University of Alberta

Geostatistics with Locally Varying Anisotropy

by

Jeffery Brian Boisvert

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mining Engineering

Department of Civil and Environmental Engineering

Spring 2010

Edmonton, Alberta

Examining Committee

Clayton V. Deutsch, Civil and Environmental Engineering

Alireza Nouri, Civil and Environmental Engineering

Dale Schuurmans, Computing Science

Hooman Askari-Nasab, Civil and Environmental Engineering

Derek Apel, Civil and Environmental Engineering

Roussos Dimitrakopolous, Mining and Materials Engineering, McGill University

To my mother

Abstract

Many geological deposits contain nonlinear anisotropic features such as veins, channels, folds or local changes in orientation; numerical property modeling must account for these features to be reliable and predictive. This work incorporates locally varying anisotropy into inverse distance estimation, kriging and sequential Gaussian simulation. The methodology is applicable to a range of fields including (1) *mining*-mineral grade modeling (2) *petroleum*-porosity, permeability, saturation and facies modeling (3) *environmental*-contaminate concentration modeling. An exhaustive vector field defines the direction and magnitude of anisotropy and must be specified prior to modeling. Techniques explored for obtaining this field include: manual; moment of inertia of local covariance maps; direct estimation and; automatic feature interpolation.

The methodology for integrating locally varying anisotropy into numerical modeling is based on modifying the distance/covariance between locations in space. Normally, the straight line path determines distance but in the presence of nonlinear features the appropriate path between locations traces *along* the features. These paths are calculated with the Dijkstra algorithm and may be nonlinear in the presence of locally varying anisotropy. Nonlinear paths do not ensure positive definiteness of the required system of equations when used with kriging or sequential Gaussian simulation. Classical multidimensional scaling is applied to ensure positive definiteness but is found to be computationally infeasible for large models, thus, landmark points are used for efficiency with acceptable losses in precision. The methodology is demonstrated on two data sets (1) net thickness of the McMurray formation in Alberta and (2) gold grade in a porphyry deposit. Integrating LVA into numerical modeling increases local accuracy and improves leave-one-out cross validation analysis results in both case studies.

Acknowledgements

I would like to thank my supervisor Dr. Clayton V. Deutsch, without whom this thesis would not have been possible. I would also like to thank the members of the Centre for Computational Geostatistics for their financial support.

I would especially like to thank John Manchuk for the countless hours of discussion and his thorough review of this thesis.

Of course, the patience of my close partner, Tammy, require note; throughout the long arduous process of writing this thesis she has continued to support my efforts.

Finally, I could not have gotten this far without the encouragement of my father.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

BLUE	best linear unbiased estimator
FMI	formation micro image
GPS	global positioning system
IDW	inverse distance weighted interpolation
ISOMAP	isometric feature mapping
L-ISOMAP	isometric feature mapping with landmark points
LLE	local linear embedding
LVA	locally varying anisotropy
MDS	multidimensional scaling
MPS	multiple point statistics
NCB	net continuous bitumen
PCA	principal component analysis
RAM	random access memory
RF	random function
RV	random variable
SAGD	steam assisted gravity drainage
SB	superblock search
SE	squared error
SIC	spatial interpolation contest
SIS	sequential indicator simulation
SNESIM	single normal equation simulation
SPD	shortest path distance
SGS	sequential Gaussian simulation
sph	spherical variogram structure
TI	training image
∇	first derivative
∇^2	second derivative
α	strike angle
β	dip angle
φ	plunge angle

$\gamma(\mathbf{h})$	variogram
$\Gamma(h)$	variogram function
ε	machine precision
λ_i	weight for the i^{th} data point
ρ	correlation
σ^2	variance
ω	inverse distance estimation power
a_{max}	range parameter in the horizontal direction with the maximum continuity
a_{min}	range parameter in the horizontal direction with the minimum continuity
a_{vert}	range parameter in the vertical direction
a_x	range parameter in the x direction
a_y	range parameter in the y direction
a_z	range parameter in the z direction
$\mathbf{A}$	modeling domain (area)
$C(\mathbf{h})$	covariance between points separated by a distance $\mathbf{h}$
$C(\mathbf{u_1},\mathbf{u_2})$	covariance between points at location $\mathbf{u_1}$ and $\mathbf{u_2}$
C	contribution factor for a variogram structure
d	distance
D	distance matrix
E	expected value
$f(Z)$	probability density function of Z
$\mathbf{h}$	distance or lag vector
h	magnitude of the distance or lag vector
H	Hessian
I	moment of inertia
J	Jacobian
L	number of landmark points
m	mean
p_{10}	10^{th} percentile
p_{50}	50^{th} percentile
p_{90}	90^{th} percentile
$P(A)$	the probability that event A will occur

q	number of dimensions for multidimensional scaling
r	number of realizations
r_1	anisotropy ratio between the major and minor directions of continuity
r_2	anisotropy ratio between the major and vertical directions of continuity
R	rotation matrix
u	a location in space
Z(**u**)	a random variable
Z*(u)	an estimate of the random variable

Chapter 1: Introduction

This chapter is organized as follows. Section 1.1 motivates the use of geostatistics for reserve estimation and motivates the use of anisotropy in these calculations for increased accuracy. Section 1.2 introduces the framework that is used throughout this dissertation to incorporate locally varying anisotropy (LVA). Sections 1.3 and 1.4 describe the problem context and summarize the contributions of this dissertation. Section 1.5 provides background on geostatistical techniques relevant to the proposed methodology. Finally, Section 1.6 reviews previous work on the subject of LVA.

1.1 Motivation

1.1.1 Motivating Geostatistics for Reserve Estimation

This dissertation is mainly concerned with the evaluation of mineral and petroleum resources. Evaluation of resources consists of calculating the volume of valuable materials (gold, copper, oil, etc.) present in a deposit. This calculation has far-reaching implications for company profit, feasibility of new operations, and the environmental impact of resource extraction; an accurate estimate of resource volume is essential. Any estimate of resource volume is made based on limited drill hole or well data and as such, all estimates contain a certain level of uncertainty (the potential for error). Thus, the primary goal of geostatistics is to make the best possible assessment of resources while obtaining the best possible understanding of the associated uncertainty.

Properties of a deposit that effect its profitable and safe extraction must be modeled based on the available sample data. Examples of properties of interest include: grade per ton; porosity; permeability; volume of oil; water saturation; contaminate concentrations; volume of radioactive elements and; density. *Modeling* these properties is simply predicting their values at all locations in a deposit including locations that have not been sampled. It would be unrealistic and too expensive to obtain sufficient drill hole data to exhaustively cover the deposit. Thus, methods have been developed that use the available drill hole data to predict values of the properties at unsampled locations. All such geostatistical methods begin with laying a grid over the deposit of interest (Figure 1.1 left) and properties are predicted in every block in the grid (Figure 1.1 middle). Calculating volumes from these predicted properties is usually a straightforward process; for example, once grade has been modeled at all locations, the number of grams in each block can be calculated and the total volume of the resource determined.

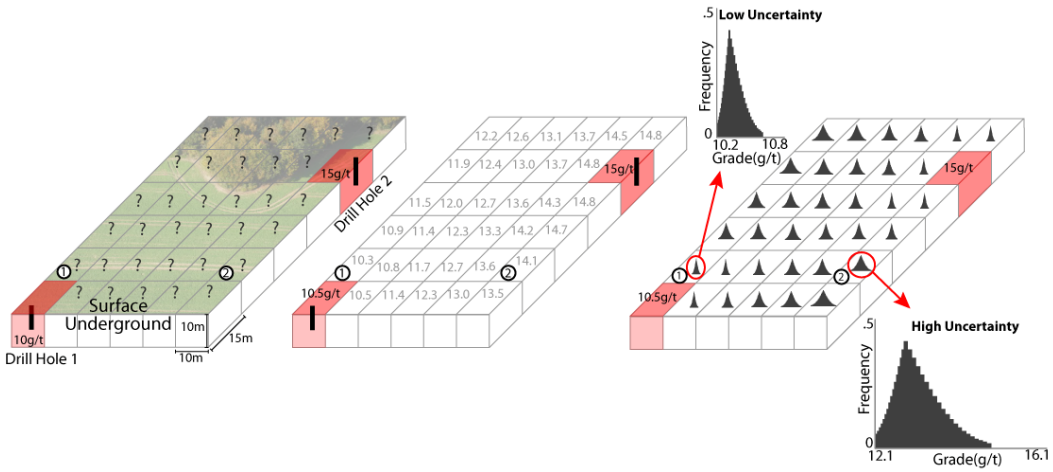


Figure 1.1: The basics of geostatistics. (image : www.surreyproperty.com). Left: A grid is placed over the deposit. Two drill holes are shown with grade values of 10 grams/ton and 15 grams/ton of gold, the grade is unknown at all other locations (?). Middle: Estimating grade at all unsampled locations. At a single block, i.e. location 1, there would be a total of $10.3g/t \cdot 10m \cdot 10m \cdot 15m \cdot 2t/m^3$=30.9kg of gold. This is repeated for all blocks and the resource volume can be calculated. Right: A distribution of values is generated at all unsampled locations. Location 1 is closer to a known drill hole so there is less uncertainty.

Any prediction at unsampled locations is uncertain. An estimate can be made but a measure of the reliability of the estimate is required. Geostatistics employs statistical tools to (1) make predictions of properties in unsampled blocks and (2) quantitatively assess the level of uncertainty in those predictions. The most common tool to describe the uncertainty of a property at an unsampled location is to predict a distribution of possible values rather than a single value. It is generally conceded that it would be too difficult and inappropriate to generate a single value at all unsampled locations (Figure 1.1 middle). As such, distributions of possible values are generated at each location (Figure 1.1 right). This allows geostatisticians to make probabilistic statements such as "*there is a 90% probability that the value at location 1 is greater than 10.3*". Looking at how this applies to calculating resource volumes, geostatisticians are able to make further probabilistic statements such as "*with a 90% confidence, there are at least 10M tons of gold ore in deposit x*" or "*with a 90% confidence, there are at least 10B barrels of oil in reservoir y*". Indeed, government regulations are moving towards these types of probabilistic reserve estimations, such as the National Instrument 51-101 Standards for Disclosure for Oil and Gas Activities in Canada (available from www.albertasecurities.com):

"Possible reserves are those additional reserves that are less certain to be recovered than probable reserves. There is a 10% probability that the quantities actually recovered will equal or exceed the sum of proved plus probable plus possible reserves."

1.1.2 Motivating Locally Varying Anisotropy

Since the origin of geostatistics with Matheron's (1962) pioneering work, numerous new techniques have been developed to generate the distributions at unsampled locations. Many of these techniques are based on the theory of kriging; introductory discussions

of kriging can be found in geostatistical text books (Journel and Huijbregts 1978; Isaaks and Srivastava 1989; Cressie 1993, Deutsch 2002). One advantage of using kriging to generate estimates is the incorporation of anisotropy into the modeling process. Anisotropy is the concept that the properties in a geological deposit are often more continuous in one direction than another. For example, coal deposits are often aerially continuous while gold vein deposits are typically discontinuous even at small scales (Figure 1.2). Eriksson and Siska (2000) present an excellent description of anisotropy and its application in geostatistics.

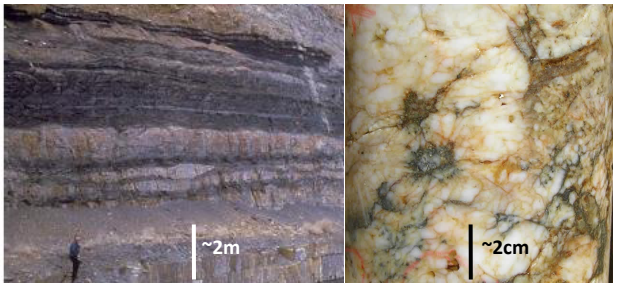


Figure 1.2: Left: A continuous coal deposit (www.scienceclarified.com). In the horizontal directions coal deposits typically show very large degrees of continuity. Right: Gold veins are typically very discontinuous and erratic as shown in this core sample (www.jpgold.com).

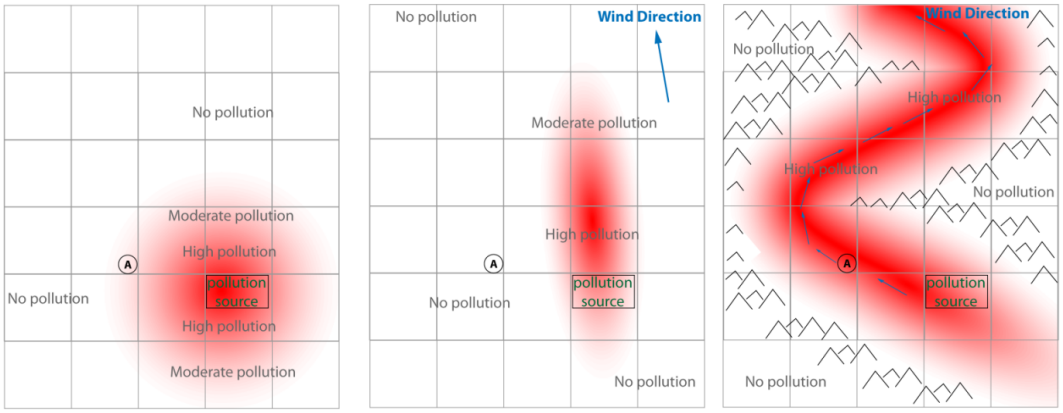


Figure 1.3: Anisotropy in geostatistics. Not to scale. Red/Gray cloud indicates increased concentration of pollution. Left: No wind - air pollution radiates from the source in all directions equally, **no anisotropy**. A moderate pollution level is expected at location A. Middle: North wind - air pollution radiates north from the source, **constant anisotropy**. No pollution is expected at location A. Right: Erratic wind - air pollution radiates from the source, **locally varying anisotropy**. A high pollution level is expected at location A.

