Methodology for Geostatistical Model of Gradational Geological Boundaries: Local Non-stationary LMC

Paula Larrondo and Clayton V. Deutsch

Department of Civil & Environmental Engineering
University of Alberta

Abstract

Geostatistical techniques, such as kriging and simulation, make strong assumptions of stationarity in the mean and the variance over the domain of interest. Unfortunately, geological nature usually does not reflect this assumption and we are forced to subdivide our model area into stationary regions that have some common geological controls and similar statistical properties. This paper addresses the significant complexity introduced by boundaries. Boundaries are often soft, that is, samples near boundaries influence multiple rock types. Due to the transitional nature of geological mechanisms is rather common that soft boundaries are characterized by a non-stationary behavior of the mean, variance or covariance.

We propose a new technique that accounts for stationary variables within rock types and additional non-stationary factors near boundaries. The technique involves the following distinct phases: (i) identification of the rock types and boundary zones based on geological modeling and the timing of different geological events, (ii) optimization of the non-stationary components of mean and variance in the boundary zone given the stationary statistical parameters of each domain and the data in the boundary region, (iii) decomposition of the covariance model into stationary and non-stationary components of a linear model of coregionalization and optimization of the latest, and (iv) estimation of grades using non-stationary cokriging. The resulting technique can be thought of as non-stationary cokriging in presence of geological boundaries. The theoretical framework and notation for this new technique is developed and illustrated with an example.

Introduction

The most common geostatistical techniques, such as kriging and Gaussian/indicator simulation, are based on strong assumptions of stationarity of the estimation domains. In particular, they are based in a second order stationary hypothesis, that is, the mean, variance and covariance remain constant across the entire domain and they do not depend on the location of the support points but only in the distance between them. Therefore, geological domains should be chosen as statistically homogeneous zones, which are geologically significant and coherent, but with enough data to allow reliable inference of the required statistics.

Once estimation domains have been selected, the nature of the boundaries between them must be established. Domain boundaries are often referred to as either ‘hard’ or ‘soft’. Hard boundaries are found when an abrupt change in the mean or variance occurs at the contact between two domains, such as coal seams or sedimentary zinc deposits. Hard boundaries do not permit the

interpolation or extrapolation across domains. In deposits where the disseminated mineralisation has a gradational nature, such as some porphyry copper deposits, and grades change transitionally across a boundary, the contact is referred to as a soft boundary (Figure 1). Soft domain boundaries allow selected data from either side of a boundary to be used in the estimation of each domain.

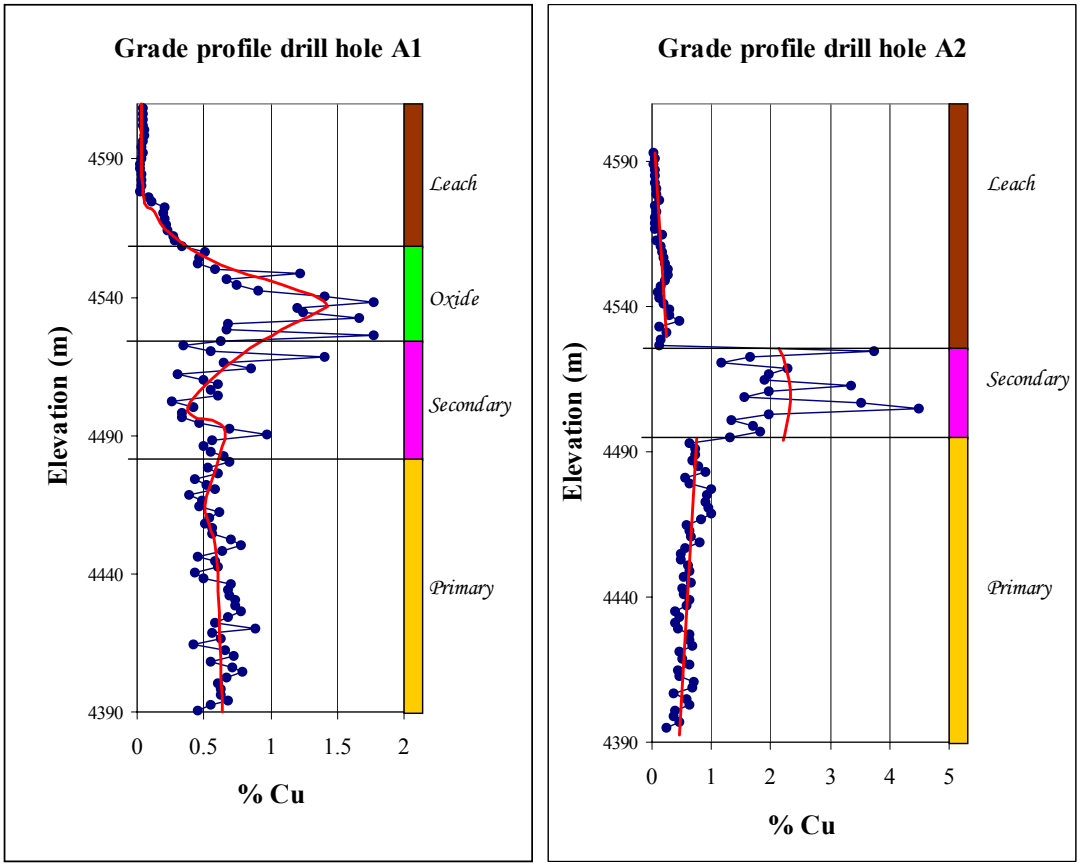

Figure 1: Grade profile of two drill holes in a porphyry copper deposit in Northern Chile, showing an example of soft boundaries (left) between mineralisation units as well as an example of hard boundaries (right). Notice how the mean and variance change across the boundaries.

It is rather common that soft boundaries are characterized by a non-stationary behavior of the variable of interest in the proximities of the boundary, that is, the mean, variance or covariance are no longer constant within a zone of influence of one rock type into the other, and their values depends on the location relative to the boundary. There could be an increase or decrease in the mean or variance towards the boundary. An example is the increased frequency of fractures towards a boundary between geological domains of structural nature. Faults or brittle zones are examples of this transition. The fractures, if mineralized, may cause the average to increase close to the boundary. Alternatively, fractures near the surface of the deposit may be leached by meteoric fluids, which may translate to a decrease in the average grade. The increase in the presence of factures will often lead to an increase in the variance closer to the boundary.

Although soft boundaries are found in several types of geological settings due to the transitional nature of the geological mechanisms, conventional estimation usually treats the boundaries between geological units as hard boundaries. This is primarily due to the limitations of current estimation and simulation procedures. We will show that non-stationary features in the vicinity of a boundary can be parameterized into a local model of coregionalization through the optimization of the non-stationary components of the mean, variance and covariance given their stationary ones. With a legitimate spatial model, estimation of grades can be performed using a form of non-stationary cokriging. This proposal provides an appealing alternative when complex contacts between different rock types exist. The correct representation of soft boundaries should ensure the reproduction of the correlation of the grades across the boundary and ensure reproduction of non-stationary variations of the mean, variance and covariance in the zone of influence of each rock type. Boundaries are of special interest in the short mine planning and improved modelling of boundaries would benefit the design and operation stage in both underground and open pit deposits. We develop the methodology in the context of mining geostatistics, but it is widely applicable in many different settings.

