

## Modeling the McMurray Formation with Locally Varying Anisotropy

J.B. Boisvert and C. V. Deutsch

*This paper presents the application of the LVA methodology to the modeling the McMurray formation thickness in the Alberta oil sands. The 2D nature of this data set allows for a clear exploration of the details of modeling with LVA. Issues such as selecting appropriate ISOMAP-L parameters, variogram modeling in high dimensions, LVA field resolution, and other model inputs are examined. Resulting errors from making approximations, such as reduced landmark points in the application of ISOMAP-L, are quantified. It is found that the modeling methodology is insensitive to ISOMAP-L parameters but sensitive to LVA field choice.*

### Introduction

The methodology applied to modeling with LVA is: (1) generate the required input parameters: LVA field; data histogram; variogram, (2) apply ISOMAP-L to ensure the positive definiteness of the kriging system of equations, and (3) perform inverse distance, kriging or simulation with LVA.

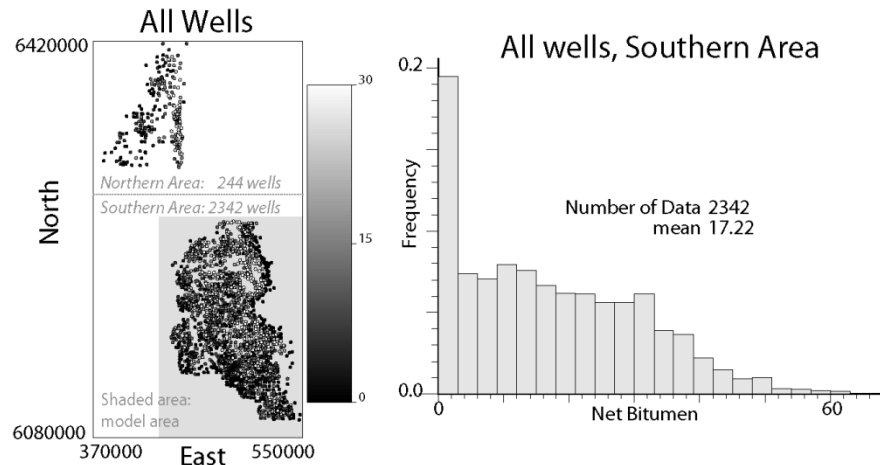
The data used in this paper can be obtained from [www.eub.ca](http://www.eub.ca). The data was used by the Alberta Energy and Utilities Board (EUB) to determine which gas pools are in pressure communication with bitumen for the assessment of risk associated with extracting hydrocarbons with steam assisted gravity drainage (Warren, 2003). The data contains the thickness of a number of oil bearing stratigraphic layers. The thickest and most significant channel is the McMurray Channel which contains a significant proportion of the oil sands in the region. The EUB study region considered two distinct regions, the northern and southern study areas. Only the southern region, where the majority of the wells are located, is considered (Figure 1). Of the 3265 wells evaluated for bitumen thickness, 2342 are located in the southern region.

### The Data

The available data cover an extent of 111km by 170km. In the southern area there are a total of 2342 wells where the net bitumen thickness has been determined (Figure 1). The drill dates for these wells are available and the data is partitioned into four subsets based on date (Table 1). The base data set contains all wells drilled prior to 2000. There are three other data sets that include wells drilled in 2000, 2001 and 2002. Such a partition allows for realistic model checking. For example, geostatistical models can be created using the pre-2000 data and checked using the wells drilled in 2000. A second model, created using the pre-2001 data, can be constructed and checked against the wells drilled in 2001, and so on. This mimics the natural progression of a deposit as models are often assessed based on future sampling. The size of the geostatistical model for this study is 575 by 900 cells with dimensions 200m x 200m, for a total of 517,500 cells.

**Table 1:** EUB data summary. Only data in the southern area are considered.

Data set	Number of wells	NCB Mean	NCB Standard Deviation
Pre-2000 wells	1673	14.3	12.1
Wells drilled in 2000	166	24.4	13.5
Wells drilled in 2001	284	25.7	14.5
Wells drilled in 2002	219	23.3	15.0
<b>All wells combined</b>	<b>2342</b>	<b>17.2</b>	<b>13.6</b>



**Figure 1:** Wells in the EUB study area. The shaded area indicates the extents of the geostatistical model for this case study.

### The LVA field

There is no available secondary data to help infer the LVA field; this is the most difficult situation as little is known about the specific geology (i.e. location of the channels) and there is no secondary data that can be used to parameterize the LVA field (i.e. geophysical data or extensively sampled secondary data). Therefore, a number of potential LVA fields are generated based on the sample data and carried through the analysis to account for the high degree of uncertainty in the LVA field.