Anisotropy is vital to this dissertation and is elaborated upon using an analogy to air pollution spread in the presence of wind. Following this example the connection between wind direction and the anisotropy in geological formations is made. Consider the simplest case of air pollution without wind (Figure 1.3 left); pollution spreads **in all directions equally** (no anisotropy). Adding a slight northerly wind (Figure 1.3 middle), the pollution is now unevenly distributed, it is **more continuous in a single direction** (constant anisotropy). If we are given the knowledge that the pollution source is located in a mountainous valley (Figure 1.3 right), the wind is now more erratic and the

pollution spread is complex, thus, the direction of anisotropy varies and the pollution is said to be **continuous in locally varying directions** (locally varying anisotropy).

Knowing the anisotropy of the pollution cloud would help in predicting pollution levels. For example, at the unsampled location A, if there is no wind we would predict moderate pollution. Adding constant northerly anisotropy in the form of wind changes this prediction to no pollution. Adding non-stationary anisotropy in the form of erratic winds further alters our estimate of the pollution level. Knowing the direction and magnitude of anisotropy helps in making better predictions at unsampled locations. This can be directly applied to modeling geological deposits. Incorporating knowledge of anisotropy in modeling improves predictions of property values at unsampled locations. If it is known that this direction changes in different areas of the deposit (Figure 1.3 right) that knowledge should be incorporated into the modeling process.

1.1.3 Locally Varying Anisotropy in Geology

In a geological setting, anisotropy refers to how continuous a deposit is in different directions. Often, deposits display the type of anisotropy described in Figure 1.3 right. Consider the LVA shown in these cross sections:

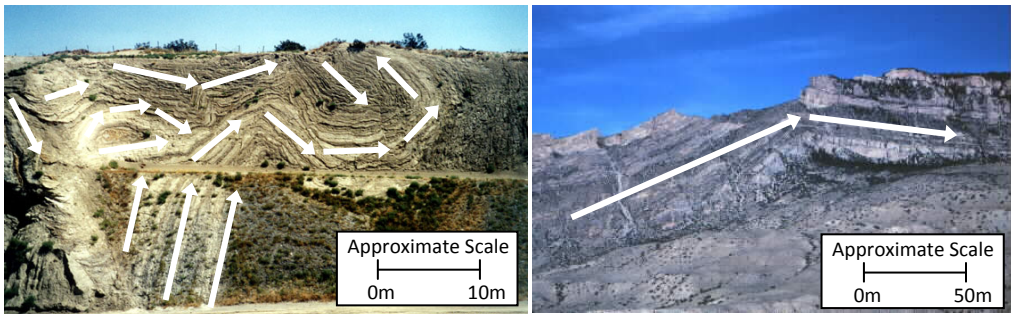


Figure 1.4: Cross sections displaying LVA. Left: Folding and faulting caused by the San Andreas Fault (http://strike-slip.geol.ucsb.edu/KESSEL/palmdaleroadcut.html). Right: Folding in the northern Rocky Mountains (http://mkutis.iweb.bsu.edu/dept/bghrn3.jpg).

Anisotropy within geological formations can be exploited to increase the accuracy of modeling. If the direction and magnitude of anisotropy are well understood, they can be transferred into modeling to improve performance. Consider an LVA field modeled after the hand drawn directions on Figure 1.4 and two drill holes through the deposit. The problem is to estimate at all the unsampled locations. Techniques, such as inverse distance, that do not normally consider anisotropy, cannot capture the horizontal continuity of the deposit (Figure 1.5 left). Other techniques, such as kriging, provide disappointing results because only a single direction of continuity can be incorporated into the modeling (Figure 1.5 middle). More geologically realistic results can be obtained by considering that the anisotropy varies locally (Figure 1.5 right); however, incorporating LVA in even this simple case is difficult. More complex anisotropy fields, such as Figure 1.4 left, are challenging.

Another example of complex LVA seen in the earth sciences is a fluvial deposits (Figure 1.6). Fluvial deposits often show LVA because they are sand that has been deposited by ancient river systems (Figure 1.6). The associated LVA map (Figure 1.6 right) is very similar to a weather map showing wind direction; however, in a geological setting the lines represent the directions in which the variable of interest is more continuous. Considering LVA in such cases would lead to more accurate numerical modeling.

The geological cross sections in Figure 1.4 and the conceptual model of Figure 1.6 provide the motivation for considering LVA; traditional geostatistical modeling only considers a single direction of anisotropy (Figure 1.5 middle). Estimates of resource volume in these situations would be improved by considering LVA (recall predicting pollution levels with LVA). This dissertation integrates LVA into geostatistical modeling.

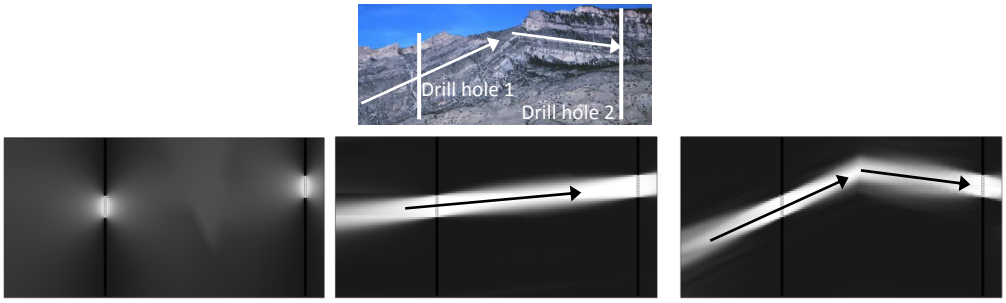


Figure 1.5: Top: cross section with the two drill holes (www.mkutis.iweb.bsu.edu). Lower Left: Inverse distance estimation. Lower Middle: Kriging with horizontal anisotropy. Lower Right: Kriging considering LVA. Plots nominally represent 300m in the x direction and 150m in the y direction, white indicates high grade while low grade is black.

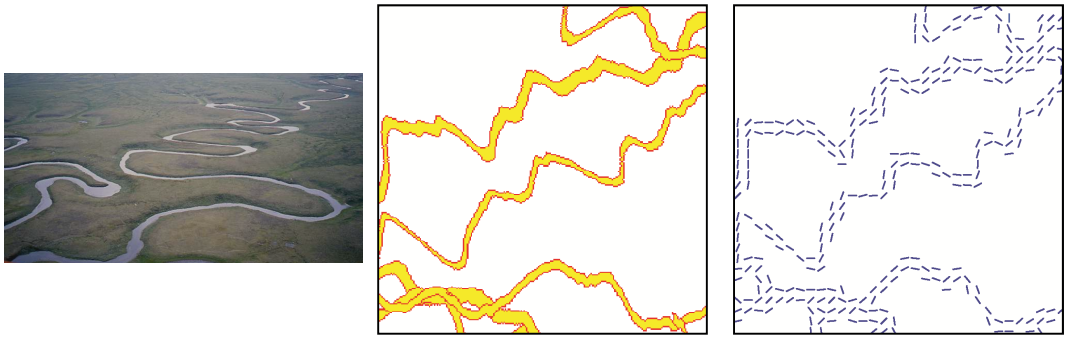


Figure 1.6: Plan view of a fluvial deposit. Left**:** A meandering channel (www.hi.is). Sand is later enriched with migrating oil to form a fluvial reservoir. Scale not available. Middle**:** Plan view of a conceptual fluvial model. Right: A map of the locally varying directions for the conceptual model. Typically, fluvial reservoirs display locally varying directions of anisotropy.

1.2 Incorporating LVA: The Shortest Path Between Points

The motivation for LVA is to incorporate complex nonlinear features to increase accuracy and generate realistic numerical models (Figure 1.3 and Figure 1.5). This section introduces the framework that will be used throughout this dissertation to model LVA.

Consider estimating at an unsampled location (?) with four surrounding data (Figure 1.7). Without anisotropy, each point receives equal weighting when applying a linear estimation technique (Equation 1.5). Considering anisotropy in different directions weights data differently. Anisotropy is incorporated in traditional geostatistics by calculating an anisotropic distance between two locations depending on the direction of anisotropy (see Appendix A for a numerical example of how this anisotropic distance is traditionally calculated). In this case, data that are closer to the location being estimated are given higher weight. The justification for this is that, geologically, two points are more related along the direction of anisotropy because during the formation of the deposit some process (sedimentation, erosion, transpiration, chemical alteration, etc) caused higher degrees of continuity in a particular direction. In the case of N-S anisotropy, data that are located to the north or south are effectively closer and, as such, have a shorter anisotropic distance (Figure 1.7 middle). In the case of E-W anisotropy, data that are located to the east or west are effectively closer and, as such, have a shorter anisotropic distance (Figure 1.7 right). This anisotropic distance between locations can be calculated as shown in Figure 1.8. The anisotropic distance between locations accounts for the preferential relationship between locations along a given direction of anisotropy.

More formally, the anisotropic distance between two points can be calculated using Equation 1.1. The range of anisotropy is specified by a_x and a_y (and a_z if considering a three dimensional case). A larger range in a particular direction effectively shortens the distance between points in that direction. The range and direction parameters fully define anisotropy (see Chapter 2 for a more in-depth discussion of anisotropy).

$$d_{12} = \sqrt{\left(\frac{d_x}{a_x}\right)^2 + \left(\frac{d_y}{a_y}\right)^2 + \left(\frac{d_z}{a_z}\right)^2} \qquad \textbf{1.1}$$

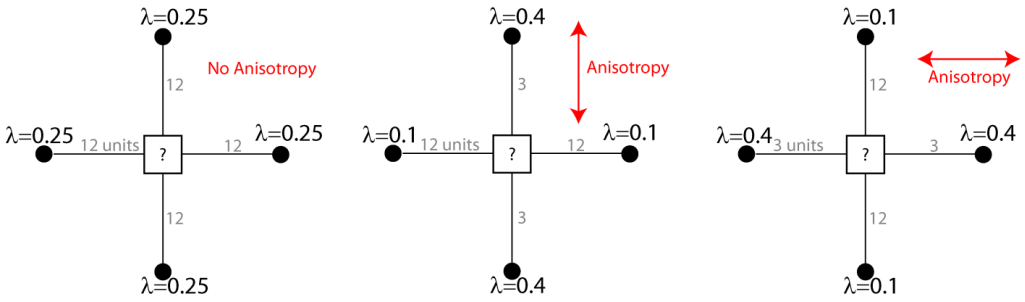
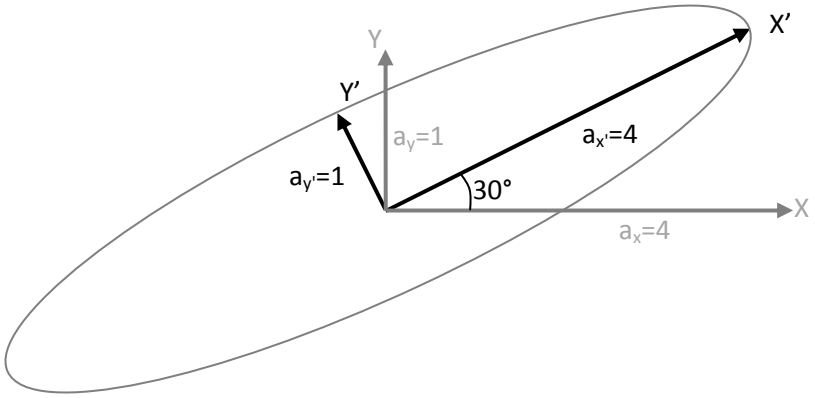
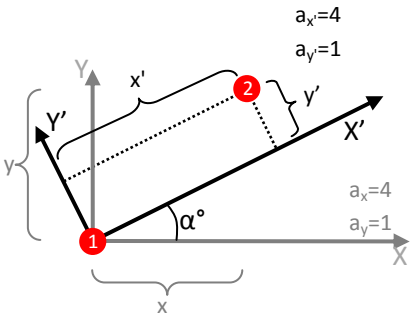


Figure 1.7: Left: all four data locations (circles) are given equal weighting when estimating at the unsampled location (?). Middle: because of the N-S anisotropy the data along the direction of anisotropy receive more weight and have a smaller anisotropic distance to the estimation location. Right: E-W anisotropy.





Case 1: No anisotropy:

$$d_{12} = \sqrt{x^2 + y^2} = \sqrt{8}$$

Case 2: 4:1 Anisotropy in the XY directions

$$d_{12} = \sqrt{\left(\frac{x}{a_x}\right)^2 + \left(\frac{y}{a_y}\right)^2} = \sqrt{\left(\frac{x}{4}\right)^2 + \left(\frac{y}{1}\right)^2}$$

Case 3: 4:1 anisotropy in the X'Y' rotated by α

$$d_{12} = \sqrt{\left(\frac{x'}{a_{x'}}\right)^2 + \left(\frac{y'}{a_{y'}}\right)^2} = \sqrt{\left(\frac{x'}{4}\right)^2 + \left(\frac{y'}{1}\right)^2}$$

Figure 1.8: Above: Anisotropy represented as an ellipse. In the original coordinate system anisotropy is 4:1 (a_x:a_y) and when rotated 30˚ it is 4:1($a_{x'}$:$a_{y'}$). Below: Calculating the distance between points 1 and 2 without anisotropy (Case 1) with anisotropy in XY (Case 2) and anisotropy in the 30˚ direction (Case 3).

