Kriging and Simulation in the Presence of Locally Varying Anisotropy

J.B. Boisvert, J. G. Manchuk, and C. V. Deutsch

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

This paper is a continuation of "Calculating Distance in the Presence of Locally Varying Anisotropy" and will discuss the incorporation of path optimization into Kriging and sequential Gaussian simulation in the presence of locally varying anisotropy. Using the optimum path between two points increases the covariance and generates more accurate estimates/simulations. Although CPU intensive, the methodology can be implemented to reproduce complex curvilinear structures in Kriged or simulated maps. Synthetic examples will highlight the ability of this methodology to reproduce complex features that would otherwise be difficult to generate with traditional Kriging or sequential Gaussian simulation. Programs are developed to use an LVA field in a geostatistical analysis: GAMV_LVA; KT3D_LVA; and SGS_LVA.

Introduction

Stationarity is required for all geostatistical modeling. It is an assumption that is necessary to infer the statistics of a location that has not yet been sampled. Some deposits can be modeled relatively easily under the umbrella of stationarity; however, some deposits display erratic features (channels, veins, etc) that make the assumption of stationarity less justifiable. There are methods of dealing with non-stationarity from using ordinary Kriging rather than simple Kriging to more complicated solutions such as multiple point statistics. We will present an alternative methodology to consider non-stationarity explicitly.

Specifically, the assumption of stationarity in traditional geostatistical modeling techniques assumes that the histogram and variogram in a modeling domain are constant. In the case of some deposits this may not be a reasonable assumption. We will present a methodology that relaxes the assumption of stationarity in the variogram and allows it to vary locally. At all locations in a model the variogram can be different which allows for the incorporation of curvilinear structures in Kriged maps (KT3D) or realizations (SGSIM) through the user provided LVA field.

This paper will discuss how locally varying anisotropy (LVA) can be incorporated into Kriging and sequential Gaussian simulation (SGS). We will require the optimization of the non-linear path between two points in the presence of LVA which is discussed in the paper, "Calculating Distance in The Presence of Locally Varying Anisotropy" in this report.

Background

The programs SGSIM, KT3D and GAMV as they appear in GSLIB90 will be modified and understanding the basic principles of Kriging, sequential Gaussian simulation and variogram interpretation is required. For details on these techniques refer to any introductory geostatistics text (Journel and Huijbregts 1978, Isaaks and Srivastava 1989, Deutsch 2002 to name a few).

Past work in the area of LVA should also be given a brief discussion. To the authors' knowledge, all past work has used the LVA for a local search only. In this framework, when determining the Kriging weights the variogram of the estimation location is applied to its local neighborhood. The variogram is assumed constant (for this one location) and the Kriging weights are calculated. Consider estimating at the gray location in Figure 1; the north-west direction would be applied to the local area and the covariance between points would be calculated with a variogram with this direction. This is an exaggerated example where the LVA field at the estimation location is drastically different than the surrounding LVA but it highlights the limitations of applying the LVA field in this way. If the LVA changes smoothly over the field and the change occurs beyond the range of the variogram, considering the local variogram to apply to the local neighborhood of the estimation location may be a valid assumption.

Methodology

As mentioned previously, the idea of Kriging with LVA is not new. Algorithms exist that will use the local direction of continuity at the point being estimated and apply it to the neighborhood of that location. Our proposed methodology will not make this assumption and will use the Newton method to maximize the covariance between two points along a possibly non-linear path. Specifically the methodology is:

1. generate the LVA field
2. search for close data to be used in estimation
3. calculate the 'true' anisotropic distance and covariance to data
4. solve the Kriging equations to determine the estimated value at the location

Step 1 would be done prior to estimation, step 2 is common to the regular implementation of Kriging and details on step 3 can be found in the paper "Calculating Distance in The Presence of Locally Varying Anisotropy". This section will focus on the practical implementation and limitations of considering LVA in Kriging and simulation.

The LVA Field

It is intended that the LVA field be generated using expert geological knowledge and understanding of the deposit. If there is sufficient data it may also be possible to directly infer the underlying directions of continuity. The LVA field may also be digitized from geological maps or obtained using secondary information.

