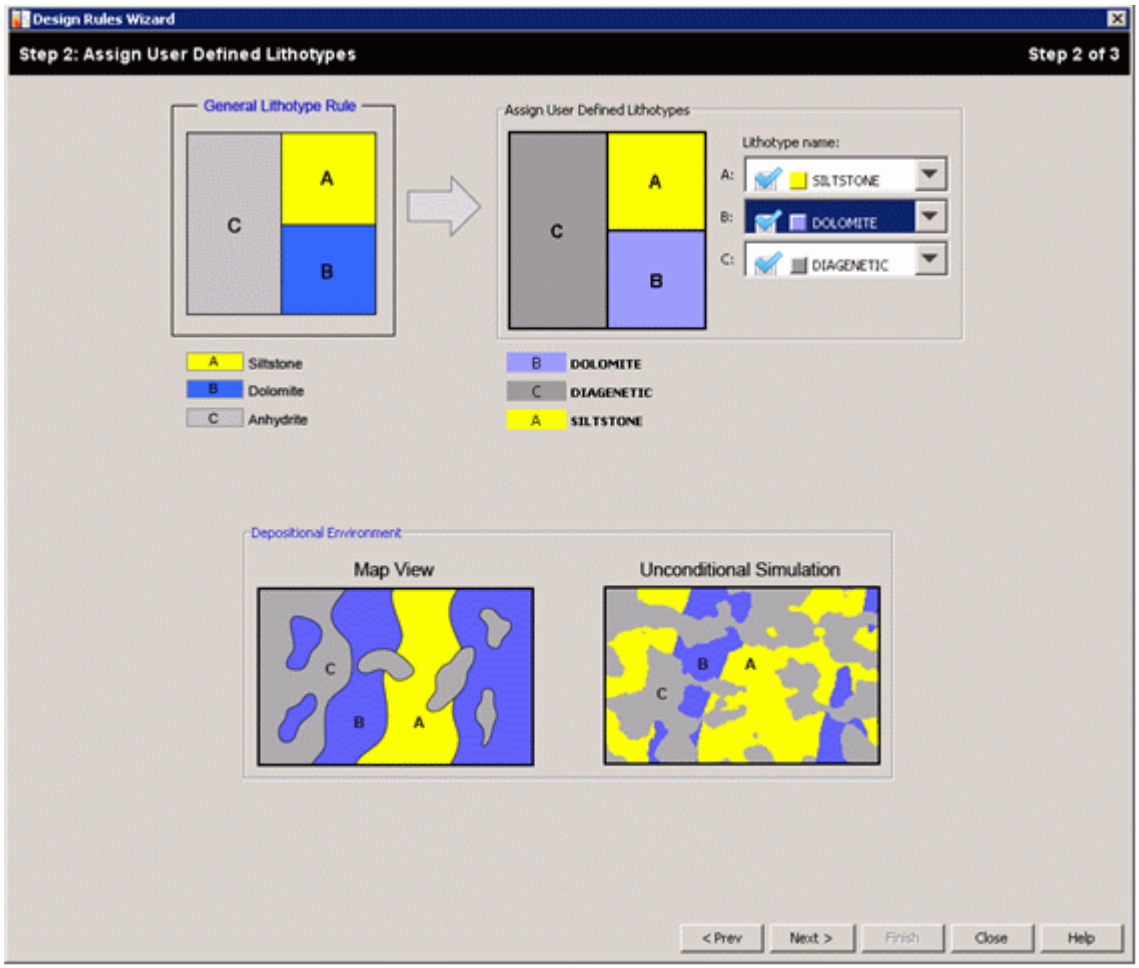
Litho Set 1 (Horizontal)