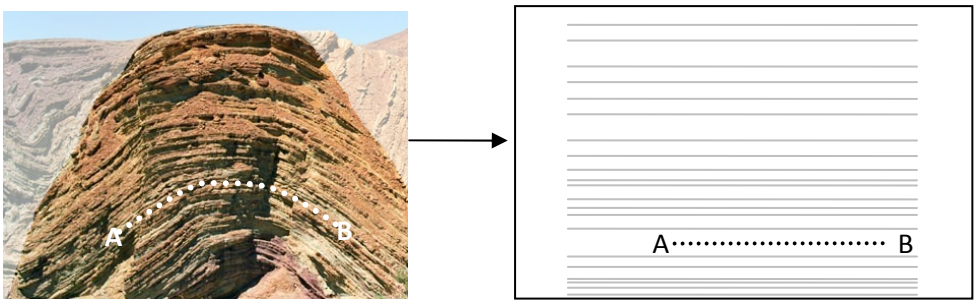


Figure 1.9: Left: An anticline with LVA (www.geology.about.com), scale not available. Points A and B are related along a nonlinear path because of the LVA. Right: Unfolding the anticline highlights the natural path between points A and B.

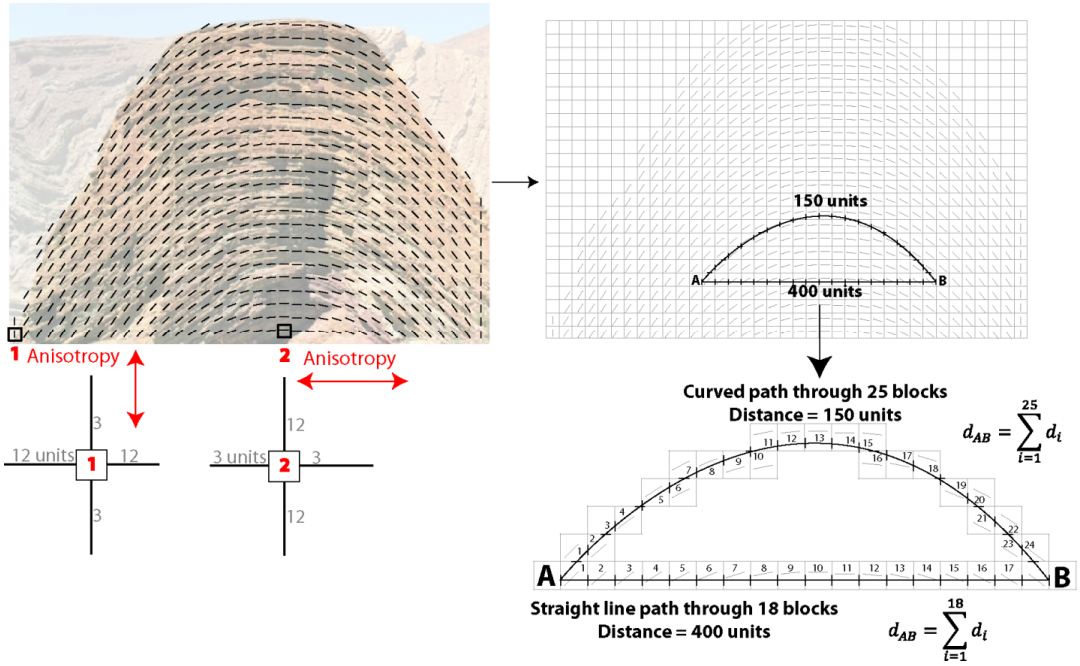


Figure 1.10: Left: An anticline with a LVA grid. Locations **1** and **2** are highlighted to show how the anisotropic distance in each cell is different and depends on the LVA field. At point 1 a horizontal path has an anisotropic distance of 3 units whereas a vertical path has an anisotropic distance of 12 units. Right: two potential paths between points A and B, each path is the summation of the distance through each cell. The shortest distance between points A and B is 150 units.

Often, it is implicitly assumed that the straight line path between points is the path along which two points are related. In the case of a single direction of continuity the straight line path corresponds to the shortest path between points. Consider a simple folded sedimentary deposit (Figure 1.9). Unfolding the deposit into its pre-fold state reveals the path along which points A and B are related. In the folded space, this path is not linear but follows the locally varying directions of continuity.

Consider a grid overlain on the anticline in Figure 1.9 with LVA shown in Figure 1.10. The anisotropy is now different in each cell of the model. Just as the *anisotropic distance* is different for a different anisotropy direction (Figure 1.7), each cell of the model (Figure 1.10) has a different direction of continuity and the anisotropic distance

then depends on the local direction of continuity. The distance between two points is a sum of the distances within each cell. For illustrative purposes consider the horizontal path between A and B. The *anisotropic distance* can be calculated in each of the 18 cells this path traverses and the distance from point A to B can be calculated as the sum. Considering the more realistic curved path between A and B intersects a total of 25 cells. Again the sum of the 25 individual distances represents the anisotropic distance between points A and B. Depending on the anisotropy field the curved path from A to B can be shorter than the straight-line path. This is the central concept used in this thesis to incorporate LVA:

Axiom 1.1: Two points in space are related by the path with the minimum *anisotropic distance*.

It is critical to note that this discussion is focused on the anisotropic distance (as calculated in Appendix A) **not** the Euclidian distance. In the case of constant anisotropy the shortest anisotropic distance is always found using the straight-line path; however this is not the case when considering LVA (Figure 1.10).

1.3 Problem Statement

The discussion thus far has broadly introduced geostatistics as a framework for calculating resource volumes and properties. The importance of considering anisotropy in these calculations has been motivated. The specific problem addressed is the incorporation of LVA into a numerical modeling framework. This dissertation is limited to the situation when sufficient *quantitative* or *qualitative* knowledge is available to infer the locally varying geological features. In such cases these features should be incorporated into modeling. Axiom 1.1 was introduced and provides the basis for the methodology used to incorporate LVA, specifically:

Thesis Statement

The use of nonlinear paths between points, generated from reliable locally varying anisotropy, improves the calculation of resources in numerical modeling.

1.4 Dissertation Contributions

The main contribution is a methodology to incorporate LVA into geological modeling in an estimation or simulation framework. The proposed method can consider complicated geological features in the form of a vector field of anisotropy (magnitude and direction of anisotropy at every location, as discussed in Chapter 2).

The second contribution is a set of techniques for LVA field inference. Often only qualitative information is known about the anisotropy field with perhaps some limited quantitative information. Generation of the anisotropy field is a crucial step and is the focus of Chapter 3.

A number of practical considerations arise when applying the proposed methodology, including: input parameter inference; uncertainty in the numerical modeling assumptions; and uncertainty in the anisotropy field. Where possible, theoretical solutions to these problems are derived, otherwise practical implementation results and guidelines are presented. These practical considerations are addressed through two case studies presented in Chapter 6.

Many of the algorithms used to incorporate LVA are computationally demanding. A major contribution of this work is the incorporation of efficient algorithms to minimize the run time required for the estimation or simulation of practically sized geomodels.

The remainder of this chapter is devoted to (1) a background discussion of traditional geostatistical and mathematical concepts that are used throughout the dissertation and (2) an overview of previous research that has attempted to integrate LVA into modeling.

1.5 Background

The specific topics summarized include: (1) the random variable formalism (2) stationarity (3) estimation techniques and (4) simulation. The *random variable* formulism discusses the basic notation and definitions required to apply the statistical tools used in geostatistics. *Stationarity* is fundamental to this work as the proposed methodology deals with deposits that display LVA. The final sections on *estimation* and *simulation* briefly introduce the techniques of inverse distance, kriging and sequential Gaussian simulation (SGS) as these are modified to incorporate LVA. These techniques are not fully derived here, interested readers are referred to any introductory geostatistics text for further detail (Goovaerts 1997; Chiles and Delfiner 1999; Deutsch 2002; Wackernagel 2003 to name a few).

1.5.1 Random Variable Formulism

The theory of geostatistics is based on random variables. Random variables are variables where the possible values the variable can be assigned are given a specific probability. A simple example of a discrete random variable would be the potential outcome of rolling a fair die. The variable is denoted by a capital letter (*Z*) and a specific outcome of that variable is denoted by a lower case (*z*). Each outcome of the variable is assigned a specific probability, *p(z)*:

$$Z = RV\ representing\ the\ outcome\ of\ rolling\ a\ fair\ die$$

$$\{Z \in [1,2,3,4,5,6]\} \qquad z = outcome\ of\ Z$$

$$p(z) = probabiliy\ of\ rolling\ a\ specific\ value$$

$$p(z) = \begin{cases} \frac{1}{6}, & if\ z = 1,2,3,4,5\ or\ 6 \\ 0, & otherwise \end{cases}$$

Random variables can also represent continuous values, such as the mineral grade in a gold deposit. Continuous random variables can take any value within a specified range. A probability density function is used to represent the probability of all possible outcomes of the continuous random variable, rather than assigning discrete probabilities as with rolling a die. Consider the following probability density function, f(x), for a random variable, X, that can vary between 2 and 12:

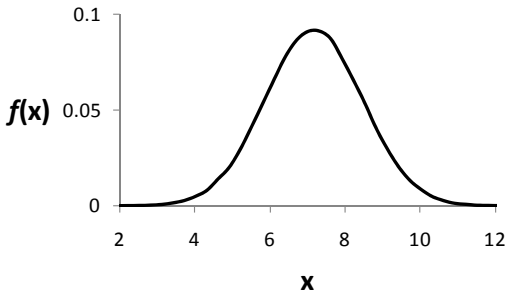


Figure 1.11: Probability density function.

All possible values of X have an associated probability density in the PDF (Figure 1.11). This random variable formulism allows for the usage of mathematical and statistical tools that can be used to estimate probability distribution functions for variables of interest at unsampled locations (i.e. the distributions discussed in Figure 1.1 and reprinted here as Figure 1.12).

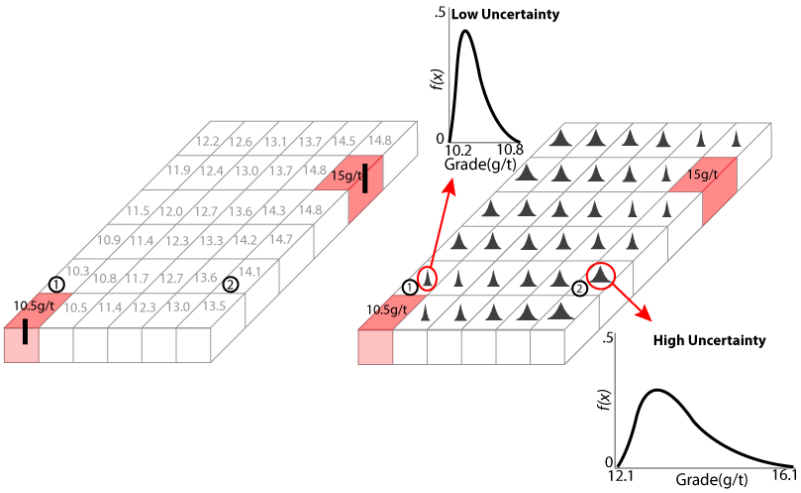


Figure 1.12: Reprint of Figure 1.1. A pdf is estimated at each unsampled location.

A random function (RF) is a set of RVs. The distribution at each location in Figure 1.12 represents a separate random variable but taken together they form a random function. More formally, consider a set of random variables, Z, representing the distributions in Figure 1.12 at each spatial location (**u**); the random function would be denoted:

$$\{Z(\mathbf{u}), \mathbf{u} \in \mathbf{A}\} \qquad \textbf{1.2}$$

where *A* is the modeling domain in Figure 1.12.

1.5.2 Stationarity

Stationarity is the decision of how to pool data for analysis. Assuming the modeling domain is first order stationary implies that the mean of a random function, Z(**u**), within a domain is constant at all locations, **u** (Equation 1.3). This allows data to be pooled to calculate the global mean of a deposit. In cases where the mean varies across the deposit the assumption of first order stationarity no longer holds.

$$E\{Z(\mathbf{u})\} = m \quad \forall \mathbf{u} \in \mathbf{A} \tag{1.3}$$

There are techniques available for considering domains that are not first order stationary. Two common approaches are to (1) separate the data into distinct populations that have different means and model each population independently or (2) introduce a trend model (an exhaustive numerical model of the mean in the domain). The trend is subtracted from the data values and the residual is modeled. McLennan (2008) describes in detail methodologies for considering data that are not first order stationary.

Second order stationarity assumes that the covariance function of a random variable is invariant under translation. The covariance between any two points separated by a lag distance, **h**, is given in Equation 1.4. Second order stationarity assumes that this covariance function, *C*(**h**), is constant within the modeling domain. This assumption has received less attention in the literature than first order stationarity.

$$E\{Z(\mathbf{u}+\mathbf{h}) \cdot Z(\mathbf{u})\} - m^2 = C(\mathbf{h}) \quad \forall \mathbf{u} \in \mathbf{A} \tag{1.4}$$

where **h** is a lag vector of length $||\mathbf{h}||$ separating two locations in space. Consider the appropriateness of the assumption of stationarity for a synthetic data set modeled after an underlying anticline (Figure 1.13). The assumption of second order stationarity is questionable as there is no single direction of continuity that can be applied. In fact, considering no anisotropy generates estimates that best follow the geological continuity of the deposit.

