

Prediction in Presence of a Nonstationary Histogram

Jason A. McLennan and Clayton V. Deutsch

Centre for Computational Geostatistics
Department of Civil and Environmental Engineering
University of Alberta

Predicting the spatial distribution of petrophysical properties is an essential aspect of natural resource characterization. Traditional geostatistical prediction without explicitly incorporating trends does not guarantee reproduction of important large scale features. There are a variety of approaches to perform estimation and simulation with a trend. These techniques are reviewed. In general, currently available techniques are undermined by the challenge of inferring the underlying spatial law of the residual RF $R(\mathbf{u})$. A new explicit approach to estimation and simulation with the trend is developed, described, and implemented. The key feature of this method is the use of locally varying transformation tables for the transformation from original units to a Gaussian distribution.

Explicit Approaches to Estimation with the Trend

The goal of geostatistical estimation is to quantify local uncertainty. The pioneering work of Danie Krige [1] during the 1950s to correct conditional biases [2] was the seed for the popular group of estimation techniques collectively referred to as kriging. From the 1960s the utility of kriging has been for both large-scale trend modeling and calculating recoverable reserves for production planning.

The techniques in this work explicitly consider a trend model during quantification of local uncertainty. These techniques can also be referred to as non-stationary kriging implementations since the mean function $m(\mathbf{u})$ used is not a constant as in stationary or simple kriging. Three non-stationary kriging implementations now reviewed.

Kriging with a locally varying mean is the first technique. Recall the kriging estimator:

$$Z_K^*(\mathbf{u}_0) - m(\mathbf{u}_0) = \sum_{s=1}^n \lambda_K(\mathbf{u}_s) [Z(\mathbf{u}_s) - m(\mathbf{u}_s)] \quad (1)$$

Instead of stationary simple kriging with stationary mean m inferred from $F(z)$, the mean $m(\mathbf{u})$ corresponds to a non-stationary trend.

Kriging with a trend (KT), originally known and developed as universal kriging (UK) by Matheron in 1969 [3], provides minimum error variance estimates of the original $Z(\mathbf{u})$ RF in the presence of a trend model $m(\mathbf{u})$. Given that $m(\mathbf{u})$ is a deterministic function of the coordinates vector \mathbf{u} , the trend is also estimated according to the same minimum error variance optimality criteria. The recent reference [4] provides a more comprehensive derivation. The convenience of KT for simultaneously and optimally estimating both the $m(\mathbf{u})$ trend and the original RF $Z(\mathbf{u})$ is offset by a number of practical implementation issues. Among others, Armstrong in 1984 [5] highlights the most important one: that of inferring the underlying spatial law.

Kriging with an external drift variable is an extension of KT where the functional form of the $m(\mathbf{u})$ trend model is limited to $V = 1$ functionals: $f_0(\mathbf{u}) = 1$ and $f_1(\mathbf{u})$ set equal to a secondary (external) variable [6]. This was first implemented successfully with secondary seismic data by Marechal, 1984 [7].

For a set of K indicator RFs, two methods are available to account for the trend. The first is the same as the approach for kriging with a prior mean except the prior mean $m(\mathbf{u})$ are the locally varying proportions of the particular k indicator variable denoted $p(\mathbf{u}_k)$. The second approach is referred to as *soft* kriging. As long as soft data is available and is amenable to a Bayesian coding into prior $p(\mathbf{u}_k)$ mean values, the Markov-Bayes model [8, 9] can be applied to explicitly integrate the trend into the resulting model of predicted geological heterogeneity and local uncertainty.

The most common and most straightforward explicit approach to geostatistical simulation with the trend model is a traditional simulation of the assumed stationary residual RF $R(\mathbf{u})$ after the modeled trend $m(\mathbf{u})$ is subtracted from the original RF $Z(\mathbf{u})$ at the available data locations [12]. The simulated residuals are then simply added back to the $m(\mathbf{u})$ trend model.

An important disadvantage of residual simulation is the inability to control the bivariate residual-mean and residual-variable distributions. In each of these distributions, a characteristic feature can possibly be present. Lueangthong [13] illustrates how each of these two features is detected:

1. Applications of traditional geostatistical simulation tools to the $R(\mathbf{u})$ SRF imply $R(\mathbf{u})$ is homoscedastic meaning the variance of the $r(\mathbf{u})$ residuals is independent of the trend values $m(\mathbf{u})$. However, virtually all residual-mean scatters of $(r(\mathbf{u}), m(\mathbf{u}))$ pairs will reveal some heteroscedastic behavior.
2. The additive dissociation of $Z(\mathbf{u})$ imposes the constraint $R(\mathbf{u}) \geq m(\mathbf{u})$ for non-negative $Z(\mathbf{u})$ variables. Simulating the spatial distribution of residuals and adding the mean back does not ensure this constraint is satisfied.

These two problems of the residual simulation approach motivate a stepwise conditional transformation technique proposed in [13]. The key idea is a stepwise transformation of the residual data conditional to the trend. This transformation assumes the following form:

$$Y_R(\mathbf{u}) = G^{-1} \left[F_{R|m} (R(\mathbf{u}) | m(\mathbf{u})) \right] \quad (2)$$

where $Y_R(\mathbf{u})$ is the Gaussian G transform of the residual random variable $R(\mathbf{u})$ conditional to local $m(\mathbf{u})$ windows. Similarly, the solution when the second of these problems persists is a stepwise transformation of the original variable conditional to the trend:

$$Y_Z(\mathbf{u}) = G^{-1} \left[F_{Z|m} (Z(\mathbf{u}) | m(\mathbf{u})) \right] \quad (3)$$

Both transformations complement conventional practice. The same decomposition in relation (3-3) is used in that the trend $m(\mathbf{u})$ and residual $R(\mathbf{u})$ is modeled separately and recombined. The pre and post processing steps are implemented in order to preserve heteroscedastic and constraint features. All necessary implementation details as well as mining and petroleum examples are given in [13].

Simulation with the Trend via Locally Varying Transformations

A new approach to accounting for the trend within a geostatistical prediction framework is now presented. This method is based on a locally varying transformation (LVT) to account for non-stationarity. The trend is built into and honored through the LVT. The theoretical framework, inference of the LVT, inference of the spatial law, estimation and simulation procedures, and implementation details are presented.

The LVT approach does not conform to traditional Gaussian theory underlying multi-Gaussian kriging and sequential Gaussian simulation approaches. Two unique theoretical aspects of the LVT approach differentiate it from traditional theory. The first aspect is the use of a transformation that is unique to each location in place of the common global transformation. The second aspect is the use of locally transformed conditioning data instead of globally transformed Gaussian values to infer the spatial law for subsequent prediction.

The first order assumption of stationarity implies all prior $F_Z(\mathbf{u}; z)$ cdfs are equivalent to a single stationary univariate distribution $F_Z(z)$ at all locations \mathbf{u} within a preset domain. The stationary $F_Z(z)$ distribution and associated first order moments are inferred with the cdf of all z sample data available within the domain where the SRF will be subsequently applied. The usual computation of the global $F_Z(z)$ cdf is performed; however, the $F_Z(z)$ distribution and associated first order mean and variance moments may be adjusted with declustering weights when preferential sampling is prevalent.

