Geostatistics with Location Dependent Moments and Distributions

David F. Machuca-Mory and Clayton V. Deutsch

Locally varying distributions can be built by weighting functions anchored at multiple points within a geologically homogeneous domain. These location dependent distributions are locally normal score transformed. The local transformation functions are approximated by Hermite polynomials to reduce the storage requirements. In the same way, other first and second order moments may be approximated. The modelling of these local continuity measures requires a robust automatic fitting algorithm. The sequential variogram fitting algorithm provides the location dependent parameters at anchor points within the domain. The use of the stable variogram model allows a locally changing variogram shape.

Introduction

In several standard geostatistical estimation and simulation techniques a single global multivariate distribution is assumed to be replicated in every point within a homogeneous domain. This assumption corresponds to the decision of strict stationarity (Chilès & Delfiner, 1999) and defines a set of stationary 1-point and multiple-point moments. Weaker decisions of stationarity require only the mean and covariance or the variogram be invariant by translation within a domain (Chilès & Delfiner, 1999). In any case, these globally stationary moments are inferred from all samples within a domain deemed statistical and geological homogeneous. This is done since local distributions, probabilities and averages cannot be constructed only with one value available at each sample location (Myers, 1989).

This approach can be unsatisfactory when dealing with phenomena that present significant local variations. Locally changing anisotropies and orientations of the local spatial correlation measures are common. Ignoring these local non-stationary features may lead to suboptimal estimates. When calculating the local moments by spatial averages and construction the global histogram, it is assumed that all samples have the same importance. Location dependent distributions and moments at a particular location can be obtained by assigning a greater importance to the closest samples to such location. This can be accomplished by weighting the available samples by a continuously decreasing function of the distance between samples and a reference point. When building a location dependent histogram the idea is that the probability contribution of each sample decreases as the distance of the sample to the reference point increases. While location dependent means, variances and other statistics can then be obtained as weighted averages. Location dependent 2-point measures of spatial correlation can also be obtained by weighted averages, but using combinations of sample pairs weights.

In the paper "Weighting criteria for Estimation of Location Dependent Moments" the properties of these weights and the methods for obtaining them are discussed. This paper focuses on the estimation of location dependent moments and the modelling of their associated parameters. Firstly, the location dependent univariate distributions and their associated moments are covered and an exhaustive definition of the local normal score transformation based in local Hermite polynomials is proposed. Secondly, several issues in the calculation and fitting of location dependent variograms are discussed, and some solutions proposed, additionally a locally changing variogram shape is proposed using stable variograms. These methodologies are illustrated using a public domain 2D dataset, for which exhaustive maps of the location dependent moments are generated. These maps can be used for locally stationary estimation and simulation.

Locally weighted distributions and local normal scores transformation

For a local univariate distribution two main issues arise when constructing them with locally weighted samples. The first is how to calculate the location dependent distributions and second how to obtain their corresponding 1-point moments. If the range of possible values at any unsampled location is modeled by a random variable $Z(\mathbf{u})$ the stationary probability of this variable being in an interval [a,b] is given by (Christakos, 2005):

$$\Pr(a \le Z(\mathbf{u}) \le b) = \int_a^b f_Z(z)dz \qquad \forall \mathbf{u} \in \mathcal{D} \tag{1}$$

Where $f_Z(z)$ is the global probability density function, PDF, of the random variable $Z(\mathbf{u})$. The stationary cumulative distribution function CDF is expressed analytically as:

$$F(z_k) = \Pr(Z(\mathbf{u}) \le z_k) = \int_{-\infty}^{z_k} f_Z(z)dz \qquad \forall \mathbf{u} \in \mathcal{D} \tag{2}$$

Under the decision of stationarity, these PDF and CDF are invariant by translation, and are inferred from all data values available at a domain $\mathcal{D}$ considered homogeneous (Deutsch & Journel, 1998). The contribution of all samples are the same, regardless their location. By other side, the location dependent PDF and CDF can be obtained by weighting the samples by a function that continuously decreases with the distance to a reference location, or "anchor point", **o**. Thus, the location dependent CDF becomes a convolution of the stationary PDF with the distance weighting function $\omega(\mathbf{o};\mathbf{u})$, centered at **o**:

$$F(z_k;\mathbf{o}) = \Pr(Z(\mathbf{o}) \le z_k) = \frac{1}{\eta}\int_{-\infty}^{z_k} \omega(\mathbf{u};\mathbf{o}) \cdot f_Z(z(\mathbf{u}))dz(\mathbf{u}) \qquad \forall \mathbf{u} \in \mathcal{D} \tag{3}$$

Where η is a scaling constant on the weights, in order to make:

$$\frac{1}{\eta}\int_{-\infty}^{\infty} \omega(\mathbf{u};\mathbf{o}) \cdot f_Z(z(\mathbf{u}))dz(\mathbf{u}) = 1 \qquad \forall \mathbf{u} \in \mathcal{D} \tag{4}$$

Thus, the nth 1-point moments are given by:

$$E(Z^n;\mathbf{o}) = \frac{1}{\eta}\int_{-\infty}^{\infty} \left(z(\mathbf{u})\right)^n \omega(\mathbf{u};\mathbf{o}) \cdot f_Z(z(\mathbf{u}))dz(\mathbf{u}) \qquad \forall \mathbf{u} \in \mathcal{D} \tag{5}$$

The location dependent mean is found by doing n =1, and the location dependent variance is given by:

$$\mathrm{var}(Z^n;\mathbf{o}) = E\left(Z - E(Z;\mathbf{o})\right)^2 = \frac{1}{\eta}\int_{-\infty}^{\infty} \left(z(\mathbf{u}) - E(Z;\mathbf{o})\right)^2 \omega(\mathbf{u};\mathbf{o}) \cdot f_Z(z(\mathbf{u}))dz(\mathbf{u}) \qquad \forall \mathbf{u} \in \mathcal{D} \tag{6}$$

These are the analytic expressions of the location dependent CDF and its most important 1-point moments. Next, it is explained how to obtain them numerically.