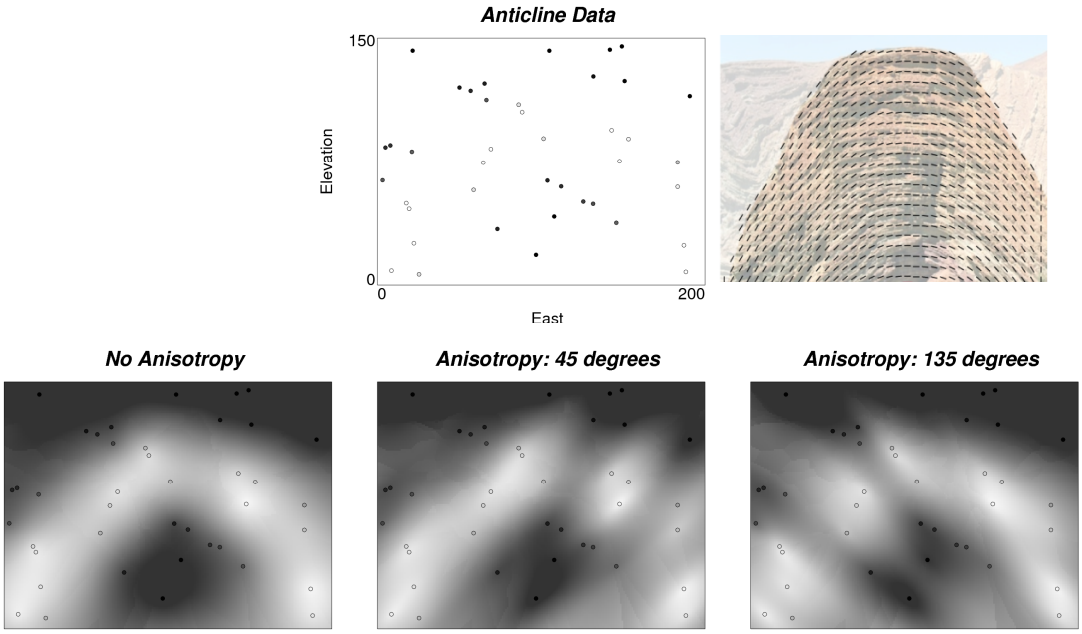


Figure 1.13: Above: Data locations for a non-stationary anticline deposit with an anisotropy field as shown in Figure 1.10 and reprinted above right (www.geology.about.com). Below: kriging with no anisotropy, 45˚ and 135˚. White indicates fabricated high grade areas for illustrative purposes, black indicates low grade.

1.5.3 Estimation Techniques

A common goal in numerical modeling is to generate a best estimate at an unsampled location given surrounding data. Two estimation techniques are modified to consider LVA, inverse distance and kriging. Both of these techniques are linear estimators as they apply a linear weight (Equation 1.5) to each of the known datum to estimate at unsampled locations (Figure 1.14). Inverse distance is considered because of its simplicity although kriging is often preferred as it minimizes the expected squared difference between the estimate and the truth.

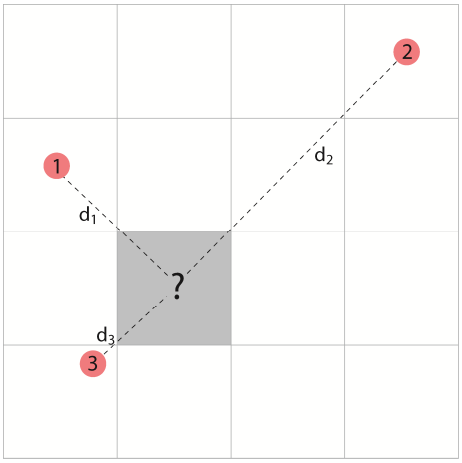


Figure 1.14: Estimating a value (Z*) at an unsampled grid location (?) using n=3 surrounding data (z_1,z_2,z_3). Weights for the three data are determined by inverse distance (Equation 1.6) and the estimate can be calculated as $z^* = z_1 \lambda_1 + z_2 \lambda_2 + z_3 \lambda_3$. Data that are closer to the estimation location (i.e. point 3) receive higher weight then data further away (i.e. point 2).

With inverse distance, the weight given to each datum is calculated as the reciprocal of the distance to the estimation location raised to a power (ω) as in Equation 1.6.

$$Z^* = \sum_{\alpha=1}^{n} \lambda_\alpha \, z_\alpha \qquad \textbf{1.5}$$

$$\lambda_\alpha = \frac{\frac{1}{d_\alpha^\omega}}{\sum_{\alpha=1}^{n} \frac{1}{d_\alpha^\omega}} \qquad \textbf{1.6}$$

Inverse distance estimation is straightforward to implement and only requires the practitioner to determine a single parameter (ω); however, this simplicity also limits the flexibility of inverse distance. Kriging is often used in geostatistics because it considers simple anisotropies, redundancy between data and the covariance in geological data (Figure 1.3 middle).

To determine estimation weights (as in Equation 1.5) kriging incorporates a model of the covariance between locations. Weights are generated by minimizing the error variance; the expected squared error (SE) between the unknown true value at a location, z(**u**), and the estimate at that location, z*(**u**) (Equation 1.7). Consider the well documented kriging equations (Equation 1.8).

$$SE = E\{[z(\mathbf{u}) - z^*(\mathbf{u})]^2\} \qquad \textbf{1.7}$$

$$\sum_{\beta=1}^{n} \lambda_\beta \mathrm{C}(\mathbf{u_\alpha}, \mathbf{u_\beta}) = \mathrm{C}(\mathbf{u_?}, \mathbf{u_\alpha}) \quad \alpha = 1, \dots, n \qquad \textbf{1.8}$$

where $\mathbf{u_\alpha}$ (α=1,2,…,n) are the data locations, $\mathbf{u_?}$ is the location to be estimated, C($\mathbf{u_\alpha}$,$\mathbf{u_\beta}$) is the covariance between data α and β, and C($\mathbf{u_?}$,$\mathbf{u_\alpha}$) is the covariance between the unsampled location and the data (Appendix A provides a numerical example of calculating this covariance). Solving this set of *n* equations generates the desired weights to apply in Equation 1.5. Derivation of these equations and further discussions on kriging can be found in many introductory geostatistics texts (Journel and Huijbregts 1978; Isaaks and Srivastava 1989; Cressie 1993; Deutsch 2002).

1.5.4 Simulation

Estimation produces numerical models that are smooth as the goal of estimation is to generate the best possible prediction at unsampled locations; such models vary smoothly from one drill hole to another. Depending on the end use of the numerical model, this smoothing may be acceptable but for many applications it is undesirable (Journel and Kyriakidis 2004). More formally, the smoothness of an estimate can be measured by its variance. Estimates have a variance that is less than the variance of the input distribution of data (Figure 1.15). The smoothing effect of kriging is well documented (Journel and Huijbregts 1978; Deutsch 2002) and the potential bias involved in using kriged maps in mine design or flow simulation necessities alternative techniques that accurately reproduce the input histogram.

Simulation is often employed to overcome the smoothness of estimators and to generate multiple possible *realities* of the variable of interest. In the framework of simulation, multiple realizations are simulated and represent equiprobable states of the variable of interest, as opposed to estimation where there is a single *best* estimate at all locations. Consider the estimated map using kriging (Figure 1.16); multiple realizations could be generated to evaluate the uncertainty in tonnage.

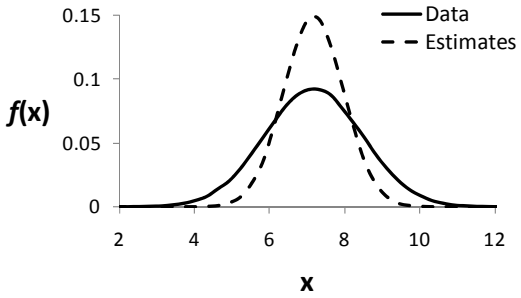


Figure 1.15: Probability density function for data (solid line) and potential estimates (dashed line). Because estimates are inherently smooth, they have less variability than the original data.

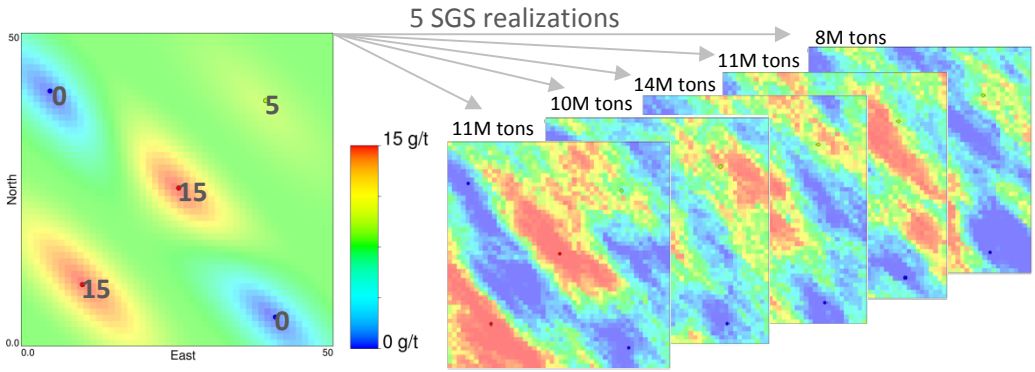


Figure 1.16: Left: A kriging map of estimates. Right: Five realizations generated with SGS. The realizations contain more realistic variability of the modeled variable. Each realization represents a potential reality of the grade and can be used to generate reserves/resources, oil volume, as input into flow simulation, etc. In this example the resource volume is shown (11, 10, 14, 11 or 8M tons are potential ore volumes).

Many geostatistical simulation algorithms use kriging. A particularly popular algorithm is sequential Gaussian simulation (Isaaks 1991). Recall the kriging estimator:

$$Z^* = \sum_{\alpha=1}^{n} \lambda_\alpha \, z_\alpha \qquad \textbf{1.9}$$

$$\sum_{\beta=1}^{n} \lambda_\beta \mathrm{C}(\mathbf{u_\alpha}, \mathbf{u_\beta}) = \mathrm{C}(\mathbf{u}_?, \mathbf{u_\alpha}) \quad \alpha = 1, \ldots, n \qquad \textbf{1.10}$$

The variance of the kriging estimate, *Z**, can also be derived (Wackernagel 2003) and is less than the variance of the input data, C(0), by a predictable amount:

$$Var\{Z^*(\mathbf{u})\} = C(0) - \sigma_{sk}^2(\mathbf{u}) \qquad \textbf{1.11}$$

where the kriging variance, $\sigma_{sk}^2(\mathbf{u})$, is defined as:

$$\sigma_{sk}^2(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n} \lambda_\alpha \mathrm{C}(\mathbf{u}_?, \mathbf{u_\alpha}) \qquad \textbf{1.12}$$

With SGS the smoothing effect of kriging (reduced variance) is overcome by adding a random residual directly into the kriging estimate, Equation 1.5 is modified to become:

$$z^*(\mathbf{u}) = R(\mathbf{u}) + \sum_{\alpha=1}^{n} \lambda_\alpha z_\alpha \qquad \textbf{1.13}$$

where the weights, λ_α, are generated from the same kriging equations as in Equation 1.8. The random component, $R(\mathbf{u})$, is added to account for the missing variance. In this way the resulting numerical models have the same variance as the input data. Moreover, by varying the random numbers used, multiple realizations with the same statistical properties can be generated, as in Figure 1.16.

A more detailed derivation and explanation of simulation can be found in (Goovaerts 1997; Deutsch 2002; Wackernagel 2003; Journel and Kyriakidis 2004). For the purposes of this dissertation only a basic understanding of estimation and simulation is required. Equations 1.5 and 1.8 provide the framework for kriging and Equations 1.11, 1.12 and 1.13 provide the framework for simulation. Equation 1.8 (reprinted here as Equation 1.14) is critical to the methodology presented in this dissertation because the covariance terms will be modified to incorporate the nonlinear path between points (Section 1.2).

$$\sum_{\beta=1}^{n} \lambda_\beta \mathrm{C}\left(\mathbf{u_\alpha}, \mathbf{u_\beta}\right) = \mathrm{C}(\mathbf{u}_?, \mathbf{u_\alpha}) \quad \alpha = 1, \dots, n \qquad \textbf{1.14}$$

1.6 Related Research

There are a number of fields that have developed techniques to incorporate locally varying features into numerical modeling. This section reviews the relevant works from the following areas:

1. Traditional Geostatistics (kriging)
2. Multiple Point Geostatistics
3. Unfolding/Stratigraphic transformations
4. Environmental Modeling
5. Generating LVA Fields

1.6.1 Related Research: Traditional Geostatistics (kriging)

Much of the past work integrating locally varying directions into geostatistical modeling has utilized kriging with a local search (Deutsch and Lewis 1992; Xu 1996; Sullivan *et al.* 2007). To determine how much weight to give each drill hole when estimating at unsampled locations, the anisotropy direction at the location being estimated is applied to its local neighborhood.