Searching for Nearby Data

Searching for conditioning data close to estimation or simulation locations will be done with an ellipsoidal search. The LVA field will not be considered in the search. A large search could be considered to ensure all relevant data are found.

Calculating Distance

Although the straight-line path will be used for searching, the optimized distance will be used once nearby data have been found. The previous paper, "Calculating Distance in the Presence of Locally Varying Anisotropy" discusses how to determine the optimum path between two points in a given LVA field. To summarize that paper, we will consider that the anisotropic distance between two points in space must consider the multiple changes in the LVA field between the points. The distance will then be the sum of the segments passing through each LVA grid block. Moreover, the shortest path between points and thus the highest covariance, may not be a straight line ($d_{total}=d_1+d_2+d_3+d_4+d_5$ in Figure 2). Because the anisotropy changes, the path is allowed to track through the LVA field and the path that results in the shortest distance between points is used. This is repeated whenever a distance calculation is necessary in either Kriging or simulation.

Calculating Covariance

Once the optimal path has been determined, the next step is to calculate the covariance between points. An isotropic covariance function will be applied with the optimized distance. The justification for an isotropic covariance function (rather than defining a function at every location) is the potential difficulty in predicting too many locally varying parameters. Another consideration in using one functional form for the covariance is the requirement for additive distances. The distance between two points in an LVA field is the sum of the distances through each grid block. This type of addition would not be valid for the covariance between the points.

Implementation

KT3D, SGSIM and GAMV were modified to use the true distance between points in the presence of an LVA field. Additional inputs include the LVA file and its grid definition. The major direction of continuity and the anisotropy ratios are required to define the LVA field. These are taken as a regular grid with the 5 required properties (α, β, γ, r_1, r_2). This grid must exhaustively cover the input data and the desired Kriging/simulation grid. Additional options include: a hash table to store the optimized distance between points which reduces CPU time by up to 80%; and a user defined stopping criteria for optimization.

SGS has two additional options to help reduce the CPU requirements of this methodology: (1) a single random path is used for all simulations and (2) Kriging is built into SGS.

Traditionally in an SGS, framework a new random path is used for each simulation. In SGS_LVA a single random path is used and n realizations are drawn as each node is visited. The distance between data and previously simulated nodes for all n realizations are identical; therefore, in a single pass through the simulation grid all n realizations can be generated. A new path for each realization would require many more distance optimizations and would increase CPU time by a factor of *n*.

The option to Krige within SGS_LVA has also been added. As each node is visited all distances to nearby data are optimized. To take advantage of this, two Kriging matrices are solved, the first is a matrix with all nearby data and previously simulated nodes. The resulting weights are used to solve for the mean and variance for sequential Gaussian simulation of all n realizations. The second Kriging matrix that is solved contains only the nearby data. The resulting weights are used to solve for the mean and variance for Kriging.

Variogram calculation has also been modified to use the underlying LVA grid in the calculation of the distance between pairs for experimental variogram calculation. The intention, as discussed above, is to use the optimized path to calculate the anisotropic distance between points and use an isotropic variogram to determine the subsequent covariance for Kriging and simulation. The resulting variogram model would be used as the isotropic variogram to generate covariance from the optimized distance and is required as input to SGS and Kriging with an LVA field.

Example 1 - Kriging

The following example will demonstrate the benefits of considering an LVA field. The objective is to reproduce non-linear features that are otherwise difficult to obtain using traditional Kriging. The two LVA fields shown in Figure 3 will be used in Kriging. One field depicts a smoothly varying anisotropy where as the other is characteristic of a channel structure. The anisotropy ratio is 10:1 in the channel, 1:1 outside the channel, and 10:1 in the smooth LVA grid. A maximum of 5 conditioning data are used in Kriging. As discussed above, the LVA field is used to calculate the anisotropic distance between two points. Normally, variogram calculation with optimized distances would be used to infer an isotropic variogram model; however, due to the small amount of data in this example a spherical structure with a range of 30 will be arbitrarily selected.