Theoretical Background and Notation

The technique involves the identification of stationary variables within each rock type and additional non-stationary components near boundaries for the mean, variance and covariance. For a geological model with K rock types or estimation domains, there are a maximum of $K(K-1)/2$ boundary zones to be defined. Then, the continuous random function $Z(\mathbf{u})$ that represents the distribution of the property of interest can be decomposed into K stationary random variables $Z_k(\mathbf{u})$ $k=1,...,K$ and a maximum of $K(K-1)/2$ non-stationary boundary variables $Z_{kp}(\mathbf{u})$, with $k,p=1,...K$ and $Z_{kp}(\mathbf{u})=Z_{pk}(\mathbf{u})$ (Figure 2). Then, at all locations, $Z(\mathbf{u})$ can be explained by the sum of a stationary component from the collocated rock type and perhaps a single non-stationary boundary variable. By definition, the non-stationary variable will take values only for locations within the maximum distance of influence of rock type k into rock type p.

The maximum distance of influence orthogonal to the boundary of rock type k into rock type p is denoted $dmax_{kp}$. A boundary zone is defined by two distances: $dmax_{kp}$ and $dmax_{pk}$, since there is no requirement that the regions on each side of the boundary are symmetric, that is, $dmax_{kp} \neq dmax_{pk}$. The modeler using all geological information available and his expertise should establish these distances. An automatic optimization algorithm is unlikely to work given that the stationary portion of the mean and variance, as well as the non-stationary factors that affect this statistics, are also unknown.

When more than two rock types converge at a boundary, two or more rock types may influence the boundary zone in the adjacent domain. In this case, precedence or ordering rules should determine the dominant boundary zone (Figure 3A). Although the behavior of a property near a boundary could be explained by the overlapping of different geological controls, the task of identifying the individuals effects of each rock type and their interactions can be quite difficult. Geological properties are not usually additive and therefore the response of a combination of different rock types is complex. Only one non-stationary factor will be considered at each location. The modeler should put together the precedence rules based on the geology of the deposit. The relative timing of intrusion, deposition or mineralisation events, geochemistry response of the protolith to an alteration or mineralisation process could be used to resolve timing

and important variables. If the geological data does not provide sufficient information to establish a geological order of events, a neutral arrangement can be chosen (Figure 3B). In this case, the precedent rock type *p* at a location will be the one to which the distance to the boundary is the minimum over all surrounding rock types.

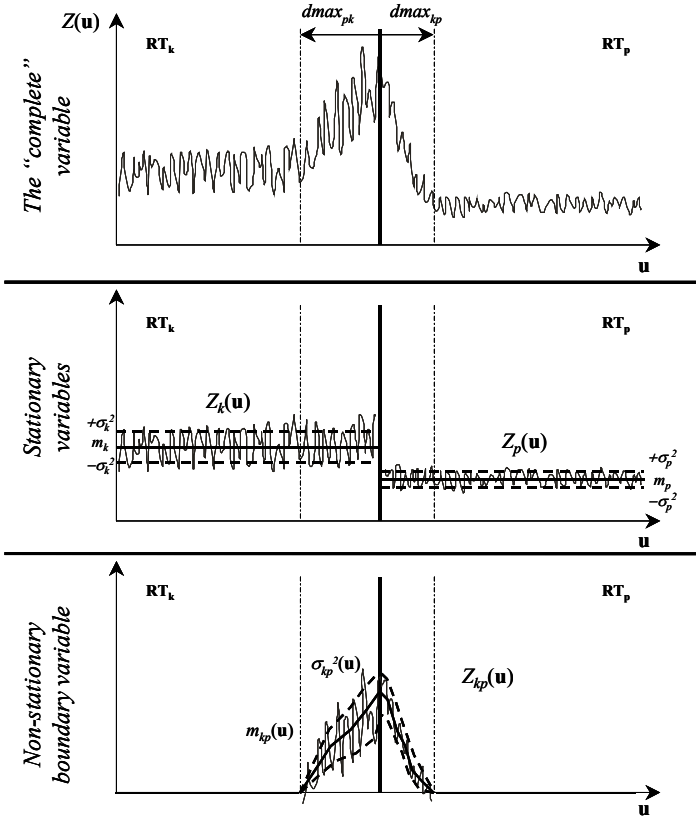

Figure 2: Decomposition of a one-dimensional random function Z(u) in two stationary variables Zk(u) and Zp(u), with constant mean and variance, and a non-stationary boundary variable Zkp(u), with a mean and variance that are functions of the distance to the boundary.

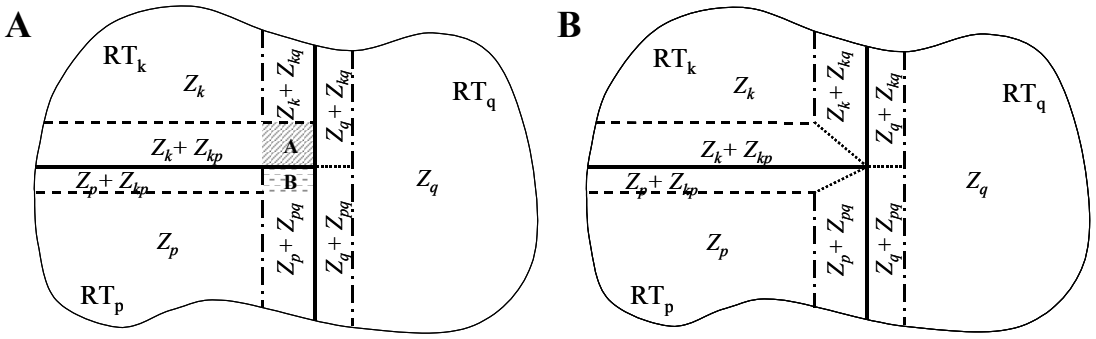

Figure 3: (A) Three rock type example where a predefined precedence rule is used to determine the precedent rock type over two possibilities. In the region A of $\mathbf{RT_k}$, $\mathbf{RT_p}$ is precedent over $\mathbf{RT_q}$, while in region B of $\mathbf{RT_p}$, $\mathbf{RT_q}$ is precedent over $\mathbf{RT_k}$. (B) A neutral arrangement for the same example.

Stationary and non Stationary Statistical Parameters

The mean function of the continuous random function $Z(\mathbf{u})$ for an specific rock type k will be the mean of the stationary variable Z_k plus the mean of any corresponding non-stationary variable $Z_{kp}(\mathbf{u})$:

$$E\{Z(\mathbf{u}_i)\} = E\{Z_k\} + E\{Z_{kp}(\mathbf{u}_i)\} = m_k + m_{kp}(\mathbf{u}_i) \quad \text{where } \mathbf{u}_i \in \mathbf{RT_k}$$

where p is the adjacent rock type that shares a boundary with rock type k.

The stationary component of the mean (m_k) is independent of location and is a constant value. The non-stationary component of the mean (m_{kp}) is a function of the distance to the boundary, $d_{pk}(\mathbf{u})$ and takes values different than zero for locations within the boundary zone defined by rock types k and p. The mean of rock type k:

$$E\{Z(\mathbf{u}_i)\} = \begin{cases} m_k & \text{, if } d_{pk}(\mathbf{u}_i) \geq dmax_{pk} \\ m_k + f(d_{pk}(\mathbf{u}_i)) & \text{, otherwise} \end{cases} \quad \text{where } \mathbf{u}_i \in \mathbf{RT_k}$$

where p is the adjacent rock type that shares a boundary with rock type k and $f(\bullet)$ is an arbitrary function that describes the mean as a function of distance to the boundary.