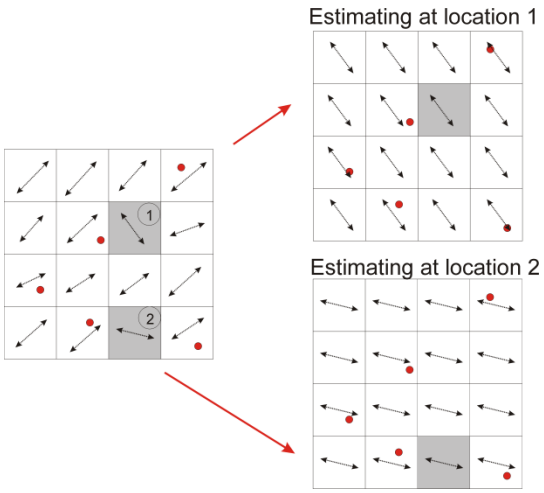


Figure 1.17: When estimating the gray locations with the surrounding data (red points) the anisotropy at the estimation location is applied everywhere and kriging is performed.

The direction of anisotropy is assumed constant for each kriging neighborhood. Consider estimating at the gray locations in Figure 1.17; the north-west direction would be applied to the local area when estimating block 1 and south-east for block 2. This is an exaggerated example where the anisotropy directions at the estimation locations are drastically different than the surrounding directions, but it highlights the limitations of accounting for LVA in this way. If the direction changes smoothly over the deposit and changes occur over large distances, considering a constant local anisotropy in this way may be reasonable. This idea has been extended to spectral methods (Borgman, Taheri and Hagan 1983) where the spectral functions are considered locally variable (Fuentes 2002a; Fuentes 2002b). However, it is still assumed that within an arbitrary region these spectral functions are stationary.

Stroet and Snepvangers (2005) have recently proposed a variant of kriging where the local anisotropy is automatically calculated from the available data. They incorporate LVA to accurately reproduce curvilinear structures (as in Figure 1.6) using an iterative image analysis technique. Unfortunately, this technique is currently limited to two dimensional data and requires sufficient data to directly infer the varying directions. If the data does not show the curvilinear structures, the method is not successful. Often in petroleum and mining applications, the data does not clearly reveal the curvilinear features due to large sample spacing but these features are known qualitatively, based on additional information. In this case, this iterative image analysis technique could not be applied.

Yao *et al*. (2007) have recently proposed a method to incorporate locally varying directions into geostatistical analysis. They suggest using a 1D spectral simulation to honor the directions and present 2D examples of their method with a proposal to extend to 3D.

Almendral et al. (2008) presented a methodology based on multidimensional scaling to incorporate LVA. They consider a nonlinear path between points as well as an embedding of the grid in three dimensions to determine the covariance between points for use in the kriging equations (Equation 1.8). This work was done concurrently and independently of this thesis.

1.6.2 Related Research: Multiple Point Geostatistics

Multiple Point Statistics (MPS) can be used as an alternative to traditional kriging based techniques. Such methods are used to generate rock type (facies) models that contain the complex nonlinear features seen in natural deposits and do not rely on kriging (Guardiano and Srivastava 1993; Journel 2004) However, MPS techniques are often limited to categorical variables, continuous variables such as ore grade, permeability, porosity, and many other important properties cannot be easily modeled. This is extremely limiting. A further limitation of the methodology is that the majority of the techniques developed to date require training images, which are difficult to obtain and select.

The relevance of MPS to the proposed research is that the goals of MPS and kriging with locally varying directions are similar; to generate models that have nonlinear and geologically realistic features. Consider the following 2D MPS model with very realistic features, including locally varying directions, that would be difficult to reproduce with kriging based techniques:

Figure 1.18: Fluvial model generated with MPS (Strebelle and Remy 2004). Model scale is typically varied to fit site specific conditions such as channel width and length.

As previously mentioned, MPS techniques rely on the availability of a training image. A training image is a rock type model that is exhaustively populated by the rock types of interest. The training image contains the complex geological features of a deposit such as locally varying directions, but is not conditioned to the available data. It is assumed that the training image has the same multiple point statistics as the deposit of interest. This is analogous to assuming that the training image contains the same geological features as the deposit of interest. All multiple point statistical algorithms extract

features from the training image and use them to generate a model conditioned to drill hole or well data. There has been little work into the development and use of training images (Boisvert *et al.* 2007) and to date training images are only widely available for limited deposit types (for vein type deposits see Boisvert *et al.* 2008 and for fluvial/deepwater reservoirs see Pyrcz *et al.* 2008).

The use of multiple point statistics is further limited by the computational effort required and the ability of the multiple point algorithm to reproduce the desired geological features in the models. Solutions to these problems are becoming available (Liu *et al.* 2004; Zhang *et al.* 2004; Arpat and Caers 2007). Even as these difficulties are overcome, multiple point statistical methods are still dependent on a training image that is representative of the deposit of interest, which may be difficult to obtain. MPS algorithms include: the use of neural networks (Caers 2001), Bayes law (Guardiano and Srivastava 1993; Strebelle 2002), simulated annealing (Deutsch 1992), updating conditional distributions with multiple point statistics (Ortiz and Deutsch 2004) and using a Gibbs sampler (Lyster and Deutsch 2006). The use of Bayes law in SNESIM (Strebelle 2002) is the most well developed and most used algorithm to date, but it still suffers from some drawbacks: difficulty obtaining a training image and difficulty modeling continuous variables. Recently, some attempts have been made to incorporate continuous variables (Wu *et al.* 2008) but training image selection and implementation decisions remain an issue. Incorporating locally varying directions into modeling as presented in this dissertation meets many of the same goals as MPS, without the need for a training image.

A recently developed technique that shows promise for incorporating complex, non-linear features into geological models of continuous variables is based on high order covariances, termed cumulants (Dimitrakopoulos 2010). Notwithstanding some initial computational issues and the difficulties surrounding the inference of statistics from training images (similar to MPS), in the near future this technique may accomplish many of the same goals as kriging with LVA if a training image is available for the area of interest.

1.6.3 Related Research: Unfolding and Stratigraphic Transformations

A technique used to model simple nonlinear features is the application of a coordinate transformation to a space where the features become linear. There are a number of transformations that can be applied depending the nature of the geological features. These transformations are most commonly applied in the case of a stratified, but folded or faulted, deposit. An assumption is made that the deposit was originally horizontal and layered but has since undergone folding and/or faulting that has resulted in predictable but nonlinear features; Mallet (2002, pages 379-442) provides an account of a number of stratigraphic transformations that can be applied depending on the nature of the folding and faulting. In the flattened space (Figure 1.19) features become linear and can be modeled with traditional geostatistics. Other transformations are also

available for straightening channel geometries (Deutsch 2002; Legleiter and Kyriakidis 2007).

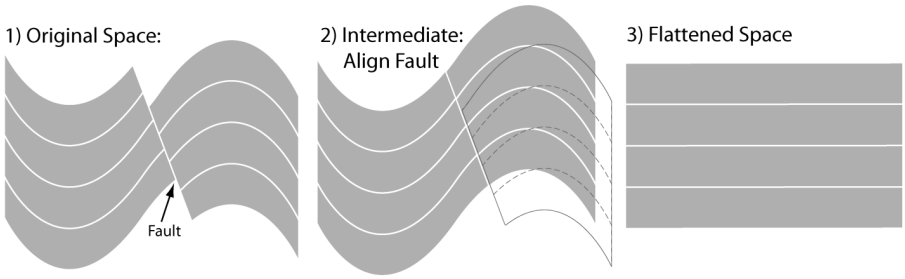


Figure 1.19: Left: Folded deposit. Middle: Aligning the fault. Right: Deposit is linear.

As Mallet (2002) highlights, applying stratigraphic transformations requires (1) complete knowledge of the geometry of the objects in their folded condition and (2) some information of the geometry of the objects in the unfolded condition. Without this detailed information, stratigraphic transformations are not possible, and even with such detailed information there is no single optimal transformation (Mallet 2002). Stratigraphic transformations are effective for highly predictable geology such as sedimentary deposits; however, they are not a solution to other situations with more complex LVA.

1.6.4 Related Research: Environmental Modeling

Environmental modeling must often consider local anisotropies because of the types of variables considered. Pollution spread, rain fall patterns, animal migration etc. exhibit very nonlinear features and display LVA in space and often in time. The pioneering work of Sampson and Guttorp (1992) which was expanded upon by multiple authors (Brown, Nhu and Zidek 1994; Guttorp and Sampson 1994; Meiring *et al.* 1997; Meiring *et al* 1998; Perrin and Meiring 1999; Damian *et al.* 2001; Schmidt and O'Hagan 2003), led the way for much of the work on LVA in an environmental context. They utilize multidimensional scaling (which will be applied in this dissertation in a similar manner) to incorporate LVA into modeling. Their approach is limited to smaller models as they use traditional MDS (Mardia, Kent and Biddy 1979) which is not practical for large grids. Moreover, they assume that there are repeated measurements at individual monitoring locations, which is rarely the case in mining or petroleum geostatistics.

Some authors have considered a kernel method or a weighted moving window average to incorporate locally varying features (Higdon *et al*. 1998; Higdon 1998; Nott and Dunsmuir 2002; ver Hoef *et al.* 2006). The parameters that define the kernel are varied locally and often non-stationary features can be reproduced. The main drawback of these kernel methods is similar to the drawback of using locally varying search parameters as the kernel is assumed locally stationary during estimation. Kernels cannot consider the case when the kernel parameters vary between data.

Considering environmental variables in an aquatic setting has lead a number of researches to consider a non-Euclidian distance metric between two locations, similar to the proposed shortest path between points (Section 1.2). Consider the path between two points along a stream, the distance traversed is not the straight line Euclidian distance; rather, it is the distance along the stream (stream distance). Incorporating this non-Euclidian distance into modeling is not straightforward as there is no guarantee that a valid covariance function can be found for the proposed distance metric, leading some authors to ignore the few indefinite systems of equations that occur (Gardner *et al.* 2003; Yuan 2004; Ganio *et al.* 2005; Boisvert *et al.* 2009). Little, Edwards and Porter (1997) use this stream distance metric to generate 2D numerical models in estuaries but do not discuss how such a metric is not positive definite. Rathbun (1998) present a similar method for modeling estuaries and simply check their metric for positive definiteness to ensure solvable kriging systems of equations; however, they do not propose a framework where positive definiteness is theoretically guaranteed.

The majority of this type of work in environmental applications focuses on using stream distances for pollution or fish migration studies. Large grids are not required for these studies; 1D grids oriented along the streams are often sufficient. In the case of geostatistical modeling for resource prediction, large 3D models containing millions of cells are required; such methods would be computationally infeasible.

Many authors have used these stream distances (Curriero 1996; Cressie and Majure 1997; Little, Edward and Porter 1997; Rathburn 1998; Kern and Higdon 1999; Loland and Host 2003; Krivoruchko and Gribov 2004; ver Hoef *et al.* 2006). A similar approach is considered in this dissertation with the *water* distance being analogous to the shortest path distance. The works listed above are important references for the techniques that will be used in this dissertation but they have a very different application as nonlinear distances are considered to incorporate physical boundaries and barriers (i.e. rivers) rather than geological anisotropies as presented in this dissertation. Guttorp and Sampson (1994) and Sampson *et al.* (2001) provide a review of the work done in this field prior to 2001.

1.6.5 Related Research: Generating LVA Fields

Determining the anisotropy field is not trivial with sparse data. Often the direction of a deposit can be determined at drill hole locations (Renard and Ruffo 1992); however, an exhaustive map of the locally varying directions is required. One possibility is to simulate the directions based on the sparse exploration data available (Xu 1996; Fuentes *et al.* 2005). Kriging or sequential Gaussian simulation could be used to generate the locally varying directions from the drill hole data (Figure 1.20).

A second method that could be used to determine the directions is to directly infer them from the data and a geological interpretation of the data. Deposits that are very

continuous tend to show directionality in the available sample data (Figure 1.21) and can be accurately interpreted by geologists to generate the locally varying directions.

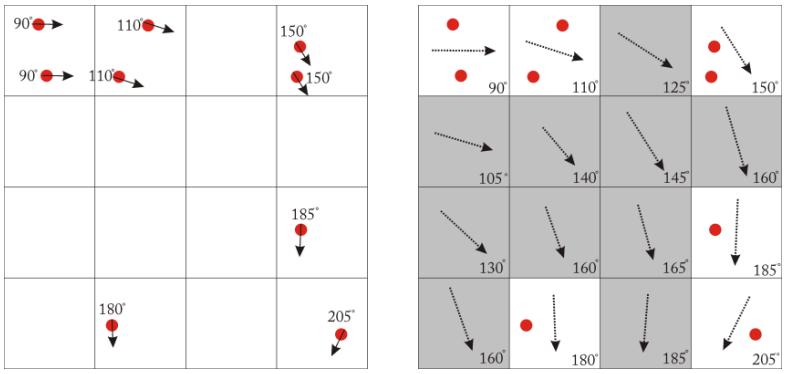


Figure 1.20: Left: Nine drill holes indicating the direction of strike of the deposit of interest. Right: Gray cells in the model contain no drill hole information, the direction must be estimated from the surrounding data.

If a rock type model of the deposit is available, the moment of inertia of the rock type can be used to determine the locally varying directions (Mohammadhassanpour 2007). This was the technique used to generate the locally varying directions of continuity shown in Figure 1.6 (also reprinted here as Figure 1.22). This technique is expanded upon in Chapter 3.