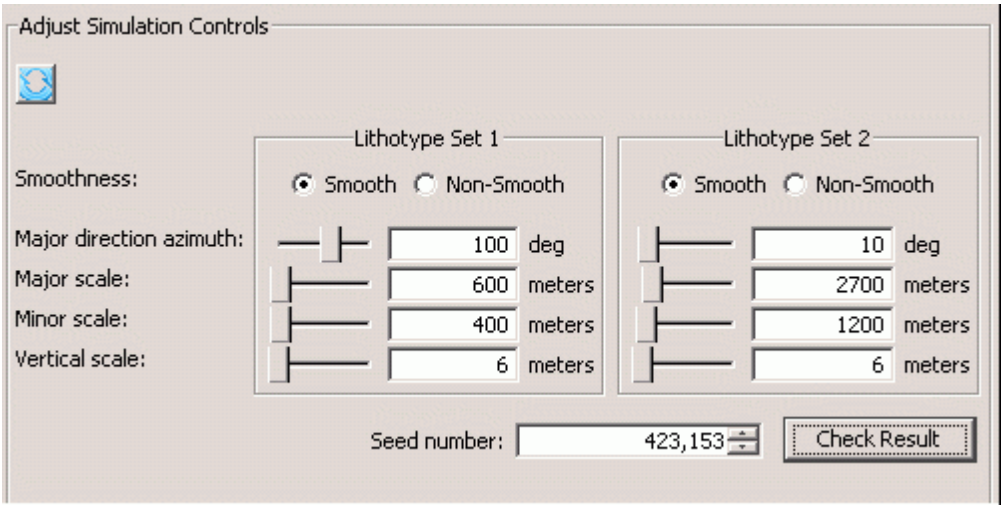
Variogram Model

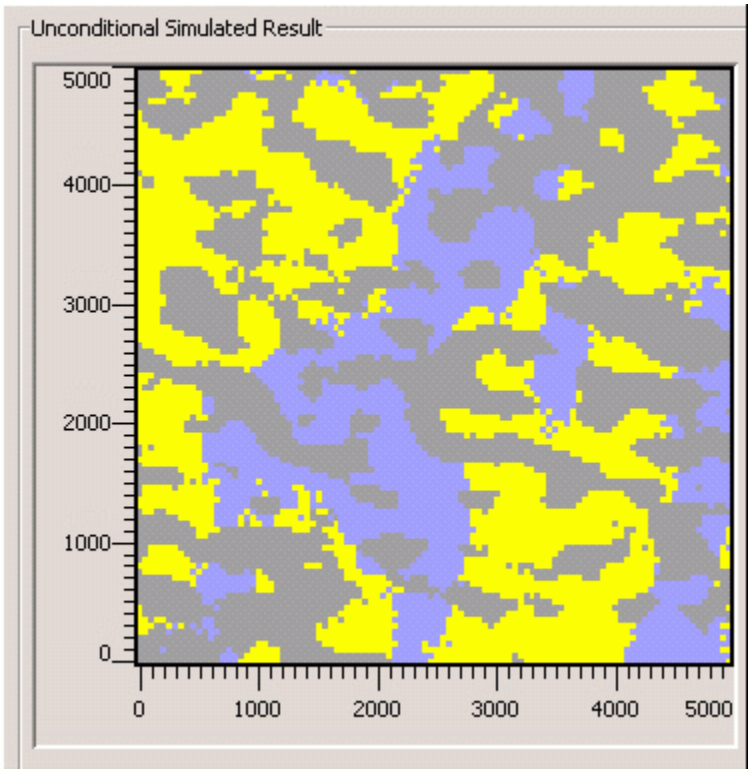
- Major Scale: 3500
- Minor Scale: 3200
- Major Axis: 0 degrees
- Model Type: Smooth

The Lithotype Rules shown in the image below are for the EM_TRAINING_01 data set. The Map view is the conceptual depositional model and the Unconditional Simulation illustrates what is expected using the above Lithotype Rule.



Once the lithotype rule is defined, the variograms for each Lithotype Set can be modeled and the parameters specified. The role played by the proportion curve maps does not mean that the choice of the variogram has no consequence. The following image illustrates the variogram parameters set for this data set and the resulting unconditional simulation.





Conclusion

The lithotype simulation method offered in DecisionSpace® Earth Modeling has proved to be very efficient in providing images reproducing the main features geological features of clastic and carbonate environments.

References

Armstrong A., Galli A., Le Loc'h G., Geffroy G. and Eschard R., Plurigaussian Simulations in Geosciences, Springer, 2003, pp 149.

Emery, X., 2004, Properties and limitations of sequential indicator simulation, in *Stochastic Environmental Research and Risk Assessment*, Vol 18, No.6, p. 414- 424

Appendix C
Terminology

The list of geostatistical terminology selected represents the most commonly encountered terms and we have not attempted to duplicate Ricardo Olea's (1991) glossary of terminology. We defined the terms as used while teaching geostatistics, so they may be slightly different from Olea's definitions.

Admissibility (of semivariogram models): for a given covariance model, the Kriging variance must be 0, this condition is also known as positive definite. There are only a limited number of functions (covariance models) that satisfy this condition. A least-squares fitting cannot ensure positive definiteness.

Anisotropy: in geostatistics, anisotropy refers to covariance models that have major and minor ranges of different distances (correlation scale or lengths). There are two types of anisotropy: Geometric anisotropic covariance models have the same sill, but different ranges; Zonal anisotropic covariance models have the same ranges, but different sills.

Auto-correlation: to compute a spatial covariance model for regionalized variable. It measures a change in variance (variogram) or correlation (correlogram) with distance and/or azimuth.

Biased estimates: there is a correlation between the standardized errors and the estimated values (see Cross-Validation). A histogram of the standardized errors is skewed. Either of these conditions suggests that there is a bias in the estimates, so that there is a chance that one area of the map with always show estimates higher (or lower) than expected.

Block Kriging: making a Kriging estimate over an area, for example estimating the average value at the size of the grid cell. The grid cell is divided into a specified number of sub-cells, a value is Kriged to each sub-cell, and then the average value is placed at the grid node.

Boolean Models: see Object Models

CoKriging: the process of estimating a regionalized variable from two or more variables using a linear combination of weights obtained from models of spatial auto-correlation and cross-correlation; this is the multivariate version of Kriging.

Conditional Simulation: a geostatistical method to create multiple equally probable images of a regionalized variable based on a spatial model. It is conditional only when the actual control data are honored. Conditional simulation is a variation of conventional Kriging or CoKriging. By relaxing some of the Kriging constraints (e.g. minimized square error), conditional simulation is able to reproduce the variance of the control data (the histogram). The final "map" captures the heterogeneity and connectivity mostly likely present in the reservoir. Post processing conditional simulations produces a measure of error (standard deviation) and other measures of uncertainty, such as iso-probability and uncertainty maps.

Correlogram: a measure of spatial dependence (correlation) of a regionalized variable over some distance. The correlogram can also be calculated with an azimuthal preference.

Covariance: the Kriging system uses covariance, rather than the variogram or correlogram values, to determine the Kriging weights. The covariance is the sill minus the variogram model (or zero minus the correlogram).

Coregionalization: the mutual spatial behavior between two or more regionalized variables and fitting the models.

Cross-correlation: to compute a spatial cross-covariance model between two regionalized variables. This provides a measure of spatial correlation between the two variables.

Cross-Validation: a procedure to check the compatibility between a data set, its spatial model and neighborhood design. It is also used to check for biased estimates produced by poor model and/or neighborhood design.

Drift: often used to describe data containing a trend. Drift usually refers to short scale trends at the size of the neighborhood.

Estimation variance: the Kriging variance at each grid node. This is a measure of global reliability, not a local estimation of error.