Similarly, the variance of $Z(\mathbf{u})$ for rock type k will be the sum of a constant stationary variance (σ_k^2) due to Z_k and the independent non-stationary variance (σ_{kp}^2) due to $Z_{kp}(\mathbf{u})$. The variance of a random function $Z(\mathbf{u})$ in a rock type k:

$$E\left\{\left(Z(\mathbf{u}_i) - E\{Z(\mathbf{u}_i)\}\right)^2\right\} = \begin{cases} \sigma_k^2 & \text{, if } d_{pk}(\mathbf{u}_i) \geq dmax_{pk} \\ \sigma_k^2 + g(d_{pk}(\mathbf{u}_i)) & \text{, otherwise} \end{cases} \quad \text{where } \mathbf{u}_i \in \mathbf{RT_k}$$

where p is the adjacent rock type that shares a boundary with rock type k and $g(\bullet)$ is an arbitrary function that describes the variance as a function of distance to the boundary.

As with the mean and variance, the covariance structure between two rock types that share a local non-stationary boundary consists of: a stationary and a non-stationary component.

$$Cov_Z(\mathbf{u}_i, \mathbf{v}_i) = E\left\{\left(Z(\mathbf{u}_i) - m(\mathbf{u}_i)\right) \cdot \left(Z(\mathbf{v}_i) - m(\mathbf{v}_i)\right)\right\} = Cov_Z^{\mathbf{S}}(\mathbf{h}) + Cov_Z^{\mathbf{NS}}(\mathbf{u}_i, \mathbf{v}_i)$$

where $\mathbf{h}=\mathbf{u}_i - \mathbf{v}_i$. Since Z_k and $Z_{kp}(\mathbf{u})$ are independent random variables, the cross terms are zero, therefore the covariance of $Z(\mathbf{u})$ is the sum of the stationary and non-stationary components. The combination of these components corresponds to a local linear model of coregionalization.

The stationary component of the covariance can be calculated and modeled from data pairs within the same internal stationary portion of a rock type, that is $\mathbf{u}_i$ and $\mathbf{v}_i$ belong to rock type k, and do not belong to any boundary zone.

To obtain the non-stationary component of the covariance model we will assume that the shape, anisotropies and relative nugget effect of the correlation for the non-stationary variable $Z_{kp}(\mathbf{u})$ $k,p=1,\ldots,K$ are stationary and that they can be specified by the modeler. Due to the non-stationary nature of variable $Z(\mathbf{u})$ at the boundary zone, this stationary spatial model shape has to by scaled at each point by a non-stationary mean and variance. The relative standardized variogram model for the boundary zone is:

$$\hat{\gamma}_{kp}(\mathbf{u}_i,\mathbf{v}_i)=\frac{1}{2}\cdot E\left\{\left[\frac{Z(\mathbf{u}_i)-m(\mathbf{u}_i)}{\sigma(\mathbf{u}_i)}-\frac{Z(\mathbf{v}_i)-m(\mathbf{v}_i)}{\sigma(\mathbf{v}_i)}\right]^2\right\}$$

where $m(\mathbf{u})=m_{kp}(\mathbf{u})+m_k$ and $\sigma(\mathbf{u})=\sigma_{kp}(\mathbf{u})+\sigma_k$. Note that the stationary component of the mean can be either m_k or m_p depending whether $\mathbf{u}_i$ or $\mathbf{v}_i$ belongs to rock type k or rock type p. The same occurs for the stationary component of the variance.

Expanding and reordering the terms of the squared difference, and since $E\left\{Z(\mathbf{u}_i)^2\right\}=\sigma(\mathbf{u}_i)^2+m(\mathbf{u}_i)$ and $E\left\{Z(\mathbf{u}_i)\right\}=m(\mathbf{u}_i)$, we can simplify the previous expression as follows,

$$2\hat{\gamma}_{kp}(\mathbf{u}_i,\mathbf{v}_i)=\frac{\sigma(\mathbf{u}_i)^2+m(\mathbf{u}_i)^2-2\cdot m(\mathbf{u}_i)^2+m(\mathbf{u}_i)^2}{\sigma(\mathbf{u}_i)^2}$$
$$-2\left(\frac{E\left\{Z(\mathbf{u}_i)\cdot Z(\mathbf{v}_i)\right\}-m(\mathbf{v}_i)\cdot m(\mathbf{u}_i)-m(\mathbf{u}_i)\cdot m(\mathbf{v}_i)+m(\mathbf{u}_i)\cdot m(\mathbf{v}_i)}{\sigma(\mathbf{u}_i)\cdot\sigma(\mathbf{v}_i)}\right)$$
$$+\frac{\sigma(\mathbf{v}_i)^2+m(\mathbf{v}_i)^2-2+m(\mathbf{v}_i)^2+m(\mathbf{v}_i)^2}{\sigma(\mathbf{v}_i)^2}$$

$$2\hat{\gamma}_{kp}(\mathbf{u}_i,\mathbf{v}_i)=2-2\left(\frac{E\left\{Z(\mathbf{u}_i)\cdot Z(\mathbf{v}_i)\right\}-m(\mathbf{v}_i)\cdot m(\mathbf{u}_i)}{\sigma(\mathbf{u}_i)\cdot\sigma(\mathbf{v}_i)}\right)$$

$$\hat{\gamma}_{kp}(\mathbf{u}_i,\mathbf{v}_i)=1-\frac{Cov_z^{\mathbf{NS}}(\mathbf{u}_i,\mathbf{v}_i)}{\sigma(\mathbf{u}_i)\cdot\sigma(\mathbf{v}_i)}$$

The non-stationary standard deviations of $\mathbf{u}_i$ and $\mathbf{v}_i$ are the ones to scaled the stationary shape of the non-stationary component of the covariance. Reordering the terms and replacing the mean and variance by the sum of their stationary and non-stationary components, we obtained an expression for the non-stationary covariance model:

$$Cov_Z^{\mathbf{NS}}(\mathbf{u}_i,\mathbf{v}_i)=E\left\{Z(\mathbf{u}_i)\cdot Z(\mathbf{v}_i)\right\}-(m_{kp}(\mathbf{v}_i)+m_k)\cdot(m_{kp}(\mathbf{u}_i)+m_k)$$
$$=(1-\hat{\gamma}_{kp}(\mathbf{u}_i,\mathbf{v}_i))\cdot(\sigma_{kp}(\mathbf{u}_i)+\sigma_k)\cdot(\sigma_{kp}(\mathbf{v}_i)+\sigma_k)$$

Since the shape, anisotropies and nugget effect of the relative standardize variogram are inputs from the modeler, the only parameter that must be established for the non-stationary covariance model is the range.

Optimization of the Statistical Parameters

We need to find the optimum $f(d_{pk}(\mathbf{u}_i))$, $g(d_{pk}(\mathbf{u}_i))$ and $Cov_Z^{\mathbf{NS}}(\mathbf{u}_i,\mathbf{v}_i)$ that fit the distribution of the random variable $Z(\mathbf{u})$ at the boundary zone given the stationary components of mean, variance and covariance, a set of precedence rules and the maximum distances of influence within the rock type model.

Although we know that the non-stationary behavior is a function of the distance of the sample to the boundary, there are several possible analytical expressions that fit the distribution of the non-stationary random variable. For the purpose of showing the proposed methodology we will consider that the non-stationary component of mean and variance follow a linear function (Figure 4) of the distance to the boundary (d_{pk}). In this scenario, the optimization of the parameter m_{kp} and σ_{kp}^2 will be equivalent to optimizing estimates of the intercepts: a_{kp} and b_{kp}.