Location dependent histograms and cdf's based on continuously varying weights

The global proportion of data values $z(\mathbf{u})$ less or equal to a cut-off value z_k is given by (Goovaerts, 1997):

$$F(z_k) = \frac{1}{n}\sum_{\alpha=1}^{n} I(\alpha;z_k) \qquad \forall \mathbf{u} \in \mathcal{D} \tag{7}$$

That is invariant by translation and do not depend of the relative location of data values, and where:

$$I(\alpha;z_k) = \begin{cases} 1 & \text{if } z(\alpha) \le z_k \\ 0 & \text{if } z(\alpha) > z_k \end{cases}$$

Is the indicator transform of the value $z(\alpha)$. The global cumulative histogram is then constructed assembling in ascending order the proportions obtained for multiple cut-off values. In a similar way, the location dependent cumulative histogram with respect to an anchor point **o** can be constructed by calculating the location dependent proportions for different cut-offs:

$$F(z_k;\mathbf{o}) = \frac{1}{\sum_{\alpha=1}^{n}\omega(\mathbf{u}_\alpha;\mathbf{o})}\sum_{\alpha=1}^{n}\omega(\mathbf{u}_\alpha;\mathbf{o}) \cdot I(z(\mathbf{u}_\alpha);z_k) \qquad \forall \mathbf{u}_\alpha \in \mathcal{D} \tag{8}$$

Where, as above, $\omega(\mathbf{o};\mathbf{u})$ are the weights obtained from a decreasing distance weighting function. Thus, the location dependent histogram within a domain $\mathcal{D}$ can be built from:

$$f(a,b;\mathbf{o}) = \frac{1}{\sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o})} \sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o}) \cdot \mathrm{I}(z(\mathbf{u}_\alpha);a,b) \qquad \forall \mathbf{u}_\alpha \in \mathcal{D} \tag{9}$$

With:

$$\mathrm{I}(z(\mathbf{u}_\alpha);a,b) = \begin{cases} 1 & \text{if } a < z(\mathbf{u}_\alpha) \le b \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

Then, the location dependent mean can be calculated by:

$$m(\mathbf{o}) = \frac{1}{\sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o})} \sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o}) \cdot z(\mathbf{u}_\alpha) \qquad \forall \mathbf{u}_\alpha \in \mathcal{D} \tag{11}$$

And the location dependent variance:

$$\sigma^2(\mathbf{o}) = \frac{1}{\sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o})} \sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o}) \cdot \left(z(\mathbf{u}_\alpha) - m(\mathbf{o})\right)^2 \qquad \forall \mathbf{u}_\alpha \in \mathcal{D} \tag{12}$$

Local normal scores transformation and local Hermite model

The assumption of multi-gaussianity is adopted by several geostatistical techniques due to the convenient characteristics of the Gaussian univariate and multivariate distributions (Goovaerts, 1997). Since non Gaussian distributions are prevalent for geological and environmental variables, in order to apply these techniques the univariate distribution needs to be transformed to a standard normal distribution.

Given a local distribution of arbitrary shape and in original units, $F(z_p;\mathbf{o})$, and a standard Gaussian distribution, $G(y_p)$, the normal scores transform (Deutsch & Journel, 1998):

$$y_p = G^{-1}\left(F(z_p;\mathbf{o})\right) \tag{13}$$

is used for transforming the original z_p values to their standard Gaussian distributed equivalents, y_p, In practice, the normal scores transform can be characterized by the correspondence between the all original z values and the corresponding normal distributed y values, such as (Deutsch & Journel, 1998):

$$F(z_p;\mathbf{o}) = G(y_p) = p \quad \forall p \in [0,1] \tag{14}$$

These location dependent correspondences are stored in lookup tables of n sorted values z_i and the corresponding y_i normal values. They can be used for computing back transformation:

$$z_i = F^{-1}\left(G(y_i);\mathbf{o}\right) = \varphi_Z(y_i;\mathbf{o}) \qquad i = 1,\ldots,n \tag{15}$$

Alternatively, the Gaussian transform function φ_Z can be approximated by a series of Hermite polynomials (Journel & Huijbregts, 1978; Wackernagel, 2003):

$$z_i = \varphi_Z(y_i;\mathbf{o}) \approx \sum_{p=0}^{P} \phi_p(\mathbf{o}) H_p[y_i] \qquad \forall \mathbf{u} \in \mathcal{D} \tag{16}$$

The coefficients $\phi_p(\mathbf{o})$ are obtaining by doing:

$$\phi_0(\mathbf{o}) = E[Z(\mathbf{u});\mathbf{o}] = m(\mathbf{o}) \tag{17}$$

and:

$$\phi_p(\mathbf{o}) = \sum_{i=2}^{n} (z_{i-1} - z_i) \cdot \frac{1}{\sqrt{p}} \cdot H_{p-1}(y_i) \cdot g(y_i) \tag{18}$$

Where g is the standard Gaussian probability density function. Note that the coefficients $\phi_p(\mathbf{o})$ are location dependent. These are expressed as:

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= -y \\ H_{p+1}(y) &= -\frac{1}{\sqrt{p+1}} y H_p(y) - \sqrt{\frac{p}{p+1}} H_{p-1}(y) \end{aligned} \tag{19}$$

The number P of Hermite polynomials must to be chosen high enough to make:

$$\sigma_Z^2(\mathbf{o}) \approx \sum_{p=1}^{P} \phi_p^2(\mathbf{o}) \tag{20}$$

which is usually between 12 and 20. The programs that implement the local normal scores transformation and Hermite polynomials are described next.

Software implementation for location cdf's and local normal scores transformation

Before proceeding with the local normal scores transformation, a number of quantiles must be generated for the local distributions. This can be done using the `Histpltsim` program (Neufeld & Deutsch, 2007). This program is aimed for plotting the cdf's of each realization in a simulated model, but can also be used for plotting the cdf's and generating multiple quantiles for differently weighted distributions. An example of parameter file for `Histpltsim` is shown in the Figure 1. Instead of a file containing the simulated realizations, the output file of the `LDW-gen` program (Machuca-Mory & Deutsch, 2008) is used as input file. It is important to specify the column with the distance based weights in this file. The number of realizations corresponds to the number of anchor points considered. A numeric output is selected and the number of cumulative quantiles recommended is at least 99.

```
START OF PARAMETERS:
../data/lithology.dat                  - file with lithology information
0   7                                  - lithology column (0=not used), code
../data/original-ncb.dat               - file with data
5   6                                  - columns for reference variable and weight
../1-LDWgenerator/histlocal-ncb.out    - file with data
5   6                                  - columns for variable and weight
0     1                                - data (0=simulation, 1=multi columns), numeric output
1  754                                 - start and finish histograms (usually 1 and nreal)
736   1   1                            - nx, ny, nz
-1.0      1.0e21                       - trimming limits
Qloc99-ncb.ps                          - file for PostScript output
Qloc99-ncb.sum                         - file for summary output (always used)
Qloc99-ncb.out                         - file for numeric output (used if flag set above)
0     75                               - attribute minimum and maximum
-1.0                                   - frequency maximum (<0 for automatic)
20                                     - number of classes
0                                      - 0=arithmetic, 1=log scaling
1                                      - 0=frequency, 1=cumulative histogram
99                                     - number of cum. quantiles (<0 for all)
4                                      - number of decimal places (<0 for auto.)
Local NCB cdf's                        - title
1.5                                    - positioning of stats (L to R: -1 to 1)
-1.1e21                                - reference value for box plot
```

Figure 1: Parameter file for `Histpltsim` program used al local quantiles generator.