External Drift: when a secondary regionalized variable (e.g. seismic attribute) is used to control the shape of the final map created by kriging or simulation. This geostatistical linear regression technique uses a spatial model of covariance.

Geostatistics: the statistics of spatially (or temporally) correlated data.

h-**Scatterplot**: select a value for separation distance, h, then plot the pairs Z(x) and Z(x+h) as the two axes of a bivariate plot, the shape and correlation of the cloud is related to the value of the variogram for distance, h.

Histogram: a plot, which shows the frequency or number of occurrences (Y-axis) of data, located in size classes of equal width (X-axis).

Iso-probability map: these maps are created by post processing conditional simulations. The maps show the value of the regionalized variable at a constant probability threshold, for example at the 10th, 50th (median), or the 90th percentiles. Note that only the 50th percentile is an acceptable solution as any other computed quantile is a fabrication and not a real solution.

Kriging: method of calculating estimates of a regionalized variable using a linear combination of weights obtained from a model of spatial correlation, this is the univariate version of CoKriging.

Kriging variance: see estimation variance

Lag: a distance parameter used during computation of the experimental covariance model. The lag distance typically has a tolerance of one-half the initial lag distance. A good initial lag is the average distance between control points.

Moving neighborhood: a search neighborhood designed to use only a portion of the control data point during kriging or conditional simulation.

Nested variogram model: a linear combination of two or more variogram (correlogram) models. For example, a short-range exponential model combined with a longer-range spherical model.

Nonconditional simulation: a method that does not use the control data during the simulation process, but only the normalized data histogram and the spatial covariance model. A method quite often used to observe the behavior of a spatial model and neighborhood design.

Nugget effect: a feature of the covariance model where the experimental points defining the model does not appear to intersect at the origin. The nugget model shows constant variance at all ranges, but is often modeled as zero variance at the control point (well location).

The Nugget effect is often assumed to be related to low data measurement precision, but is more likely related to a sampling bias where the sample spacing is much larger than the feature sampled, such that it is never possible to have two sampled locations within the same feature.

Object Models: this is a stochastic simulation technique whose function is to model objects such channels, splays, bars, or any geological feature. The Marked-point Process can model any shape provided the geometries can be specified (height, length, width, thickness, sinuosity, wavelength, etc.), with tolerances as to the mean and standard deviation. Such modeling is used only for facies.

Ordinary (Co-)Kriging: the local mean varies and is re-estimated based on the control points in the search neighborhood ellipse (moving neighborhood).

Outliers: from a statistical definition, these are data points falling outside about 2.5 standard deviation of the mean value of the sample population. These could be the result of bad data values or local anomalies.

Parallel Layering: when defining the layering or bedding scheme during the design of a geocellular grid, parallel layering requires the User to select a reference surface, typically the top or base horizon defining the reservoir interval and a layer thickness. Parallel to a top produces onlap or baselap, with lower layer pinching out, whereas Parallel to a base creates erosion or toplap with truncated layers.

Pluri-Gaussian Simulation (PGS): an extension of the Truncated Gaussian Simulation method which overcomes the limitation of simulating a strict transition of facies. PGS can model complex facies relationship and uses two variogram ones, one for each facies set.

Point Kriging: making a Kriging estimate at a specific point, for example at a grid node, or a well location.

Positive Definite: see admissibility

Proportional Layering: when defining the layering or bedding scheme during the design of a geocellular grid, Proportional layering divided the reservoir interval into an equal number of layers, locally thinning or thickening based on the local isochore thickness.

Random function: the random function has two components: (1) a regional structure component manifesting some degree of spatial auto-correlation (regionalized variable) and lack of independence in the proximal values of Z(x), and (2) a local, random component (random variable).

Random variable: created by some random process, whose values follow a probability distribution, such as a normal distribution.

Range: the distance where the variogram reaches the sill or when the correlogram reaches zero correlation. Also known as the correlation range or correlation scale.

Regionalized variable: a variable that has some degree of spatial auto-correlation and lack of independence in the proximal values of Z(x).

Risk Map: see Uncertainty Map

Simple Kriging: the global mean is constant over the entire area of interpolation and is based on all the control points used in a unique neighborhood (or is supplied by the user).

Sequential Simulation: geostatistical algorithms used to simulate continuous properties (sequential Gaussian) or discrete properties, such as facies (sequential indicator). These methods use previously simulated values to update the local CDF and requires a careful neighborhood design to avoid shortscale artifacts in the simulated result.

Semivariogram: a measure of spatial dependence (dissimilarity or increasing variability) of a regionalized variable over some distance. The semivariogram can also be calculated with an azimuthal preference. The semivariogram is commonly called the variogram. See also correlogram.

Sill: the level of variance where the variogram reaches its correlation range. The variance of the sample population is the theoretical sill of the variogram.

Stationarity: the simplest definition is that the data do not exhibit a trend. This implies that a moving window average shows homogeneity in the mean and variance over the study area.

Stochastic modeling: used interchangeably with conditional simulation, although not all stochastic modeling applications necessarily use control data.

Turning Bands Simulation: performs a Gaussian simulation along a series of one-dimensional lines (bands) in 1-2- or 3-D. There is no restriction on the number of random fields (bands) to simulate or the basic type of covariance models used. The simulation can be conditional or unconditional. This is the first simulation algorithm used for petroleum applications and is a very fast and efficient algorithm.

Truncated Gaussian Simulation: this method is used to simulate facies when there is a strict transition of one facies to another, for example the ABC, where facies A contacts B and will not transition into C, unless B is missing from the sequence. This method converts facies indicators into probabilities and used the Turning Bands simulation algorithm to "map" probabilities. Once the probabilities of each facies type is simulated, the Gaussian variable are truncated based on the proportion of the input facies data.