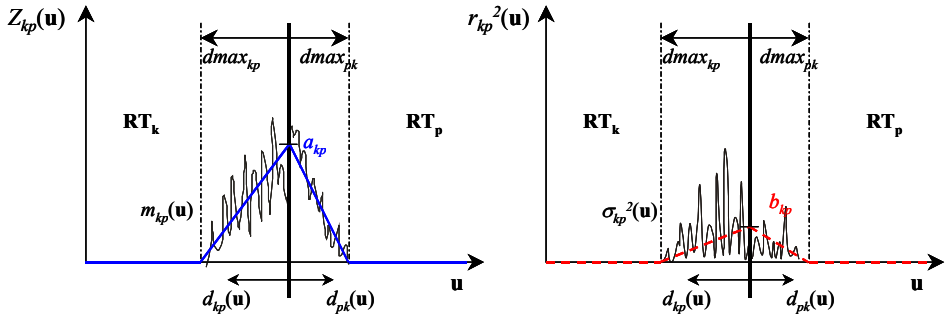

Figure 4: Mean and variance of the random variable Z_{kp} (**u**), modeled by a linear function of the distance to the boundary.

The mean m_{kp} is optimized given that m_k is known from the experimental average of data within rock type k, outside any boundary zone. The objective function is:

$$O_m = \sum_{k=1}^{K}\sum_{p=1}^{P}\sum_{i=1}^{N_{kp}}\left[z(\mathbf{u}_i) - (\hat{m}_k + m_{kp}(\mathbf{u}_i))\right]^2$$

where z($\mathbf{u}_i$) is the outcome value at every location in the boundary zone, N_{kp} is the total number of data in zone k-p, $\hat{m}_k$ is the experimental average of all data in $\mathbf{RT_k}$ and outside any boundary zone, and $m_{km}(\mathbf{u}_i)$ is the non-stationary mean at location $\mathbf{u}_i$ calculated as:

$$m_{kp}(\mathbf{u}_i) = \begin{cases} \dfrac{\left(dmax_{kp} - d_{kp}(\mathbf{u}_i)\right)}{dmax_{km}} \cdot a_{kp} & \text{for } 0 \le d_{kp}(\mathbf{u}_i) \le dmax_{kp} \\ \dfrac{\left(dmax_{pk} - d_{pk}(\mathbf{u}_i)\right)}{dmax_{pk}} \cdot a_{kp} & \text{for } 0 \le d_{pk}(\mathbf{u}_i) \le dmax_{pk} \\ 0 & \text{for } d_{kp}(\mathbf{u}_i) \ge dmax_{kp} \text{ and } d_{pk}(\mathbf{u}_i) \ge dmax_{pk} \end{cases}$$

Although the experimental mean of rock type k should include all samples in this geological unit, the samples in any boundary zone are excluded since their non-stationary component is yet to be determined.

The optimization of the mean can be achieved by iteratively modifying a_{kp} $\forall k,p$, in a random fashion while accepting all changes in a_{kp} that reduce objective function. This is a simplified version of the simulated annealing formalism.

Since the variance is a statistic of second order that depends on the mean, once the optimum of m_{kp} is found, we can proceed to optimize ${\sigma_{kp}}^2$ assuming the ${\sigma_k}^2, k=1,...,K$ values are known from the experimental variance of data within the internal stationary portion of rock type k.

The optimum ${\sigma_{kp}}^2$, will be the one that minimizes the following objective function:

$$O_{\sigma^2} = \sum_{k=1}^{K}\sum_{p=1}^{P}\sum_{i=1}^{N_{kp}}\left[\mathrm{r}(\mathbf{u}_i)^2 - (\hat{\sigma}_k^{\,2} + {\sigma_{kp}}^2(\mathbf{u}_i))\right]^2$$

where $\mathrm{r}(\mathbf{u}_i)$ is the residual value at every location in the boundary zone, that is, $r(\mathbf{u}_i) = z(\mathbf{u}_i) - (m_k + m_{kp}(\mathbf{u}_i))$. $\hat{\sigma}_k^{\,2}$ is the experimental variance of all data within the stationary region of rock type k, and ${\sigma_{kp}}^2(\mathbf{u}_i)$ is the non-stationary variance at location $\mathbf{u}_i$ calculated from:

$${\sigma_{kp}}^2(\mathbf{u}_i) = \begin{cases} \dfrac{\left(dmax_{kp} - d_{kp}(\mathbf{u}_i)\right)}{dmax_{kp}} \cdot b_{kp} & \text{for } 0 \le d_{kp}(\mathbf{u}_i) \le dmax_{kp} \\ \dfrac{\left(dmax_{pk} - d_{pk}(\mathbf{u}_i)\right)}{dmax_{pk}} \cdot b_{kp} & \text{for } 0 \le d_{pk}(\mathbf{u}_i) \le dmax_{pk} \\ 0 & \text{for } d_{kp}(\mathbf{u}_i) \ge dmax_{kp} \text{ and } d_{pk}(\mathbf{u}_i) \ge dmax_{pk} \end{cases}$$

To find the optimum covariance model we minimize the following objective function:

$$O_{Cov} = \sum_{i=1}^{N}\left[\hat{C}\left(z(\mathbf{u}_i), z(\mathbf{v}_i)\right) - C_{MOD}\left(z(\mathbf{u}_i), z(\mathbf{v}_i)\right)\right]^2$$

where $\hat{C}$ denotes the experimental covariance of the pair located at $\mathbf{u}_i$ and $\mathbf{v}_i$, which is just the multiplication of the two residual values: $r(\mathbf{u}_i) \cdot r(\mathbf{v}_i)$, and C_{MOD} the modeled boundary covariance, which is the sum of the stationary and non-stationary component.

Find the optimum covariance model of a boundary zone is equivalent to optimizing the range of the relative standardized variogram scaled by the non-stationary standard deviation (Figure 5). This assumption provides some advantages over a full optimization algorithm to find the non-stationary covariance structure. Simplicity and fewer artifacts are the main advantages. The range is iteratively modified by a random amount until the difference between the experimental and modeled covariance is minimized.

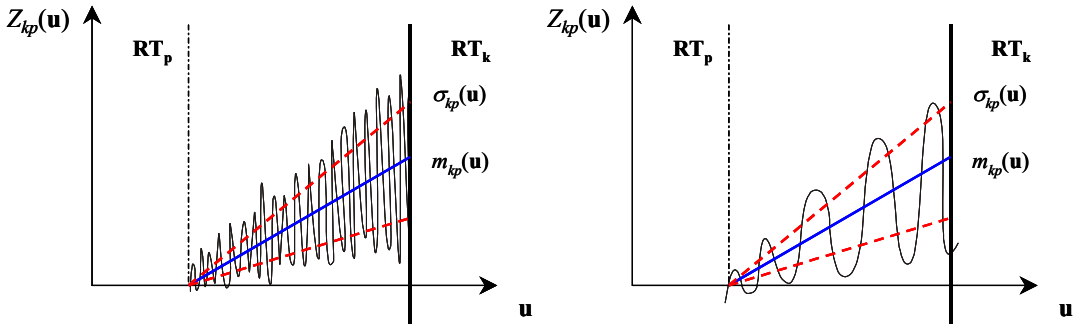

Figure 5: The non-stationary covariance of $Z_{kp}(\mathbf{u})$ is defined by its non-stationary mean and variance and the shape of the correlation, which for the same mean and variance can be different as represented by the outcomes profile. The range of correlation is higher in the right.