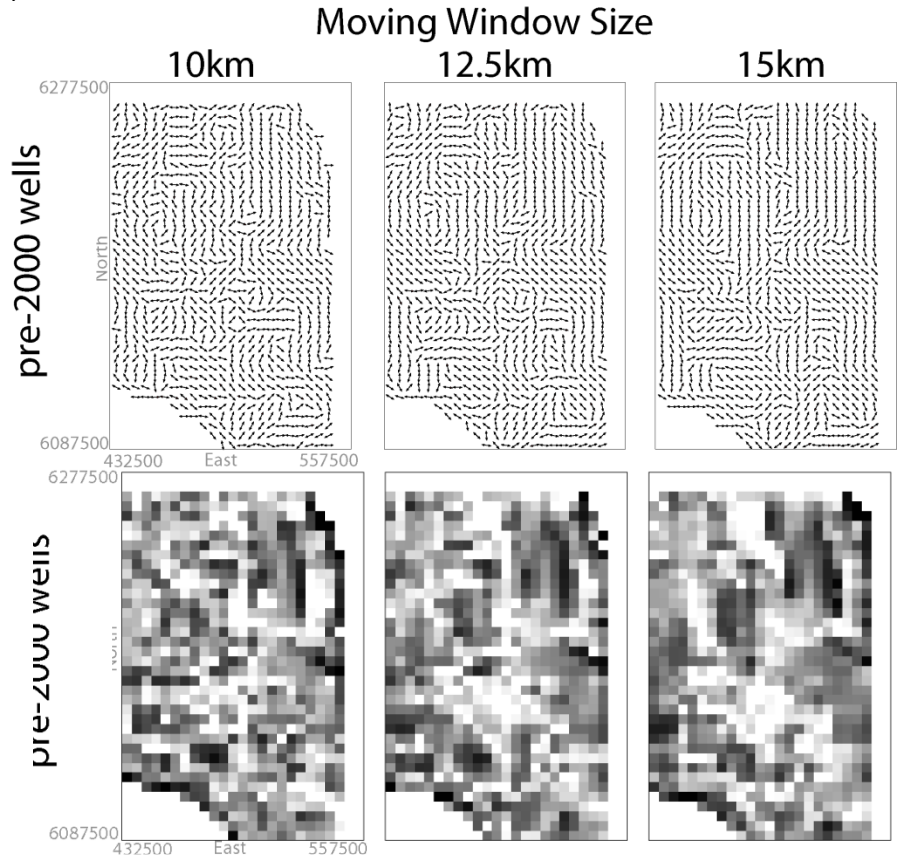
The first methodology used to generate the LVA field is the moment of inertia method (paper 403 in CCG report 10, 2008). This methodology provides both a direction of continuity as well as a magnitude of continuity. The only parameter required to implement this technique is the moving window size, used to calculate the local variogram volume maps. Window sizes of 10km, 12.5km and 15km are used (Figure 2). The window sizes were selected based on a visual inspection of the resulting LVA fields; window sizes larger than 15km produce LVA fields that are too smooth and do not capture the locally varying directions of continuity while window sizes smaller than 10km are too erratic.

In addition to generating the LVA field from the moment of inertia method, two manual LVA fields are generated (Figure 3). This is accomplished by first block kriging the NCB data with a long range variogram and analyzing the resulting smoothed map. Block kriging is implemented to produce a smoother map which highlights the large scale features. A normal score transform is used before kriging as simulation is performed in Gaussian space and requires input parameters such as the LVA field and the variogram in normal units. Moreover, the normal score transform reduces the effect of outliers on the apparent directions of continuity.

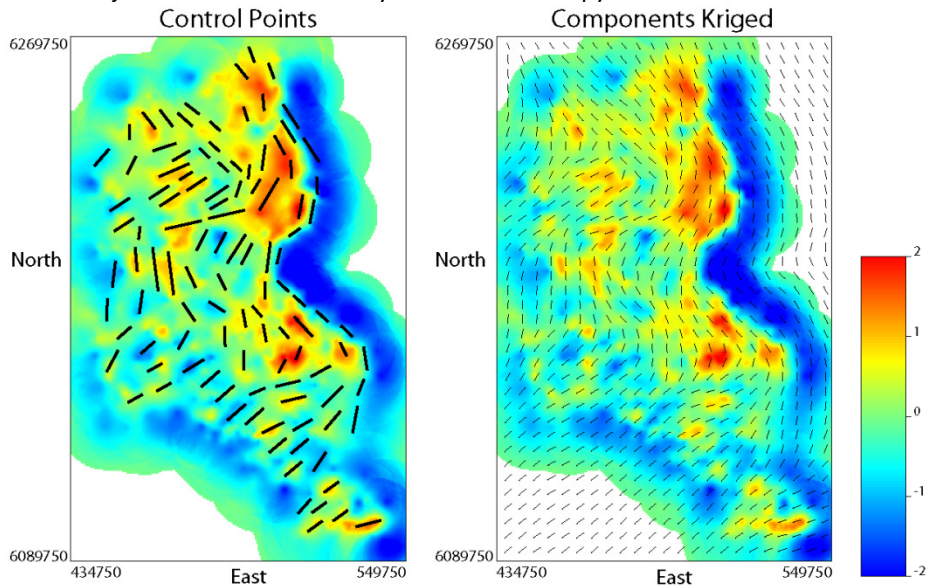
Directions of continuity are manually selected based on the block kriged map (Figure 3) and the x and y components of the directions are kriged to generate an exhaustive directional field. For simplicity, the anisotropy ratio for the first manual LVA field is constant. The value of the constant anisotropy is determined by performing LVA kriging with multiple anisotropy ratios and selecting the one that performs best with crossvalidation (Figure 4). The covariance between the true values and the estimated values is used as a measure of accuracy. A range of anisotropy values are implemented, and a constant value of 1:2.2 or 0.45 is selected. This is based on the plateau reached after an anisotropy of 0.45, beyond 0.45 there is very little change in the cross validation results (Figure 4).

A second manual LVA field is generated with the same directions of continuity but with a variable anisotropy ratio. The variable anisotropy ratio is determined by further analyzing the cross validation results. For each datum, the anisotropy ratio that generated the minimum error is assigned to that location and the anisotropy ratio is block kriged to generate the locally varying anisotropy ratio map for the second manual LVA field.