The ordered cumulative local quantiles in the numeric output of `Histpltsim` program are normal scores transformend using a modified version of the `nscore` program (Neufeld & Deutsch, 2007). This version, called `nscore-loc`, requires that the number of quantiles and anchor points be specified. The corresponding parameter file is presented in Figure 2.

```
START OF PARAMETERS:
../2-quantile_loc/Qloc99-ncb.out          - file with data
1  2  0                                   - columns for anchor point number, variable and weight
99  754                                   - number of quantiles and anchor points
0       1.0e21                            - trimming limits
0                                         - 1=transform according to specified ref. dist.
../histsmth/histsmth.out                  - file with reference dist.
1   2                                     - columns for variable and weight
nscore-loc-g-qu.out                       - file for output
nscore-loc-g-qu.trn                       - file for output transformation table
```

Figure 2: Parameter fire for local normal scores transformation program, `nscore-loc`.

The `Herco-loc` program uses the transformation table provided by `nscore-loc`, to generate the location dependent Hermite coefficients for each anchor point. The output file with the Hermite coefficients also contains the coordinates of their corresponding anchor points. The parameter file for `Herco-loc` program is shown in Figure 3.

```
START OF PARAMETERS:
../3-nscore-loc/nscore-loc-g-qu.trn          - file with input transformation table
../1-LDWgenerator/histlocal-ncb.dbg          - file with anchor point coordinates
1 2 3                                        - X,Y,Z coordinates for anchor points
20                                           - number of Hermite polynomials to use (e.g. np=100)
herco_loc-qug.out                            - file for point scale output to check the anamorphosis
herco_loc-qug.dbg                            - file with Hermite Coefficients, fi(p)
```

Figure 3: Parameter fire for the local Hermite coefficients calculation algorithm, `Herco-loc`.

Calculation and fitting of Location dependent variograms

When obtaining 2-point moments the 1-point distance based weights $\omega(\mathbf{o};\mathbf{u}_\alpha)$ corresponding to the sample pair endpoints can be combined making use of mixture rules (Korvin, 1982; Machuca-Mory & Deutsch, 2008), such as the arithmetic average:

$$\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})=\frac{\omega(\mathbf{u}_\alpha;\mathbf{o})+\omega(\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})}{2} \tag{21}$$

Or the geometric average:

$$\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})=\sqrt{\omega(\mathbf{u}_\alpha;\mathbf{o})\cdot\omega(\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})} \tag{22}$$

Where $\mathbf{u}_\alpha$ and $\mathbf{u}_\alpha+\mathbf{h}$ are two sample locations separated by a vector **h**. Experimentation show that any of these two average types yield to similar results (Machuca-Mory & Deutsch, 2008).

Location dependent variograms, correlograms and covariances

Using the 2-points distance based weights, location dependent experimental semivariogram is calculated as:

$$\gamma(\mathbf{h};\mathbf{o})=\frac{1}{2\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})}\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})\cdot\left(z(\mathbf{u}_\alpha)-z(\mathbf{u}_\alpha+\mathbf{h})\right)^2 \tag{23}$$

While the location dependent covariance can be obtained from:

$$C(\mathbf{h};\mathbf{o})=\frac{1}{\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})}\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})\cdot z(\mathbf{u}_\alpha)\cdot z(\mathbf{u}_\alpha+\mathbf{h})-m_{-\mathbf{h}}(\mathbf{o})\cdot m_{+\mathbf{h}}(\mathbf{o}) \tag{24}$$

$$\begin{aligned}m_{-\mathbf{h}}(\mathbf{o})&=\frac{1}{\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})}\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})\cdot z(\mathbf{u}_\alpha)\ ,\\ m_{+\mathbf{h}}(\mathbf{o})&=\frac{1}{\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})}\sum_{\alpha=1}^{N(\mathbf{h})}\omega(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})\cdot z(\mathbf{u}_\alpha+\mathbf{h})\end{aligned} \tag{25}$$

And the location dependent experimental correlogram is given by:

$$\rho(\mathbf{h};\mathbf{o}) = \frac{C(\mathbf{h};\mathbf{o})}{\sqrt{\sigma^2_{-\mathbf{h}}(\mathbf{o}) \cdot \sigma^2_{+\mathbf{h}}(\mathbf{o})}} \tag{26}$$

$$\sigma_{-\mathbf{h}}(\mathbf{o}) = \frac{1}{\sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o})} \sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot \left[z(\mathbf{u}_\alpha) - m_{-\mathbf{h}}(\mathbf{o}) \ \right]^2 ,$$

$$\sigma_{+\mathbf{h}}(\mathbf{o}) = \frac{1}{\sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o})} \sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot \left[z(\mathbf{u}_\alpha + \mathbf{h}) - m_{+\mathbf{h}}(\mathbf{o}) \right]^2 \tag{27}$$

In practical applications the experimental location dependent correlogram is preferred over the location dependent variogram. The interpretation of correlogram sill value is straightforward. The location dependent correlograms are clearly interpreted as the coefficient of correlation between sample values at a given separation vector **h**. Whereas the value of variogram sill is can be lower in certain directions due to zonal anisotropy or can show an increasing drift that reflects a trend (Gringarten & Deutsch, 2001). However, location dependent variograms are sensitive to the sample weighting

Semiautomatic fitting of local model parameters and shape

Location dependent experimental variograms, correlograms and covariances are calculated in relation of multiple anchor points conveniently located within a domain. The number of these anchor points is usually around a few hundreds. Fitting the models for such a large number of spatial continuity measures is only practical with the help of automatic or semiautomatic fitting algorithms. Semiautomatic fitting can be performed iteratively by slightly changing the variogram model parameters in order to minimize the next function objective:

$$O(\mathbf{h};\mathbf{o}) = \min \left[\lambda(\mathbf{h}) \cdot \left(\gamma(\mathbf{h};\mathbf{o}) - \hat{\gamma}(\mathbf{h};\mathbf{o}) \right)^2 \right] \tag{28}$$

Where $\gamma(\mathbf{h};\mathbf{o})$ is the experimental variogram value at a lag h, and $\hat{\gamma}(\mathbf{h};\mathbf{o})$ is the corresponding variogram model value. $\lambda(\mathbf{h})$ is the weight assigned to the lag **h** proportional to its inverse distance to the origin, the number of pairs used in the calculation of its corresponding variogram value, or both (Larrondo, Neufeld, & Deutsch, 2003).