Traditionally, a single stationary transformation between original z sample data values and subsequently transformed Gaussian y values is defined using the single stationary $F_Z(z)$ distribution as follows:

$$y = G_Y^{-1}(F_Z(z)) \quad (4)$$

Original z data is forward-transformed according to (10), predictions are performed using conditioning y values, and then the predicted y values are back-transformed into original z unit predictions with the reverse of (4):

$$z = F_Z^{-1}(G_Y(y)) \quad (5)$$

Quantiles, not estimates, are back-transformed according to (5). Estimates are derived then by numerical integration. The forward and back transformation in (4) and (5) converts the original $Z(\mathbf{u})$ SRF to a standard normal Gaussian $Y(\mathbf{u})$ SRF for the purpose of prediction and then converts $Y(\mathbf{u})$ predictions back to $Z(\mathbf{u})$ predictions for post processing and visualization. All that is needed to perform these transformations is an assumption of multi-Gaussianity and the representative stationary $F_Z(z)$ cdf.

The theoretical framework for the LVT technique is fundamentally different than the traditional theory underlying geostatistical prediction in that an assumption of first order stationarity is not employed. That is, the local $F_Z(\mathbf{u}; z)$ cdfs are not equivalent to a single stationary univariate distribution $F_Z(z)$ built from all z sample data available within a fixed domain \mathbf{D} :

$$F_{Z(\mathbf{u};z)}(z) \neq F_Z(z) \quad \forall \mathbf{u} \in \mathbf{D} \quad (6)$$

The inequality in relation (6) is the foundation of the LVT approach and the following theoretical developments and prediction procedures.

The local $F_Z(\mathbf{u}; z)$ cdfs are calculated at each \mathbf{u} location by applying locally varying weights $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ to each available conditioning sample data:

$$F_{Z(\mathbf{u};z)}(z) = \frac{1}{\sum_{s=1}^S w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)} \sum_{s=1}^S w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s) \cdot x(z(\mathbf{u}_s); z) \quad \forall z \quad (7)$$

where the notation $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ indicates a location-dependent set of weights applied to the conditioning data available within the determined domain \mathbf{D} . The indicator transform $x(z(\mathbf{u}_s); z)$ is:

$$x(z(\mathbf{u}_s); z) = \begin{cases} 1, & \text{if } z(\mathbf{u}_s) \leq z \\ 0, & \text{otherwise} \end{cases} \quad \forall z \quad (8)$$

The $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ weights are not declustering weights that adjust $F_Z(z)$ to be representative of the domain \mathbf{D} in the presence of biased sampling. The weights in relation (7) adjust the global $F_Z(z)$ distribution to be representative at that particular location \mathbf{u} .

The $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ weights are calculated with a smooth kriging or inverse distance scheme parameterized similar to that for building a smoothly varying trend model. A different set of weights and different local $F_Z(\mathbf{u}; z)$ cdf are calculated at different \mathbf{u} locations since the configuration of available sample data relative to the \mathbf{u} location changes depending on \mathbf{u} .

All univariate summary statistics can be calculated to summarize each local $F_Z(\mathbf{u}; z)$ distribution. The locally varying transformed mean $m(\mathbf{u})$ for each of the $F_Z(\mathbf{u}; z)$ cdfs is calculated as:

$$m(\mathbf{u}) = \frac{1}{\sum_{s=1}^S w^{\text{LVT}}(\mathbf{u}; \mathbf{u}_s)} \sum_{s=1}^S w^{\text{LVT}}(\mathbf{u}; \mathbf{u}_s) z(\mathbf{u}_s) \quad (9)$$

And the locally varying transformed variance $\sigma^2(\mathbf{u})$ for each of the $F_Z(\mathbf{u}; z)$ cdfs is:

$$\sigma^2(\mathbf{u}) = \frac{1}{\sum_{s=1}^S w^{LVT}(\mathbf{u}; \mathbf{u}_s)} \sum_{s=1}^S w^{LVT}(\mathbf{u}; \mathbf{u}_s) (z(\mathbf{u}_s) - m(\mathbf{u}))^2 \quad (10)$$

Prediction utilizing an underlying assumption of Gaussianity is still employed in the LVT approach. However, $Z(\mathbf{u})$ is no longer assumed a SRF. It is a RF without first order stationarity since there is no longer a single stationary $F_Z(z)$ distribution from which a global transformation from original $Z(\mathbf{u})$ space to a Gaussian SRF $Y(\mathbf{u})$ space for prediction as in relation (4) can be defined.

An intermediate RF $T(\mathbf{u})$ is used to convert from original $Z(\mathbf{u})$ space to a Gaussian SRF space. A locally varying transformation between original z sample data values and subsequently transformed t values is defined using the standard Gaussian distribution G_Y and the previously defined local $F_Z(\mathbf{u}; z)$ cdfs as follows:

$$t = G_Y^{-1}\left(F_{Z(\mathbf{u}; z)}(z)\right) \quad (11)$$

Since the z sample data are transformed to t from a set of different $F_Z(\mathbf{u}; z)$ cdfs, the resulting global distribution of transformed t values denoted by $F_T(t)$ is Gaussian but not standard Gaussian with zero mean and unit variance. The expected value of the $T(\mathbf{u})$ RF is zero, and the variance will be less than unity for increased variability in the $w^{LVT}(\mathbf{u}; \mathbf{u}_s)$ weights.

Unlike $Z(\mathbf{u})$, the $T(\mathbf{u})$ RF will be assumed to have stationary first order mean and second order covariance because the non-stationary trend model $m(\mathbf{u})$ is built into the locally varying transformation through the $F_Z(\mathbf{u}; z)$ cdfs calculated from the $w^{LVT}(\mathbf{u}; \mathbf{u}_s)$ weights. Traditional Gaussian estimation and simulation procedures can be performed using this non-standard Gaussian $T(\mathbf{u})$ SRF. The $T(\mathbf{u})$ SRF is converted to a standard Gaussian SRF for prediction.

A transform of the t values to standard Gaussian y_{LT} values is defined using the stationary $F_T(t)$ distribution:

$$y_{LT} = G_Y^{-1}\left(F_T(t)\right) \quad (12)$$

Original z data is forward-transformed according to the locally varying transformation in relation (11), the transformed t values are forward-transformed according to (12), and predictions are made with conditioning y_{LT} values.

The predicted y_{LT} quantile values are then back-transformed to the intermediate $T(\mathbf{u})$ SRF space with the reverse of (12):

$$t = F_T^{-1}\left(G_Y(y_{LT})\right) \quad (13)$$

And then the back-transformed t quantiles from (13) are back-transformed to original unit values using the reverse of (11):

$$z_{LT} = F_{Z(\mathbf{u}; z)}^{-1}\left(G_Y(t)\right) \quad (14)$$

Estimates are derived by numerically integrating over a series of quantiles.

The stepwise forward transformation procedure converting from the non-standard Gaussian non-stationary RF $Z(\mathbf{u})$ to the intermediate non-standard Gaussian stationary RF $T(\mathbf{u})$ to the standard Gaussian stationary RF $Y_{LT}(\mathbf{u})$ space can be summarized as:

$$\begin{aligned} t &= G_Y^{-1}\left(F_{Z(\mathbf{u}; z)}(z)\right) \\ y_{LT} &= G_Y^{-1}\left(F_T(t)\right) \end{aligned} \quad (15)$$

Similarly, the stepwise backward transformation procedure converting from a standard Gaussian stationary Gaussian RF $Y_{LT}(\mathbf{u})$ to the intermediate non-standard Gaussian stationary RF $T(\mathbf{u})$ to the original non-standard Gaussian non-stationary RF $Z(\mathbf{u})$ space can be summarized as:

$$\begin{aligned}
t &= F_T^{-1}(G_Y(y_{LT})) \\
z_{LT} &= F_{Z(\mathbf{u};z)}^{-1}(G_Y(t))
\end{aligned} \tag{16}$$

Notice that even though both $Y_{LT}(\mathbf{u})$ and $Y(\mathbf{u})$ are standard Gaussian SRFs, the $Y_{LT}(\mathbf{u})$ SRF is different than its traditional analog $Y(\mathbf{u})$ due to the locally varying transformation. The y conditioning data are calculated only once; they are the same regardless of the \mathbf{u} location being estimated.