In both LVA fields the resulting Kriged maps display the anticipated non-linear features of the LVA field. It would be difficult to reproduce these structures using a single direction of continuity unless there were a large number of samples to control the features. The number of data used for these synthetic examples was intentionally small to highlight the effect of the LVA field; the data alone do not show the features of the underlying LVA. The structure seen in the Kriged maps (Fig. 3) is heavily influenced by the LVA fields.

Two issues were encountered: firstly, positive definiteness is not guaranteed with the optimized distance measure proposed. It has been the authors experience that the number of poorly conditioned Kriging matrices is low when ranges less than twice the field size and spherical or exponential variogram structures are used with simple Kriging. Figure 4 shows the number of indefinite systems that arise if a spherical variogram structure is used; an exponential variogram resulted in no indefinite systems. The smoothness of the LVA grid is also a factor; no indefinite systems occurred for the smooth LVA grid. Ideas for dealing with the indefinite systems are proposed in the future work section below. A robust solver can be used to solve these systems.

The second issue is CPU time. As shown in Figure 4 the CPU time for LVA Kriging is high when compared to traditional Kriging with a maximum of 15 data. The increase in CPU time for larger grids is linear; therefore, the time required for larger grids can be estimated from Figure 4. A simple hash table has been added to store the distance calculation between data points. Storing the distance between points in a hash table can significantly reduce the run time of the program at the cost of increased memory requirements.

Example 2 – Simulation

The two LVA fields used in the Kriging example will also be used to demonstrate simulation with LVA. The following parameters will be used in simulation: 4 close data and 4 previously simulated nodes; an isotropic variogram with a spherical structure and a range of 30; and a search of 30.

The realizations generated by SGS_LVA (Fig. 5 and Fig. 6), show the variability typical of SGS but also show the structure inherent in the underlying LVA field. As a check, Figure 7 and Figure 8 show the average of the 100 realizations generated in comparison with the Kriged result. They are very similar.

CPU time and poorly conditioned matrices are also issues with simulation. Because of the implementation of a single path and simulation of all realizations simultaneously, the CPU speed of SGS_LVA (Fig. 9) is similar to the CPU speed of Kriging given the same number of data used; however, increased redundancy and screening effects result in more indefinite systems.

General Guidelines

The main difficulty with incorporating LVA into Kriging and SGS is occurrence of indefinite systems. A few ideas for eliminating them will be discussed in Future Work below, but in most situations the effect of poorly conditioned matrices can be minimized with parameter selection. The following practical guidelines should be considered to minimize the number and effect of indefinite systems:

- Use simple rather than ordinary Kriging. Ordinary Kriging has been found to result in a larger number of indefinite matrices.
- Using less conditioning data and/or previously simulated nodes will reduce the size of the Kriging matrices and thus reduce the number of indefinite systems. Because much of the structure of the resulting Kriged maps or realizations will be dictated by the LVA grid, a small number of data can be used to generate acceptable maps.
- When determining what isotropic variogram to use: a shorter range will result in more stable Kriging matrices; including a nugget effect (even a small nugget effect) will increase matrix stability; and exponential and spherical variogram structures are more stable.
- In general, a smoother, less erratic and lower anisotropy ratio LVA grid will result in fewer indefinite systems.

In order to assist with debugging an extra file ‘indef_mat.out” is written out when running SGS_LVA and KT3D_LVA. This file contains the location of the indefinite matrices and poorly conditioned matrices, see Manchuk and Deutsch (2007) for a description of matrices that are fixed because of poor weights or Kriging variances.

Conclusions and Future Work

Algorithms to Krige and simulate with an LVA field have been presented. Past work with LVA has assumed that the straight line path between points is the path that will generate the ‘correct’ covariance; however, this is not the case. A Newton method optimization has been implemented that finds an optimum non-linear path between points and uses an isotropic variogram to calculate covariance.

Lately, in geostatistics, there has been a focus on geological realism and using qualitative geological knowledge in the modeling process. Often the curvilinear structures that are known to exist cannot be incorporated into the modeling process because there are few methodologies that can easily integrate such information without having a large number of data available. We present a methodology that uses an LVA field and can reproduce curvilinear features using Kriging and sequential Gaussian simulation techniques. Features that would otherwise be poorly approximated can be transferred into geostatistical predictions through the LVA field.