The parameters are chosen randomly and iteratively. A random small perturbation is applied to the chose parameter, and the impact of the modified variogram model is evaluated according the objective function (28). Unless fixed to a certain value, the variogram model parameters that are considered in this optimization are the nugget effect, the ranges, the anisotropy orientation and the sill. By contrast, the variogram shape is fixed in order to avoid switching one model to another between adjacent locations. However highly continuous and smooth areas require a continuous variogram model, such as the Gaussian, and discontinuous zones are better modelled with a spherical or exponential variogram. To address this issue locally changing variogram shape can be used, such as the stable model (Chilès & Delfiner, 1999):

$$\hat{\gamma}(\mathbf{h};\mathbf{o}) = c(\mathbf{o}). \left[1 - \exp\left(- \left(\frac{3\mathbf{h}}{a(\mathbf{o})} \right)^{\beta(\mathbf{o})} \right) \right] \qquad 0 < \beta(\mathbf{o}) \le 1 \tag{29}$$

The shape of this model is controlled by the power value $\beta(\mathbf{o})$, it is the Gaussian model when $\beta(\mathbf{o}) = 1$ and it is equal to the exponential model for $\beta(\mathbf{o}) = 2$ (see figure 4). Note that this and the other model parameters, such as the sill $c(\mathbf{o})$ and the range $a(\mathbf{o})$ are referred and depend of the location **o**. By incorporating the perturbations of the power value parameter, a continuously changing variogram shape can be fitted at the anchor points within the domain.

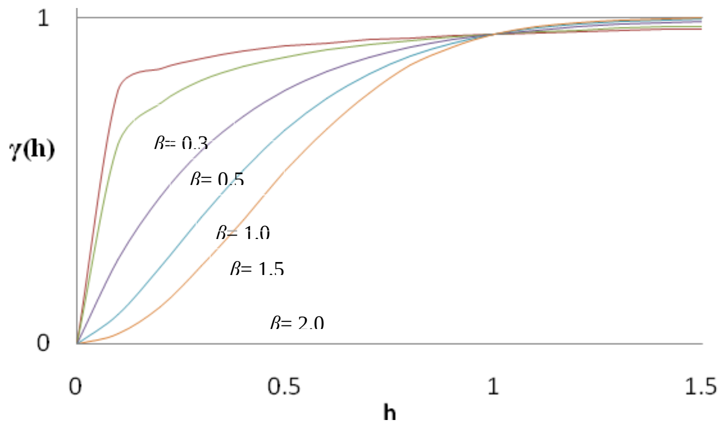

Figure 4: stable model shape according to different power values.

In theory multiple structures could be fitted, in practice instead, not more than two structures are recommended. The reason for this is that the higher the number of variogram structures fitted, the higher are the degrees of freedom in the fitting algorithm, which is translated in an increased instability of fitted parameters. Nevertheless, anomalous parameter values may be fitted using this semi-automatic algorithm at some of the several anchor points. This may occur even when restricting the number of variogram structures. Anomalous parameter values can be reduced by introducing penalties for the extremely high or extremely low values. The penalty function can take different forms, a simple one is quadratic:

$$K_{\alpha,i} = b \cdot (k_{\alpha,0}(\mathbf{o}) - k_{\alpha,i}(\mathbf{o}))^2 \quad i \in \mathbb{N}^+ \tag{30}$$

This is, for a parameter *α*, the corresponding penalty value $K_\alpha(i)$ is proportional to the square difference between the base parameter value $k_{\alpha,0}(\mathbf{o})$ and the value assigned to this parameter at the iteration *i*, $k_{\alpha,i}(\mathbf{o})$. This square difference can be multiplied by a positive factor, *b*, in order to strengthen the penalty applied to divergent parameter values. The base parameter value can be taken from the globally stationary variogram model. The objective function (29) becomes:

$$O(\mathbf{h};\mathbf{o}) = \min\left[\lambda(\mathbf{h}) \cdot \left(\gamma(\mathbf{h};\mathbf{o}) - \hat{\gamma}(\mathbf{h};\mathbf{o})\right)^2 + \sum_{\alpha=1}^{Np} K_\alpha\right] \tag{31}$$

Where *Np* is the number of penalized parameters. These penalties, however, do not completely eliminate the occurrence of anomalous parameter values, unless the strengthening factor *b* is considerably increased at expense of fitting flexibility. Posterior identification and correction of anomalous local parameter values are often necessary.

There are several methods for the identification of spatial outliers; these can be grouped in graphic and quantitative test (Lu, Chen, & Kou, 2003). Graphic approaches rely on maps and diagrams that highlight the anomalous values. Quantitative algorithms evaluate if the difference or the ratio between a sample value and a statistic calculated from neighbouring samples falls outside a tolerance interval. This can be done iteratively until no other ratio or difference outside the tolerance interval can be identified (Lu, Chen, & Kou, 2003). The statistic of local neighbouring values can be equally weighted or weighted by their inverse distance to the location where a value is evaluated (Kou, Lu, & Chen, 2006).

Once the locally anomalous variogram parameter values are identified, several alternatives are available for the parameter correction. The first one is the manual fitting of the variogram for the anchor point **o**. This alternative is not recommended since it easily can turn tedious and the final manually fitted parameters may diverge of the semi-automatically fitted parameters at the neighbouring anchor points. A straightforward option is just to replace the local variogram models that present anomalous parameters by new local variogram models whose parameters have been interpolated from neighbouring anchor points. Finally, a robust local variogram fitting algorithm can be used. This algorithm is described next.

Robust fitting of location dependent variograms

The steps proposed for this algorithm are the following:

1. Perform semi-automatic fitting of the local variograms at all anchor points with default initial values for local parameters. Penalty functions can be applied at this point.
2. Check sequentially for locally anomalous parameter values. These can be identified by one of the spatial outlier detection methods indicated before. The tolerance interval and other constraints for outlier identification can be modified by the user.
3. If an anomalous variogram parameter is identified the local variogram model at the corresponding anchor point. Now the initial parameter values for the fitting algorithm are the spatially weighted averages of the parameter values fitted at neighbouring anchor points. If no outlier value is found at this anchor point, move to the next and repeat from 2.
4. Check the new local parameters for values exceeding the tolerance interval. If none is found, move to the next anchor point and repeat from 2. Conversely, if anomalous parameters are still present, proceed with the second part of the algorithm.
5. Fit again the local variogram models alternating the lag weighting criteria. Keep the fits that yield the minimum error. Continue from 2 until no improvement is possible or all the anchor points are visited.

This algorithm is still in development. Currently the programs available for location dependent variograms correspond to the location dependent variogram calculation, `Gamvlocal`, and semiautomatic variogram fitting in a single pass, `Varfit-loc`. These programs are described next.