Notice that the local $F_Z(\mathbf{u}; z)$ cdfs need to be defined at all sample data and potential prediction locations \mathbf{u} . To implement the forward stepwise transformation in relation (15), $F_Z(\mathbf{u}; z)$ is only needed at the sample data locations \mathbf{u} . To visualize the trend model $m(\mathbf{u})$ and to implement the backward stepwise transformation in relation (16), $F_Z(\mathbf{u}; z)$ is needed at all prediction locations \mathbf{u} .

It is through the $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ kriging or inverse distance weights and resulting local $F_Z(\mathbf{u}; z)$ cdfs that the non-stationary trend model $m(\mathbf{u})$ is built into prediction. The transformation procedure in relation (15) including the locally varying transformation in relation (12) removes the trend to create a standard normal Gaussian SRF where traditional geostatistical prediction can be performed. The back transformation in relation (16) preserves the trend model $m(\mathbf{u})$ or non-stationarity.

Consider a small example where a location \mathbf{u} resides within a high potential area inside the domain \mathbf{D} . High valued sample data will receive higher $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ weights, low sample data values will receive lower $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ weights, and the local $F(\mathbf{u}; z)$ cdf will be lower for lower z values and higher for higher z values. The forward transformed y_{LT} conditioning values using relation (15) and the y_{LT} predictions will tend to be lower than forward transformed y conditioning values using relation (4) and y predictions. The back transformation using relation (16) to z_{LT} predictions will then tend to be higher than z predictions effectively accounting for the trend model $m(\mathbf{u})$ or locally varying mean in this high valued area.

The Underlying Spatial Law

The spatial law or SRF is defined by a multivariate Gaussian distribution and a covariance or variogram function. The original $Z(\mathbf{u})$ SRF is transformed to a standard normal Gaussian SRF $Y(\mathbf{u})$ for prediction in a multivariate Gaussian context. The Gaussian variogram $\gamma_Y(\mathbf{h})$ identifies the spatial law used for subsequent prediction. The $\gamma_Y(\mathbf{h})$ variogram in integral notation and incorporating the transform in (10) is:

$$\gamma_Y(\mathbf{h}) = \frac{1}{2} \int_{\mathbf{u} \in \mathbf{D}} [y(\mathbf{u}) - y(\mathbf{u} + \mathbf{h})]^2 d\mathbf{u} = \frac{1}{2} \int_{\mathbf{u} \in \mathbf{D}} [G_Y^{-1}(F_Z(z(\mathbf{u}))) - G_Y^{-1}(F_Z(z(\mathbf{u} + \mathbf{h})))]^2 d\mathbf{u} \tag{17}$$

where the integration volume \mathbf{u} includes all possible \mathbf{h} lag vector locations within the domain \mathbf{D} . The $\gamma_Y(\mathbf{h})$ variogram is required by theory.

The inference of $\gamma_Y(\mathbf{h})$ is straightforward. $\gamma_Y(\mathbf{h})$ is inferred by calculating:

$$\gamma_Y(\mathbf{h}) = \frac{1}{2P(\mathbf{h})} \sum_{s=1}^{P(\mathbf{h})} (y(\mathbf{u}_s) - y(\mathbf{u}_s + \mathbf{h}))^2 \tag{18}$$

$Y_{LT}(\mathbf{u})$ is a standard Gaussian SRF. The first order mean and variance moments are zero and one, respectively. The spatial law of the $Y_{LT}(\mathbf{u})$ SRF can be identified with the variogram $\gamma_{Y_{LT}}(\mathbf{h})$. In integral notation and incorporating the transform in (21), the $\gamma_{Y_{LT}}(\mathbf{h})$ variogram is written:

$$\gamma_{Y_{LT}}(\mathbf{h}) = \frac{1}{2} \int_{\mathbf{u} \in \mathbf{D}} [y_{LT}(\mathbf{u}) - y_{LT}(\mathbf{u} + \mathbf{h})]^2 d\mathbf{u} \tag{19}$$

One approach to infer $\gamma_{Y_{LT}}(\mathbf{h})$ is with the direct calculation:

$$\gamma_{Y_{LT}}(\mathbf{h}) = \frac{1}{2P(\mathbf{h})} \sum_{s=1}^{P(\mathbf{h})} (y_{LT}(\mathbf{u}_s) - y_{LT}(\mathbf{u}_s + \mathbf{h}))^2 \tag{20}$$

This approach results in the same problems as attempting to calculate the residual $\gamma_R(\mathbf{h})$ variogram directly with $r(\mathbf{u})$ data with a non-stationary mean previously. These problems were discussed earlier.

A procedure for calculating and inferring the $\gamma_{\text{LVT}}(\mathbf{h})$ variogram is developed below. This offers a distinct advantage over other conventional methods for predicting with the trend.

LVT and $\gamma_{\text{LVT}}(\mathbf{h})$ Inference

There are two inference problems with the LVT approach. The first is inference of the locally varying $F(\mathbf{u}; z)$ cdfs with $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ weighting. The second is that of inferring the $\gamma_{\text{LVT}}(\mathbf{h})$ variogram and spatial law.

Inference of Locally Varying Transformation Tables

The local non-stationary prior $F_Z(\mathbf{u}; z)$ cdf representative of the area surrounding the \mathbf{u} location must be constructed. All available sample data $z(\mathbf{u}_s)$ within the predetermined domain \mathbf{D} are used to construct each local $F_Z(\mathbf{u}; z)$ cdf. The weights $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ assigned to the $z(\mathbf{u}_s)$ data at a particular \mathbf{u} location are based on either global kriging or global inverse distance schemes. The spatial distribution of $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ weights are used to generate a set of local $F_Z(\mathbf{u}_n; z)$ cdfs containing a first order mean moment $m(\mathbf{u})$ that varies smoothly through \mathbf{D} representing a deterministic understanding of the trend or non-stationarity.

Figure 1 illustrates the LVT idea using a simple 2D schematic example with 10 $z(\mathbf{u}_s)$ sample data shaded lighter for lower values and darker for higher values. The example is conceptual. Two hand-drawn contour lines separate the domain \mathbf{D} into high, medium, and low valued locations. The conventional approach to prediction, invoking the assumption of stationarity, assumes all local $F_Z(\mathbf{u}; z)$ cdfs are equivalent to the stationary $F_Z(z)$ cdf built from the 10 sample data weighted equally ($w(\mathbf{u}_s) = 1/S = 1/10$) or by weighted by declustering ($w(\mathbf{u}_s) = w^{\text{D}}(\mathbf{u}_s)$). This cdf is shown in Figure 2. The proposed LVT approach to prediction does not assume first order stationarity and instead calculates each local $F_Z(\mathbf{u}; z)$ cdf differently by weighting the 10 $z(\mathbf{u}_s)$ data. A low, medium, and high valued location $F_Z(\mathbf{u}; z)$ cdf is shown in Figure 2 with light, medium, and dark lines, respectively. When \mathbf{u} is located within high valued zones, $F(\mathbf{u}; z)$ shifts to the right (high values) since lower valued and further away sample data receive lower $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$.

An essential guideline to trend modeling is to avoid the tendency to model too much spatial variability. This principle is relevant here when calculating the $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ weights to determine each $F_Z(\mathbf{u}; z)$ cdf, since these locally varying cdfs have the trend $m(\mathbf{u})$ imbedded within them. For both the kriging and inverse distance schemes, a global search is used retaining all S sample data at each \mathbf{u} location. A relatively low inverse distance power and significant nugget constant c are used for an inverse distance approach while a relatively high nugget and long range is combined with a block discretization for the kriging approach.