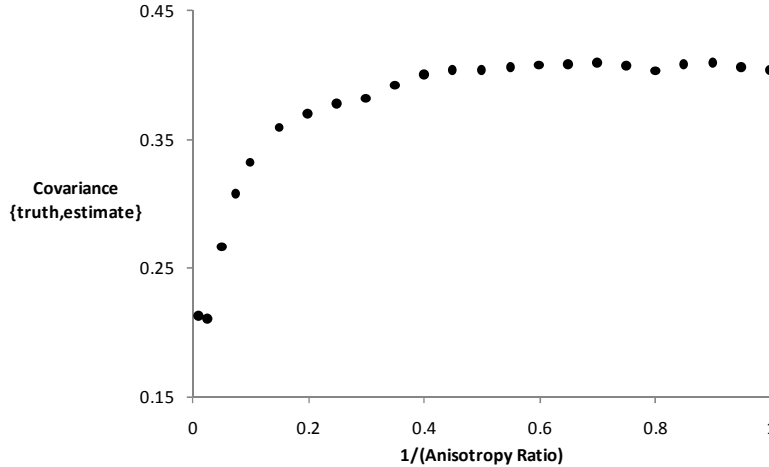
There are a total of three LVA fields generated from the moment of inertia method for each data set considered with different moving window sizes and there are two LVA fields manually generated based on the pre-2000 data. All LVA fields are carried forward in the modeling process to account for the uncertainty in the LVA field. In other situations where the LVA field is known with more certainty a single LVA field may suffice.



**Figure 2:** Above: Major direction of continuity. Below: Anisotropy ratio.



**Figure 3:** Left: Lines are manually selected based on the block kriged NCB value. These control points become the data used to krig the direction of anisotropy. Right: Resulting direction of anisotropy.



**Figure 4:** Covariance between the truth and the estimate. 1/ratio is plotted, i.e. an anisotropy ratio of 4:1 would be 0.25 (1/4).

### ISOMAP-L

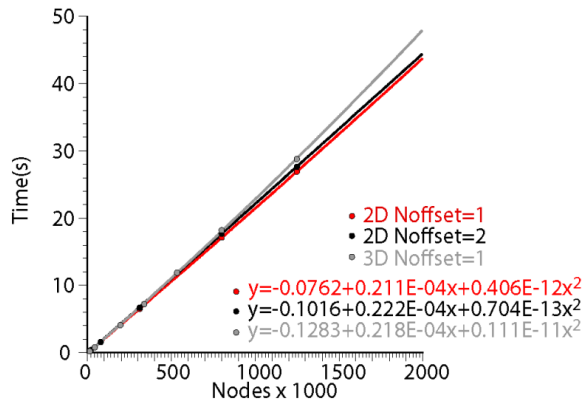
The mathematical requirement for positive definiteness of the kriging system of equations requires that the covariance matrix between data used in kriging to be positive definite. ISOMAP-L (paper 111 in CCG report 10, 2008) is applied to meet this constraint. The only input to ISOMAP-L is the location of  $L$  landmark points that are used to embed the geostatistical grid into a high dimension Euclidian space. The number and location of these landmark points must be determined.

The most straightforward approach to defining the  $L$  landmark points is based on the required run time. Recall that the SPD (Chapter 4) from the  $L$  landmark points to each grid cell is required for embedding with ISOMAP-L. Ideally, every cell in the geostatistical grid would be considered a landmark point, however, this may require prohibitive CPU time to determine the required SPD with the Dijkstra algorithm. Figure 5 can be used to determine the necessary run time given a 2D or 3D geostatistical modeling grid with  $n$  cells. The practitioner can select  $L$  such that the algorithm effectively runs in the allotted time available. For example, if a 5,000,000 cell 2D model is required and 2 offsets for the Dijkstra algorithm are selected, the CPU time to calculate the distance from one landmark point to all grid cells is 113s. If the practitioner is willing to wait 12 hours for this algorithm the maximum number of landmark points is  $L = 380$ .

A more theoretical determination of  $L$  requires the use of an error measure such as the stress criterion (Equation 1).  $L$  could be selected such that the error criterion is below an acceptable threshold. The drawback of this methodology is that the threshold can vary depending on the LVA field used; it is not possible to select a single acceptable stress level to determine  $L$  without knowing the lower bound on the stress, which is too CPU intensive to determine for large models. Moreover, the stress is not actually a measure of importance in a geostatistical application; consider increasing every  $\hat{d}_{ij}^q$  in Equation 1 by a factor of 10. This would dramatically increase the stress but the resulting geostatistical realizations would be unchanged (only the variogram range would be increased by a factor of 10). Therefore, it is recommended that a case specific error measure be used to judge the number of landmark points. Alternatively, the more practical solution based on the necessary CPU time can be applied.

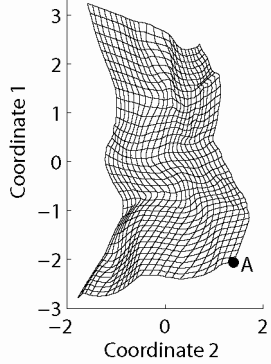
$$Stress(q) = \left\{ \frac{\sum_{j=1}^N \sum_{i=1}^{j-1} (d_{ij}^q - \hat{d}_{ij}^q)^2}{\sum_{j=1}^N \sum_{i=1}^{j-1} (d_{ij}^q)^2} \right\}^{1/2} \quad \mathbf{1}$$

where  $d_{ij}^q$  is the original distances,  $\hat{d}_{ij}^q$  is the mapped differences in  $q$  dimensions and  $N$  is the number of grid nodes