Software implementation for location dependent variograms

The `Gamvlocal` program allows the calculation of location dependent experimental variograms for untransformed and locally normal scores transformed data values. The parameter file for the current version of `Gamvlocal` is presented in figure 5. The input file is the output file containing the original values or the locally normal scores transformed values and their corresponding weights assigned with respect to the P anchor points defined. Thus, this file must to have $N \times P$ entries, where N is the total number of samples, including those with values beyond the trimming limits. Two options for the mixing rule of 1-point weight to 2-points weights are available: geometric or arithmetic average, see (21) and (22). The program reads the data values and weights, and combines the required pairs for each experimental variogram direction and tolerances. The current version allows the calculation of location dependent variograms, covariances and correlograms, a future version will include all the variogram types available at the `Gamv2004` program (Neufeld & Deutsch, 2007). The output file consists of the experimental location dependent variogram, covariance or correlogram in the `Gamv2004` output format. In order to facilitate the automatic model fitting of local correlograms the value $1-\rho(\mathbf{h};\mathbf{o})$ is reported in the output file.

```
START OF PARAMETERS:
nscore-local.out               - weights generator or local normal scores transformation output file
1  2  3  0  6                  - columns for AP number, X, Y, Z coordinates and weight
1   7                          - number of variables,column numbers
-1.0e21     1.0e21             - trimming limits
736  249                       - number of samples and anchor points
0                              - anchor points in file = 0, in grid = 1
26   493500   1000.0           - nx,xmn,xsiz
29   6204500  1000.0           - ny,ymn,ysiz
1    0.0      1.0              - nz,zmn,zsiz
cap1600a.clp                   - file with anchor points location
1   2   3                      - columns for X, Y, Z coordinates
0                              - Weights mixing rule: 0->geometric average, 1->arithmetic average
Local-correl.out               - file for variogram output
Local-correl.dbg               - file for debugging
2                              - number of directions
0 20 10000 0 15 10000          - Dir 01: azm,atol,bandh,dip,dtol,bandv
13  800  600                   - nlag,xlag,xtol
90 20 10000 0 15 10000         - Dir 01: azm,atol,bandh,dip,dtol,bandv
13  800  600                   - nlag,xlag,xtol
1                              - standardize sills? (0=no, 1=yes)
1                              - number of variograms
1 1 4                          - tail var., head var., variogram type

type 1 = traditional semivariogram
     3 = covariance   (-3 calculates variance-covariance)
     4 = correlogram  (-4 calculates 1-correlation)
```

Figure 5: parameter file for the `Gamvlocal` program.

The `Varfit-loc` program is based in the `Varfit` program (Neufeld & Deutsch, 2004) for globally stationary variogram modelling. It reads sequentially for every anchor point the experimental variogram, covariance or correlogram values from a `Gamvlocal` program output file. The parameter file (see fig 6) is similar to `Varfit`'s parameter file. In the main parameters block, the only additions are the lines for specifying the number of anchor points, for the folder and prefix of the graphs of local model fitted at each anchor point, and for a summary of the fitted parameters at each anchor point. The `Varfit-loc` program allows the fitting of stable variogram models for the first structure. If this model is not used, it is recommended to use the model(s) used for the globally stationary variogram as fixed the structure type(s). Penalties are imposed only if the corresponding parameter is not fixed; in this case a penalty factor is required. The higher this factor is the stronger will be the penalty for diverging parameter values.

```
START OF MAIN PARAMETERS:
249                              - number of anchor points
2                                - number of variograms
1                                - number of nested structures
0                                - constant angle between structures
1                                - inverse distance weighting (0=no, 1=yes)
0                                - number of pairs weighting (0=no, 1=yes)
250                              - minimum number of pairs to use
../LDV/vf-wl-g30-1si-1           - folder and prefix for PostScript output
vf-wl-g30-1si.var                - file for variogram models
vf-wl-g30-1si.sum                - file for-1s-1.summary file
Local Semivariograms             - project title


START OF VARIOGRAMS SPECIFICATION:
wl-dec-var-g30.out               - variogram #1 file
   1                             - variogram order in file
wl-dec-var-g30.out               - variogram #1 file
   2                             - variogram order in file

START OF ADVANCED OPTIONS:
0    0    0                      - zonal Anis: Hmax, Hmin, Vert (0=no, 1=yes)
0    0    0                      - cyclicity:  Hmax, Hmin, Vert (0=no, 1=yes)
1   1.0                          - fix sill
0   0.1   0.1                   - fix nugget effect, penalty factor
1                                - number of structure types to fix
1 6                              - structure number and structure type (6 for stable model)
0                                - number of Hmax ranges to fix
0                                - number of Hmin ranges to fix
0                                - number of Vert ranges to fix
0  0.0                              - number of azimuth angles to fix, Penalty factor
1  0.0                              - number of dip angles to fix, Penalty factor
1    0                           - structure number and dip angle
1  0.0                             - number of plunge angles to fix, Penalty factor
1      0                         - structure number and plunge angle
0                                - number of variogram preferences
0      10.0    0.1              - fix Hmax/Vert anis. (0=no, 1=yes), Penalty factor
0       0.33   0.1              - fix Hmin/Hmax anis. (0=no, 1=yes), Penalty factor


Insert the required lines (from below) for the options chosen
above (must be placed below the option flag).  Some parameters
can be fixed for each structure (range) while
others (hmax, hmin, hvert, anisotropy) are fixed for all structures
1  2                       -   structure number and structure type
1  1000                    -   structure number and Hmax range
1   200                    -   structure number and Hmin range
1    10                    -   structure number and Vert range
1    45                    -   structure number and azimuth angle
1   -30                    -   structure number and dip angle
1    30                    -   structure number and plunge angle
1     5                    -   variogram number, and scale factor for preference
```

Figure 6: Parameter file for the `Varfit-loc` program.

Example

The datasets for illustrating the location dependent moments approach have been taken from the well known Walker Lake dataset (Isaaks & Srivastava, 1989). For the sake of comparison a clustered and a gridded data set are considered (see figure 7). In the gridded data set samples are arranged in a semi-regular grid of 10m x 10m. Anchor points were located in a regular mesh of 20m by 20m, adding up to 195 points. All the moments and parameters were obtained first at each anchor point by weighting all the samples inversely proportional to the distance to them. The distance weighting function used was a Gaussian kernel with a standard deviation of 20. Once obtained the location dependent moments and parameters at each anchor point, these values were smoothly interpolated at high resolution within the entire study area. The interpolation technique chosen was global ordinary kriging with a spherical isotropic variogram model of range 100m and a nugget effect of just 1% of the sill value.

The local means and variances calculated for both data sets at each anchor point are shown in figure 8 overlying their corresponding interpolated values. In Figure 9 the local P25, median and P50 quantiles are plotted. The 195 local cdf's are shown in Figure 10, there it can be observed a significant variation of the local cdf's in comparison to the global cdf. The proportional effect is reflected by the positive correlation between location dependent means and variances (see figure 11). The local normal scores transformation functions were modelled using 30 Hermitian coefficients at each anchor point. Figure 12 shows the interpolated values of the three first coefficients.

