# Sequential Gaussian and Indicator Simulation with Location-Dependent Distributions and Statistics

David F. Machuca-Mory and Clayton V. Deutsch

*The use of location-dependent distributions and statistics is proposed for geostatistical simulation under the assumption of local stationarity. The local distributions and statistics are obtained using distance weighting functions. For Sequential Gaussian Simulation, the Gaussian transformation of each local distribution embeds the local changes in the local mean, variance and histogram shape. The same weights used for inferring the local distribution modify the local measures of spatial continuity, which adapt to local variations informed by data. The local Gaussian transformations are modelled by Hermite polynomial series and the resulting coefficients are stored. The local measures of correlation are fitted semiautomatically and, as for the Hermite coefficients, the resulting parameters are stored at the resolution of the simulation grid. The sequential simulation algorithms read these local parameters, update the local distribution, retransform the data, and recalculate the covariance at every location. This increases considerably the demand of computer resources. However, the resulting models are richer in local information. These models are accurate and can be more precise compared to those generated by stationary techniques. Locally stationary indicator simulation models offer a better reproduction of the connectivity observed in categorical data.*

## Introduction

Standard geostatistical simulation techniques are constrained by the assumption of strict stationarity. This is, the cdf and its statistics remain invariant within a domain deemed homogeneous. The main reason for this restriction is that under a multiGaussian framework the mean and variances obtained by Simple Kriging (SK), a strictly stationary technique, are identical to the local conditional mean and variance given a set of surrounding data (Journel, 1980). Several approaches have been already proposed for relaxing this requirement and dealing with different aspects of non-stationarity in Gaussian based simulation techniques. The use of SK with explicit local prior means or implicit local means obtained from Ordinary Kriging is a common approach for dealing with trends in the mean (Deutsch & Journel, GSLIB. Geostatistical Software Library and User's Guide, 1998). Simulation using histograms and variograms inferred within contiguous sub-domains has been also developed (Deutsch, 2002; Leuangthong, Prins, & Deutsch, 2006). Local normal scores transformation of non-stationary distributions has been proposed recently for dealing with trends in the mean and other local changes in the histogram (McLennan & Deutsch, 2008). Similarly, for indicator-based simulation, the use of local proportions is common when modelling categorical variables (Deutsch, 2002; Lyster & Deutsch, 2004).

This paper proposes an alternative approach based distance weighted distributions, and their statistics, used under a decision of local stationarity for sequential Gaussian and indicator simulation. Under the decision of local stationarity, the local cdf and their statistics are defined in relation to reference or anchor point. These are deemed invariant on translation only if the reference point remains unchanged. This stationary model is expressed as:

$$\operatorname{Prob}\{Z(\mathbf{u}_\alpha) < z_1, \ldots, Z(\mathbf{u}_n) < z_K; \mathbf{o}_i\} = \operatorname{Prob}\{Z(\mathbf{u}_\alpha + \mathbf{h}) < z_1, \ldots, Z(\mathbf{u}_n + \mathbf{h}) < z_K; \mathbf{o}_j\}$$
$$\forall\ \mathbf{u}_\alpha, \mathbf{u}_\beta + \mathbf{h} \in D, \text{ and only if } i{=}j \qquad (1)$$

One way of obtaining the local distributions and statistics is weighting the data inversely to their distance to the each anchor point. For a discussion on the inference of distance weighted statistics refer to the papers “Optimal Weights for Location Dependent Moments” (Machuca-Mory & Deutsch, 2008a) and “Location Dependent Moments and Distributions Based in Continuously Varying Weights” (Machuca-Mory & Deutsch, 2008b). The focus of this paper is the development of locally stationary sequential Gaussian and indicator simulation. The next two sections discuss the theory behind these algorithms. The following section presents some details of their software implementation that are important for the correct application of the two proposed techniques. Their performance is illustrated with the help of 2-D continuous and categorical realistic data. The last section discusses the advantages and disadvantages of the proposed approach.

**Locally Stationary Sequential Gaussian Simulation**

Local normal scores transformation of inverse distance weighted cdfs have been already proposed to account for trends in Sequential Gaussian Simulation (SGS) (Gonzales, McLennan, & Deutsch, 2006). The idea is to modify the global cdf by the inverse distance weights at each simulated location, perform the normal score transformation of the weighted cdf keeping the transformation table, draw a simulated value on the conditional distribution, and backtransform it using the local transformation table. By contrast, the algorithm proposed in this paper requires that the Gaussian transformation function be defined prior to the simulation. This is done so in order to decrease the processing demand of rebuilding the complete transformation tables at each location and because the weights used for locally weighting the cdf are also used for inferring the local measures of spatial continuity. Using a single set of weights for the local cdfs their 1-point and 2-point statistics assures the mutual consistency between them; 2-point statistics become 1-point statistics when the sample separation is zero. Since the same weights modify locally all statistics, the variogram of the residuals is no longer required.

An efficient way to approximate these functions is by a series of Hermite polynomials (Journel & Huijbregts, 1978; Wackernagel, 2003):

$$z = \varphi_Z(y;\mathbf{o}) \simeq \sum_{q=0}^{Q} \phi_q(\mathbf{o}) H_q[y] \tag{2}$$

Where $z$ is the value in original units, $y$ is standard normal transform, $\varphi_Z(y;\mathbf{o})$ is the local transformation function at point $\mathbf{o}$, $\phi_q(\mathbf{o})$ are the local Hermite coefficients and $H_q[y]$ are the Hermite polynomials. A more detailed discussion of the local Gaussian transformation function modelling using Hermite polynomials is given in the paper "Location Dependent Moments and Distributions Based in Continuously Varying Weights" (Machuca-Mory & Deutsch, 2008b). A limited number of Hermite expansions can be used for fitting the transformation function at the price of loss of accuracy. The fitting is improved if it is performed on 100 or more percentiles calculated on despiked values instead of the original values.