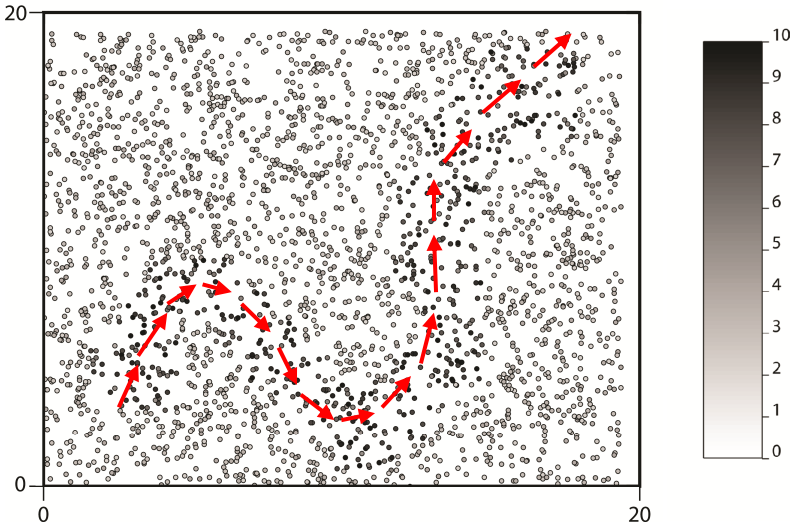


Figure 1.21: A situation where the locally varying directions are easily inferable from the available drill hole data. Circles represent drill hole data with a clearly visible curvilinear structure. Locally varying directions are indicated by arrows.

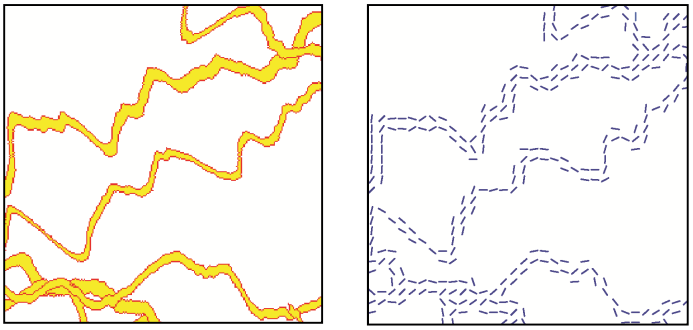


Figure 1.22: Locally varying directions for a fluvial deposit (reprint of Figure 1.6). Left: Conceptual fluvial deposit. Right: A map of the locally varying directions of anisotropy.

1.7 Outline

The focus of this dissertation is the presentation of a methodology to incorporate non-stationary anisotropy into geological modeling using the nonlinear path between points. To determine this path, multiple optimization techniques will be presented in Chapter 4; however, first Chapter 2 discusses the physical meaning of anisotropy and its incorporation in traditional modeling methodologies. This chapter provides a better understanding of anisotropy and emphasizes the need to incorporate non-stationary features and the difficulty in doing so. Chapter 3 then describes techniques to generate the anisotropy field from quantitative and/or qualitative data. The nature of the LVA field affects the resulting numerical models and is an important inference step in the proposed methodology.

Estimation and simulation using the nonlinear paths generated from Chapter 4 will be explained in Chapter 5, providing the main contribution of this work. Inverse distance estimation, kriging and sequential Gaussian simulation are modified to incorporate LVA in a mathematically sound manner. In Chapter 6 the application of the proposed methodology is demonstrated on two case studies based on net bitumen thickness data from an oil sands application as well as gold grade in a porphyry deposit. Practical issues are addressed, such as the extension from 2D to 3D, searching for nearby locations when distance is defined nonlinearly, generating the LVA field and model validation.

Chapter 2: On the Nature of Anisotropy

The notion of anisotropy is central to the research presented in this thesis. As such, the concept of anisotropy must be clearly defined. Section 2.1 quantifies anisotropy using the variogram. Section 2.2 further explores the mathematics used to incorporate anisotropy into geostatistical modeling. Section 2.3 treats anisotropy as a geological concept and defines the types and sources of anisotropy common in geological deposits. Finally, Section 2.4 discusses how the nature of anisotropy is dependent on the scale at which it is considered. While this chapter does not provide an original contribution, the nature of anisotropy is central to the proposed methodologies developed and requires clarification.

2.1 Definition of Anisotropy

Anisotropy is defined as "*having properties that differ according to the direction of measurement*" (Morris 1975). The term is used in many scientific fields with similar connotation; in the field of geostatistics anisotropy describes the directional dependence of the continuity of variables such as rock type, facies type, porosity, permeability, mineral grade, concentration, etc. Natural geological processes can spatially elongate these properties in preferential directions. In such cases the properties are said to be anisotropic as their continuity differs when examined in different directions.

2.1.1 Qualitative Assessment of Continuity

The definition of continuity is not as clear as the definition of anisotropy. Continuity can be qualitatively assessed by visual examination, for example, the coal deposit has a higher degree of continuity when compared to the gold vein deposit (Figure 2.1). The continuity is higher in the coal seam because points separated by tens of meters, in the horizontal direction, are similar while in the gold vein points separate by only a few centimeters appear unrelated.

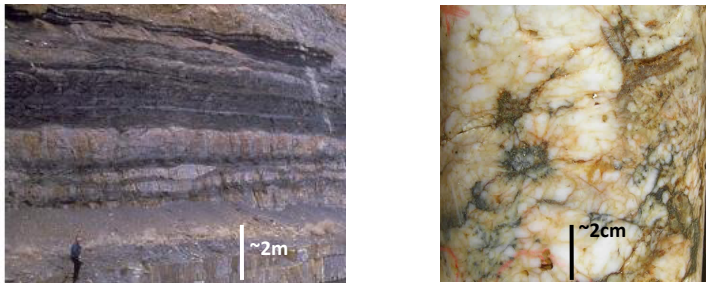


Figure 2.1: Left: A continuous coal deposit (www.scienceclarified.com). Right: Gold veins are typically discontinuous and erratic (www.jpgold.com).

The coal seam has a higher degree of continuity in the horizontal direction but the stratigraphic deposition of coal results in less vertical continuity. Thus, the coal grade is anisotropic because the degree of continuity in the coal seam depends on direction.

This anisotropic behavior is due to the geological processes that result in aerially continuous sheets of coal. Such visual assessments of continuity are important during initial data exploration; however, a more quantitative assessment of the degree of continuity is required if anisotropy is to be introduced into numerical modeling.

2.1.2 Quantitative Assessment of Continuity: The Variogram

Continuity can be assessed quantitatively by considering the separation distance at which locations become unrelated. The most common measure of continuity used in geostatistics is the covariance (Equation 2.1) between two locations separated by a lag vector, **h** (Figure 2.2); higher covariance between locations separated by **h** implies a higher degree of continuity. By definition, anisotropy exists when the covariance between locations separated by **h** depends not only on ||**h**|| but also on the orientation of **h**. This covariance measure can be estimated from the available samples by considering different lags, **h**, and calculated as the experimental covariance (Equation 2.2) between sample locations *approximately* separated by **h**. Normally, tolerance parameters are introduced to define points that are separated by **h** as it is unlikely that there are many pairs of data separated by exactly **h**. These tolerance parameters are explained in detail in Deutsch and Journel (1998).

$$C(\mathbf{h}) = E\{Z(\mathbf{u}) \cdot Z(\mathbf{u}+\mathbf{h})\} - [E\{Z(\mathbf{u})\}]^2 \quad \textbf{2.1}$$

$$\gamma(h) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \left(Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})\right)^2 \quad \textbf{2.2}$$

$\gamma(h)$ is technically a *semi*-variogram while $2\gamma(h)$ is the variogram. For convenience the term *semi* is dropped for the remainder of this dissertation with the understanding that $\gamma(h)$ refers to a *semi*-variogram.

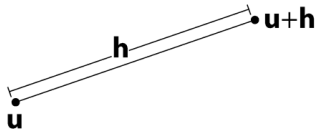


Figure 2.2: A lag vector, **h**, between locations **u** and **u+h.**

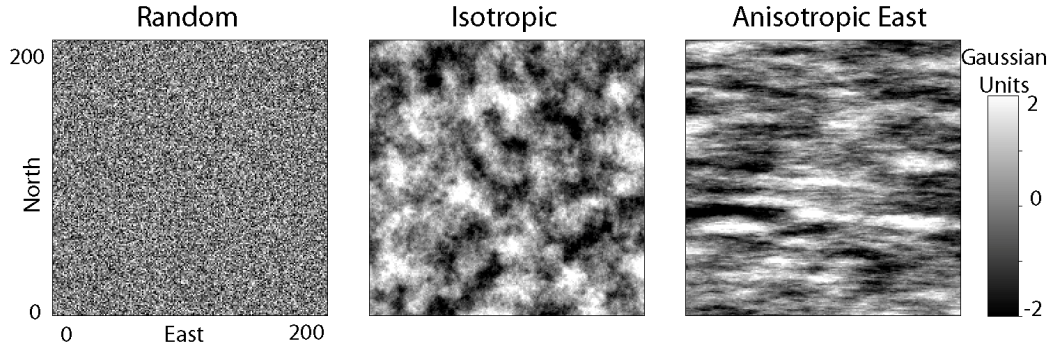
Consider an illustrative example of three fields with different degrees of continuity (Figure 2.3). Isotropy exists when the covariance between locations depends only on ||**h**|| (Figure 2.3 middle). The covariance between points is identical regardless of the direction considered and reduces to nearly 0 after a range of approximately 20 units. In contrast, the range for the random case is nominally zero at a very small lag distance. Because the isotropic case has a larger range, it is said to be more continuous than the random case. In this way, the continuity of a deposit is quantified.

The notion of covariance representing the continuity of a deposit can be extended to the anisotropic case where the degree of continuity depends on direction. The tables in Figure 2.3 are useful for identifying various directions of anisotropy. In the anisotropic

case, there is more East-West continuity than North-South because for all lag distances, ||**h**||, the covariance in the East-West direction is larger.

Examining the continuity in multiple directions is more conveniently done by graphing the variance in the form of a variogram. The variogram (Figure 2.4) is a representation of the variance between points separated by **h** and is related to the covariance by Equation 2.3. The benefit of examining the variogram is that the continuity in many different directions for anisotropic variables can be visualized simultaneously (Figure 2.5). Note that, by definition (Equation 2.3), when $\gamma(\mathbf{h})=C(\mathbf{0})$ the data are uncorrelated, C(**h**)=0.

$$C(\mathbf{h}) = C(0) - \gamma(\mathbf{h}) \qquad \textbf{2.3}$$



||h|| (unit)	Covariance (East)	Covariance (North)
0	1.00	1.00
2	**0.01**	**0.01**
4	0.01	0.02
6	0.01	0.02
8	0.02	0.02
10	0.02	0.01
15	0.02	0.01
20	0.02	0.02
25	0.00	0.02
30	0.01	0.01
35	0.02	0.01
40	0.02	0.02
45	0.01	0.01
50	0.01	0.02
55	0.01	0.02
60	0.01	0.00

Range>0 uncorrelated

||h|| (unit)	Covariance (East)	Covariance (North)
0	1.00	1.00
2	0.85	0.85
4	0.71	0.71
6	0.58	0.58
8	0.46	0.45
10	0.34	0.33
15	0.13	0.09
20	**0.04**	**-0.03**
25	0.01	-0.06
30	-0.03	-0.04
35	-0.08	0.01
40	-0.11	0.04
45	-0.08	0.04
50	-0.03	0.03
55	0.00	0.01
60	-0.01	-0.01

Range≥20 uncorrelated

||h|| (unit)	Covariance (East)	Covariance (North)
0	1.00	1.00
2	0.94	0.71
4	0.88	0.44
6	0.82	0.22
8	0.77	0.07
10	0.72	**0.01**
15	0.59	0.05
20	0.47	0.07
25	0.36	-0.04
30	0.26	-0.01
35	0.18	0.04
40	0.11	0.08
45	0.07	0.02
50	**0.05**	-0.04
55	0.03	0.05
60	0.01	0.01

Range≥10 uncorrelated (North)

Range≥50 uncorrelated (East)

Figure 2.3: Left: A random continuous variable. Middle: An isotropic variable, continuity depends only on ||**h**||. Right: An anisotropic variable, where continuity depends on the orientation and magnitude of **h**. Distances are unit distances.

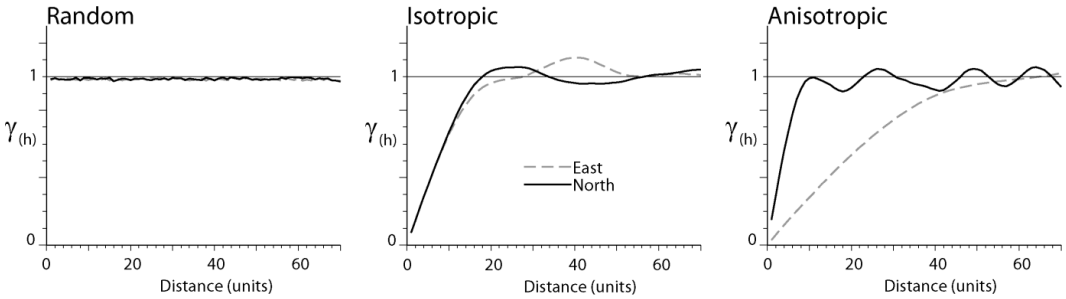


Figure 2.4: Variograms for the variables in Figure 2.3. C(0)=1.0.