Future work should be directed at stabilizing the Kriging matrices and decreasing CPU requirements. One idea to address the problem of indefinite matrices is to increase the degrees of freedom of the problem. Kriging systems developed with LVA may be geometrically infeasible in the dimensionality of the given problem, thus they may be indefinite. A solution would be to consider the data to exist in a higher dimension. A variogram that is positive definite in higher dimensions would remove the indefinite systems (Curriero, 2005). Consider Kriging a location with 5 conditioning data in 3D. Because of the optimization of the distances, the geometry (orientation of data and distances between data) may be physically impossible; however, if these data were considered in 6 dimensions the orientation would be physically possible. If a variogram is used that is positive definite in 6 dimensions the resulting Kriging system would be solvable (Curriero, 2005). A second idea to reduce the number of indefinite matrices is to jointly minimize the path between all of the points in a system. Rather than optimizing the path between points independently, a global optimization of the configuration would ensure that if one path is modified all other paths are modified to keep the geometry plausible. This would eliminate the presence of indefinite Kriging matrices.

The source of the LVA field also requires further investigation. It would be reasonable to obtain a single model of anisotropy from a geological interpretation of the deposit; however, it would be difficult to assess the degree of uncertainty in this model. A methodology to generate multiple LVA fields for a set of data would be extremely useful for implementation of this methodology.

References

Curriero, F. C., 2005. On the use of non-Euclidean isotropy in geostatistics. Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 94.

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

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

Journel, A. G., and Huijbregts, Ch. J., 1978. Mining Geostatistics. New Jersey, The Blackburn Press. 600 p.

Manchuk, J. and Deutsch, C. V., 2007. Robust Solution of Normal (Kriging) Equations. In CCG Report 9.

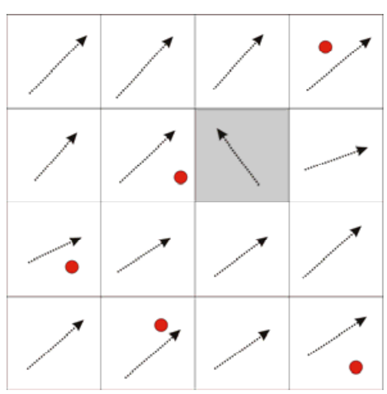

Figure 1: An example LVA field with the major direction of continuity shown as arrows and data as circles. Consider estimating the gray location without using the local change in anisotropy.

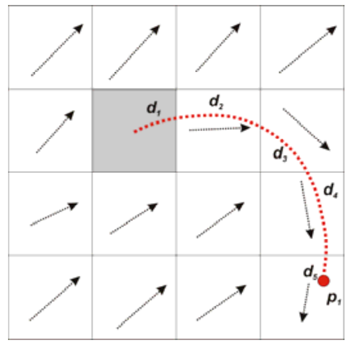

Figure 2: Calculating the true distance between the gray estimation location and a data point p_1. The dashed path shows the optimized path.