The Locally Stationary Sequential Gaussian Simulation (LSSGS) requires as input, besides the data file, the gridded Hermite coefficients and local variogram parameters at the same resolution of the simulation grid. The current design of the LSSGS algorithm proceeds in the following steps:

i. Read and store the local Hermite coefficients and the location-dependent variogram parameters for every simulation node in a previously defined grid.
ii. Visit each simulation node in a random path. Search for the surrounding conditioning data and previously simulated grid nodes.
iii. Construct the location-dependent transformation function with the local Hermite coefficients. Perform the local Gaussian transformation of surrounding data and previously simulated nodes.
iv. Obtain the mean and variance of the local ccdf by locally stationary Simple Kriging with the location-dependent variogram model informed by the corresponding local variogram parameters.
v. Perform Monte Carlo simulation for obtaining a simulated value from that ccdf.
vi. Back transform the simulated value according to the local Gaussian transformation function. Add the simulated values in original units to the data set.
vii. Go to the next node in the random path and loop from step iii until all nodes are simulated.

**Locally Stationary Sequential Indicator Simulation**

Locally Stationary Sequential Indicator Simulation (LSSIS) requires the local category proportions and local indicator measures of spatial continuity. Given a category $s_k$, its local proportion is obtained by:

$$F(\mathbf{u};s_k;\mathbf{o}) = \text{Prob}\{Z(\mathbf{u}) = s_k;\mathbf{o}\} = \sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o}) \cdot I(\mathbf{u}_\alpha;s_k) \ \in [0,1] \qquad \forall \mathbf{u}_\alpha \in D,\ k = 1,...,K \tag{3}$$

Where $\omega(\mathbf{u}_\alpha;\mathbf{o})$ are the distance weights assigned to a sample located at $\mathbf{u}_\alpha$ in relation to the anchor point **o**. While the categorical indicator function is defined as:

$$I(\mathbf{u}_\alpha;s_k) = \begin{cases} 1, & \text{if } \ z(\mathbf{u}_\alpha) = s_k \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

This indicator function replaces the continuous values $z(\mathbf{u}_\alpha)$ and $z(\mathbf{u}_\alpha + \mathbf{h})$ in the calculation of the location-dependent experimental indicator semivariograms and covariances. Thus, the location-dependent indicator semivariogram is obtained from:

$$\gamma_I(\mathbf{h};s_k;\mathbf{o}) = \frac{1}{2} \sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot \left[ I(\mathbf{u}_\alpha;s_k) - I(\mathbf{u}_\alpha + \mathbf{h};s_k) \right]^2 \tag{5}$$

With $\omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o})$ as the 2-point weight assigned to the pair of samples located in $\mathbf{u}_\alpha$ and $\mathbf{u}_\alpha + \mathbf{h}$. In the same way, the location-dependent covariance anchored at a point **o** is calculated as:

$$C(\mathbf{h};s_k;\mathbf{o}) = \sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot I(\mathbf{u}_\alpha;s_k) \cdot I(\mathbf{u}_\alpha + \mathbf{h};s_k) - F_{-\mathbf{h}}(s_k;\mathbf{o}) \cdot F_{+\mathbf{h}}(s_k;\mathbf{o}) \tag{6}$$

With the tail and head local proportions given by:

$$F_{-\mathbf{h}}(s_k;\mathbf{o}) = \sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot I(\mathbf{u}_\alpha;s_k) \ , \qquad F_{+\mathbf{h}}(s_k;\mathbf{o}) = \sum_{\alpha=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot I(\mathbf{u}_\alpha + \mathbf{h};s_k) \tag{7}$$

Among the local measures of spatial continuity, the location-dependent correlogram is preferred due to its straightforward interpretation and robustness, in its indicator form this is calculated by:

$$\rho(\mathbf{h};s_k;\mathbf{o}) = \frac{C(\mathbf{h};s_k;\mathbf{o})}{\sqrt{\sigma^2_{-\mathbf{h}}(\mathbf{o};s_k) \cdot \sigma^2_{+\mathbf{h}}(\mathbf{o};s_k)}} \in [-1,+1] \tag{8}$$

With the tail and head local indicator variances given by:

$$\sigma^2_{-\mathbf{h}}(s_k;\mathbf{o}) = F_{-\mathbf{h}}(s_k;\mathbf{o}) \left[ 1 - F_{-\mathbf{h}}(s_k;\mathbf{o}) \right], \qquad \sigma^2_{+\mathbf{h}}(s_k;\mathbf{o}) = F_{+\mathbf{h}}(s_k;\mathbf{o}) \left[ 1 - F_{+\mathbf{h}}(s_k;\mathbf{o}) \right] \tag{9}$$

When samples of different categories appear in different densities, location-dependent indicator covariances and correlograms can reach zero value at a short lag. This is because at short lags the both locations of each sample pairs are located within the same densely sampled category; therefore, the sum in the covariance becomes equal to one, as well as the local tail and head proportions. This artifact may cause artificially short ranges in the semiautomatic fitting of the local variogram models. Thus when the sample density varies with the category, location-dependent variograms may be preferred over covariances and correlograms.