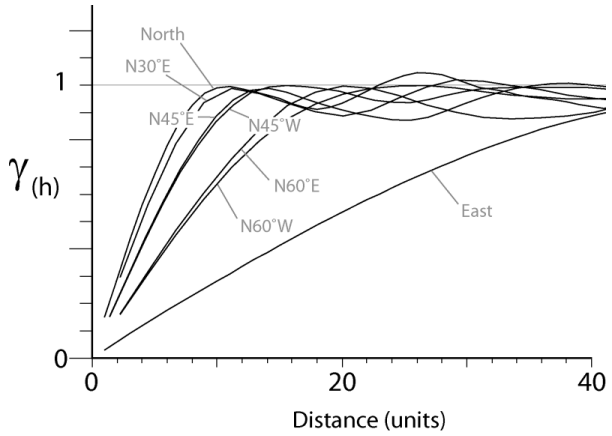


Figure 2.5: Variograms in multiple directions for the anisotropic map in Figure 2.3. Unit distances are considered.

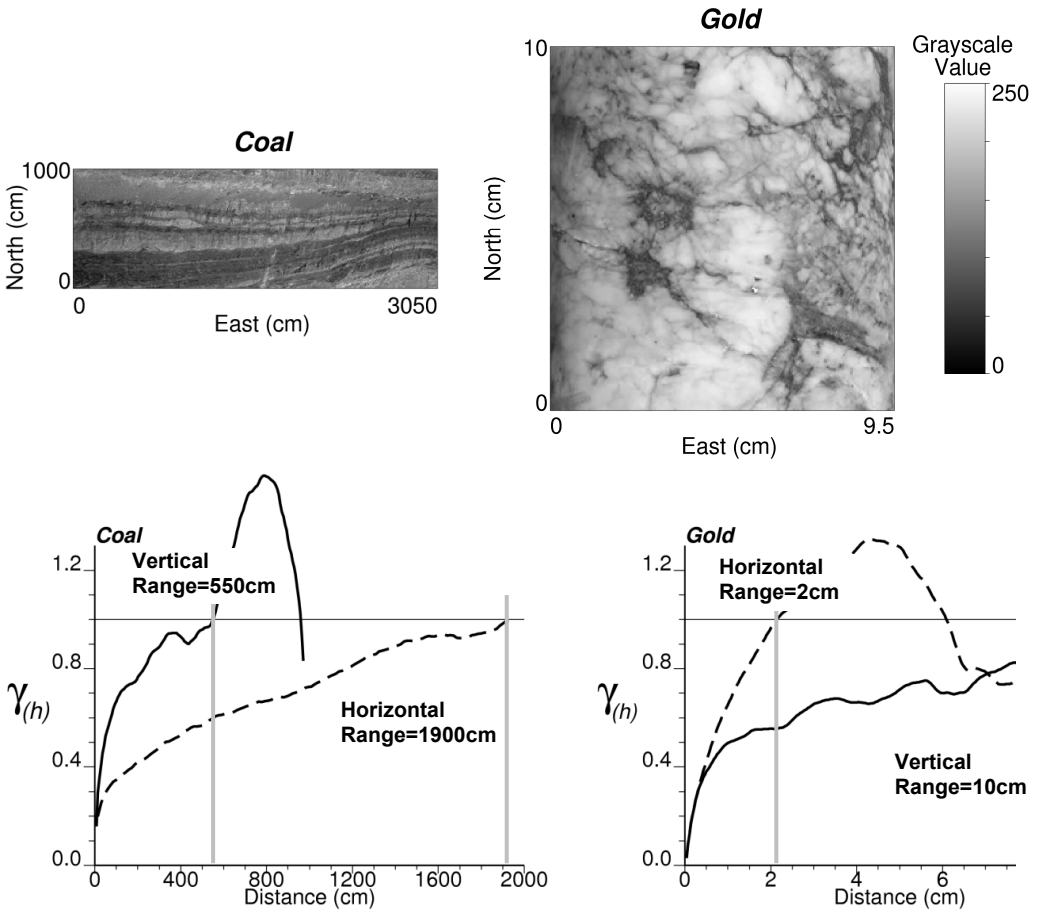


Figure 2.6: Variograms for the coal and gold deposit images. The data has been standardized by a normal score transform to allow for a comparison of the variograms of the two images.

Previously, it was claimed that coal (Figure 2.1) has a *higher degree of continuity* than gold. Using the variogram the degree of continuity can be quantified and compared. The images of coal and gold are converted to grayscale (Figure 2.6) and the grayscale value is assumed to be proportional to the grade of the deposits; associated variograms are provided (Figure 2.6). Coal is more continuous as points separated by up to 19m in

the horizontal direction are correlated, while correlation approaches 0.0 by 10cm in the gold deposit.

This example also highlights the concept of anisotropy. Both examples show that the covariance between locations depends on distance as well as orientation. In the coal example covariance is higher in the horizontal direction while it is higher in the vertical when considering the gold sample.

2.1.3 Incorporating Anisotropy in Classical Geostatistics

To model the anisotropy in a geological setting, two definitions are required: the direction of continuity; and the anisotropy ratio. Consider the variograms in Figure 2.5, the major direction of continuity is the direction in which there is the greatest continuity (East). It is assumed that the minor direction is orthogonal to the major direction (North) even if this does not coincide with the direction of minimum continuity. In three dimensions a third intermediate angle is required to fully define the direction of continuity. The magnitude of anisotropy is captured as an anisotropy ratio. Literature is inconsistent in defining the anisotropy ratio. If the anisotropy ratio is smaller than 1.0 it is assumed that Equation 2.4 holds. When the anisotropy ratio is reported to be larger than 1.0 the reciprocal of Equation 2.4 is assumed. A third form of the anisotropy ratio is a true ratio, such as 10:1, where the larger number represents the relative range of anisotropy in the major direction (a_{max}), and the smaller number represents the relative range in the minor direction (a_{min}).

$$anisotropy\ ratio = \frac{range\ in\ the\ minor\ direction}{range\ in\ the\ major\ direction} = \frac{a_{min}}{a_{max}} \quad \textbf{2.4}$$

The variogram provides a quantitative measure of the existence and degree of continuity for any data type: continuous; categorical; exhaustive; or sample. This is commonly referred to as the *experimental variogram*. The experimental variogram cannot be used when implementing kriging or SGS (Chapter 1) rather, a modeled variogram must be generated that results in a positive definite set of kriging equations. There are a number of mathematical functions that are known to produce positive definite kriging systems of equations (Christakos 1984; Cressie 1993) in 2D or 3D and can be used to model the experimental variogram (Table 2.1). The shape of these common variogram functions are shown in Figure 2.7. These functions can be combined in an additive manner to increase the flexibility of modeling (Equation 2.5).

$$\gamma(h) = \sum_{i=1}^{n} C_i \cdot \Gamma_i(h) \quad \textbf{2.5}$$

where C_i is the variance contribution assigned to the i^{th} function, $\Gamma_i(h)$. Note that $\sum C_i = C(0)$ and is denoted as the *sill* of the variogram.

Table 2.1: Common functions used to model variograms. These are shown on Figure 2.7 with the exception of the nugget effect which acts as a step function. *h* is the standardized distance discussed in Chapter 1 (Equation 1.1).

Function	Equation	Comments
Nugget Effect	$\Gamma(h) = \begin{cases} 0 & if\ h = 0 \\ 1 & if\ h > 0 \end{cases}$	Used to incorporate small scale variability.
Spherical	$\Gamma(h) = \begin{cases} 1.5h - 0.5(h)^3 & if\ h < 1 \\ 1 & if\ h \geq 1 \end{cases}$	Most commonly used structure.
Hole Effect	$\Gamma(h) = 1.0 - \cos(\pi h)$	Used to incorporate cyclic oscillations in variogram modeling.
Exponential	$\Gamma(h) = 1 - e^{-h}$	Similar to the spherical model but rises faster at the origin and asymptotically approaches its maximum value.
Gaussian	$\Gamma(h) = 1 - e^{-h^2}$	

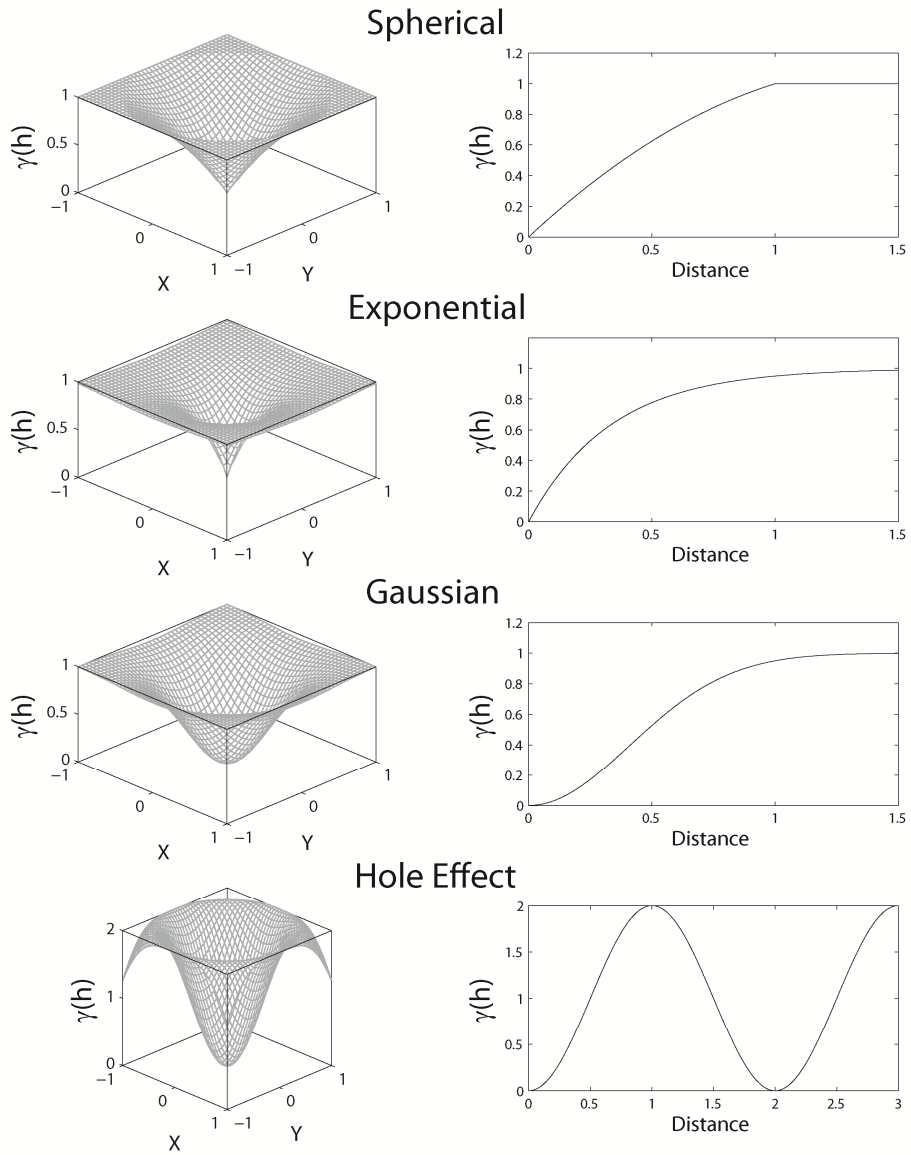
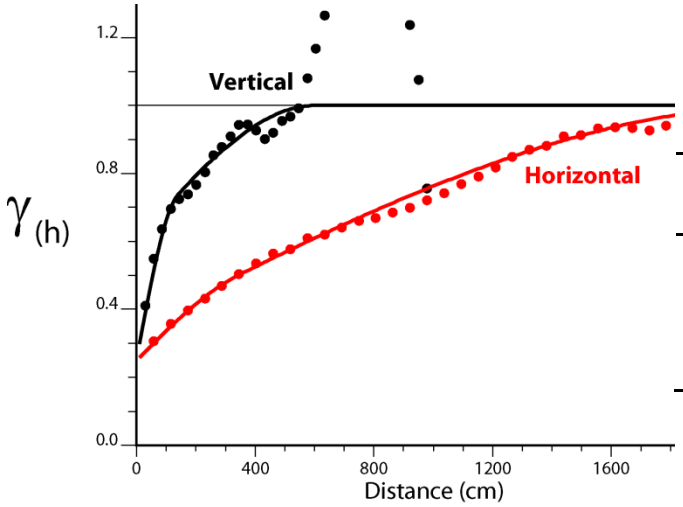


Figure 2.7: Common variograms used to model the spatial relationship between locations. Left: Variograms shown in a standardized 2D X-Y space. Right: Variograms shown as a function of the distance between locations. All variograms have a sill of 1.0 with an effective range of 1.0.

The variogram model for a given deposit is fit to the experimental data but may also be influenced by additional expert knowledge of the deposit (Gringarten and Deutsch 2001). The process for calculating an experimental variogram consists of searching the available sample data for samples separated by the desired lag, **h**, within some user supplied tolerances. The experimental variance (Equation 2.2) of these pairs provides the circular points on Figure 2.8. Often a number of directions are considered for the orientation of **h**, and the directions that display the maximum and minimum continuity are modeled. The appropriate functions, $\Gamma_i(h)$, contributions, C_i, and range parameters are iteratively selected to fit the variogram model, $\gamma(h)$, to the experimental variogram. For further guidance on variogram modeling, the interested reader is referred to Isaaks and Srivastava (1989) where a detailed example is provided.