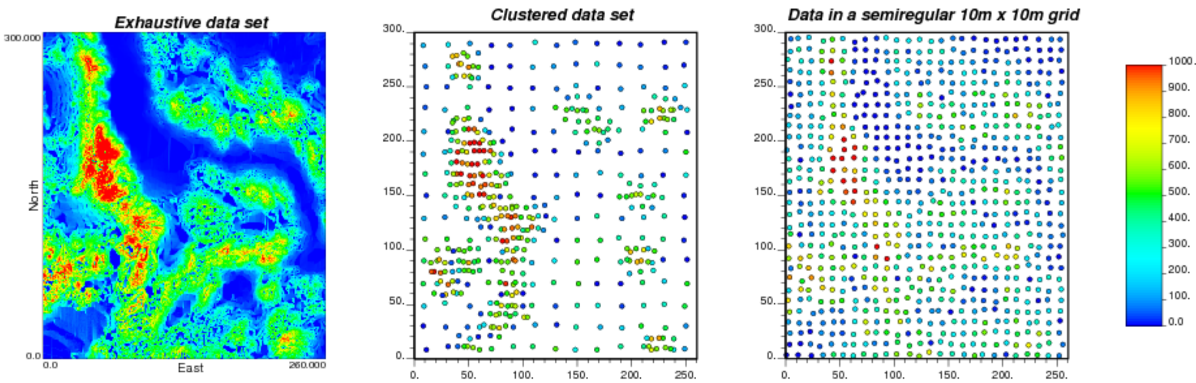

Figure 7: Exhaustive, clustered and gridded datasets of the Walker Lake site.

The weights used for experimental local correlograms calculation were obtained from the geometric average of the distance based weights assigned to each end of the sample pairs. Geometric averages of these weights yield to similar experimental variogram values as arithmetic averages for short to medium distance lags. The difference is appreciable only at the longest lags. With the purpose of allow a proper functioning of the semi-automatic variogram fitting algorithm the $1-\rho(\mathbf{h};\mathbf{o})$ values were used for modelling the variograms instead of the correlogram values. For the clustered data set, the local experimental $1-\rho(\mathbf{h};\mathbf{o})$ values were fitted using a single exponential structure. The interpolated local correlogram parameters are shown in Figure 13. For the gridded data set were used alternatively an exponential and a stable variogram models (see Figure 14). Note in the plots of location dependent variogram parameter that, for both the clustered and gridded data sets, reflect the local features of data. This is particularly clear for the anisotropy orientations and ratios.

Discussion and Conclusions

The location dependent distributions are constructed by considering distance weighting functions with the global stationary distribution. Since it would be very demanding to define these weighting functions at every location, they are anchored at a limited number of points. The location dependent distributions incorporate not only the modeling of a locally varying mean but also of a non-stationary variance. These local distributions may be considerably different of the globally stationary distribution.

An exhaustive definition of the locally varying distributions may be obtained by interpolating the percentiles of the local distributions between anchor points. Alternatively, the distributions can be modelled by limited number of Hermite polynomials, which reduce the dimensionality of working with 100 or more quantiles. This approach is applied for the exhaustive modeling of the local normal scores transformation function.

The local normal scores transformation of the location dependent distributions allow the application of Gaussian based algorithms in a locally stationary framework. Modelling the local normal scores transformations by local Hermite coefficient reduce the dimensionality compared to using local transformation tables. The smooth interpolation of the local Hermite coefficients allows the exhaustive definition of the functions of local transformation and back-transformation

Location dependent measures of spatial continuity are obtained using the same weights used for location dependent distributions and 1-point moments. Correlograms are preferred for their robustness and straightforward interpretation. Modelling is performed sequentially and semi-automatically at every anchor point. The local data fluctuations and outliers may cause anomalous values in fitted parameters. A robust fitting algorithm has been devised, but it is still in development. However, overall the fitted variogram parameters reasonably reflect the local features of the data values, particularly the local anisotropy ratios and orientations.

References

Chilès, J.-P., & Delfiner, P. (1999). *Geostatistics, modeling spatial uncertainty.* New York: John Wiley & Sons, INC.

Christakos, G. (2005). *Random Field Models in Earth Sciences.* Mineola, New York: Dover Publications, INC.

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

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation.* New York: Oxford University Press.

Gringarten, E., & Deutsch, C. V. (2001). *Teacher's aide. Variogram interpretationa and modelling*. *Mathematical Geology , 33* (4), 507-534.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics.* New York: Oxford University Press.

Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics.* New Jersey: The Blackburn Press.

Korvin, G. (1982). Axiomatic characterization of the general mixture rule. *Geoexploration* (19), 267-276.

Kou, Y., Lu, C.-T., & Chen, D. (2006). Spatial weighted outlier detection. *SIAM International Conference on Data Mining* (pp. 613-617). Bethesda, Maryland: Society for Industrial and Applied Mathematics (SIAM).

Larrondo, P. F., Neufeld, C. T., & Deutsch, C. V. (2003). *VARFIT: A program for semi-automatic variogram modelling - Paper 61.* Centreo for Computational Geostatistics. Report 5.

Lu, C.-T., Chen, D., & Kou, Y. (2003). Algorithms for Spatial Outlier Detection. *Third IEEE International Conference on Data Mining (ICDM'03)*, (pp. 597-600). Melbourne, Florida.

Machuca-Mory, D. F., & Deutsch, C. V. (2008). *Optimal weights for location dependent moments - Paper 108.* Centre for Computational Geostatistics. Report 10.

Myers, D. E. (1989). To be or not to be ... Stationary? That is the question. *Mathematical Geology , 21* (3), 347-362.

Neufeld, C., & Deutsch, C. V. (2007). *CCG software catalogue.* University of Alberta, Department of Civil and Environmental Engineering. Centre for Computational Geostatistics.

Neufeld, C., & Deutsch, C. V. (2004). *Developments in semiautomatic variogram fitting - Paper 404.* Centre for Computational Geostatistics. Report 6.

Wackernagel, H. (2003). *Multivariate Geostatistics: An introduction with applications* (3rd Edition ed.). Berlin Heidelberg: Springer-Verlag.