There is no theoretical impediment in applying LSSIS to continuous variables; however in practice the task of calculating and modelling the location-dependent indicator variograms for different cut-offs at different locations may result tedious. Although the same can be said when the number of categories is high, in several cases, the non-stationary modelling using LSSIS can be restricted to a few categories.

The LSSGS algorithm is similar to its stationary counterpart (Deutsch & Journel, 1998). The main differences consist in the continuous updating of the local proportions and local variogram model parameters at each simulated location. The local cdf is built by locally stationary Indicator Kriging (LSIK), which is nothing else that the traditional indicator kriging, but with using local parameters specific of each location.

**Implementation**

Besides the incorporation of location-dependent transformations, variogram models and local proportions of categorical variables, another difference in the implementation of the locally stationary simulation technique is that covariance lookup tables are not used. Stationary techniques increase their efficiency by calculating all the covariances needed in the kriging equations and storing them in a covariance lookup table, which is read during the sequential simulation. This is feasible only if a global variogram model is used. Previously proposed sequential Gaussian simulation algorithms with locally varying angles (Leuangthong, Prins, & Deutsch, 2006), group the variogram azimuth and dip angles within a limited number of classes. The different combinations of azimuth and dip angle class averages and their corresponding rotation equations and covariance tables are stored in 5-D matrices. In the proposed approach, by contrast, not only the angles vary locally, but also the complete set of variogram parameters. Although a covariance lookup table of higher dimensions could be used to accommodate different local variogram parameters grouped in classes, it was opted to prescind from it. Thus, all the required covariances need to be recalculated at each simulation node using the corresponding local variogram parameters. As it is shown below, this results in an increase of the computation time, which is, however, within tolerable limits.

LSSGS is implemented as a modification of the ultimatesgsim program (Deutsch & Zanon, 2002). The new parameter file contains an additional block for the specification of the files containing the local Hermite coefficients and local variogram parameters (see Figure 1). These are grid files with the same resolution and arrangement as the defined by the grid parameters for simulation. The memory requirements for storing all these arrays of local parameters may be excessive for big grids. An alternative would be to keep in memory only the parameters at anchor point locations and interpolate them within a small neighbourhood at every simulation node.

LSSIS is implemented in the program SISIM_loc. This is a modification of the GSLib program SISIM_lm, which allows local proportions (Deutsch & Journel, 1998). The only modification in the parameter file of this program is the inclusion of a block for the definition of the local variogram parameters grid files (see Figure 2). As for LSSGS, if this block is absent or erroneous, the program will simulate using the parameters of the global variogram model for building a covariance lookup table.

**2D example**

The performance of LSSIS is illustrated using categorical data from the Walker Lake dataset (Isaaks & Srivastava, 1989). The first category corresponds to the very low-grade domain and the second, to the remaining areas. Figure 1, left, presents the locations of the categorical data. A Gaussian kernel weighting function with 20 units bandwidth was used for inferring the required local statistics. Figure 1, right, shows the resulting local proportions for the low-grade category. The location-dependent experimental indicator semivariograms were calculated at anchor points separated by 20 units. Figure 4 presents the parameters of the stable variogram models fitted semiautomatically to the experimental local variograms. The local nugget effect was fixed to zero at all locations. Stable variogram models (Chilès & Delfiner, 1999) permit a locally changing variogram shape. As the upper right side of Figure 4 shows, the exponent of the stable model tends to a value of two at areas where there is little or no variation in the categories. Higher values of the stable model exponent give the variogram model a shape close to the Gaussian variogram model. The local anisotropy orientations (Figure 4, upper right) reflect satisfactorily the continuity observed in data. The global variogram model fitted to the same categorical dataset has the following parameters:

$$\gamma_I(\mathbf{h}) = Exp_{\substack{a_{Az150}=196 \\ a_{Az60}=60}}(\mathbf{h}) \tag{10}$$

A global proportion of 0.22 for the low-grade category was used in the traditional SIS. LSSIS was performed using the gridded local proportions and local indicator variogram parameters. Figure 5 shows equivalent examples of the 100 SIS and LSSIS realizations. Using location-dependent indicator variograms in indicator simulation produce realizations that reproduce better the changes of spatial continuity observed in categorical data. Contrarily, SIS realizations reflect the uniform spatial continuity defined by the global indicator variogram model. E-type estimates of the 100 LSSIS realizations (Figure 6, right) show probability contours that are more continuous and geologically appealing than those produced by SIS

(Figure 6, left). Running the 100 SIS realizations of this 78000 cells model took 8.6 min, while LSSIS took 37.2 min. This considerable increment in the computation time is because the LSSIS recalculates the complete set of required covariances at each location and for each realization.