Unique neighborhood: a neighborhood search ellipse that uses all available data control points. The practical limit is 100 control points. A unique neighborhood is used with simple Kriging.

Uncertainty map: these are maps created by post processing conditional simulations. A threshold value is selected, for example, 8% porosity, an uncertainty map shows at each grid node, the probability that porosity is either above or below the chosen threshold.

Variogram: see semivariogram

Weights: values determined during an interpolation or simulation, which are multiplied by the control data points in the determination of the final estimated or simulated value at a grid node. To create a condition of unbiasedness, the weights, sum to unity for geostatistical applications.

Reference

Olea, R. A., 1991, Geostatistical Glossary and Multilingual Dictionary, New York, Oxford University Press, 192 p.

Appendix D
Selected References

In compiling a bibliography on geostatistics, we are faced with the difficult decision of selecting only a few of the many excellent books and articles that have been published. The small set chosen represents some of the more classical papers and those that have been helpful in our endeavor to understand geostatistics. Some of these are difficult reading, because they are laced with jargon and mathematical notation.

For a very comprehensive bibliography see the following reference:

Coburn, T.C, 2006, Bibliographic Resources for Geostatistics, Spatial Analysis and Stochastic Modeling in Petroleum Exploration and Development., Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp.325-409.

Books

Agterberg, F. P., 1974, Geomathematics, New York, Elsevier Scientific Publishing, 596 p.

Armstrong, M., A. G. Galli, G. LeLoc'h, F. Geffroy, and R. Eschard, 2003, Plurigaussian Simulations in Geosciences, Springer-Verlag, New York, 149 p.

Chilès, J.-P. and P. Delfiner, 1999. Geostatistics: modeling spatial uncertainty. New York: John Wiley & Sons, 695 p.

Clark, I., 1977. Practical kriging in three dimensions. Computers & geosciences, vol. 3, pp. 173-180.

Clark, I., 1979, Practical Geostatistics, London, Applied Science Publishers, 129 p.

Cosentino, L., 2001, Integrated Reservoir Studies, Editions TECHNIP, Institut Français Du Pétrole Publications, Paris, France, 310 pp.

Cressie, N., 1991, Statistics for Spatial Data, New York, Wiley, 900 p.

David, M., 1977, Geostatistical Ore Reserve Estimation, New York, Elsevier Scientific Publishing, 364 p.

Davis, J. C., 2002 (Third Edition), Statistics and Data Analysis in Geology, Second Edition, John Wiley & Sons, New York, 637 p.

Deutsch, C. V. and A G. Journel, 1997, GSLIB: Geostatistical Software Library and User's Guide, Second Edition, Oxford University Press, New York, Oxford, with software diskettes, 369 pp.

Deutsch, C.V., 2002, Geostatistical Reservoir Modeling, Applied Geostatistics Series, Oxford University Press, New York, 376, pp.

Dubrule, O., 1998, Geostatistics in Petroleum Geology, AAPG Continuing Education Course Notes Series #38, Pub. By AAPG, Tulsa, OK.

Dubrule, O., 2003, Geostatistics for Seismic Data Integration in Earth Models, 2003 Distinguished Instructor Short Course No. 6, Pub. By EAGE, and SEG, Tulsa, OK.

Goovaerts, Pierre, 1997. Geostatistics for natural resources evaluation. New York: Oxford University Press, 483 p.

Henley, S., 1981, Nonparametric Geostatistics, Elsevier Applied Science Publishers LTD., Essex, UK, 145 p.

Hohn, M. E., 1999, Geostatistics and Petroleum Geology, Kluwer Academic Publishers, Dordrecht, The Netherlands, 235 pp.

Isaaks, E. H. and R. M. Srivastava, 1989, An Introduction to Applied Geostatistics, Oxford University Press, Oxford, 561 p.

Jensen, J.L, L. Lake, P. Corbett and D. Goggin, (2003), Statistics for Petroleum Engineers and Geoscientists, 2nd Edition, Elsevier, 338 p.

Journel, A. G., 1989, Fundamentals of Geostatistics in Five Easy Lessons, Short course in Geology, Vol. 8, American Geophysical Union, 40 p.

Journel, A. G. and Huijbergts, 1978, Mining Geostatistics, Orlando, Florida, Academic Press, 600 p.

Leuangthong, O., K. D. Khan and C. V. Deutsch, 2008, Solved Problems in Geostatistics, John Wiley & Sons, Inc., New Jersey, 207 pp.

Matheron, G., 1962, Traite de Geostatistique Appliquee, tome 1, Editions Technip, Paris

Matheron, G., 1971. The Theory of Regionalized Variables and Its Application. Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau, 5.

Olea, R. A., 1991, Geostatistical Glossary and Multilingual Dictionary, New York, Oxford University Press, 177 p.

Stochastic Modeling and Geostatistics, (1994), J. M. Yarus and R. L. Chambers, Eds. AAPG Computer Applications in Geology 3, 379, pp.

Stochastic Modeling and Geostatistics, Volume II, (2006), T. C. Coburn, J. M. Yarus and R. L. Chambers, Eds. AAPG Computer Applications in Geology 5 , 409, pp.

Wackernagel, H., 1995, Multivariate Geostatistics: An Introduction with Applications, Springer-Verlag, Berlin, 256 p.

Papers

Abry, C. G., 1975. Geostatistics model for predicting oil: Tatum Basin, New Mexico. AAPG Bulletin, vol. 59, pp. 2111-2127.

Ahmed, U., S. F. Crary and G. R. Coates, 1991, Permeability estimation: the various sources and their interrelationships, JPT, May.

Alabert, F. 1989, Constraining description of randomly heterogeneous reservoirs to pressure test data: a Monte Carlo study, Annual Technical Conference and Exhibition, SPE 19600.