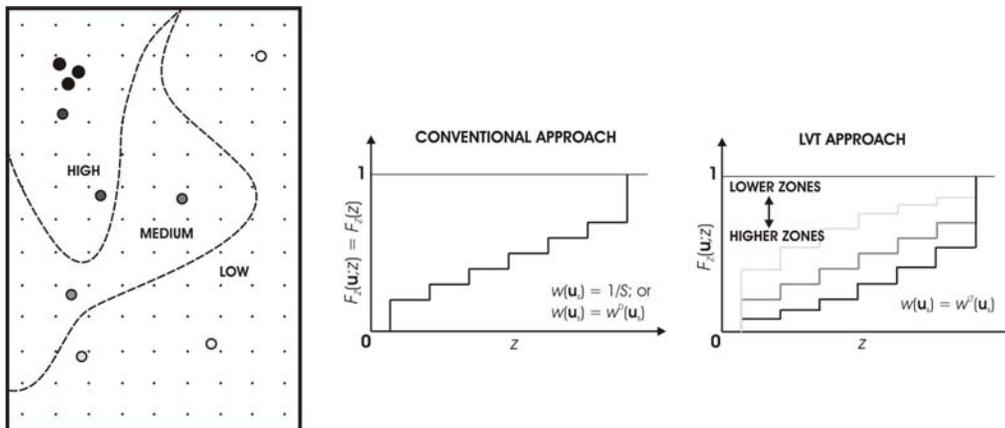


Figure 1: An illustration of the stationary cdf used for conventional prediction and a low, medium, and high case local $F_Z(\mathbf{u}; z)$ cdf used for the proposed LVT prediction approach within a schematic 2D domain.

The spatial variability of the local $F_Z(\mathbf{u}; z)$ cdfs and trend model $m(\mathbf{u})$ is sensitive to the inverse distance power and variogram parameters. Figure 2 (top) schematically shows these parameters for a low and high variability trend model. The resulting local $F_Z(\mathbf{u}_n; z)$ cdfs are also shown (bottom). For inverse distance powers near zero or kriging with high nugget effect variograms, the set of $w^{\text{LT}}(\mathbf{u}; \mathbf{u}_s)$ weights are nearly

equal to $1/S$ at all \mathbf{u} locations; the assumption of stationarity is strong, and the variation between $F_Z(\mathbf{u}; z)$ cdfs is small. For increasing inverse distance powers and decreased nugget effect variograms, the set of $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ weights increase from one \mathbf{u} location to another, the assumption of stationarity is relaxed, and the variation between local $F_Z(\mathbf{u}_n; z)$ cdfs increases.

Calculating locally varying $F_Z(\mathbf{u}; z)$ cdfs as described above will account for trends in the variable of interest. The $F_Z(\mathbf{u}; z)$ cdfs are implemented within prediction as locally varying transformation tables.

Inference of the Spatial Law

The variogram for the $Y_{LT}(\mathbf{u})$ RF is not as straightforward as calculating $\gamma_{Y_{LT}}(\mathbf{h})$ according to relation (19) using y_{LT} data. There are several qualitative solutions to approximate the spatial law of $R(\mathbf{u})$ and $Y_{LT}(\mathbf{u})$; the one proposed earlier was modeling no more spatial variability in $m(\mathbf{u})$ than is offered deterministically; however, there is no robust calculation of the residual variogram when predicting with a trend.

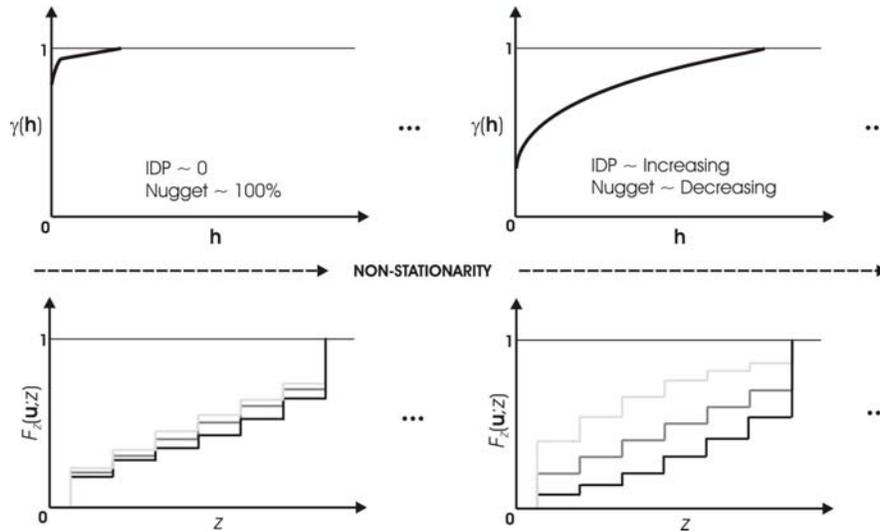


Figure 2: An illustration of the effect of increasing the variability in the trend with increasing inverse distance powers (IDP) or more continuous variograms has on the spatial variation of the local $F_Z(\mathbf{u}; z)$ cdfs.

An advantage of the LVT approach is that the variogram of the $Y_{LT}(\mathbf{u})$ RF can be calculated directly and quantitatively. Calculating the $\gamma_{Y_{LT}}(\mathbf{h})$ variogram is done with a mapping procedure starting from the $\gamma_Y(\mathbf{h})$ Gaussian variogram model and transforming through the locally varying $F_Z(\mathbf{u}; z)$ cdfs to the $\gamma_Y^{LT}(\mathbf{h})$ variogram model. There are two main steps required to map a γ_Y variogram value to its corresponding $\gamma_{Y_{LT}}$ variogram value for a given \mathbf{h} lag vector.

The conventional Gaussian transformation in relation and associated Gaussian variogram model $\gamma_Y(\mathbf{h})$ is required. Figure 3 shows a calculated (open bullets) and modeled (line) $\gamma_Y(\mathbf{h})$ variogram.

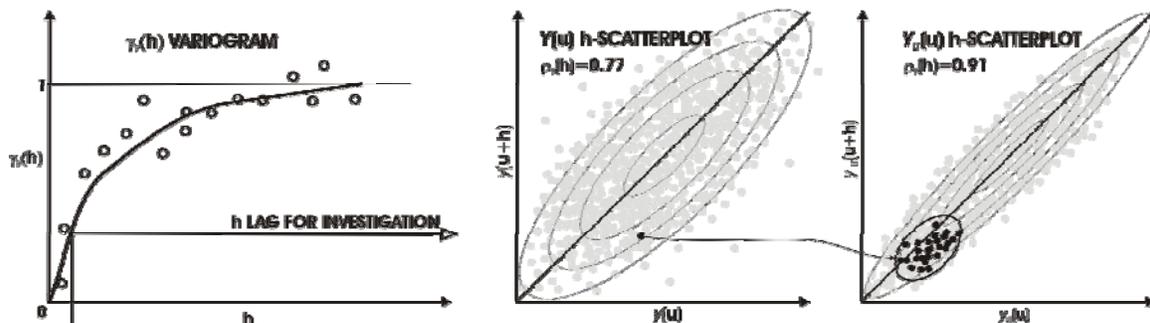


Figure 3: An illustration of the procedure implemented to map the single $\gamma_Y(\mathbf{h})$ variogram value indicated to its corresponding $\gamma_Y^{LT}(\mathbf{h})$ variogram value.