The LSSGS example uses the continuous variable of the same dataset. Figure 7 presents maps of exhaustive reference data and the clustered dataset that is used for inferring the location-dependent distributions and their statistics. A Gaussian kernel with the same parameters as for the indicator variable was used for calculating the local distributions and statistics of the continuous variable. The parameters of the location-dependent variograms have been presented before (Machuca-Mory & Deutsch, Locally stationary multiGaussian kriging with local change of support - Paper 104, 2009). The local Gaussian transformation functions were modelled using 41 expansions of the Hermite polynomials series. Figure 8, left, shows the local mean, which correspond to the first Hermite coefficient, and the local variance, at the right side, approximated by the sum of the squares of the 40 remaining coefficients. The global correlogram model used in SGS is given by:

$$1-\rho(\mathbf{h}) = 0.2 + 0.8 \times Exp_{\substack{a_{Az165}=112 \\ a_{Az75}=27}}(\mathbf{h}) \tag{11}$$

As for the SIS and LSISS, 100 realizations were generated for SGS and LSSGS. Figure 9, left, shows an example of the globally stationary SGS realizations and an equivalent example of the LSSGS realizations (right). Although the local mean and local variance are embedded in the local transformation function, the posterior local distributions in LSSGS realizations are not required to reproduce the local distributions. The posterior local means (Figure 10) honour the large scale features of the local prior means (Figure 8, left), but in the case of LSSGS the posterior means are considerably modified by the input of location dependent variograms. Their effect is also noticeable in the posterior local variances (Figure 11). Compare, by instance, the high prior variances zone in the west side of Figure 8, right, with the posterior SGS (Figure 11, left) and LSSGS (Figure 11, right) variances. SGS realizations honour the high variance zone, while the LSSGS posterior variances are lowered by the presence of locally low relative nugget effect values and long local ranges at this zone.

While the reproduction of local prior distributions is not a requirement of locally stationary simulation techniques, the reproduction of the global cdf is. LSSIS realizations reproduce satisfactorily the global category proportions (Figure 12, left). For LSSGS, the realizations cdfs show a slight bias in the lower values (Figure 12, right). The overestimation of low values is caused by several factors. One of them is the abundance of zero values, which cause a spike in cdf, which is difficult to model by Hermite polynomials even after applying the program despike. The numerical instability of the Hermitian Gaussian transformation model causes order relation problems that may translate in the overestimation of the spikes when backtransforming. Additionally, if the zero values are contiguous, as it is the case of this dataset, the sequential simulation using previously simulated nodes produce zones of very low values, which are extended due to local cdfs with low means and variances, and local variograms of high continuity. Contrarily, traditional SGS realizations with a global cdf show higher values within zero and very low-grade zones, allowing a correct reproduction of the input cdf. The solution to this problem is to work within separate domains modelled using LSSIS.

The use of location dependent proportions and variograms improves the connectivity in the LSSIS realizations. For this example, the connectivity between the locations marked as blue circles in Figure 3 and Figure 5 is checked. Half of the LSSIS realizations are connected between these points, while only two of the 100 SGS realizations are. Figure 13, left, shows the distribution of the SIS and LSSIS realizations according the number of connected cells. The accuracy plot (Figure 13, right), calculated using the exhaustive dataset as reference, shows that LSSGS realizations are accurate and more precise than SGS realizations. The increased precision of LSSGS realizations comes with the price of increased computation time. Running 100 SGS realizations took 9.6min, while it took 23.5min for the same number of LSSGS realizations.

**Concluding remarks**

Locally stationary SIS realizations reproduce satisfactorily the global category proportions and the local spatial continuity informed by location-dependent indicator variograms. The increased local information provided by the local variogram models result in not only more geologically realistic models than those produced by traditional SIS, but also in a better reproduction of the connectivity within categories. This also impacts the domain boundary uncertainty contours, which appear more continuous in areas were the local spatial continuity diverges from the global spatial continuity model. These benefits come with the price of an increased computation time; LSSIS can be more than 4 times slower than SIS. Additionally, the effort required for modeling the location-dependent variograms at several anchor points may discourage its application for more than a few categories, even using semiautomatic variogram fitting algorithm.

LSSGS realizations reproduce the non-stationary features informed by the local variogram models. These models also impact the posterior conditional variances. At locations were the local variogram models show high continuity the conditional variances are lowered. Since the same distance weights modify both the local cdfs and experimental local correlograms, there is no need of using the trend model residuals for inferring the later.

Using Hermite polynomials is an efficient way of modelling and storing the Gaussian transformation functions at all required locations. However, if spikes are present in data, fitting the transformation functions by Hermite polynomial series may be difficult and order relation problems may appear.

The global input cdf is reasonably reproduced by the LSSGS realizations. However, an overestimation of low-grade values may occur if samples with zero value are abundant and contiguous. In such cases, the sequential simulation using previously simulated nodes produce zones of very low values, which are expanded due to local cdfs with low means and variances, and local variograms of high continuity. The solution to this problem is to work within separate domains. Despite this drawback LSSGS can produce accurate realizations, which have higher precision than those generated by traditional SGS. As for LSSIS, LSSGS is much slower than its globally stationary counterpart is. This is because instead of using a global covariance lookup table, the covariances need to be recalculated for each location and realization using the location dependent parameters.

Locally stationary simulation techniques may noticeably improve the models of categorical and continuous variables. However, this comes at the price of increased effort in the inference of the location-dependent cdfs and their statistics, as well as an increased demand of computer resources. Thus, the use of these techniques can only be justified if information is abundant enough to allow a reliable inference of all the required local statistics.

## References

Chilès, J.-P., & Delfiner, P. (1999). *Geostatistics, Modeling Spatial Uncertainty.* New York: John Wiley & Sons, INC.

Deutsch, C. V. (2002). *Geostatistical Reservoir Modeling.* New York: Oxford University Press.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB. Geostatistical Software Library and User's Guide* (2nd Edition ed.). New York: Oxford University Press.