Anguy, Y., R. Ehrlich, C. M. Prince, V. L. Riggert, and D. Bernard, 1994, The Sample Support Problem for Permeability Assessment in Sandstone Reservoirs, in Stochastic Modeling and Geostatistics, (1994), J. M. Yarus and R. L. Chambers, Eds. AAPG Computer Applications in Geology, No. 3, pp. 273-286.

Araktingi, U. G., W. M. Bashore, T. T. B. Tran, and T. Hewett, 1993, Integration of Seismic and Well Log Data in Reservoir Modeling, in Reservoir Characterization III, B. Linville, Ed., Tulsa, Oklahoma, Penwell Publishing, pp. 515-554.

Bahorich, M., L. Peyton, A. Vassiliou, R. L. Chambers, G. Partyka, and J. M. Ross (1996), The Use of Seismic Attributes for a Geostatistical Based Integrated Reservoir Description and Uncertainty Assessment, 17th Annual Research Conference, Gulf Coast Section, SEPM.

Bashore, W. M., U. G. Araktingi, M. Levy, and W. J. Schweller, 1994, Importance of a Geological Framework and Seismic Data Integration for Reservoir Modeling and Subsequent Fluid-Flow Predictions, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 159-176.

Behrens, R. A. and T. T. Tran, 1998, Incorporating seismic data of intermediate resolution into 3D reservoir models, 1998 SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, Sept. 27-30, SPE 49143, pp. 497-507.

Beucher-Darricau, H. B. Doligez and J. M. Yarus, 2006, Modeling Complex Reservoir with Multiple Conditional Techniques: A Practical Approach to Reservoir Characterization, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 289-298.

Blanc, G., D. Guerillot, D. Rahon, F. Roggero, and the HELIOS Reservoir Group, 1996, Building geostatistical models constrained by dynamic data - a posteriori constraints, SPE/NPF European Conference, Stavanger, Norway, April 16-17, 1995, SPE 35478, pp. 19-33.

Botton-Dumay, R., Cogrel, Y.M., Massonat, G.J. and Eberle, N., 1997,"Relistic methodology for permeability modeling used for conserving heterogeneity during assisted history matching - applied to turbiditic reservoir field case", Annual Technical Conference and Exhibition, San Antonio, Texas, SPE 38677.

Burns, C. S., 1997, Integrating 3D seismic into the reservoir model, and its impact on reservoir management, Fifth Latin American and Caribbean Petroleum Engineering Conference, Rio de Janeiro, Aug. 30 - Sept 3, 1997, SPE 38996.

Caers, J. B.G. Arpat and C.A. Garcia, 2006, Featured-based Probabilistic Interpretation of Geobodies from Seismic Amplitudes, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 91-108.

Chambers, R. L., M. A. Zinger and M. C. Kelly, (1994), Constraining Geostatistical Reservoir Descriptions with 3-D Seismic Data to Reduce Uncertainty, in J. M. Yarus and R. L. Chambers, eds., Stochastic Modeling and Geostatistics, AAPG Computer Applications in Geology, No. 3, pp. 143-158.

Chambers, R. L. and J. M. Yarus, (2002), Quantitative Use of Seismic Attributes for Reservoir Characterization, RECORDER, Canadian SEG, Vol. 27, pp. 14-25, June.

Chambers, R. L. and J. M. Yarus, (2007), Geologically Based, Geostatistical Reservoir Modeling, Volume VI, Emerging and Peripheral Technologies, H.R. Warner, Jr., editor, in Petroleum Engineering Handbook, L. W. Lake, Editor-in-Chief, Chapter 2, pp. VI-51 to VI-111.

Chambers, R. L., J. M. Yarus and K. B. Hird (2000), Geostatistics for the Non-Geostatisticians- Part 1, published in The Leading Edge, May.

Chambers, R. L., J. M. Yarus and K. B. Hird (2000), Geostatistics for the Non-Geostatisticians- Part 2, published in The Leading Edge, June.

Christakos, G., 1992, Random fields models in the earth sciences, San Diego, CA, Academic Press, 474 p.

Clark, I., 1979, The Semivariogram - Part 1: Engineering and Mining Journal, Vol. 180, No. 7, pp. 90-94.

Clark, I., 1979, The Semivariogram - Part 2: Engineering and Mining Journal, Vol. 180, No. 8, pp. 90-97.

Cressie, N., 1990, The Origins of Kriging, Mathematical Geology, Vol. 22, No. 3, pp. 239-252.

Cressie, N. and D. M. Hawkins, 1980, Robust Estimation of the Variogram, Mathematical Geology, Vol. 12, No. 2, pp. 115-125.

Coburn, T.C. J. M. Yarus and R.L. Chambers., 2006, Geostatistics and Stochastic Modeling: Bridging the 21st Century, in Stochastic Modeling and Geostatistics, Volume 2,T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp.3-10.

Coburn, T.C, et al.., 2006, A Virtual Conversation on the Impact of Geostatistics in Petroleum, Production and Reservoir Engineering, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp.23-32.

Cosentino, L., A Pelgrain de Lestang and J.E. Gonzalez, 2006, Geostatistics: A Faset-response Tool for Reservoir Management, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 301-312.

Datta-Gupta, Akhil, Larry W. Lake, and Gary A. Pope, 1995. Characterizing heterogeneous permeable media with spatial statistics and tracer data using sequential simulated annealing. Mathematical geology, vol. 27, pp. 763-788.

Delfiner, P., J. P. Delhomme, and J. Pelissier-Combescure, 1983. Application of geostatistical analysis to the evaluation of petroleum reservoirs with well logs. Proceedings of the 1983 SPWLA Annual Logging Symposium, Calgary, July 27-30.