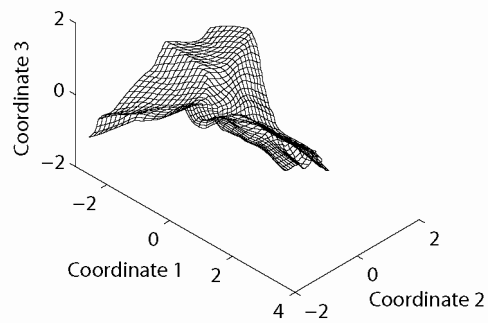


**Figure 5:** Time required to calculate the distance from  $n$  grid nodes to a single landmark point.

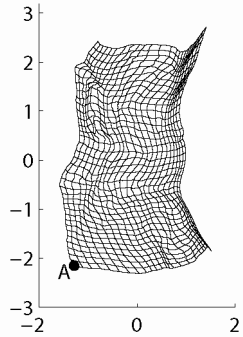
2D - Manual, Constant Anisotropy Ratio



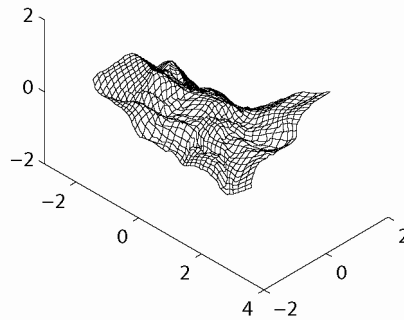
3D - Manual, Constant Anisotropy Ratio



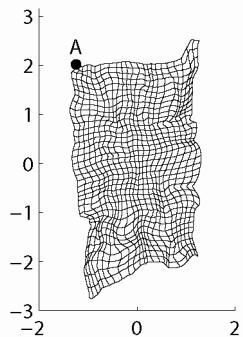
2D - Manual, Variable Anisotropy Ratio



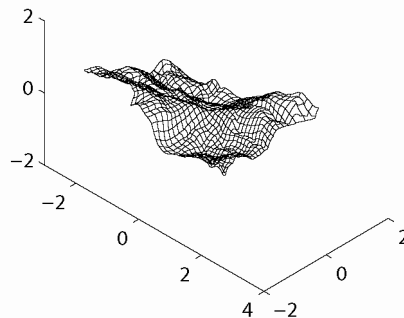
3D - Manual, Variable Anisotropy Ratio



2D - Moment of Inertia



3D - Moment of Inertia



**Figure 6:** Grids embedded in 2 and 3 dimensions for select LVA fields. The moment of inertia LVA field was generated with a window size of 12.5km. Note that the direction of each coordinate is arbitrary.

Once the number of landmark points has been selected their location must be determined. The most straightforward implementation of ISOMAP-L would be regularly spaced pattern of points using the maximum number of landmark points as determined above. The alternative is to optimize the location of the landmark points to minimize an error. The stress measure may not be an ideal error measure because the lower bound on the stress is not known a priori. A deposit specific error function is preferred.

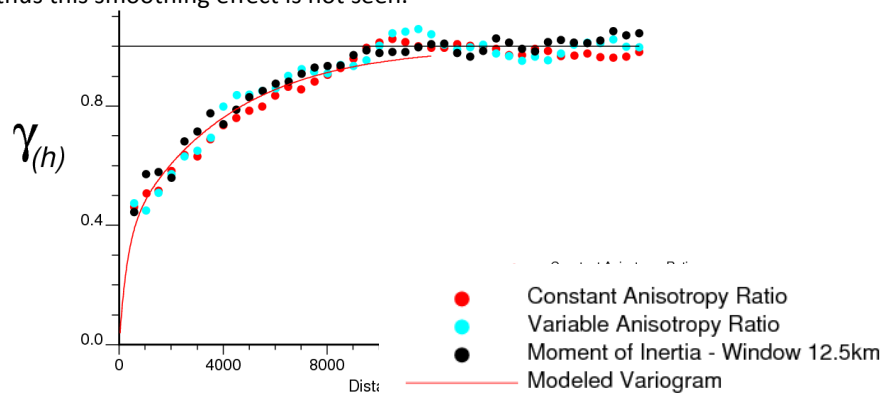
The effect of ISOMAP-L on the input grid can be visualized by plotting each grid in the first 2 or 3 dimensions after applying ISOMAP (Figure 6). Because the input grid is 2D, the resulting grid is also a plane, even in 3D (Figure 6). The smoothness of the manual LVA field is evident, while the moment of inertia method generates more erratic grids. Moreover, the grids may become rotated relative to each other as the direction of the axis is irrelevant to the reproduction of the input distance matrix, which is the goal of L-ISOMAP. Further, note the length of the grid in the first three coordinate systems. The dimensions are ordered by importance (i.e. variance of the linear combinations from L-ISOMAP), therefore, the range of the coordinates are arranged in decreasing magnitude.