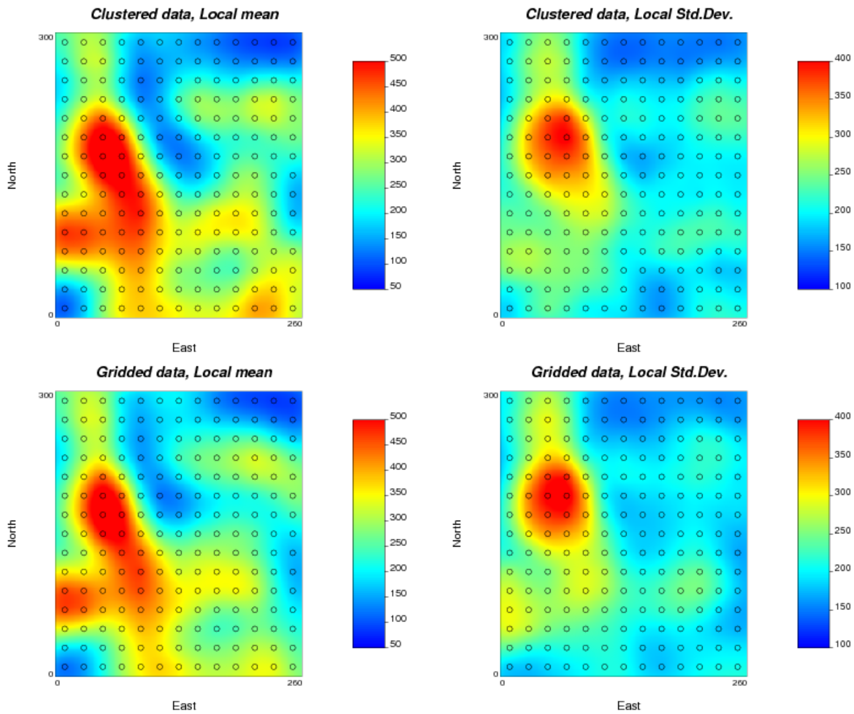

Figure 8: Local means and standard deviations for the clustered (above) and gridded data sets (below) calculated at anchor points (circles) and interpolated between them.

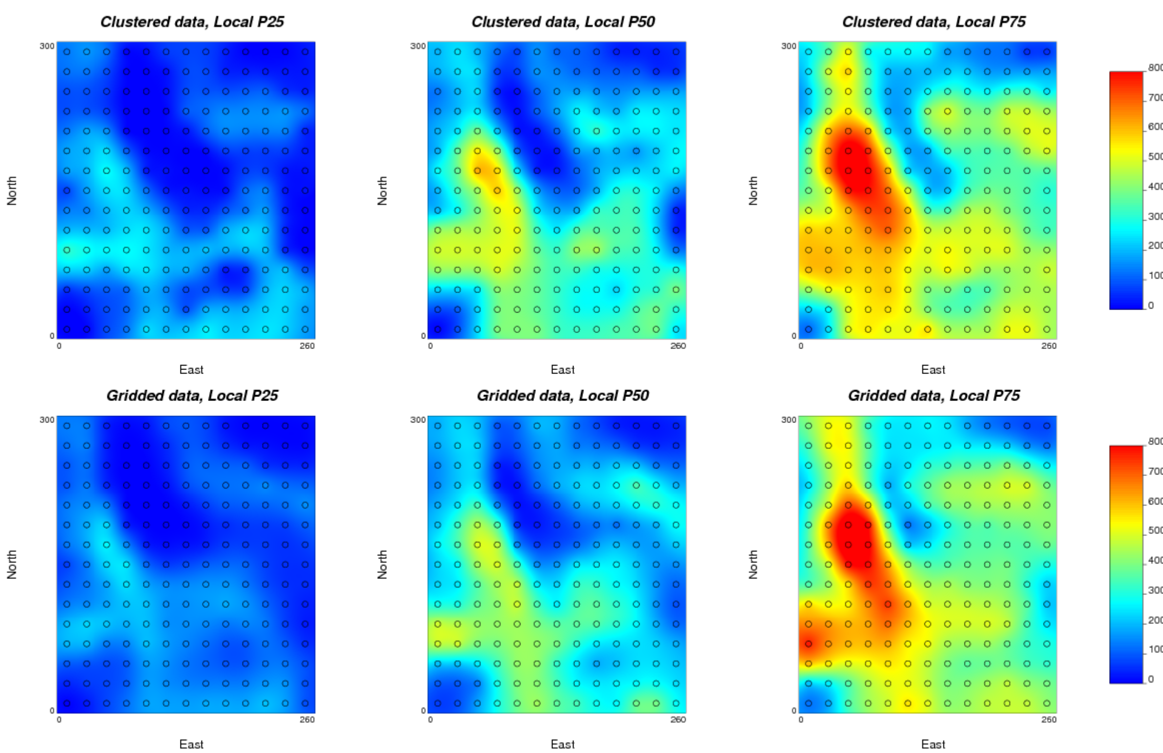

Figure 9: Local quartiles for the clustered (above) and gridded (below) data sets

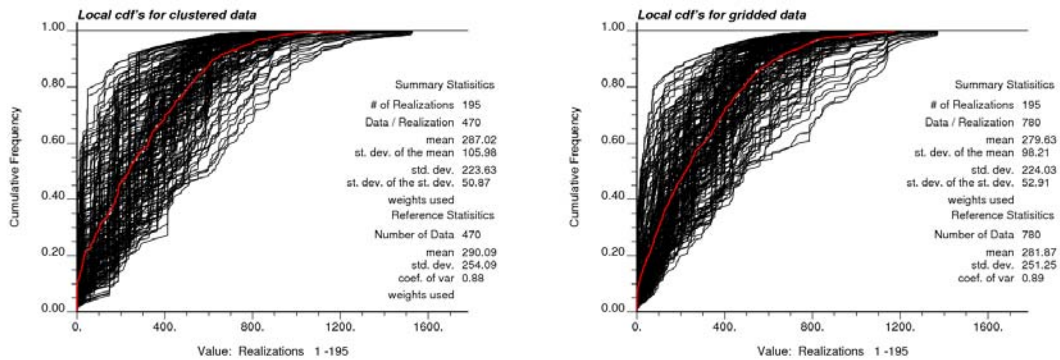

Figure 10: global (bold red curve) and local cdfs (black curves) for clustered (left) and gridded (right) data.

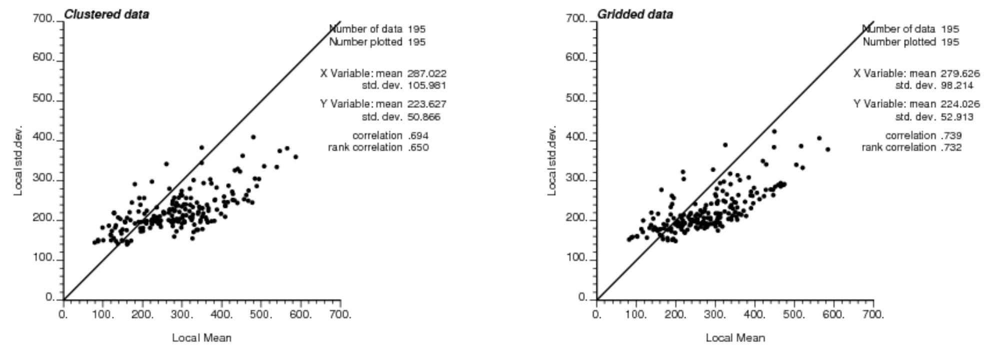

Figure 11: Proportional effect reproduced by the location dependent mean and standard deviation.