Deutsch, C. V., 1993, Conditioning reservoirs to well test information, Transaction of the 1992 Convention on Geostatistics, Troia, Kluwer Academic Pub.

Deutsch, C. V. and A. G. Journel, 1994, Integrating well test-derived effective absolute permeabilities in geostatistical reservoir modeling, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 131-142.

Deutsch, Clayton V. and Sanjay Srinivasan, 1996. Improved reservoir management through ranking stochastic reservoir models, in Proceedings, SPE/DOE tenth symposium on improved oil recovery, vol. 2. Tulsa, OK: Society of Petroleum Engineers, pp. 105-112 (SPE/DOE 35411).

Doyen, P. M., 1988, Porosity From Seismic Data: A Geostatistical Approach, Geophysics, Vol. 53, pp. 1263-1295.

Doyen, P. M. and T. M. Guidish, 1992, Seismic Discrimination of Lithology and Porosity, a Monte Carlo Approach, in Reservoir Geophysics: Investigations in Geophysics, R. E. Sheriff, Ed., Vol. 7, Tulsa, Oklahoma, Society of Exploration Geophysicists, pp. 243-250.

Doyen, P. M., D. E. Psaila and S. Strandenes, 1994, Bayesian Sequential Indicator Simulation of Channel Sands in the Oseberg Field, Norwegian North Sea, SPE Annual Technical Conference and Technical Exhibition, New Orleans, Louisiana, SPE 28382.

Dubrule. O., 1989, A review of stochastic models for petroleum reservoir, in Geostatistics, Armstrong, ed., Kluwer publishers, p. 493-506.

Dubrule, Olivier and Helge H. Haldorsen, 1986. Geostatistics for permeability estimation, in Reservoir Characterization (Larry W. Lake and Herbert B. Carroll, Jr., eds.). New York: Academic Press, Inc., pp. 223-247.

Englund, E. J., 1990, A Variance of Geostatisticians, Mathematical Geology, Vol. 22, No. 4, pp. 313-341.

Froidevaux, R., 1992, Probability field simulation, in A. Soares, ed., Geostatistics Troia 1992, Proceedings of the Fourth Geostatistics Congress: Kluwer Academic Publishers, Dordrecht, Holland, p. 73-84.

Frykman, P., 2006, Spatial Variablity of Porosity in Chalk: Comparison of an Outcrop and a North Sea Reservoir, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 279-288.

Haas, A. and B. Noetinger, 1999. Stochastic reservoir modelling constrained by well test permeabilities, in Geostatistics Wollongong '96 (E. Y. Baafi and N. A. Schofield, eds.). Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 501-511.

Haldorsen, H. H. and E. Damsleth, 1990, Stochastic modelling, JPT, April, p. 404-412.

Hatløy, A. S., 1994, Numerical Modeling Combining Deterministic and Stochastic Methods, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 109-120.

Herweijer, J.C., J.M. Yarus and M.J.V. Suana, 2006, Process-based Knowledge Management: Overview and Application of Iterative 3-D Modeling Workflows, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 313-324.

Hird, K. B. and M. G. Kelkar, 1992. Conditional simulation method for reservoir description using spatial and well performance constraints. SPE Paper 24750, 67th Annual Technical Conference and Exhibition of the SPE, Washington, D.C.

Isaaks, E. H. and R. M. Srivastava, 1988, Spatial Continuity Measures for Probabilistic and Deterministic Geostatistics, Mathematical Geology, Vol. 20, No. 4, pp. 239-252.

Goovaerts, P., 1999, Impact of the simulation algorithm, magnitude of ergodic fluctuations and number of realizations on the space of uncertainty of flow properties: Stochastic Environmental Research and Risk Assessment, v. 13, no. 3, p. 161-182.

Goovaerts, P., 2006, Geostatistical Modeling of the Spaces of Local, Spatial, and Response Uncertainty for Continuous Petrophysical Properties, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp.59-80.

Meyers, D. E.., 2006, Reflections on Geostatistics and Stochastic Modeling, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 11-22.

Galli, A, G. Le Loc'h and R. Eschard, 2006, An Application of the Truncated Pluri-gaussian Method for Modeling Geology, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 109-122.

Gribov, A., K. Krivoruchko and J. M. Ver Hoef., 2006, Modeling the Semivariogram: New Approach, Methods Comparison, and Simulation Study, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 45-58.

Guerillot, D.R. and I.F. Morelon, 1992, Sorting equiprobable geostatistical images by simplified flow calculations, 67th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, SPE 24891, pp. 327-342.

Jensen, J. L., P. W. M. Corbett, G. E. Pickup, and P. S. Ringrose, 1996. Permeability semivariograms, geological structure, and flow performance. Mathematical geology, vol. 28, pp. 419-436.

Journel, A. G., 1983. Nonparametric estimation of spatial distributions. Mathematical Geology, vol. 15, pp. 445-468.

Journel, A. G., 1986, Geostatistics: Models and tools for earth sciences. Math. Geol.18, 119-140

Journel, A. G., 1988, Non-parametric Geostatistics for Risk and Additional Sampling Assessment, in Principles of Environmental Sampling, L. Keith, Ed., American Chemical Society, pp. 45-72.

Journel, A. G., 1990. Geostatistics for reservoir characterization. SPE Paper 20750. 65th Annual Technical Conference and Exhibition of the SPE, New Orleans

Journel, A.G., 1994, Geology and reservoir geology, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 19-20.

Journel, A. G. and Z. Ying, 2001, The theoretical links between sequential Gaussian, Gaussian truncated simulation, and probability field simulation, Mathematical Geology, v. 33, No. 1, p. 31-40.

Kalkomey, C. T., 1997, Potential risks when using seismic attributes as predictors of reservoir properties. The Leading Edge, March.