The next step entails constructing the \mathbf{h} -scatterplot for the particular \mathbf{h} lag of interest and sampling many y pairs. These pairs are denoted:

$$(y(\mathbf{u}), y(\mathbf{u} + \mathbf{h}))^l \quad l = 1, \dots, L \quad (21)$$

where L is on the order of 1,000 to 10,000 pairs. In the space of $Y(\mathbf{u})$, the distribution of pairs in (27) is bivariate Gaussian and fully parameterized by the correlation $\rho_Y(\mathbf{h})$, which is calculated:

$$\rho_Y(\mathbf{h}) = 1 - \gamma_Y(\mathbf{h}) \quad (22)$$

An underlying assumption is that a reliable variogram model and correlation in (22) can be derived from the available y data.

The L sample pairs in (21) are then drawn from a bivariate Gaussian \mathbf{h} -scatterplot with a Monte Carlo Simulation procedure. Two values, $d(\mathbf{u})$ and $y(\mathbf{u} + \mathbf{h})$ are drawn randomly from each marginal distribution; the correlation between $y(\mathbf{u})$ and $y(\mathbf{u} + \mathbf{h})$ is then imparted with the following equation:

$$y(\mathbf{u})^l = \rho_Y(\mathbf{h}) \cdot d(\mathbf{u})^l + \sqrt{(1 - \rho_Y(\mathbf{h})^2)} \cdot y(\mathbf{u} + \mathbf{h})^l \quad l = 1, \dots, L \quad (23)$$

Relation (23) is equivalent to sequential Gaussian simulation with a single conditioning datum. Figure 4 shows an \mathbf{h} -scatterplot of the L $y(\mathbf{u})$ and $y(\mathbf{u} + \mathbf{h})$ pairs for the \mathbf{h} lag indicated on the $\gamma_Y(\mathbf{h})$ variogram plot.

The other main step is transforming these L pairs from the $Y(\mathbf{u})$ Gaussian SRF space to a new set of L pairs within the $Y_{LT}(\mathbf{u})$ Gaussian SRF space where $\rho_{Y_{LT}}(\mathbf{h})$ can be calculated directly. This new set of pairs is denoted:

$$(y_{LT}(\mathbf{u}), y_{LT}(\mathbf{u} + \mathbf{h}))^l \quad l = 1, \dots, L \quad (24)$$

There is, however, no single stationary $F_Z(z)$ cdf from which a unique transformation to $Y_{LT}(\mathbf{u})$ can be made since, by definition of the local $F_Z(\mathbf{u}; z)$ cdfs, the transformation is locally varying. Therefore, a manageable number N of randomly chosen \mathbf{u}_n head locations are used to identify N \mathbf{u} and $\mathbf{u} + \mathbf{h}$ head-tail paired locations where the transformation from $y(\mathbf{u})$ to $y_{LT}(\mathbf{u})$ and $y(\mathbf{u} + \mathbf{h})$ to $y_{LT}(\mathbf{u} + \mathbf{h})$ is performed as:

$$\begin{aligned} z(\mathbf{u}_n) &= F_Z^{-1}(G_Y(y(\mathbf{u}_n))) \\ t(\mathbf{u}_n) &= G_Y^{-1}(F_{Z(\mathbf{u}; z)}(z(\mathbf{u}_n))) \quad n = 1, \dots, N \\ y_{LT}(\mathbf{u}_n) &= G_Y^{-1}(F_T(t(\mathbf{u}_n))) \end{aligned} \quad (25)$$

for all head locations and similarly for all $\mathbf{u}_n + \mathbf{h}$ tail locations (replace \mathbf{u}_n with $\mathbf{u}_n + \mathbf{h}$). Relation (25) is a three stage stepwise transformation. The three transforms are: (1) from y values within $Y(\mathbf{u})$ space to z values within $Z(\mathbf{u})$ space, (2) from z values within $Z(\mathbf{u})$ space to t values within $T(\mathbf{u})$ space, and then (3) from t values within $T(\mathbf{u})$ space to y_{LT} values within $Y_{LT}(\mathbf{u})$ space. The number of head-tail location pairs N is on the order of 10 to 100 depending on the size of domain.

The stepwise transformation in relation (31) is repeated for all L pairs to obtain a total of $L \times N$ $(y_{LT}(\mathbf{u}), y_{LT}(\mathbf{u} + \mathbf{h}))$ pairs. One $(y(\mathbf{u}), y(\mathbf{u} + \mathbf{h}))$ pair and the corresponding set of N $(y_{LT}(\mathbf{u}), y_{LT}(\mathbf{u} + \mathbf{h}))$ pairs are shown in Figure 4. The full bivariate distribution of $(y_{LT}(\mathbf{u}), y_{LT}(\mathbf{u} + \mathbf{h}))$ pairs is then used to construct an \mathbf{h} -scatterplot in bivariate Gaussian $Y_{LT}(\mathbf{u})$ SRF space from which the correlation $\rho_{Y_{LT}}(\mathbf{h})$ is calculated. Since the $Y_{LT}(\mathbf{u})$ SRF has unit variance, the $\gamma_{Y_{LT}}(\mathbf{h})$ variogram is calculated as:

$$\gamma_{Y_{LT}}(\mathbf{h}) = 1 - \rho_{Y_{LT}}^{LT}(\mathbf{h}) \quad (26)$$

This mapping procedure is then repeated for all \mathbf{h} lags from the $\gamma_{Y_{LT}}(\mathbf{h})$ variogram model until the full $\gamma_{Y_{LT}}(\mathbf{h})$ variogram model is built. The mapped $\gamma_{Y_{LT}}(\mathbf{h})$ variogram is then used to interpolate conditioning y_{LT} values within $Y_{LT}(\mathbf{u})$ space.

Prediction

The traditional procedures for prediction must be modified in order to integrate the LVT approach. A step-by-step methodology is now presented for geostatistical estimation and simulation with the trend using the LVT approach. Estimation is implemented within a multi-Gaussian framework. There are nine major steps to the methodology:

1. Collect all relevant hard $z(\mathbf{u}_s)$ sample data of the attribute of interest subsequently used for conditioning the estimation;
2. At each sample location \mathbf{u} , perform global kriging or global inverse distance to determine the $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ weights, calculate the local $F(\mathbf{u}; z)$ cdf using relation (7), and then transform the $z(\mathbf{u}_s)$ values to their corresponding $y_{LT}(\mathbf{u}_s)$ values using the stepwise forward transformation in relation (15);
3. Establish the local $F(\mathbf{u}; z)$ cdfs at all subsequent simulation locations \mathbf{u} by applying global kriging or global inverse distance to determine the $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ set of weights and calculating the local $F(\mathbf{u}; z)$ cdfs using relation (7);
4. Construct and display the trend model $m(\mathbf{u})$ at all subsequent estimation locations;
5. Determine the spatial law of $Y_{LT}(\mathbf{u})$ by transforming the original $z(\mathbf{u}_s)$ sample data to Gaussian $y(\mathbf{u}_s)$ data using (4), calculating the $\gamma_Y(\mathbf{h})$ variogram using relation (18), performing Monte Carlo Simulation from the $Y(\mathbf{u})$ space \mathbf{h} -scatterplot with (23), transforming the $Y(\mathbf{u})$ space \mathbf{h} -scatterplot to the $Y_{LT}(\mathbf{u})$ space \mathbf{h} -scatterplot with (25), and calculating the resulting correlation $\rho_{Y_{LT}}(\mathbf{h})$ and $\gamma_{Y_{LT}}(\mathbf{h})$ variogram with (26) for all \mathbf{h} ;
6. Establish a regular path through the network of subsequent estimation locations \mathbf{u} ;
7. At an estimation location \mathbf{u} , build the conditional cumulative distribution function (ccdf) with SK using the standard normal $y_{LT}(\mathbf{u}_s)$ conditioning data values and $\gamma_{Y_{LT}}(\mathbf{h})$ variogram established in steps 2 and 3. The resulting Gaussian ccdf is parameterized by the SK estimate $y_{LT}^*(\mathbf{u})$ and standard deviation $\sigma_{LT}^*(\mathbf{u})$;
8. Univariate summaries of the local ccdfs are calculated by back transforming a suitable number of quantiles:

$$z_{LT}^{(j)}(\mathbf{u}) = F_{Z(\mathbf{u}; z)}^{-1} \left(F_T \left(G_Y^{-1} \left(\sigma_{LT}^*(\mathbf{u}) \cdot \left(G_Y \left(p^{(j)} \right) \right) + y_{LT}^*(\mathbf{u}) \right) \right) \right) \quad j=1, \dots, J \quad (27)$$

where the p^j probabilities are evenly spaced between zero and one discretizing a standard Gaussian distribution. The resulting conditional distributions of $z_{LT}^{(j)}(\mathbf{u})$ values can be used to calculate univariate summaries such as the mean and variance;