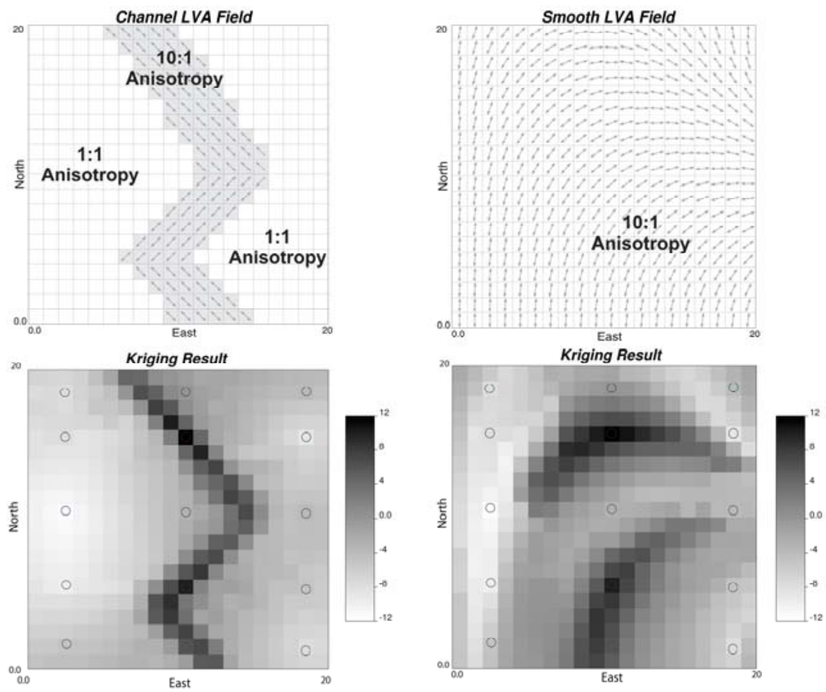

Figure 3: Above: Major direction of the LVA field. Below: Kriging result with the appropriate LVA field showing the 15 conditioning data used. Kriging was done with a maximum of 5 data.

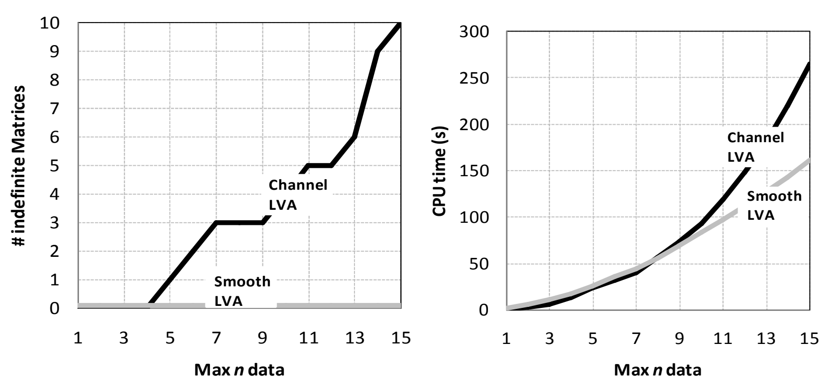

Figure 4: *KT3D_LVA* - Left: Number of indefinite matrices for both LVA fields. Right: CPU time using a hash table. The grid is 20x20 blocks and requires the solution to 400 Kriging matrices. ‘Max *n* data’ refers to using a maximum of *n* data in Kriging.