Kelley, V. A. and S. Mishra, 2006, An Efficient Approach fro Quantifying the Uncertainty in Groundwater-model Predictions , in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 151-158.

Krige, D. G., 1951, A Statistical Approach to Some Basic Mine Evaluation Problems on the Witwatersrand, J. Chem. Metall. Min. Soc. South Africa, vol. 52, pp. 119-39.

Krige, D. G., 1984. Geostatistics and the definition of uncertainty. Inst. Min. Met. Trans., Vol. 93, Sect. A, pp. A41-47.

Krivoruchko, K, A. Gribov, and J.M. Ver Hoef, 2006, A New Method for Handling the Nugget Effect in Kriging, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 81-90.

King, M.J., and M. Mansfield, 1999, Flow simulation of geologic models, SPE Reservoir Eval. And Eng., SPE 57469, Vol. 2, No. 4, pp. 351-367

Lake, Larry W., Allan J. Scott, and Gary A. Kocurek, 1986. Reservoir characterization for numerical simulation-final report. Washington, DC: U.S. Department of Energy, Bartlesville Project Office, DOE/BC/ 10744-8, 226 p.

Lake, Larry W., Herbert B. Carroll, and Thomas Wesson, 1991. Second international reservoir characterization conference, in Reservoir characterization II (Larry W. Lake, Herbert B. Carroll, Jr., and Thomas C. Wesson, eds.). San Diego: Academic Press, pp. 478-491.

Lia, O., H. More, H. Tjelmeland, L. Holden, and T. Egeland, 1997, Uncertainties in reservoir production forecasts, AAPG Bulletin, V. 81, No. 5, May 1997, pp.775-802.

Lecour, M. A. P. Thore and Cognot, 2006, Consistent Structural Model Simulations for Volumetric Uncertainty Study, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 159-172.

Lines, L., H. Tan, S. Treitel, J. Beck, R.L. Chambers, J. Eager, C. Savage, J. Queen, W. Rizer, P. Buller, D. Cox, J. Sinton, J. Ballard, G. Kokkoros, A. Track, P. Guerendel, J. Harris, S. Lazaratos, and M. Van Shaak, (1995), Integrated Reservoir Characterization: Beyond Tomography, GEOPHYSICS, Vol. 60, No. 2, pp. 354-364. This paper received the 1996 SEG Best Paper Award.

Matheron, G., 1963, Principles of Geostatistics, Economic Geology, vol. 58, pp. 1246-66.

Matheron, G., 1970, Random functions and their application in geology, in Geostatistics, a Colloquium, D. F. Merriam (editor), Plenum, New York, pp. 79-88.

Matheron, G., H. Beucher, C. de Fouquet, A. Galli, D. Guerillot, and C. Ravenne 1987, Conditional simulation of the geometry of fluvio-deltaic reservoirs, 1987 SPE Annual Technical Conference and Exhibition, Dallas, Texas, Sept. 27-30, SPE 16753, pp. 591-599.

Mathieu, G., F. Verdier, P. Houel, J. Delmas, and H. Beucher, 1993, Reservoir heterogeneity in fluviatile Keuper facies: A subsurface and outcrop study, in Subsurface Reservoir Characterization from Outcrop Observations, Eds. Eschard and Doligez, Technip Publication, Paris 1993, p. 145-160.

McKenna, S.A. and P.C. Reeves., 2006, Fractured Continuum Approach to Stochastic Permeability Modeling, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 173-186.

Meyers, D. E., 1983. Estimation of linear combinations and co-kriging. Mathematical geology, vol. 15, no. 5, pp. 633-637.

Meyers, D. E., 1984. Co-kriging: new developments, in Geostatistics for natural resources characterization, pt.1 (G. Verly et al., eds.)., pp. 295-305.

Meyers, D. E.., 2006, Reflections on Geostatistics and Stochastic Modeling, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 11-22.

Olea, R. A., 1975, Optimum mapping techniques using regionalized variable theory: series on spatial analysis no. 3, Lawrence, Kansas, Kansas State Geological Survey, 137 p.

Olea, R. A., 1977, Measuring Spatial Dependence with Semivariograms, Lawrence, Kansas, Kansas Geological Survey, Series on Spatial Analysis, No. 3, 29 p.

Olea, R. A., 1994, Fundamentals of Semivariogram Estimation, Modeling and Usage, in Stochastic Modeling and Geostatistics, (1994), J. M. Yarus and R. L. Chambers, Eds. AAPG Computer Applications in Geology, No. 3, pp. 27-36.

Olea, Ricardo A., 1999. Geostatistics for engineers and earth scientists. Norwell, MA: Kluwer Academic Publishers, 303 p.

Oliver, Dean S., 1996. On conditional simulation to inaccurate data. Mathematical geology, vol. 28, pp. 811-818.

Patterson, P. E., T.A. Jones, A. D. Donovan, J.D. Ottmann and C.J. Donofrio, 2006, Hierarchical Description and Geologic Modeling of a Braided-stream Depositional System, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 221-236.

Pyrcz, M. J., E. Gringarten, P Frykman and C. V. Deutsch, 2006, Representative Input Parameters for Geostatistical Simulation, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 123-137.

Ravenne, C. and H. Beucher, 1988, Recent developments in description of sedimentary bodies in a fluvio-deltaic reservoir and their 3D conditional simulation, Annual Technical Conference and Exhibition, SPE 18310.

Royle, A. G., 1979, Why Geostatistics?, Engineering and Mining Journal, Vol. 180, pp. 92-101.

Samson, P., J-M. Casaux, B. Cavaillès, E. Larribau, F. Morandini, P. Biondi, S. Calassou, D. Janodet, J-M. Chautru, X. Legrand, and J.-F. Rainaud, 1998, 3D modeling from seismic to fluid flow simulation: a case study, 1998 SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, Sept. 27-30, SPE 49293, pp. 741-750.