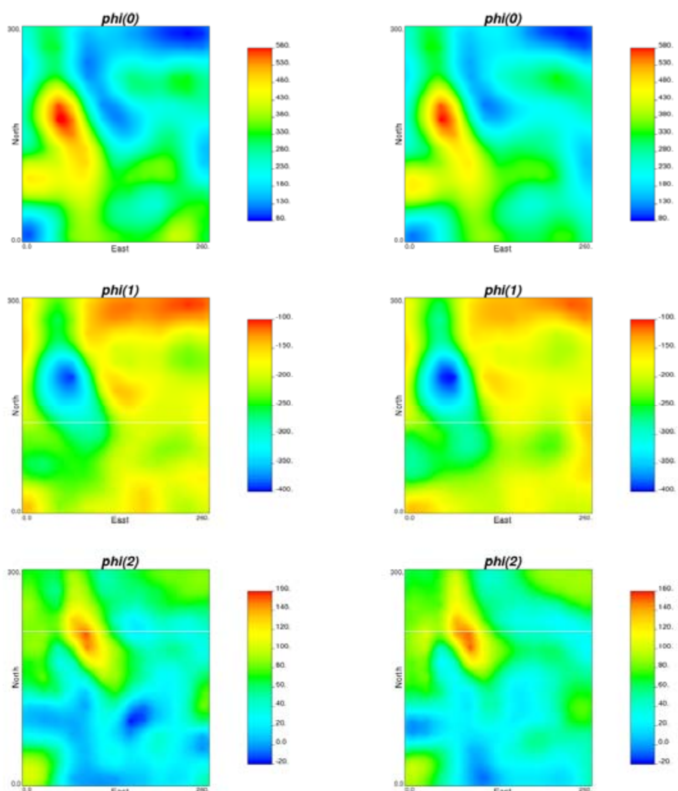

Figure 12: interpolated values of the three first coefficients of thirty used for modelling the local normal scores transformation function.

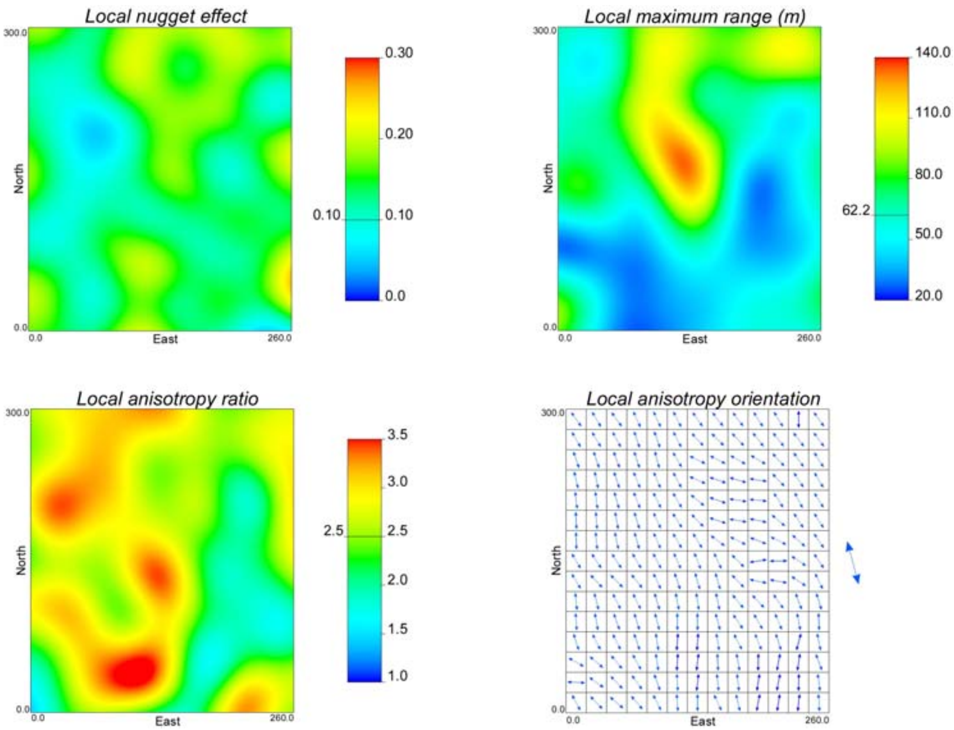

Figure 13: Local exponential model parameters for the clustered data set

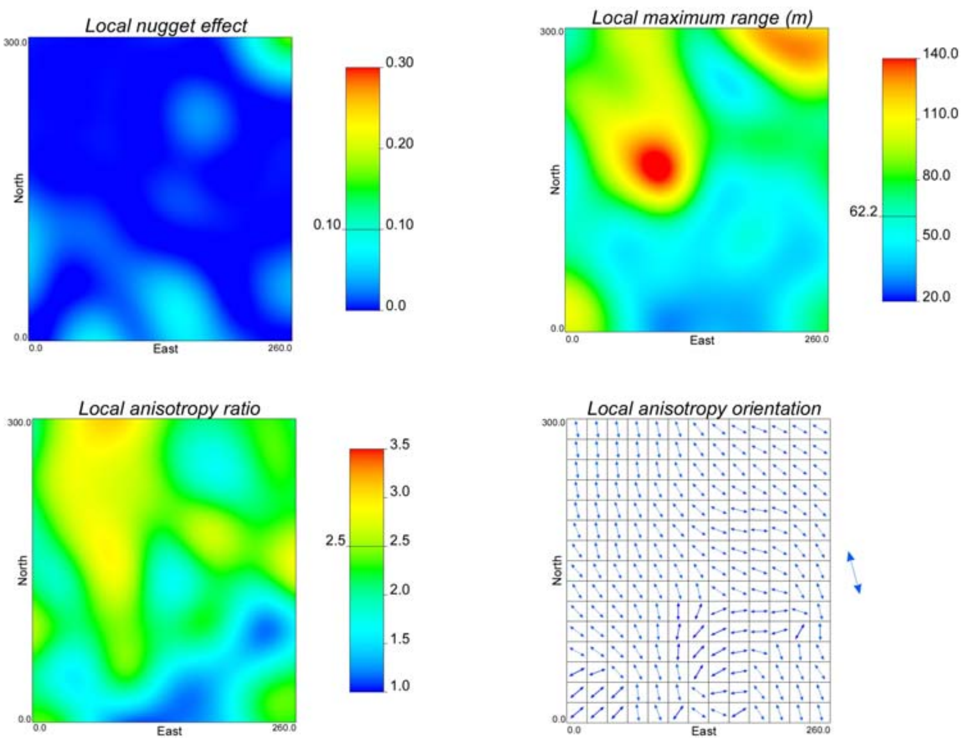

Figure 14: Local exponential model parameters for the gridded data set