Deutsch, C. V., & Zanon, S. (2002). Ultimatesgsim: Non-stationary sequential gaussian simulation by rock type - Paper 51. *Report 4.* Centre for Computational Geostatistics.

Gonzales, E., McLennan, J. A., & Deutsch, C. V. (2006). A new approach to sequential gaussian simulation with a trend: Non-stationary transformation tables - Paper 120. *Report 8.* Centre for Computational Geostatistics.

Isaaks, E. H., & Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics.* New York: Oxford University Press.

Journel, A. G. (1980). The lognormal approach to predicting local distributions. *Mathematical Geology , 12* (4), 285-303.

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

Leuangthong, O., Prins, C., & Deutsch, C. V. (2006). SGSIM_LVA: Gaussian simulation with locally varying angles - Paper 408. *Report 8.* Centre for Computational Geostatistics.

Lyster, S., & Deutsch, C. V. (2004). Short note on sequential indicator simulation with local probabilities derived from a deterministic rock type model - Paper 108. *Report 6.* Centre for Computational Geostatistics .

Machuca-Mory, D. F., & Deutsch, C. V. (2008b). Geostatistics with location dependent moments and distributions - Paper 109. *Report 10.* Centre for Computational Geostatistics.

Machuca-Mory, D. F., & Deutsch, C. V. (2009). Locally stationary multiGaussian kriging with local change of support - Paper 104. *Report 11.* Centre for Computational Geostatistics.

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

McLennan, J. A., & Deutsch, C. V. (2008). A new approach to SGS with a trend: a non stationary gaussian transformation. *Proceedings of the Eighth International Geostatistics Congress, Geostats 2008* (pp. 419-428). Santiago, Chile: GECAMIN Ltd.

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

```
START OF L-D Parameters
1                              - use location dependent statistics? (0=no, 1=yes)
Herpol-wlcdes.dat              - file with local Hermite polynomials (in columns)
30 1                           - number of hermite polynomials and column for phi(0)
kgl-cf-wlc-g30-1sexp-c0.out    - file for local nugget effect (same grid)
1                              - column for local nugget effect
exp.out                        - file for local stable model power (same grid)
4                              - column for local stable model power
kgl-cf-wlc-g30-1sexp-c1.out    - file for local sill contributions (same grid)
1 0                            - columns for sill contributions for the 1st and 2nd structure
ah.out                         - file for local ranges (same grid)
1  2  3                        - columns for hmax, hmin, hvert for the 1st structure
ang1.out                       - file with local angles
1  2  3                        - columns for ang1, ang2, ang3 for the 1st structure
```

**Figure 1:** Additional block in the ultimateSGSIM parameter file for coefficients of the location-dependent Hermitian Gaussian transformation models and the parameters of the location-dependent variogram models.

```
                  Parameters for SISIM_loc
                  ************************

START OF PARAMETERS:
0                              - 1=continuous(cdf), 0=categorical(pdf)
2                              - number thresholds/categories
0.0  1.0                       - thresholds / categories
0.7813 0.2187                  - global cdf / pdf
rdcDHx200.out                  - file with data
2   3 0  8                     - columns for X,Y,Z, and variable
LDM_IFac.dbg                   - file with gridded indicator prior mean
0    1.0e21                    - trimming limits
0.0   1.0                      - minimum and maximum data value
1      1.0                     - lower tail option and parameter
3      1.0                     - middle     option and parameter
1      1.0                     - upper tail option and parameter
rdcDHx200.out                  - file with data
8   9                          - columns for variable, weight
3                              - debugging level: 0,1,2,3
sisim_loc.dbg                  - file for debugging output
sisim_loc.out                  - file for simulation output
1                              - number of realizations
22   900.0  200.0              - nx,xmn,xsiz
24  -800.0  200.0              - ny,ymn,ysiz
1    1.0   10.0                - nz,zmn,zsiz
69069                          - random number seed
6                              - maximum original data  for each kriging
6                              - maximum previous nodes for each kriging
1                              - assign data to nodes? (0=no,1=yes)
0     3                        - multiple grid search? (0=no,1=yes),num
0                              - maximum per octant    (0=not used)
500.0  250.0  2.0              - maximum search radii
130.0   0.0   0.0              - angles for search ellipsoid
50    25     1                 - size of covariance lookup table
0    1.0                       - 0=full IK, 1=median approx. (cutoff)
1    0.00   1.0                - One   nst, nugget effect, stable model exp
2    1.00 130.0   0.0   0.0    -      it,cc,ang1,ang2,ang3
         500.0  200.0  10.0    -      a_hmax, a_hmin, a_vert
1    0.00   1.0                - Two   nst, nugget effect, stable model exp
2    1.00 130.0   0.0   0.0    -      it,cc,ang1,ang2,ang3
         500.0  200.0  10.0    -      a_hmax, a_hmin, a_vert

START OF LDVs
vfLIC0.sum                     - Cat 1: file for variogram parameters (same grid)
5   7                          - columns:   nugget effect, exp
8   9   0   0                  -            cc,ang1,ang2,ang3
12  13   0                     -            a_hmax, a_hmin, a_vert
vfLIC1.sum                     - Cat 2: file for variogram parameters (same grid)
5   7                          - columns:   nugget effect, exp
8   9   0   0                  -            cc,ang1,ang2,ang3
12  13   0                     -            a_hmax, a_hmin, a_vert
```

**Figure 2:** Parameter file of SISIM_loc program including a block for the parameters of  location-dependent variogram models.