### Estimation with Inverse Distance and Kriging

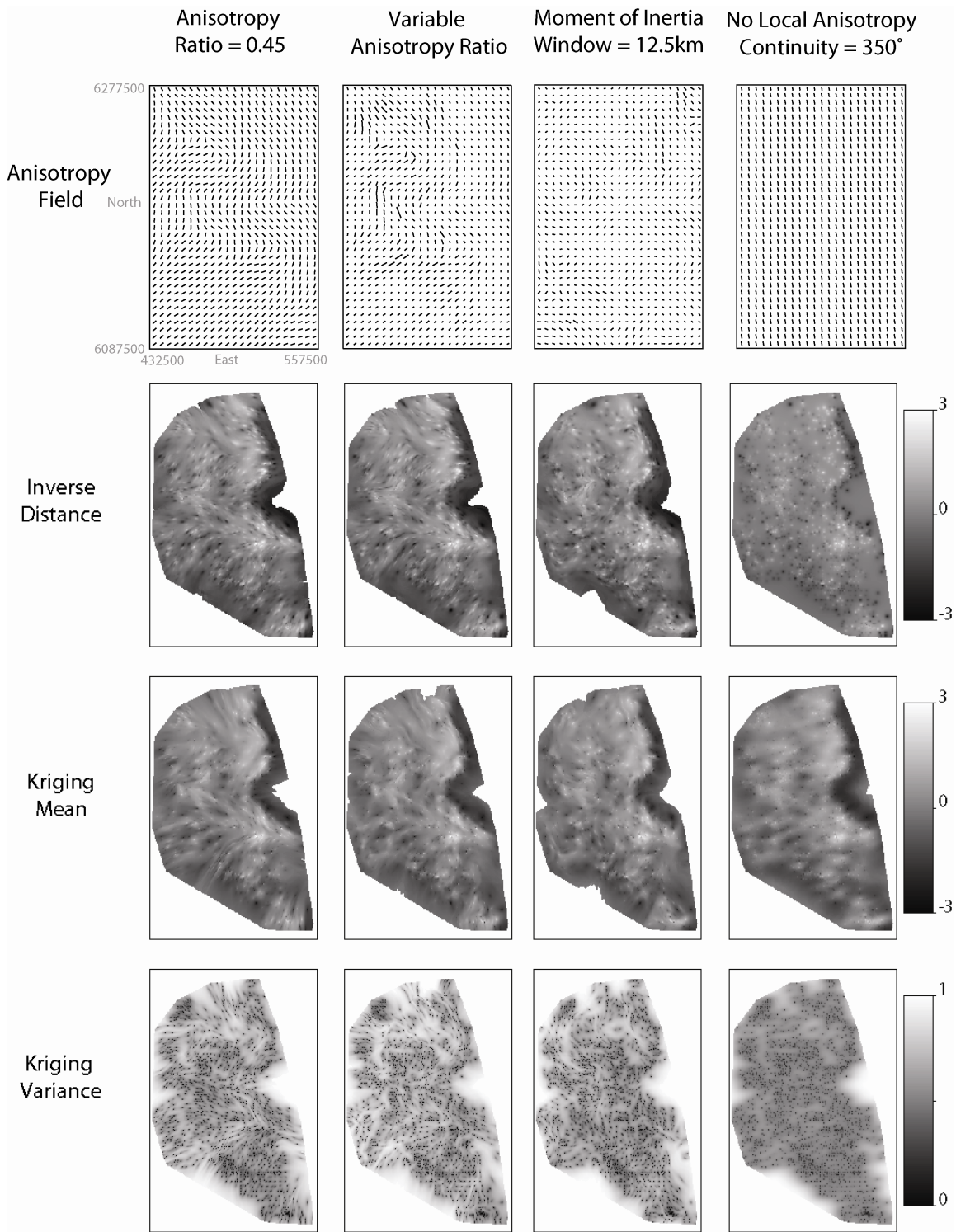
With the previous determination of the LVA field and the location of the landmark points for ISOMAP-L, kriging can be applied to map the NCB. Figure 8 presents the kriging and inverse distance maps using the two LVA fields generated manually as well as one of the moment of inertia LVA fields. The same modeled variogram was used for all maps (Figure 7) as the experimental variograms are very similar for each LVA field considered. Thus, the only difference in the kriging maps (Figure 8) is the LVA field. The modeled variogram consists of two exponential structures with ranges of 900m and 12000m and contributions of 0.35 and 0.65 respectively. A total of 100 landmark points were used with a spacing of 11.5km in the x direction and 18km in the y direction.

A few comments can be made about the maps in Figure 8:

- 1) The LVA field has a large effect on the local directions of anisotropy seen in the resulting maps, such as the NE direction of continuity in the southern area and the NW direction of continuity in the northern area.
- 2) The effect of the local directions of continuity on the kriging variance is to elongate the low variance parallel to the local directions of continuity and to increase the variance perpendicular to the direction of continuity. With traditional kriging the variance is low only in the global direction of continuity (350°).
- 3) The kriging maps are smoother than the inverse distance maps. This is because the error in applying dimensionality reduction algorithms in this application is realized as an increase in the distance between locations, particularly for shorter distances. This increase in distance results in a lower weight assigned to data and a higher weight assigned to the global mean, effectively smoothing out the maps. Dimensionality reduction algorithms are not required for inverse distance thus this smoothing effect is not seen.