Estimation in presence of local non-stationary boundaries

The basic linear regression equation for non-stationary simple kriging is:

$$z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha(\mathbf{u}) \cdot \left[z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha) \right]$$

where $z^*(\mathbf{u})$ is the estimate at unsampled location $\mathbf{u}$, $m(\mathbf{u})$ is the stationary plus the non-stationary mean value at location $\mathbf{u}$, $\lambda_\alpha(\mathbf{u})$ is the weight assigned to datum $z(\mathbf{u}_\alpha)$, n is the number of close data to the location $\mathbf{u}$ being estimated, and $m(\mathbf{u}_\alpha)$ are the n stationary plus the non-stationary mean values at the data locations.

To find the optimal weights $\lambda_\alpha(\mathbf{u})$, $\alpha=1,\ldots,n$ the kriging system must be solved:

$$\sum_{\beta=1}^{n} \lambda_\beta(\mathbf{u}) \cdot Cov(\mathbf{u}_\alpha, \mathbf{u}_\beta) = Cov(\mathbf{u}, \mathbf{u}_\alpha) \quad \text{with } \alpha, \beta = 1,\ldots,n$$

where $\lambda_\alpha(\mathbf{u})$, $\alpha=1,..,\ n$ are the simple kriging weights, $Cov(\mathbf{u}_\alpha, \mathbf{u}_\beta)$, $\alpha,\beta=1,..,\ n$ correspond to the data-to-data covariances, and $Cov(\mathbf{u}, \mathbf{u}_\alpha)$, $\alpha=1,..,\ n$ are the data-to- unknown location covariances. In the presence of local non-stationary boundaries, the terms of the data covariance matrix and the vector of data-to-estimate covariances are obtained combining the stationary and non-stationary covariance model components. If both locations are in the same rock type and one or both locations are in the same boundary zone, the covariance is the stationary plus the non-stationary covariances; otherwise, it is only the stationary component. If they are in different rock types and both samples are in the same boundary zone the covariance is the non-stationary component only. The covariance is zero in all other cases.

Simple kriging estimator is unbiased and provides the minimum estimation variance estimate,

$$\sigma_E^{\,2}(\mathbf{u}) = \sigma^2(\mathbf{u}) - \sum_{\alpha=1}^{n} \lambda_\alpha(\mathbf{u}) \cdot Cov(\mathbf{u}, \mathbf{u}_\alpha)$$

where $\sigma^2(\mathbf{u})$ is the variance, which in our case has a stationary and a non-stationary component as well.

The greatest disadvantage of kriging is that the estimates are smooth and the joint variability of the kriged estimates is incorrect. The amount of missing variability is the kriging variance. Sequential Gaussian simulation overcomes this problem by adding back this missing variability, as it adds a random residual to the estimate, drawn from a normal distribution with zero mean and variance equal to the kriging variance. In the case described in this thesis, the estimation variance has also a non-stationary component that makes the implementation of sequential Gaussian simulation in the presence of local non-stationary boundaries delicate. This implementation is part of the proposed future work.

1-D Example

To illustrate the concepts and methodology described above will use a small 1-D synthetic example. Three independent unconditional simulations, with different variogram models were used to build the variable $Z(\mathbf{u})$, that will represent the metal grade across a boundary between two rock types. Two of the simulations were transformed to a non-standard normal distribution to reflect different average grade and variability across a boundary. The third simulation was transform to a non-standard normal distribution but with mean and variance as functions of distance to the boundary. The mean was assumed to follow a linear function with a symmetric maximum distance of influence of 20 meters and an intercept a_{12}=2.0 while the variance follows a linear function as well, with the same maximum distance of influence and an intercept b_{12}=1.0. The final random variable was obtained by joining the two first simulations to obtain a 1-D array of 200 points and adding the third simulation to the values from locations 80.5 up to 199.5 (Figure 6).

The first step in this methodology is to infer all stationary and non-stationary statistical parameters for each rock type and boundary zone. For this the FORTRAN optimization programs: `opt_mean`, `opt_var` and `opt_cov` were used (Larrondo and Deutsch, 2004). This programs mimic the optimization algorithm describe above.

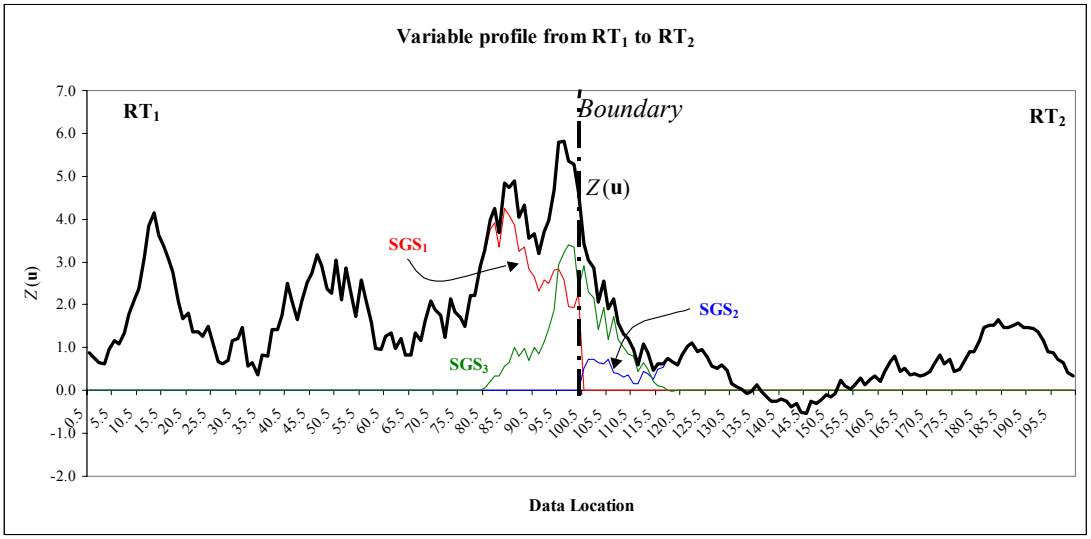

Figure 6: Dataset for 1-D example. Random variable profile of metal content against location along the X-coordinate. **SGS_1**, **SGS_2** and **SGS_3** correspond to the underlying simulations used to generate this synthetic example.

The optimum intercept value a_{21} is 3.22 (Figure 7), slightly higher than the target value of 2.0, used to create this synthetic data, since the 20 values closest to the boundary are relatively higher than the rest of the values of rock type 1, due to ergodic fluctuations in the simulation. The discontinuity in the optimum mean profile after the boundary is a consequence of the difference between the stationary means of rock type *1* and rock type *2*. The reference stationary means are well reproduce by the optimization algorithm within the ergodic fluctuations of the simulations used to build the reference.