9. Proceed to the next estimation location \mathbf{u} and loop over steps 7 and 8 until all estimation locations have been visited.

Simulation: SGS

Simulation is implemented within a sequential Gaussian simulation framework. There are twelve major steps to the methodology. The first five steps of the methodology are the same as the first five steps in the multi-Gaussian methodology above. There are seven additional steps:

6. Establish a random path through the network of subsequent simulation locations \mathbf{u} ;
7. At a simulation location \mathbf{u} , build the conditional cumulative distribution function (ccdf) with SK using the standard Gaussian $y_{LT}(\mathbf{u}_s)$ conditioning data values and $\gamma_{Y_{LT}}(\mathbf{h})$ variogram established in steps 2 and 3. The resulting Gaussian ccdf is parameterized by the SK estimate $y_{LT}^*(\mathbf{u})$ and standard deviation $\sigma_{LT}^*(\mathbf{u})$;
8. Draw a simulated value $y_{LT}^{(r)}(\mathbf{u})$ from the local ccdf;
9. Perform the stepwise backward transformation of $y_{LT}^{(r)}(\mathbf{u})$ to $z_{LT}^{(r)}(\mathbf{u})$ using (16);

10. Add the previously simulated $z_{LT}^{(r)}(\mathbf{u})$ value to the pool of conditioning data;
11. Proceed to the next simulation location \mathbf{u} according to the previously established random path established in step 6 and loop over steps 7 to 10 until all simulation locations have been visited;
12. Repeat steps 6 through 11 with a different random path and random number seed to construct $r = 1, \dots, R$ realizations.

Application Example

An example is presented to illustrate the LVT approach to simulation with the trend. The data used in this example are from a vein-type gold deposit. There are 67 samples (g/t) located on a 2D easting-elevation section. Figure 4 shows the distribution of equally weighted gold data spatially in a location map (left) and statistically in a cdf (top right). Since clustered samples were taken from higher gold grade areas, polygonal declustering is performed to obtain a representative cdf (bottom right). Although this cdf is not used directly in the LVT approach, it will be referenced later for validation of the LVT simulation results. The representative gold grade distribution is positively skewed with a mean and variance of 0.86 and 1.32, respectively.

The location map in Figure 4 reveals a strong trend of decreasing gold concentration with depth that should be honored in simulation. Since there is an inadequate amount of data for reproducing the trend automatically, the trend must be incorporated explicitly. The LVT simulation approach is now presented.

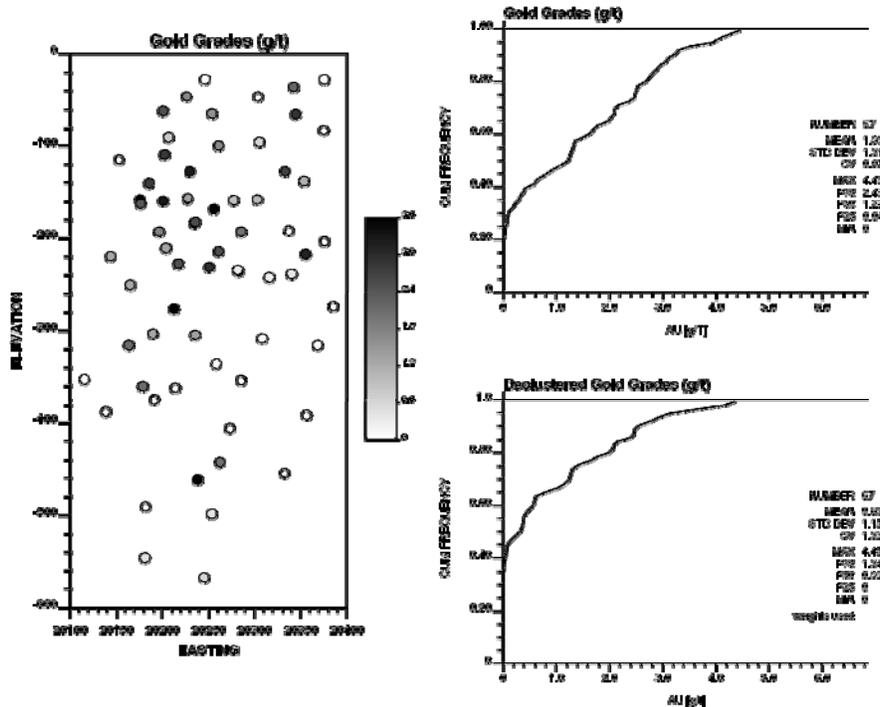


Figure 4: The distribution of original z sample data shown spatially in the location map (left) and as a cdf (top right) and the declustered distribution shown as a cdf (bottom right). The same grayscale is used throughout the remainder of the application example.

At each of the 67 sample data locations, global ordinary kriging is performed to determine the $F_Z(\mathbf{u}; z)$ cdfs. A 3×3 discretization is combined with the 40% nugget variogram with 500m isotropic range. Figures 12 and 13 show the distributions of t and y_{LT} values, respectively, with a location map (left) and cdf (right). The standard normal cdf of transformed y values is also calculated and shown with both cdfs as a shaded

line. Notice in Figure 5 that although both the $F_T(t)$ and $F_Y(y)$ cdfs are normal with a mean of zero, $F_T(t)$ is non-standard normal with a variance of 0.69. Both distributions in Figure 6 are standard normal.

To show the variation in the local $F_Z(\mathbf{u}; z)$ cdfs, a high valued and low valued location are considered for observation. Figure 7 shows the $F_Z(\mathbf{u}; z)$ and $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ spatial distribution of weights for a low (left) and high (right) valued location. The sample locations are enclosed in a small square. The equally weighted $F_Z(z)$ cdf in Figure 4 is also shown with a shaded line. Notice the increase of $F_Z(\mathbf{u}; z)$ in low valued areas and increase of $F_Z(\mathbf{u}; z)$ in high valued areas. The means $m(\mathbf{u})$ of the low and high valued \mathbf{u} location $F_Z(\mathbf{u}; z)$ cdfs are 0.74 and 1.91, respectively. Now the local $F_Z(\mathbf{u}; z)$ cdfs are calculated for each \mathbf{u} location on a 150 x 300 grid of simulation locations. There are then a total of 45,000 $F_Z(\mathbf{u}; z)$ cdfs that are calculated. The locally varying mean $m(\mathbf{u})$ or trend is extracted as the expected value of each of these local cdfs. Figure 8 shows the spatial distribution and cdf of the $m(\mathbf{u})$ trend values. Notice the smooth deterministic nature of the trend in the spatial distribution and low variance in the cdf. Also notice that the trend model lowers the equal weighted mean from 1.36 to 1.19 accounting for the clustered high grade gold samples.