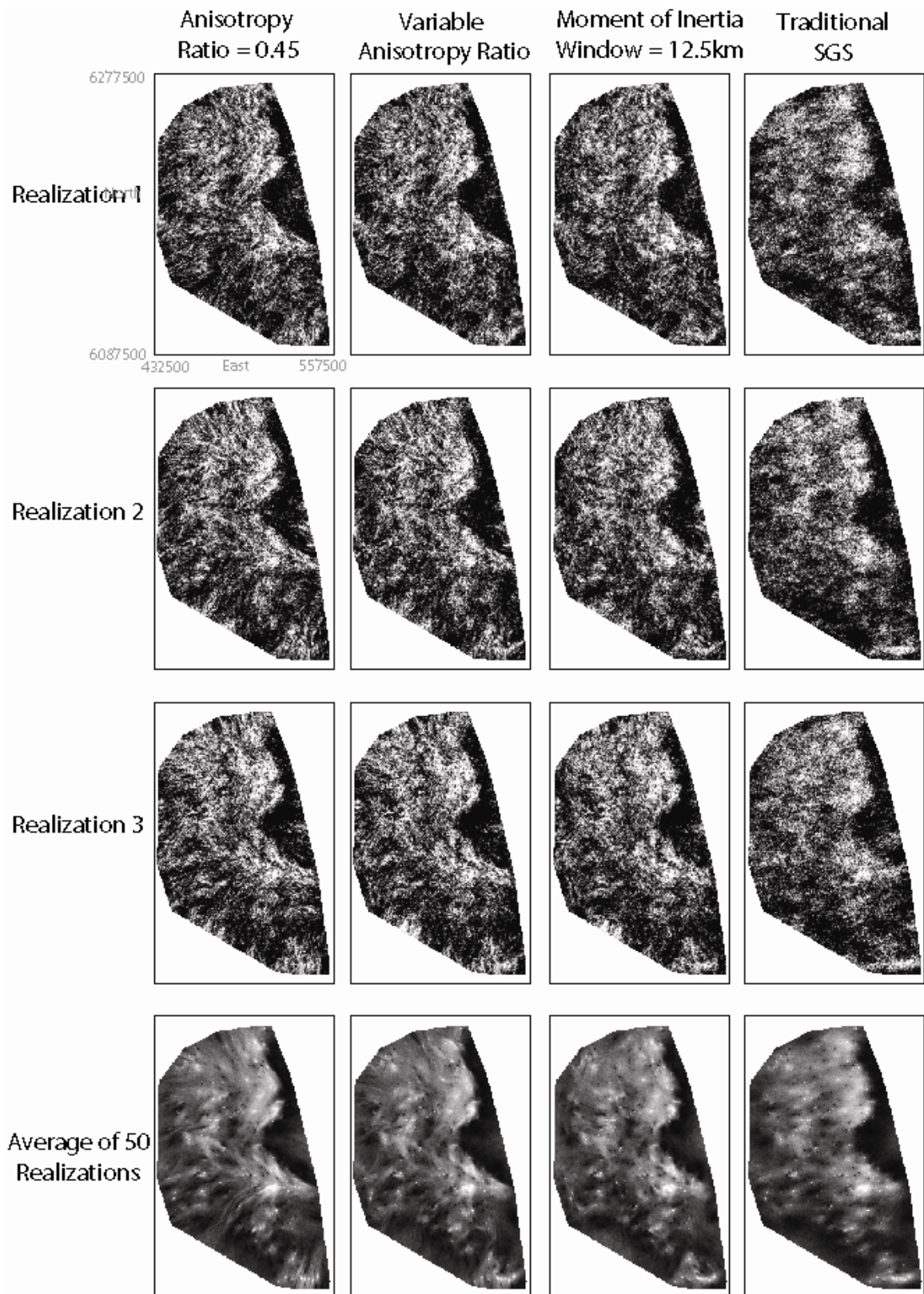


**Figure 7:** The experimental variogram for different LVA fields is different as the distance is calculated using the LVA field.



**Figure 8:** Simple kriging and inverse distance in normal score units using three different LVA fields. A search range of 7000 and a maximum of 30 data were used in the kriging. The pre-2000 data is used as conditioning data for these maps. Models are trimmed to the convex hull of the conditioning data.