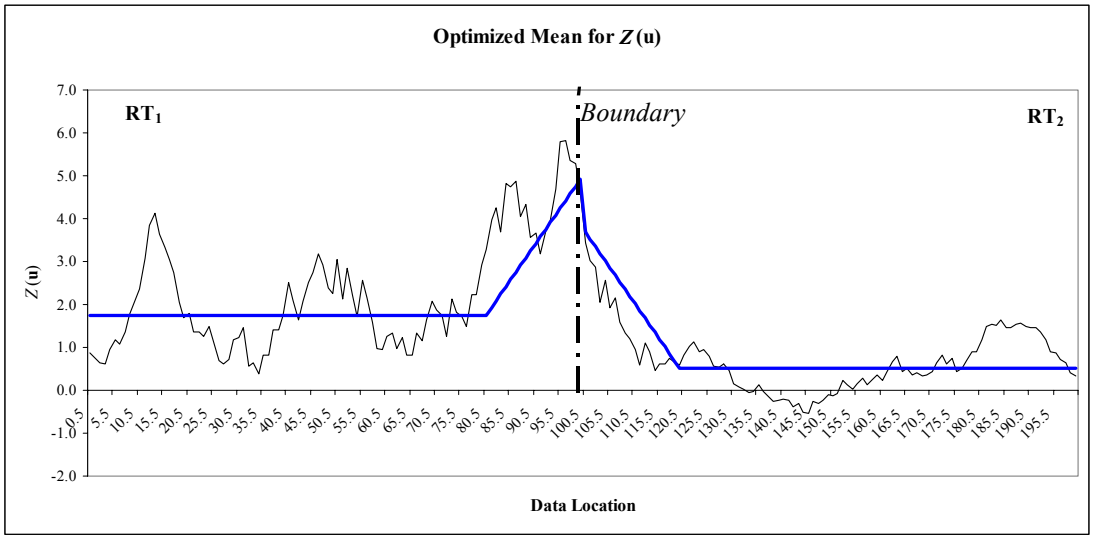

Figure 7: Optimized mean obtained for dataset of 1-D example.

The stationary and non-stationary components of the mean are used as an input to find the stationary variances of each rock type and the optimum non-stationary component of the boundary zone. The objective function in this case is the squared difference between the stationary plus the non-stationary variance and the residuals squared. The residuals are obtained using the already optimized expression for the mean:

$$m(\mathbf{u}_i) = \begin{cases} 1.75 & \text{where } 0.5 \le \mathbf{u}_i \le 79.5 \\ 1.75 + \dfrac{(20 - d(\mathbf{u}_i))}{20} \cdot 3.22 & \text{where } 80.5 \le \mathbf{u}_i \le 99.5 \\ 0.51 + \dfrac{(20 - d(\mathbf{u}_i))}{20} \cdot 3.22 & \text{where } 100.5 \le \mathbf{u}_i \le 199.5 \\ 0.51 & \text{where } 120.5 \le \mathbf{u}_i \le 199.5 \end{cases}$$

The optimum intercept value of b_{21}, in this example is attained at 0.50 (Figure 8). The optimized intercept is slightly different to the one used to created this synthetic dataset, due to statistical fluctuations from the mean and variance near the boundary zone. The reference stationary variances of rock type 1 and 2 are well reproduced by the optimization algorithm within the ergodic fluctuations of the simulations used to build the reference.

The optimization algorithm used to find the optimum range of the non-stationary covariance structure that fits the experimental covariance (Figure 9) calculated for each pair within the 40 meter zone of influence of the boundary, require as an input from the modeler: the shape, anisotropies and nugget effect of the relative variogram model. In this case a spherical isotropic model with a nugget effect of 0.0 was adopted. The optimum range obtained is 6.37; acceptably similar to the range of 10 used to build this synthetic dataset given the natural variations we have seen for mean and variance.

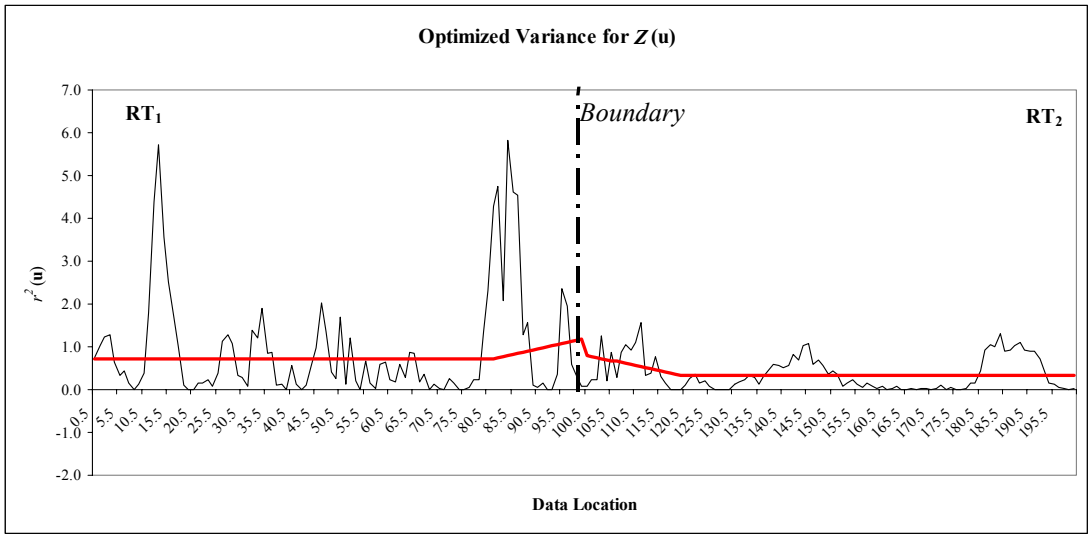

Figure 8: Optimized variance obtained for 1-D example.

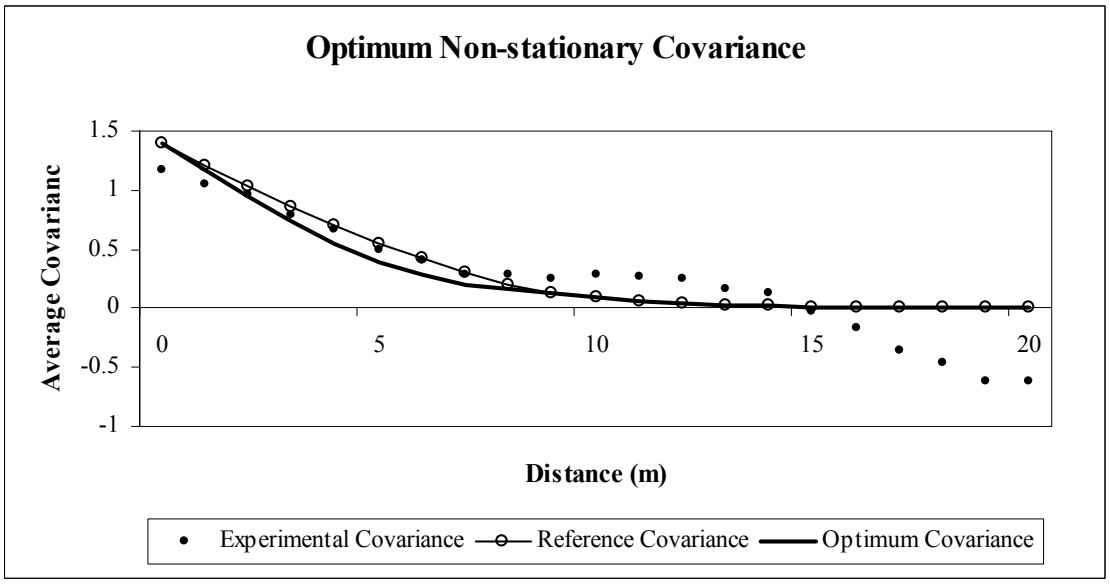

Figure 9: Experimental covariance from pairs within the boundary zone, optimum non-stationary covariance obtained from `opt_cov` and original covariance of the non-stationary component, $\mathbf{SGS}_3$, used to build the synthetic dataset $Z(\mathbf{u})$.