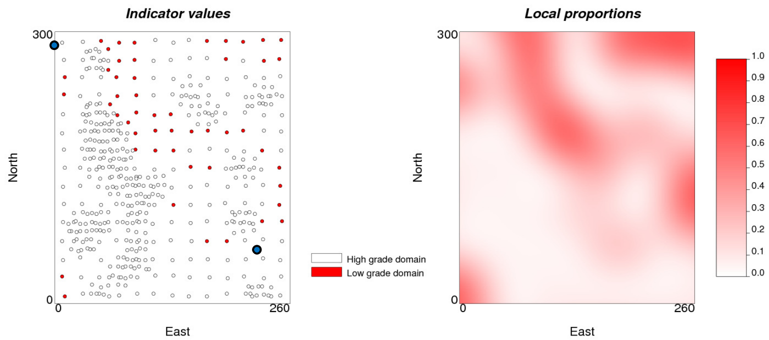


**Figure 3:** Location of categorical samples of the Walker Lake dataset (left). Local proportions of the low-grade category inferring using a Gaussian kernel distance function with 20units bandwidth (right). The blue dots in the left figure indicate the locations for checking the reproduction of connectivity.

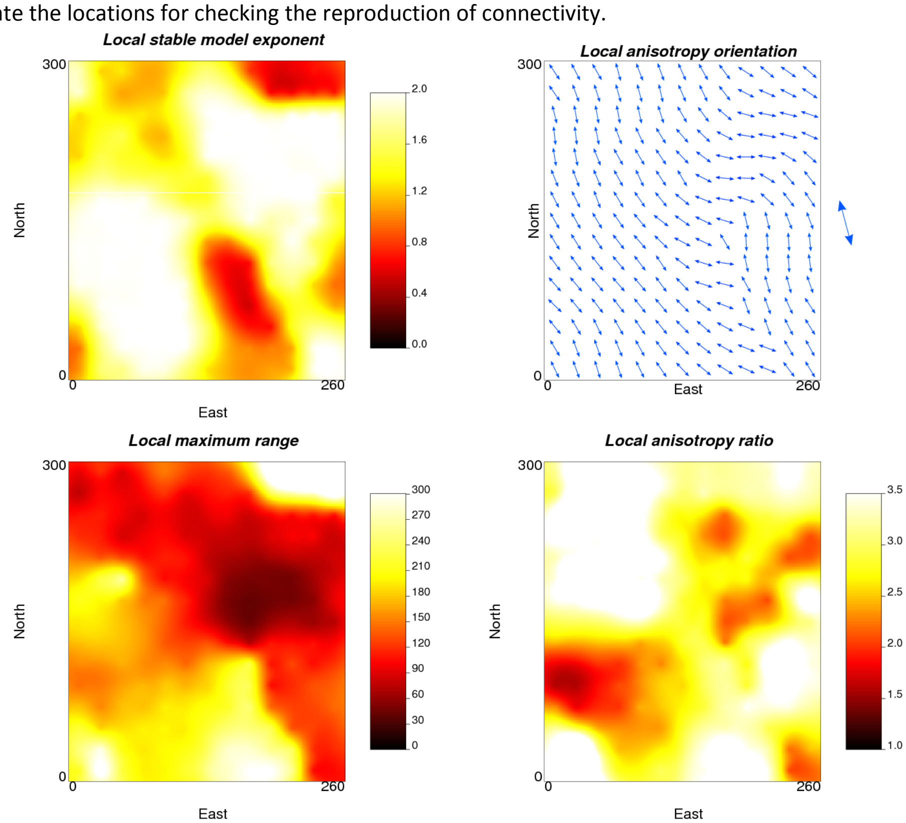


**Figure 4:** Local parameters for the location-dependent indicator variogram models.

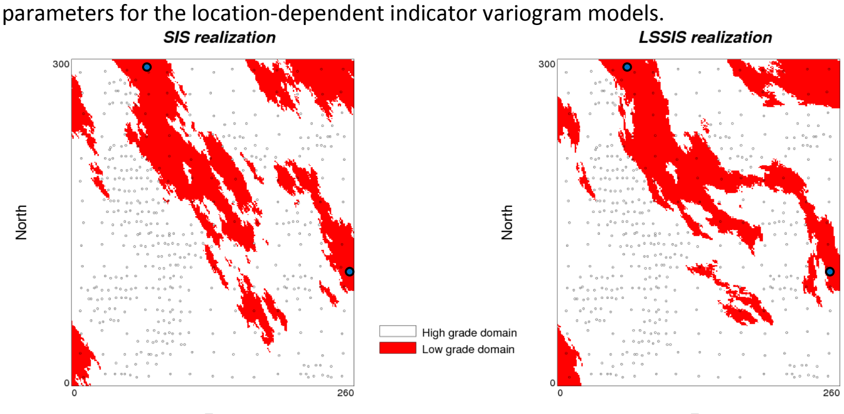


**Figure 5:** Examples of a SIS realization (left) and a equivalent LSSIS realization (right). The blue dots in the figures indicate the locations for checking the reproduction of connectivity.