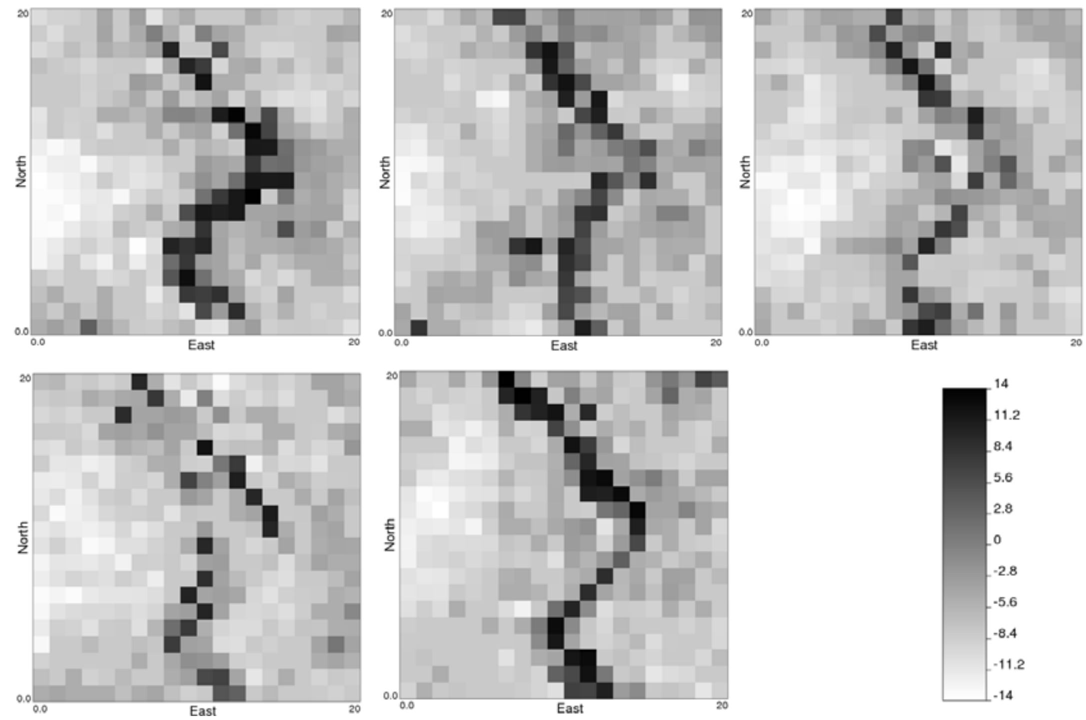

Figure 5: 5 typical realizations using the channel LVA field. The 15 conditioning data shown on Figure 3 were used to generate the realizations.

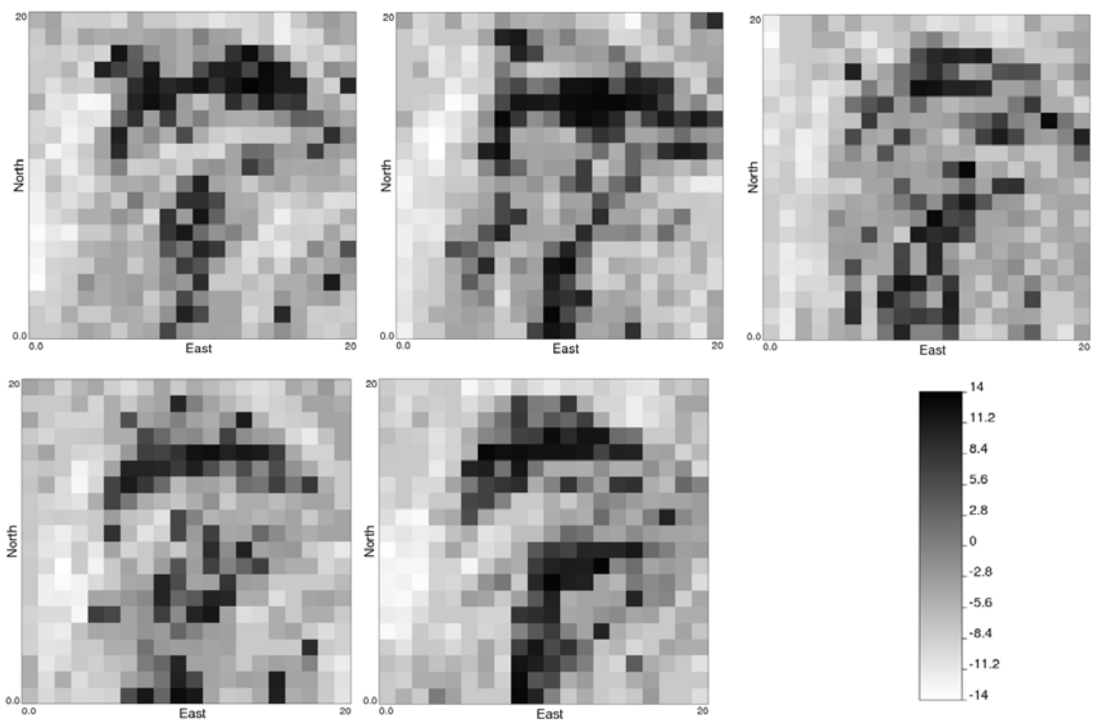

Figure 6: 5 typical realizations using the smooth LVA field. The 15 conditioning data shown on Figure 3 were used to generate the realizations.