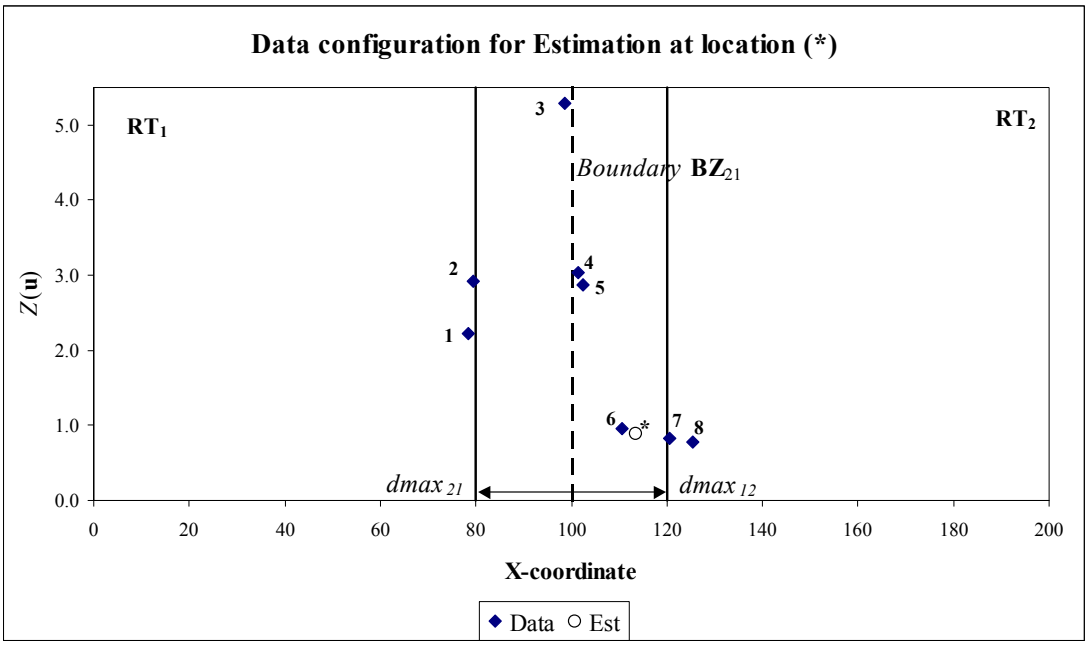

Figure 10: Data configuration for the estimation of an unknown location with the 1-D example.

For the estimation we also need the stationary covariance models for rock type *1* and rock type *2*, we will assume that the variogram models are the ones used to generate the underlying unconditional simulations.

Estimation in the presence of soft boundaries was performed using `kt3d_bound` (Larrondo and Deutsch, 2004) using the mean, variance and covariance previously obtained from optimization. First we will show how estimation is performed with 8 surrounding data (Figure 10) at a single location, and then we will review the results of estimation considering a conditioning dataset of one out of four grid nodes from the reference.

The data covariance matrix in terms of the stationary and non-stationary component, for the example configuration is:

$$\begin{bmatrix}
Cov(0)^{\mathbf{S}}_{\mathbf{Z_1}} & & & & & & & \\
Cov^{\mathbf{S}}_{Z_1} & Cov(0)^{\mathbf{S}}_{Z_1} & & & & & & \\
Cov^{\mathbf{S}}_{Z_1}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_1}+Cov^{\mathbf{NS}}_{Z12} & Cov(0)^{\mathbf{S}}_{Z_1}+Cov(0)^{\mathbf{NS}}_{Z12} & & & & & \\
0.0 & 0.0 & Cov^{\mathbf{NS}}_{Z12} & Cov(0)^{\mathbf{S}}_{Z_2}+Cov(0)^{\mathbf{NS}}_{Z12} & & & & \\
0.0 & 0.0 & Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov(0)^{\mathbf{S}}_{Z_2}+Cov(0)^{\mathbf{NS}}_{Z12} & & & \\
0.0 & 0.0 & Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov(0)^{\mathbf{S}}_{Z_2}+Cov(0)^{\mathbf{NS}}_{Z12} & & \\
0.0 & 0.0 & 0.0 & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov(0)^{\mathbf{S}}_{Z_2} & \\
0.0 & 0.0 & 0.0 & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2}+Cov^{\mathbf{NS}}_{Z12} & Cov^{\mathbf{S}}_{Z_2} & Cov(0)^{\mathbf{S}}_{Z_2}
\end{bmatrix}$$

Similarly, the data-to-estimate covariance vector in this case is:

$$\begin{bmatrix} 0.0 \\ 0.0 \\ Cov_{Z_{12}}^{\mathbf{NS}} \\ Cov_{Z_2}^{\mathbf{S}} + Cov_{Z_{12}}^{\mathbf{NS}} \\ Cov_{Z_2}^{\mathbf{S}} + Cov_{Z_{12}}^{\mathbf{NS}} \\ Cov_{Z_2}^{\mathbf{S}} + Cov_{Z_{12}}^{\mathbf{NS}} \\ Cov_{Z_2}^{\mathbf{S}} + Cov_{Z_{12}}^{\mathbf{NS}} \\ Cov_{Z_2}^{\mathbf{S}} + Cov_{Z_{12}}^{\mathbf{NS}} \end{bmatrix}$$

Completing the data covariance matrix and the data–to-estimate covariance vector, calculating the stationary and non-stationary component as shown before, the resultant kriging system is:

$$\begin{bmatrix} 1.0 & 0.9 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.9 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 2.149 & 0.382 & 0.227 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.382 & 1.027 & 0.842 & 0.141 & 0.044 & 0.014 \\ 0.0 & 0.0 & 0.227 & 0.842 & 1.003 & 0.152 & 0.052 & 0.019 \\ 0.0 & 0.0 & 0.0 & 0.141 & 0.152 & 0.805 & 0.130 & 0.078 \\ 0.0 & 0.0 & 0.0 & 0.044 & 0.052 & 0.130 & 0.25 & 0.188 \\ 0.0 & 0.0 & 0.0 & 0.014 & 0.019 & 0.078 & 0.188 & 0.25 \end{bmatrix} \bullet \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \\ \lambda_8 \end{bmatrix} = \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \\ 0.108 \\ 0.119 \\ 0.422 \\ 0.164 \\ 0.108 \end{bmatrix}$$

Given that the data covariance matrix is invertible, and calculating the mean at each data location and the mean at the estimate location, as the stationary plus the non-stationary mean, the estimated value is 1.15. The "true" value at this point was 0.89.

The reproduction of the reference values using a conditioning dataset of one out of four grid nodes is fairly good as shown in Figure 11 and 12. The only problem arises at the edges of the boundary zone, where unusual kriging weights occur leading to discrepancies between the estimate and reference. These weights are originated because the covariance of the estimate to data is higher than the covariance of the data to itself due to a non-stationary component in the first one, but not for the data-to-data covariance; in this case the estimate is inside the boundary zone, while the sample is outside. For example, the kriging system for the block at 80.5 meters is,

$$\begin{bmatrix} 1.0 & 0.748 & 0.610 & 0.276 & 0.276 & 0.056 & 0.056 & 0.0 \\ 0.748 & 1.793 & 0.276 & 0.765 & 0.056 & 0.276 & 0.0 & 0.056 \\ 0.610 & 0.276 & 1.0 & 0.056 & 0.610 & 0.0 & 0.276 & 0.0 \\ 0.276 & 0.765 & 0.056 & 1.892 & 0.0 & 0.783 & 0.0 & 0.276 \\ 0.276 & 0.056 & 0.610 & 0.0 & 1.0 & 0.0 & 0.610 & 0.0 \\ 0.056 & 0.276 & 0.0 & 0.783 & 0.0 & 1.992 & 0.0 & 0.802 \\ 0.056 & 0.0 & 0.276 & 0.0 & 0.610 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.056 & 0.0 & 0.276 & 0.0 & 0.802 & 0.0 & 2.091 \end{bmatrix} \bullet \begin{bmatrix} \lambda_1 = 1.929 \\ \lambda_2 = -0.171 \\ \lambda_3 = -0.575 \\ \lambda_4 = -0.020 \\ \lambda_5 = -0.017 \\ \lambda_6 = 0.028 \\ \lambda_7 = 0.086 \\ \lambda_8 = -0.003 \end{bmatrix} = \begin{bmatrix} 1.447 \\ 0.967 \\ 0.566 \\ 0.351 \\ 0.208 \\ 0.097 \\ 0.026 \\ 0.0 \end{bmatrix}$$

In this case the kriging weight for the closest data sample is unusually high, although the mismatch between the estimate and the reference it is not large. There are other examples for which the differences are more dramatic (see Larrondo and Deutsch (2004) in this volume). The origin of this non-physical covariance model needs to be reviewed in the future. These relatively rare problems in kriging will become more important for the implementation of this technique in simulation where the correct estimate and estimation variance are essential for the reproduction of the conditioning data and its variability.