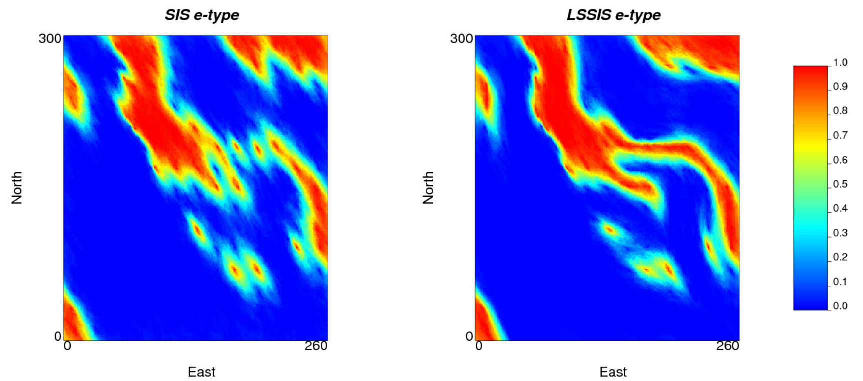


**Figure 6:** E-type estimates calculated from 100 SIS realizations (left) and 100 LSSIS realizations (right)

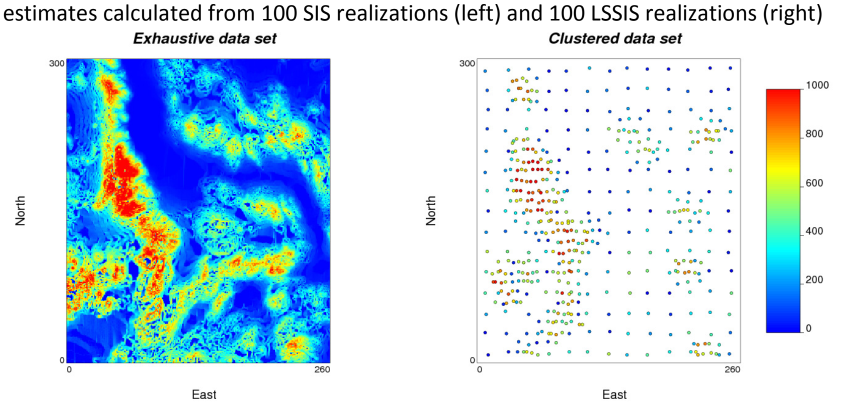


**Figure 7:** Walker Lake data set. Exhaustive reference map of the continuous variable (left) and locations of the clustered samples (right)

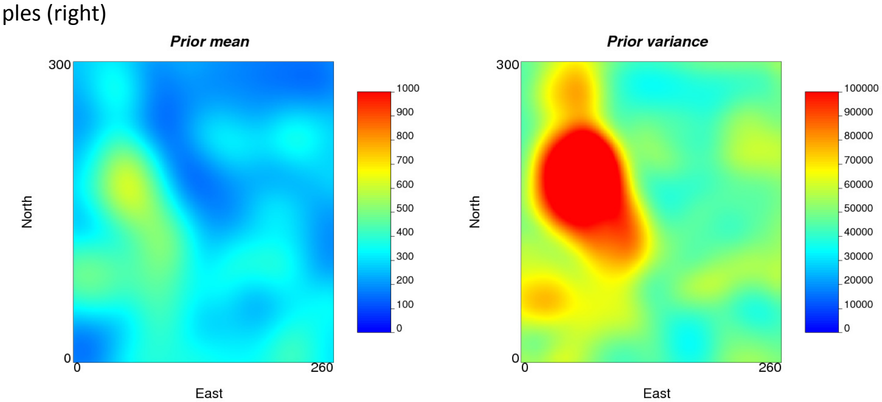


**Figure 8:** Prior local means (left) and variances (right) approximated by the first local Hermite coefficient and by the sum of the squares of the following 40 local Hermite coefficients, respectively.