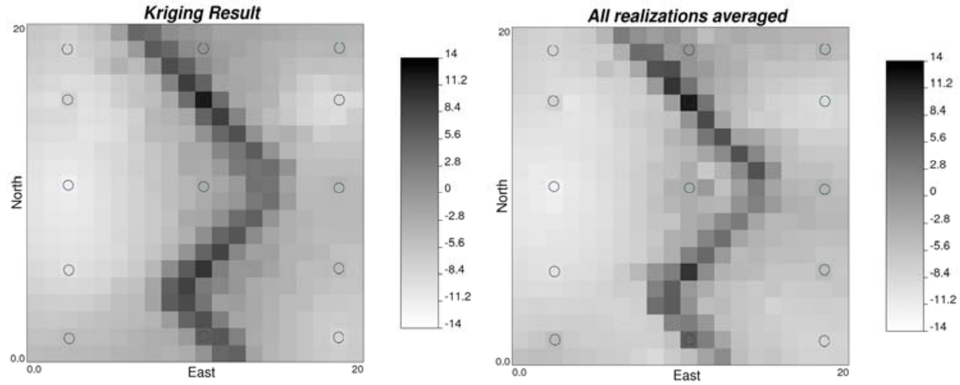

Figure 7: Channel LVA Grid. Left – Kriging result using 4 conditioning data. Right – Average of 100 realizations using 4 conditioning data and 4 previously simulated nodes.

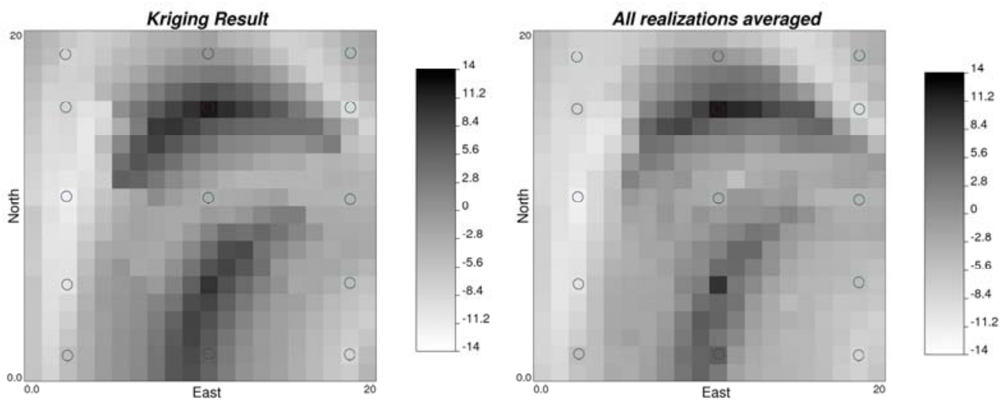

Figure 8: Smooth LVA Grid. Left – Kriging result using 4 conditioning data. Right – Average of 100 realizations using 4 conditioning data and 4 previously simulated nodes.

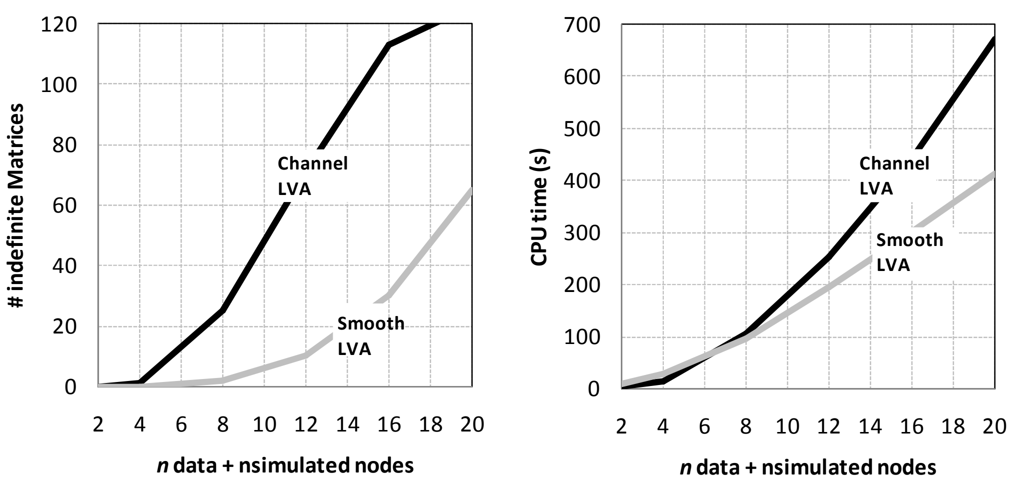

Figure 9: *SGS_LVA* - Left: Number of indefinite matrices for both LVA fields considering an exponential variogram structure. Right: CPU time. The grid is 20x20 blocks and requires the solution to 400 Kriging matrices.