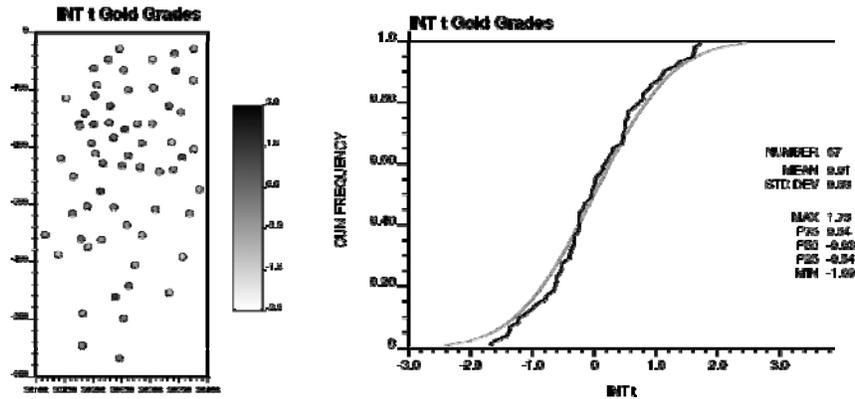


Figure 5: The location map (left) and cdf (right) of the 67 calculated t values. The distribution of y values is also shown with a shaded line for comparison.

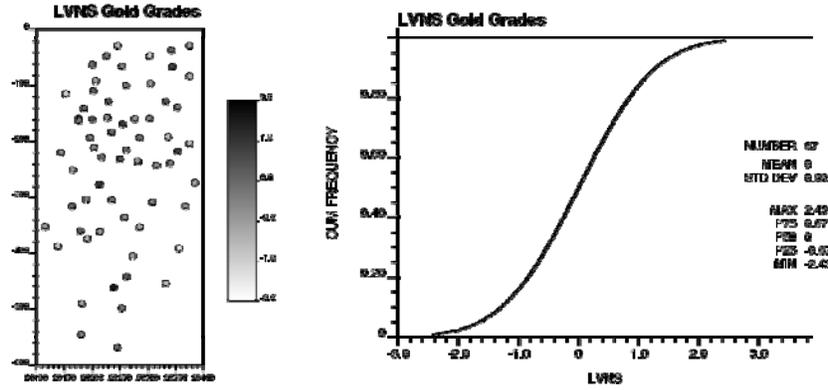


Figure 6: The location map (left) and cdf (right) of the 67 calculated y_{LT} values. The distribution of y values is essentially identical.

The variogram mapping procedure is now performed to convert the $\gamma_Y(\mathbf{h})$ variogram to the full $\gamma_Y^{LT}(\mathbf{h})$ variogram needed for the interpolation of y_{LT} data. A variography study was conducted on the normal score y values of the 67 original z gold sample data. The highest correlation and principle variogram direction is at 45° in the easting-elevation plane; the minor variogram direction is then at a 135° direction. Figure 8 shows the final calculated $\gamma_Y(\mathbf{h})$ variogram points and model line for both the 45° (dark) and 135° (shaded) direction. The analytical form of the $\gamma_Y(\mathbf{h})$ variogram model is:

$$\gamma_Y(\mathbf{h}) = 0.20 + 0.80 \cdot Sph(\mathbf{h}) \begin{matrix} a_{45^\circ} = 145\text{m} \\ a_{135^\circ} = 90\text{m} \end{matrix} \quad (28)$$

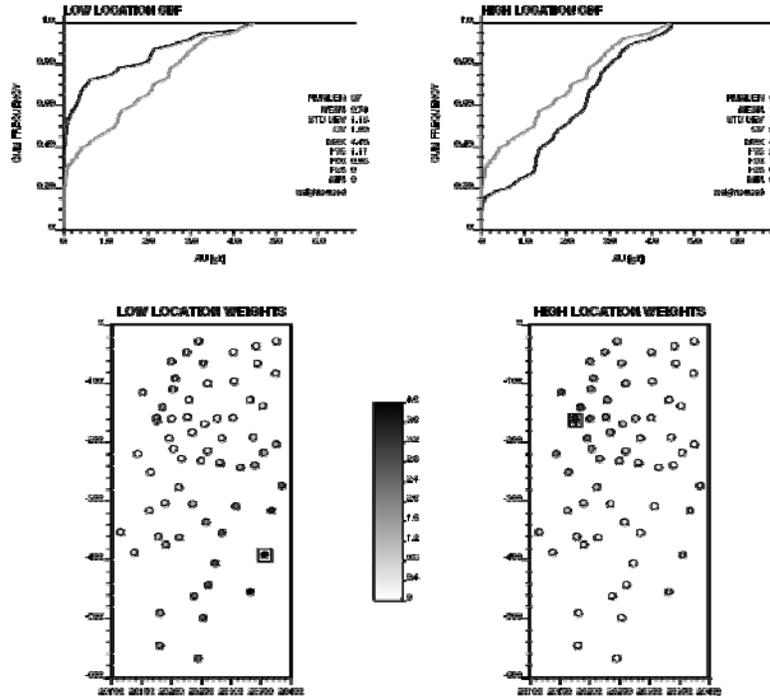


Figure 7: The local $F_Z(\mathbf{u}; z)$ cdf and spatial distribution of $w^{LT}(\mathbf{u}; \mathbf{u}_s)$ weights for the low valued \mathbf{u} location (left) and high valued \mathbf{u} location (right) indicated by the squares.

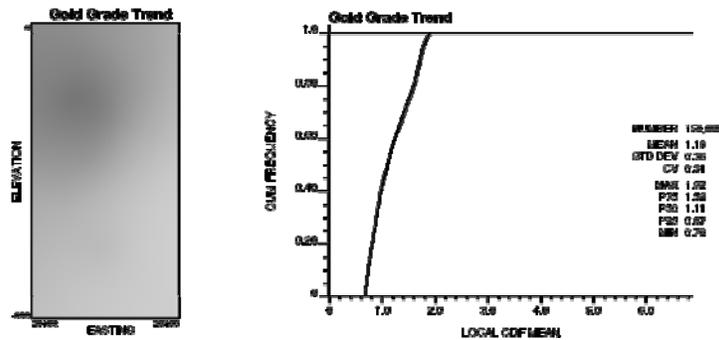


Figure 8: The trend $m(\mathbf{u})$ map (left) and cdf calculated from all 150,000 local $F_Z(\mathbf{u}; z)$ cdfs. The grayscale is the same as in Figure 5.

The variogram model in relation (28) is then mapped to the $\gamma_{Y^{LT}}(\mathbf{h})$ variogram model. The $\gamma_{Y^{LT}}(\mathbf{h})$ variogram model is shown in Figure 9 with the broken dark and shaded lines. Notice the spatial correlation of the y_{LT} values is much greater than that for the y values. The analytical form of the $\gamma_{Y^{LT}}(\mathbf{h})$ variogram model is:

$$\gamma_{Y^{LT}}(\mathbf{h}) = 0.10 + 0.69 \cdot Sph(\mathbf{h})_{\substack{a_{45^\circ} = 140\text{m} \\ a_{135^\circ} = 75\text{m}}} + 0.21 \cdot \mathbf{h}_{\substack{a_{45^\circ} = \infty \\ a_{135^\circ} = 295\text{m}}} \quad (29)$$

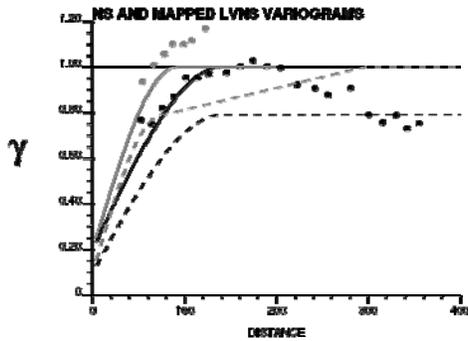


Figure 9: The γ values normal score $\gamma_Y(\mathbf{h})$ experimental (points) and model (line) variogram shown with the mapped $\gamma_{Y_{LT}}(\mathbf{h})$ model (broken line) variogram. The principle 45° direction is shown in dark and the minor 135° direction is shaded.