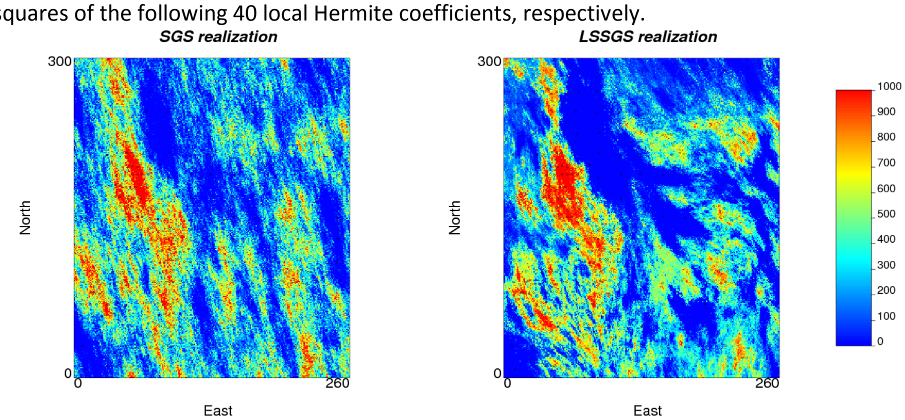


**Figure 9:** Examples of a SGS realization (left) and a equivalent LSSGS realization (right).

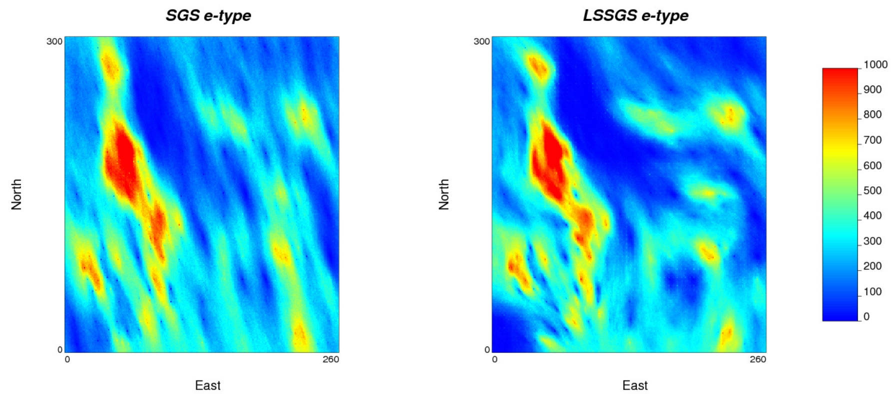


**Figure 10:** E-type estimates calculated from 100 SGS realizations (left) and 100 LSSGS realizations (right)

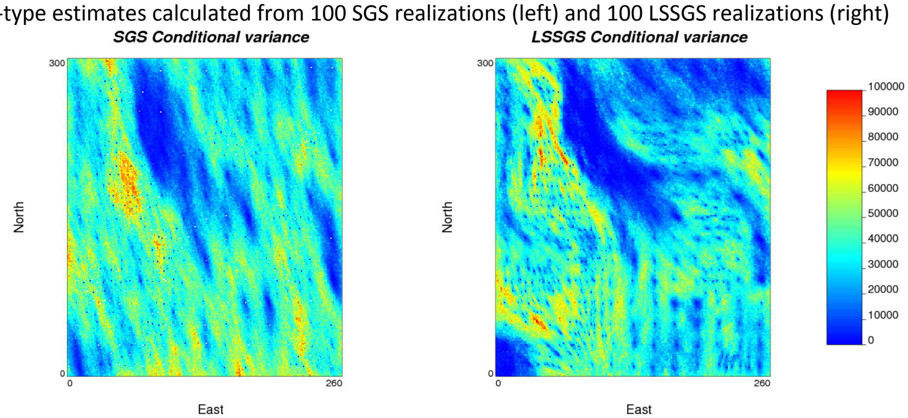


**Figure 11:** Posterior conditional variances calculated from 100 SGS realizations (left) and 100 LSSGS realizations (right)

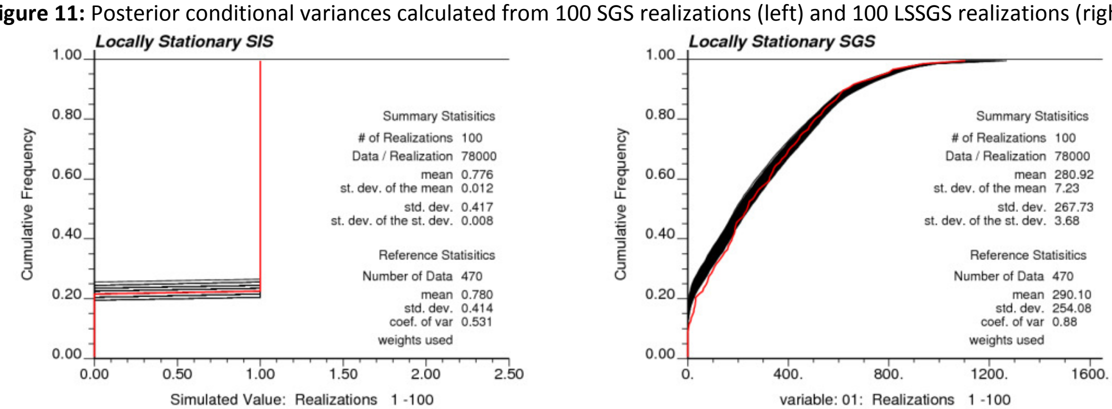


**Figure 12:** Categorical cdf reproduction of 100 LSSIS realizations (left). Continuous cdf reproduction of 100 LSSGS realizations (right)

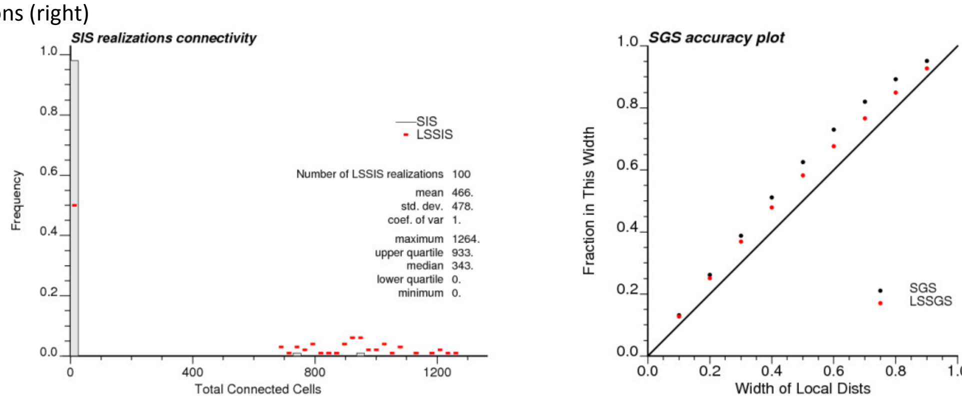


**Figure 13:** Connectivity histograms for SIS and LSSIS realizations (left). Accuracy plots of SGS vs. LSSGS realizations (right).