**Figure 9:** 3 SGS realizations using various LVA fields. Gray scale ranges from 0m NCB (black) to 40m (white).

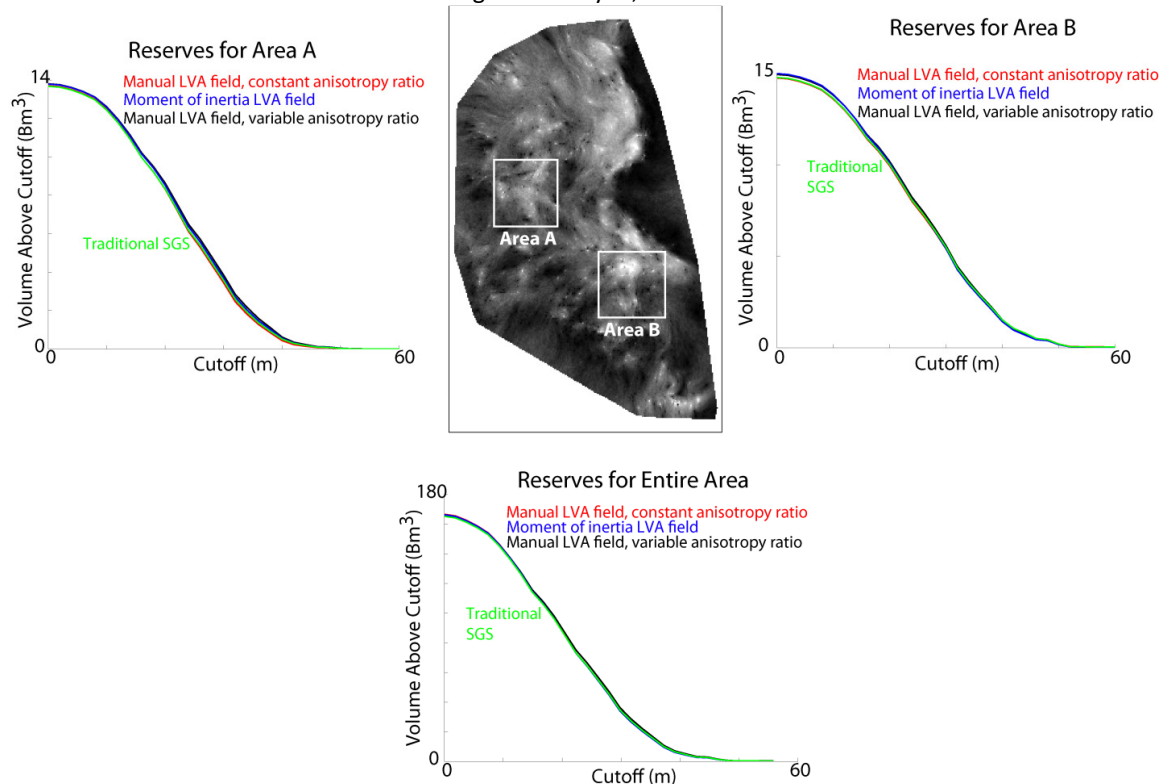


### Sequential Gaussian Simulation

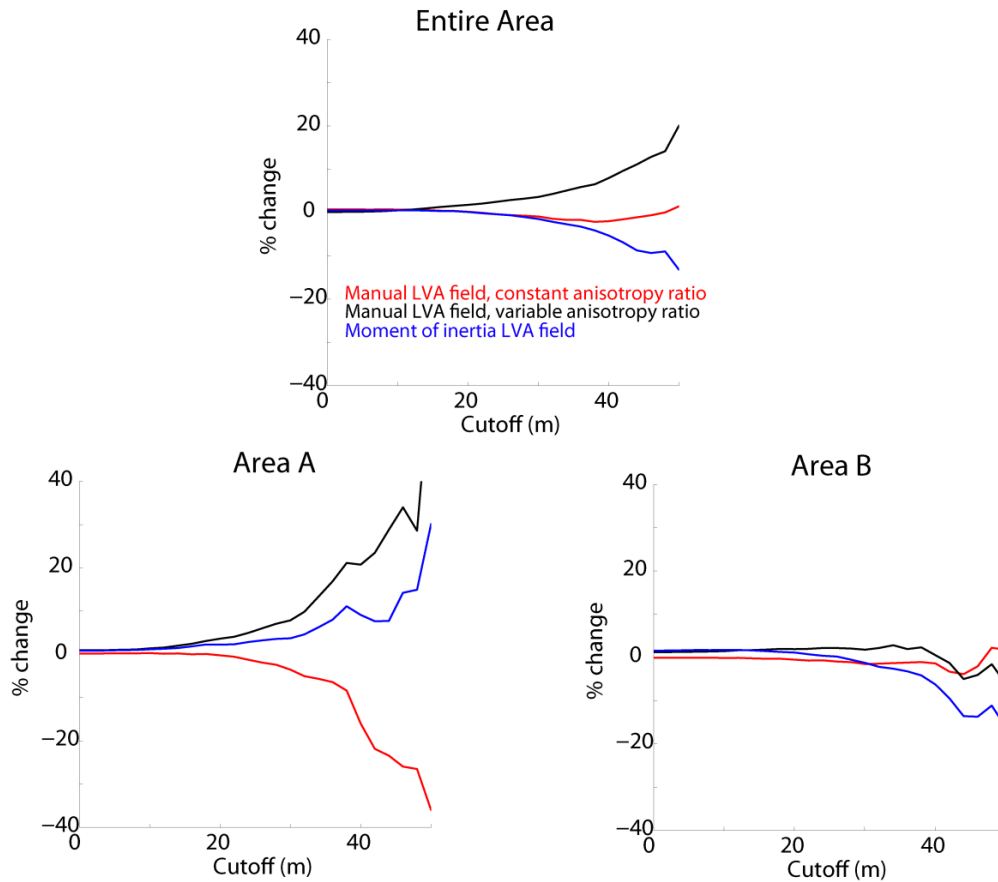
Kriging maps are useful as they provide a local estimate that is the most likely value at a given location; however, they are not appropriate for resource and reserve calculations because estimates do not generally reproduce the input histogram and variogram (Journel and Kyriakidis 2004) and reserves are often biased. Sequential Gaussian simulation (SGS) has emerged as a practical solution to the biased reserves calculated with kriging. The same parameters are required for SGS as with kriging and the variogram in Figure 7 is also used here for SGS. Figure 9 shows a number of realizations generated for each LVA field. A total of 100 landmark points were used with a spacing of 11.5km in the x direction and 18km in the y direction.

The LVA is not as visually dominant as with kriging (Figure 8) because of the addition of the variability inherent in SGS. The LVA is somewhat masked; however, when the realizations are averaged (Figure 9) the features are more distinguishable. Of interest when applying SGS is the resulting reserve calculations (Figure 10). Although difficult to perceive, there is a significant difference in the reserves calculated with different LVA fields; consider the % change in the reserves calculated for the entire model, area A and area B (Figure 11). The reserves are provided for the smaller areas to highlight that while the global reserves can be similar to traditional SGS, locally there may be larger differences.

The reserves calculation is sensitive to the choice of LVA field. Consider the difference in the reserves for area A when using the two manual LVA fields (Figure 11). Both LVA fields consider the same directions of anisotropy; however, there is considerable difference in the calculated reserves, most notably for cutoffs above 20m. It must be stressed that the selection of the LVA field is a modeling decision. Ideally, accurate geological knowledge of the area can be relied upon to construct reasonable LVA fields. If such information is unknown, the LVA field should be treated as an additional unknown variable and various LVA fields carried through the analysis, as shown here.



**Figure 10:** Reserve calculations using various LVA fields. Reserves for areas A and B are also shown. Reserves were calculated for 50 realizations and averaged.



**Figure 11:** Change in reserves in comparison to considering traditional SGS with a single direction of continuity of  $350^\circ$ . Reserves were calculated for 50 realizations and averaged.

### Sensitivity Studies

This section presents the effect of various input parameters on geomodeling with LVA. The various data sets are used to assess the quality of the models generated. There are three cases considered for modeling:

Case 1: Wells drilled pre-2000 are used to generate geomodels and assessed with wells drilled in 2000.

Case 2: Wells drilled pre-2001 are used to generate geomodels and assessed with wells drilled in 2001.

Case 3: Wells drilled pre-2002 are used to generate geomodels and assessed with wells drilled in 2002.

In each case the models are assessed by examining the appropriate set of 'new' wells. A cross validation analysis provides the measure of model quality. Specifically, the covariance between the truth from the 'new' wells and the model values is used. The sensitivity of the covariance measure for the following parameters are assessed:

- 1) The number of landmark points
- 2) The number of dimensions to retain after applying ISOMAP-L
- 3) The grid resolution of the LVA field.
- 4) The number of offsets used in the calculation of the SPD with the Dijkstra algorithm.