Seifert, D. and J. L. Jensen, 2000. Object and pixel-based reservoir modeling of a braided fluvial reservoir. Mathematical Geology, vol. 32, pp. 581-603.

Shrestha, R.K., S. Faruqi, M. Boeckmann and M.K. Wang, 2006, Three-dimensional Seismic Inversion and Geostatistical Reservoir Modeling of the Rotliegend Gas Field, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 237-252.

Sichel, H. S., 1952, New methods in the statistical evaluation of mine sampling data, Trans. Inst. Min. Metall., vol. 61, no. 6, pp. 261-8.

Soares, A, J.A. Almeida and L. Guerreiro, 2006, Incorporating Secondary Information Using Direct Sequential Cosimulation, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 35-44.

Srinivasan, S. and Caers, J. (2000), Conditioning reservoir models to dynamic data -- a forward modeling perspective, Annual Conference and Technical Exhibition, Dallas, Texas, Production operations and Engineering, SPE 62941.

Srivastava, R. M., 1990. An application of geostatistical methods for risk analysis in reservoir management. Proceedings, 65th Annual technical conference and exhibition. Tulsa, OK: Society of Petroleum Engineers, pp. 825-834 (SPE 20608).

Srivastava, R. M., 1992, Reservoir characterization with probability field simulation: 1992 SPE Annual Technical Conference and Exhibition, Washington, D.C., Oct. 5-8, SPE Paper 24753.

Srivastava, R. M., 1994, An overview of stochastic methods for reservoir simulation, Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds. AAPG Computer Applications in Geology, No. 3, pp. 3-16.

Srivastava, R. M., 1994, The visualization of spatial uncertainty, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 339-346.

Strebelle, S.B., 2006, Sequential Simulation for Modeling Geological Structures from Training Images, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 139-150.

Thiele, M. R., S. E. Rao, and M. J. Blunt, 1996. Quantifying uncertainty in reservoir performance using streamtubes. Mathematical Geology, vol. 28, pp. 843-856.

Torres-Verdín, C. A. Grijalba-Cuenca and H. W. J. Debeye, 2006, A Comparison Between Geostatistical Inversion and Conventional Geostatistical-simulation Practices for Reservoir Description, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 187-206.

Truscott, S. J., 1929, The Computation of the probable value of ore reserves from assay results, Trans. Inst. Min. Metall., vol. 38, pp. 482-96.

Tyler, N. and R. J. Finley, 1991, Architecture controls on the recovery of hydrocarbons from sandstone reservoirs. In: SEPM Concepts in Sedimentology Paleontology. Vol. 3, pp. 1-5.

Tyler, N. and J. K. Gholston, 1988, Heterogeneous submarine fan reservoirs, Permian Spraberry Trend, West Texas: The University of Texas at Austin, Bureau of Economic Geology Report of Investigations No. 171, 37p.

Tyler, K., A. Henriquez, and T Svanes, 1994, Modeling Heterogeneities in Fluvial Domains: A Review of the Influences on Production Profiles, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 77-90.

Van Riel, P. P. Mesdag, H. Debeye and M. Sams, 2006, Full Integration of Seismic Data into Geostatistical Reservoir Modeling, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 207-218.

Wang, J. and A. C. MacDonald, 1997, Modelling channel architecture in a densely drilled oilfield in east China, Annual Technical Conference and Exhibition, SPE 38678.

Watermeyer, G. A., 1919, Applications of the theory of probability in the determination of ore reserves, J. Chem. Metall. Min. Soc. South Africa, vol. 19, pp. 97-107.

Weber, K.J., 1986 - "How heterogeneity affects oil recovery" - Reservoir Characterization, L.W. Lake and H.B. Carroll Jr. (editors), Academic Press, Orlando, Florida, pp 487-544.

Wingle, W. L., and E. P. Poeter, 1993, Uncertainty associated with semi variograms used for site simulation: Ground Water, v. 31, p. 725-734.

Williams, D. B. and A.A.M. Aqrawi, 2006, Utility of Using Acoustic Impedance Data in the Stochastic Modeling of a Carbonate Reservoir, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 253-268.

Wolf, D. J., K. D. Withers, and M. D. Burnaman, 1994, Integration of Well and Seismic Data Using Geostatistics, in Stochastic Modeling and Geostatistics, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, No. 3, pp. 177-200.

Xu, W., T. T. Tran, R. M. Srivastava and A. G. Journel, 1992, Integrating Seismic Data in Reservoir Modeling: the Collocated CoKriging Alternative, 1992 SPE Annual Technical Conference and Exhibition, Washington, D.C., Oct. 5-8, SPE 24742, pp.833-842.

Xu, W., K. Wrolstad, D. Kramer, P. Dooley, K. Dominique and D.T. Vo, 2006, Brazos A-105 D-Sand Reservoir Modeling by Integration of Elastic Inversion Resutls with Geostatistical Techniques, in Stochastic Modeling and Geostatistics, Volume 2, T.C. Coburn, J. M. Yarus and R. L. Chambers, Eds., AAPG Computer Applications in Geology, pp. 269-278.

Yarus, J. M. and R. L. Chambers, 2006, Practical Geostatistics-An Armchair Overview for Petroleum Reservoir Engineers, SPE Paper 103357, Distinguished Author Series, JPT, November, pp. 78-88.

Yarus, J.M. K. Yang and K. Kramer, 2002, Practical Workflows for Reservoir Modeling, Proceedings, Geostatistics Rio 2000, 6-17 August, ed. M. Armstrong, et al. Dordrecht, The Netherlands, Kluwer Academic Publishers, pp. 69-84.