All of the information required to perform simulation with the trend imbedded within locally varying $F_Z(\mathbf{u}; z)$ cdfs is ready. The z data have been collected and transformed to y_{LT} data with the 67 local $F_Z(\mathbf{u}; z)$ cdfs at sample \mathbf{u} locations, the 150,000 simulation location $F_Z(\mathbf{u}; z)$ cdfs have been calculated, the trend $m(\mathbf{u})$ model has been extracted and visualized, and the $Y_{LT}(\mathbf{u})$ spatial law has been determined. The remaining seven steps in the SGS simulation procedure are now implemented for $R = 30$ realizations. The first four realizations are shown in Figure 10 spatially and in the form of a cdf. The declustered cdf and original cdf in Figure 4 are also shown with the shaded and broken lines, respectively, for comparison. The declustered distribution is honored via incorporating the trend through the locally varying transformation. The original cdf without declustering is not reproduced.

The LVT approach to simulation with the trend should honor the large-scale features of the trend. This check is performed by comparing the etype from the 30 realizations to the trend model in Figure 8. As Figure 11 illustrates, the large-scale features of the trend model are honored. The most important advantage of the LVT approach is the ability to reproduce key features in the trend with an underlying spatial law that can be calculated directly.

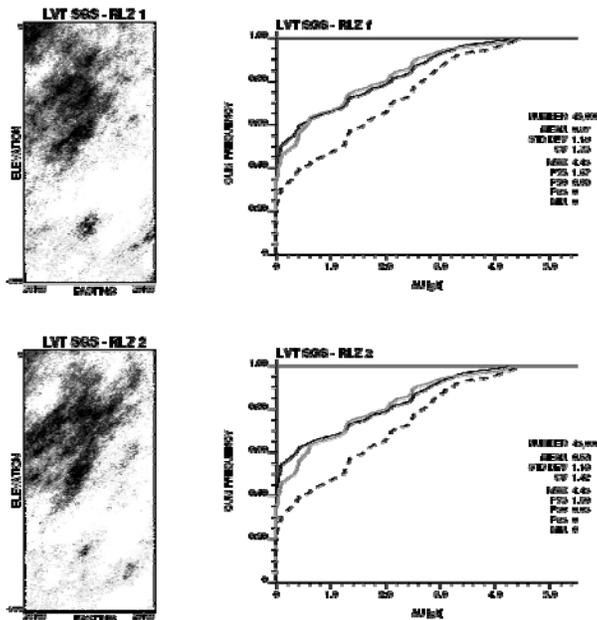


Figure 10: The first two LVT SGS simulated realization maps (left) and cdfs (right). The declustered distribution of z values is also shown with a shaded line for comparison. The grayscale is as in Figure 5.

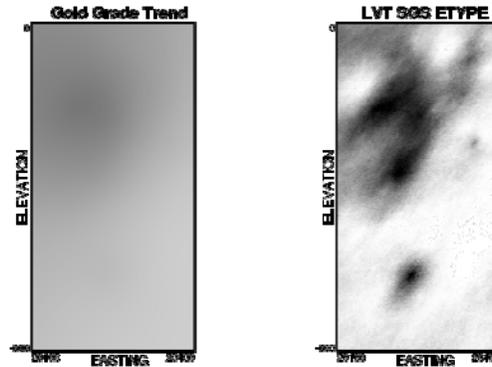


Figure 11: The comparison between the input $m(\mathbf{u})$ trend model imbedded within the local $F_Z(\mathbf{u}; z)$ cdfs and the etype from 30 SGS realizations using the LVT approach. The grayscale is as in Figure 5.

Remarks

Prediction with the trend can be an important aspect of improving predictions. Regardless of the chosen domain size, the SRF formalism may be incapable of accounting for large-scale continuity characteristic of the trend model when there is an inadequate amount of conditioning data. There is a large variety of estimation and simulation techniques available to incorporate trends. Underlying these techniques, however, is an ambiguous definition of the variogram and spatial parameters needed. This problem can be addressed with techniques that combine separate RFs or by modeling the trend with no more variability than is offered deterministically. Still, there is no objective quantitative procedure for inferring the spatial law underlying prediction with the trend. The motivation for the LVT approach to predicting with the trend was an alternative prediction technique for which the underlying variogram can be calculated directly. The fundamental basis of the LVT technique is the use of a locally varying transformation to and from standard normal space where conventional prediction can be performed. The local non-stationary cdfs used for the transformations effectively account for the trend and allow the correct spatial law to be calculated directly.

References

- Alabert, F. *Stochastic Imaging of Spatial Distributions Using Hard and Soft Information*. Master's Thesis, Stanford University, California, 1987.
- Armstrong, M. *Problems with Universal Kriging*. *Mathematical Geology*, vol.16, no. 1, 1984.
- Deutsch, C.V. *Geostatistical Reservoir Modeling*. Oxford University Press, New York, 2002.
- Deutsch, C.V. *Stepwise Conditional Transformation in Estimation Mode*. Center for Computational Geostatistics (CCG), University of Alberta, Report Eight, 2006.
- Dimitrakopoulos, R. *Conditional Simulation of Intrinsic Random Functions of Order k*. *Mathematical Geology*, vol. 22, no. 3, 1990.
- Journel, A.G. *Constrained Interpolation and Qualitative Information*. *Mathematical Geology*, vol. 18, no. 3, 1986.
- Journel, A.G. *Fundamentals of Geostatistics in Five Lessons*. American Geophysical Union, 1989.
- Krige, D.G. *A Statistical Approach to some Mine Valuations and Allied Problems at the Witwatersrand*. Master's Thesis, University of Witwatersrand, South Africa, 1951.
- Leuangthong, O. and Deutsch, C.V. *Transformation of Residuals to Avoid Artifacts in Geostatistical Modeling with a Trend*. *Mathematical Geology*, vol. 36, no. 3, 2004.
- Matheron, G. *Le Krigeage Universel: Fascicule 1*. Cahiers du CMM, 1969.
- Marechal, A. *Kriging Seismic Data in Presence of Faults*. *Geostatistics for Natural Resources Characterization*, vol. 2, Reidel Publishing Company, Netherlands, 1984.
- McLennan, J.A., Leuangthong, O. and Deutsch, C.V. *Kriging with a Trend*. Center for Computational Geostatistics (CCG), University of Alberta, Report Eight, 2006.
- McLennan, J.A., Leuangthong, O. and Deutsch, C.V. *A Closer Look at the Kriging Equations*. Center for Computational Geostatistics (CCG), University of Alberta, Report Eight, 2006.
- McLennan, J.A. and Deutsch, C.V. *Conditional Non-Bias of Geostatistical Simulation for Prediction of Recoverable Reserves*. *Canadian Institute of Mining and Metallurgy (CIM) Bulletin*, vol. 97, no. 1080, 2004.
- Zhu, H. and Journel, A.G. *Formating and Integrating Soft Data: Stochastic Imaging via the Markov-Bayes Algorithm*. Fourth International Geostatistics Conference: Troia 1992, Soares, A. (Ed), vol. 1, 1993.