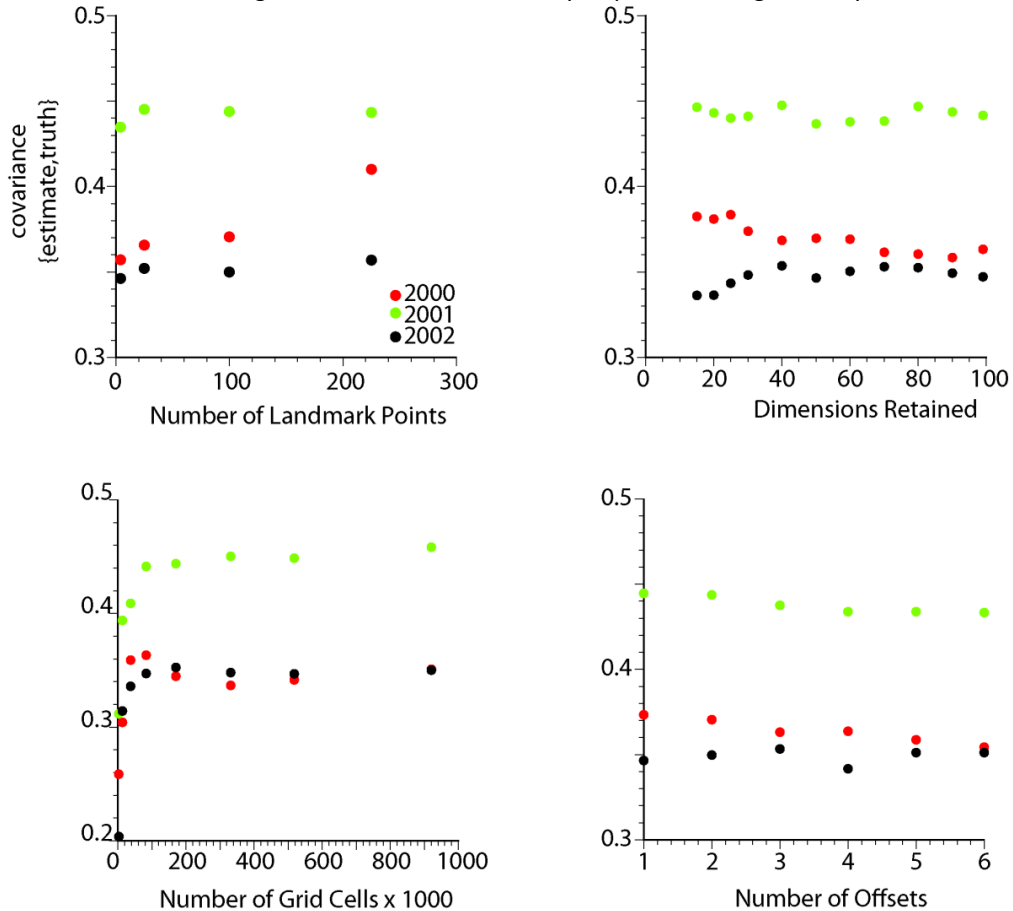
Sensitivities are assessed for input parameters 1-4. Each parameter is varied from a minimum value to a maximum value (Table 2). Because these parameters affect the implementation of ISOMAP-L, the variogram for each model is slightly different. Therefore, automatic variogram fitting software VARFIT (Pardo-Iguzquiza) was applied for each variogram and visually inspected to ensure reasonable fits.

**Table 2:** Summary of first 4 parameters assessed for model sensitivity.

Parameter	Minimum Value	Maximum Value	Default Value
Number of landmark points	2	200	100
Number of dimensions to retain	2	199	99
Grid resolution of the LVA field	4 x 4	460 x 720	230 x 360
Number of offsets used with Dijkstra	1	5	2

The grid resolution of the LVA field (Table 2) refers to the discretization of the LVA field. A very coarse grid does not allow for flexible paths with the Dijkstra algorithm. Increasing the grid resolution generates more accurate SPD's but requires additional CPU time.

The number of dimensions to retain (Table 2) refers to the number of dimensions after MDS to retain in the embedding of the grid in a high dimensional space. The maximum number of dimensions that can be retained is  $L-1$ ; however, fewer dimensions can be used to reduce the memory requirements of the methodology. For each grid cell location,  $d$  dimensions must be stored in memory for easy access. For large grids of size  $N$  and large numbers of landmark points, it may not be possible to store all  $(L-1)N$  coordinate locations. Selecting  $d \ll L-1$  reduces the memory requirements significantly.



**Figure 12:** Effect of various parameters on the covariance measure. Note the different scale in the lower left figure.

The methodology is fairly insensitive to the majority of the ISOMAP-L and Dijkstra parameters. Thus, the parameters, such as the number of landmark points, can be reduced significantly without adverse effects.

## **Conclusions**

The methodology proposed to model with LVA was applied to model the net continuous thickness of an oilsands deposit in Northern Alberta. The large extents of this deposit indicated that there were numerous directions of continuity. These locally varying features were effectively incorporated into geostatistical maps, estimations and simulations.

The main purpose of this example was to demonstrate the selection of the various input parameters required to implement the methodology. Specifically, the generation of the LVA field and the selection of the necessary parameters for SPD calculation and ISOMAP-L implementation were discussed. To review, the number of landmark points selected should be based on the allotted CPU time available. The orientation of the landmark points can be a regular grid for simplicity. The variogram is modeled based on the available data but requires the LVA field for the calculation of the experimental variogram. A sensitivity to the selection of ISOMAP-L parameters showed that they have a small effect on the resulting model. In contrast, selection of the LVA field was shown to have a drastic effect on important reserve calculations, especially when local areas are analyzed.

## **References**

- Pardo-Iguzquiza E. 1999. VARFIT: a fortran-77 program for fitting variogram models by weighted least squares. *Computers & Geosciences*. 25(3):251-261.
- Warren A. 2003. Report 2003-A: EUB Athabasca Wabiskaw-McMurray Regional Geological Study, December 31, 2003. Alberta Energy and Utilities Board. 187 p.
- Journel, A. and Kyriakidis, P. 2004 *Evaluation of Mineral Reserves: A Simulation Approach* Oxford University Press 232p.