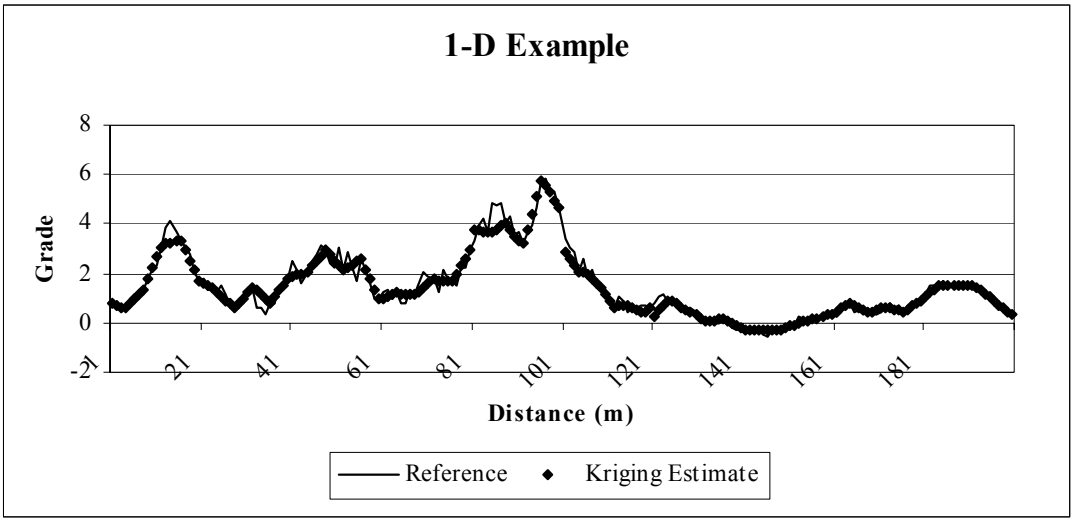

Figure 11: Grade reproduction profile along the X-coordinate. Reference values versus kriging estimates.

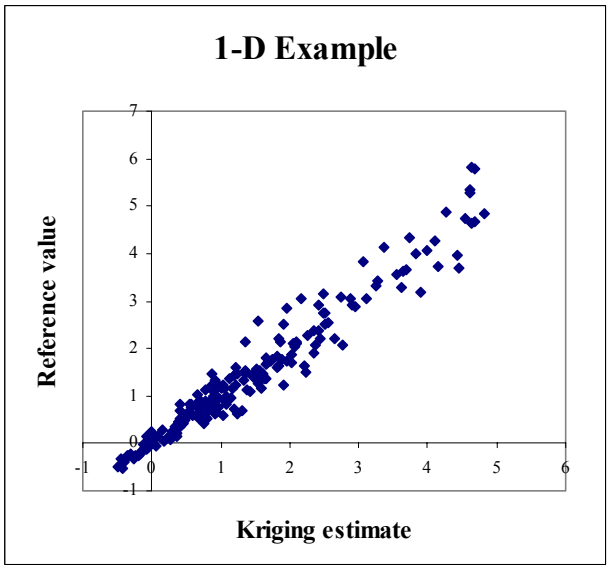

Figure 12: Scatter plot reference values versus kriging estimates, for the 1-D example.

For a larger grid, 2000 meters instead of 200, the reproduction of the reference improves (Figure 13) as the ergodic fluctuations have less influence in the underlying unconditional simulations use to build the reference and more samples are available to find the stationary and non-stationary components of mean, variance and covariance.

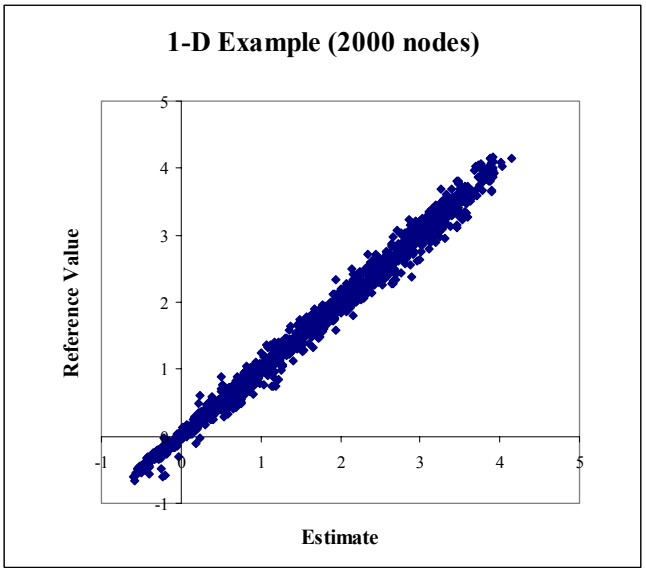

Figure 13: Reference values versus kriging estimates considering 2000 nodes instead of 200.

Conclusions and Future Work

The geological mechanisms involved in the formation of a deposit are in most cases transitional in nature, which yields contacts between domains that are diffuse or gradational.

This new technique provides a theoretically robust methodology to handle non-stationary soft boundaries. To apply this methodology the user must distinguish between stationary regions within each rock type and boundary zones where the statistical parameters such as the mean, variance or covariance are no longer constant. The non-stationary components of the mean and variance are optimized assuming a linear relationship with the distance to the boundary. All the optimization algorithms are a simplified version of the simulated annealing formalism, where only perturbations that minimize the objective function are accepted.

This work has considered that the mean and variance increase towards the boundary. A decreasing mean near a boundary could be handled by a negative non-stationary mean, but this could lead to negative grade estimates. Decreasing variance near a boundary cannot be handled with this formalism. The non-stationary variance must be positive. A decreasing mean plus the proportional effect could decrease the variance, but it is a limitation of the presented methodology.

The non-stationary features of the mean, variance and covariance are parameterized into a legitimate local model of coregionalization. Through this spatial model a non-stationary form of cokriging accounts for the changes in mean and variance at the vicinity of boundaries. The kriging estimates reproduce the non-stationary behavior of a reference distribution at the boundary zone

By construction, the kriging variance also has a non-stationary component. Since the kriging variance is the missing variability that is reintroduced in simulation, its implementation in the presence of local non-stationary boundaries will be delicate and is part of the future work.

Acknowledgements

We would like to acknowledge the industry sponsors of the Centre for Computational Geostatistics at the University of Alberta for supporting this research.

References

P. F. Larrondo and C. V. Deutsch. Application of Local Non-stationary LMC for Gradational Boundaries. In *Centre For Computational Geostatistics*, volume 6, Edmonton, AB, 2004.