C_i	Type of Model	Horizontal Range (a_{max})	Vertical Range (a_{min})
0.25	Nugget	N/A	N/A
0.10	Spherical	350cm	140cm
0.25	Spherical	2200cm	140cm
0.40	Spherical	2200cm	600cm

Figure 2.8: Experimental variogram (points) and modeled variogram (lines) for the vertical and horizontal directions in the coal image (Figure 2.6). Right: Parameters for the modeled variogram.

For computational simplicity it is useful to standardize the distance vector, **h**, to a unit-less scalar distance, *h*, that considers the modeled anisotropy in each of the principle directions (Equation 2.6). This is done by calculating an effective anisotropic distance between two locations (Chapter 1) as the decomposition of **h** into three directions (3D) or two directions (2D). The range parameters, a_{max}, a_{min} and a_{vert} are specified by the practitioner and fit to the available data.

$$h = \sqrt{\left(\frac{\Delta h_{max}}{a_{max}}\right)^2 + \left(\frac{\Delta h_{min}}{a_{min}}\right)^2 + \left(\frac{\Delta h_{vert}}{a_{vert}}\right)^2} \qquad \textbf{2.6}$$

2.1.4 Summary of the Definition of Anisotropy

The previous section introduced the concepts of anisotropy, isotropy and continuity. Continuity of a deposit is quantified by considering how correlated two locations are. If this correlation is directionally dependent, the property of interest is anisotropic and if the correlation does not depend on direction the variable is isotropic.

The variogram (Equations 2.2 and 2.5) was introduced as a graphical plot that can be used to visualize how the continuity of a variable changes with distance and direction defined by **h**. The range of anisotropy is most often considered to be the distance at which the variogram reaches a value of $C(0)$, or 95% of $C(0)$ when the variogram model reaches the sill asymptotically.

Normally, the variogram is modeled with a function that is known to produce positive definite kriging equations (Chapter 1) for a mathematically valid implementation of kriging. Thus, the experimental variogram must be fit with a function (Table 2.1). The anisotropy ratio is used to quantify how much larger the continuity is in a major direction compared with the minor direction.

2.2 Mathematics of Anisotropy

Anisotropy of a stationary deposit is defined in 2D by one angle (the major direction of continuity) and two range parameters. Anisotropy can be thought of as an ellipse (Figure 2.9) with the major direction of anisotropy corresponding to the orientation of the semi-major axis and the orthogonal minor direction of continuity corresponding to the semi-minor axis of the ellipse. The radius of the ellipse in these directions is given by the appropriate ranges (Figure 2.9). In 3D, three angles and three range parameters are required to fully define anisotropy.

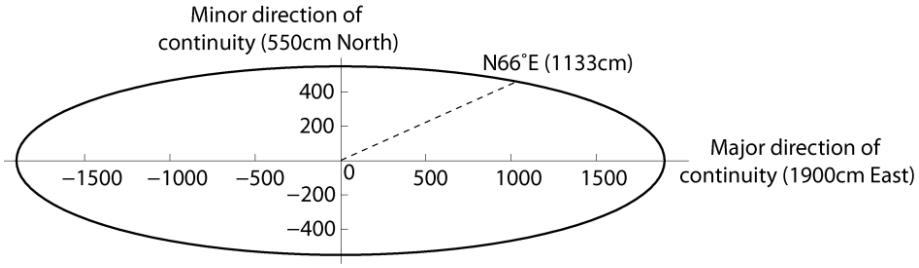


Figure 2.9: Anisotropy visualized as an ellipse. This example shows a major direction of continuity in the East (a_{max} =1900cm) and a minor direction of continuity in the North (a_{min}=550cm). The range in any direction is calculated from the ellipse, as shown for N66˚E.

2.2.1 Data Transformations

The transformation from a coordinate system in which the variable of interest displays anisotropy to an orthogonal coordinate system where the variable is isotropic is useful for (1) computational purposes and (2) to capture anisotropy in the form of a single matrix (Equation 2.9).

Transformation to an isotropic space for 2D data is much simplified over 3D data. First the direction of anisotropy must be determined by either a visual assessment of the available data or an examination of the variogram in a number of different directions to select the direction with the largest range. Once the major direction of anisotropy is found, the X and Y axes are rotated to align with the major direction of anisotropy and scaled to obtain an isotropic coordinate system in which the anisotropy ratio is

accounted for by the transformations (Figure 2.10). The data in the transformed space can be fit with an isotropic function (Table 2.1) that only depends on the distance, *h*.

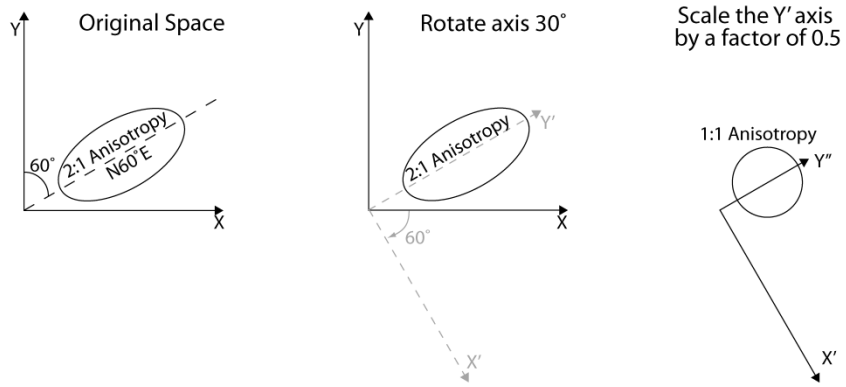


Figure 2.10: Left: Consider 2:1 anisotropy in N60˚E. To obtain isotropy, rotate the X axis to align with the direction of continuity (middle) and scale the rotated Y' axis by a factor of 0.5.

This transformation is obtained by multiplying the data in original coordinates (x,y) by the rotation matrix in Equation 2.7 to obtain the rotated coordinates (x',y'). These matrices are obtained using simple geometry (Eriksson and Siska 2000).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1/a_{max} & 0 \\ 0 & 1/a_{min} \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \textbf{2.7}$$

where α is a clockwise rotation of the azimuth measured from North (Y-axis in Figure 2.10). The following is obtained for the example in Figure 2.10:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1/a_{max} & 0 \\ 0 & 1/a_{min} \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} =$$

$$= \begin{bmatrix} 1/2 & 0 \\ 0 & 1/1 \end{bmatrix} \begin{bmatrix} \cos 60 & -\sin 60 \\ \sin 60 & \cos 60 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \textbf{2.8}$$

These transformations are only used in the calculation of the covariance between locations required for implementing kriging or SGS. The transformed space is convenient as the scalar distance, *h*, can be used to calculate distance and ultimately the covariance between locations using the variogram model. The extension of these transformations to three dimensions is accomplished with three rotations and three range parameters (Figure 2.11). The resulting orthogonal coordinate system (see orthogonality proof in Newham 2005, page 11) is isotropic, thus the covariance is a function of only the scalar distance, *h*. This transformation allows for the straightforward calculation of the anisotropic distance between any two locations in the presence of anisotropy.

Again, these rotations can be obtained with the following rotation matrix:

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \frac{1}{a_{max}} & 0 & 0 \\ 0 & \frac{1}{a_{min}} & 0 \\ 0 & 0 & \frac{1}{a_{vert}} \end{bmatrix} \begin{bmatrix} \cos\alpha \cdot \cos\varphi - \sin\alpha \cdot \sin\beta \cdot \sin\varphi & -\sin\alpha \cdot \cos\varphi - \cos\alpha \cdot \sin\beta \cdot \sin\varphi & \cos\beta \cdot \sin\varphi \\ \sin\alpha \cdot \cos\beta & \cos\alpha \cdot \cos\beta & \sin\beta \\ -\cos\alpha \cdot \sin\varphi - \sin\alpha \cdot \sin\beta \cdot \cos\varphi & \sin\alpha \cdot \sin\varphi - \cos\alpha \cdot \sin\beta \cdot \cos\varphi & \cos\beta \cdot \cos\varphi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \textbf{2.9}$$

where α is a strike rotation about the Z axis, β is a dip rotation about the X axis, and φ is a plunge rotation about the Y axis, with the magnitude of anisotropy defined by a_{max} in the X'' direction, a_{min} in the Y'' direction and a_{vert} in the Z'' direction. The rotation matrix in Equation 2.9 is obtained by considering three successive 2D rotations (Equation 2.7) in the X-Y, Z-Y' and X'-Z' planes (Figure 2.11).

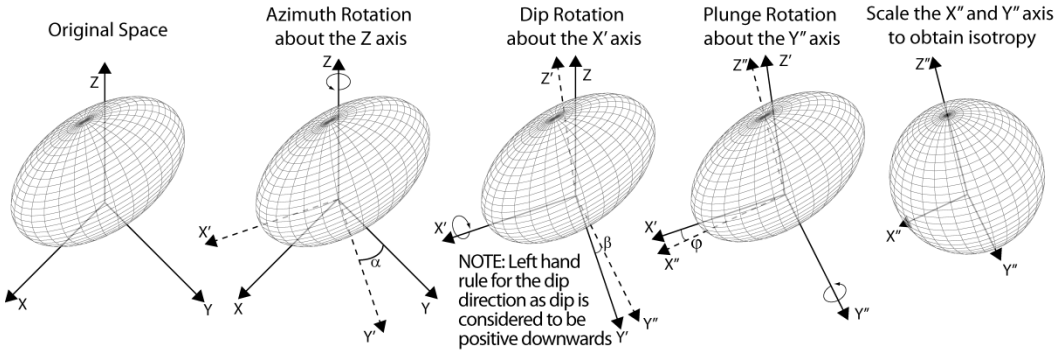


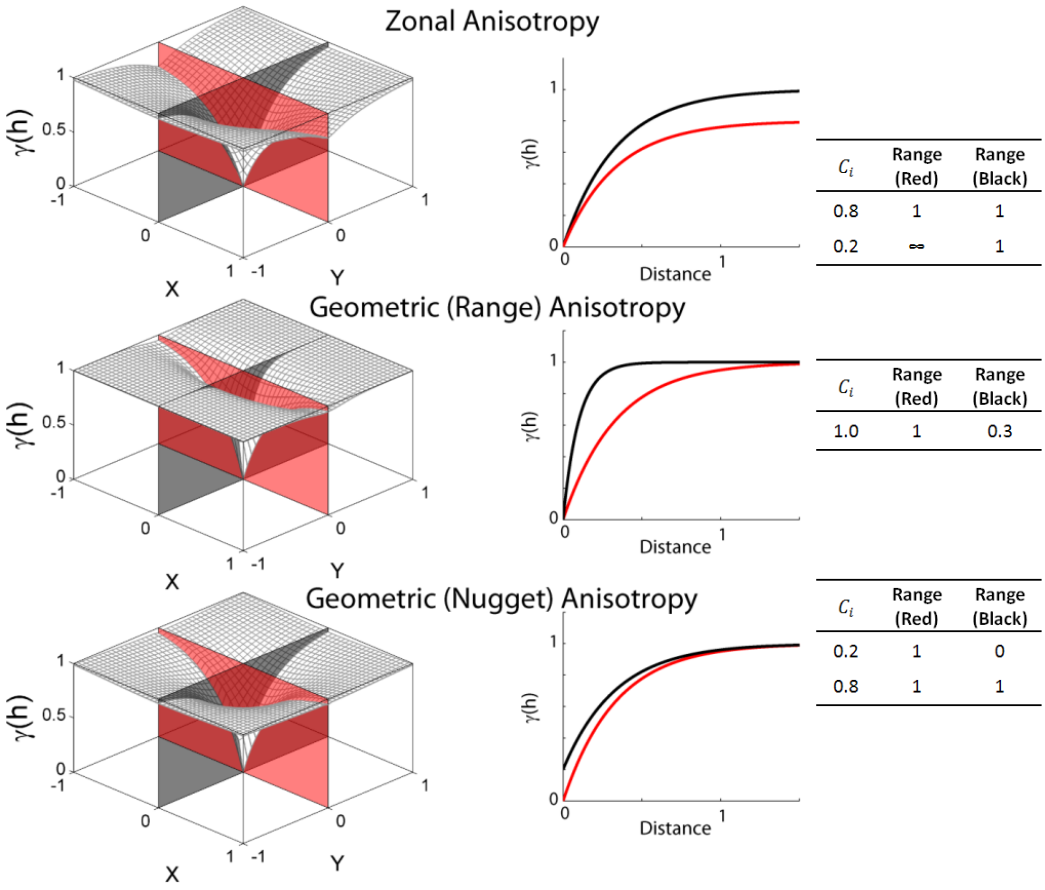
Figure 2.11: Left: Consider a_{max}=a_{min}=2 while a_{vert}=1. To obtain isotropy, rotate three times, first about the Z-axis, then about the X' axis and finally about the Y'' axis. Scaling the X'' and Y'' axes results in isotropy.

2.2.2 Types of Anisotropy

In Section 2.2.1 rotation and scaling of the coordinate system was presented to facilitate the calculation of the scalar distance between locations, *h*, that is used in covariance calculations (Table 2.1). The variogram, along with the range parameters (a_{max}, a_{min} and a_{vert}) and anisotropy directions (α, β and φ) can be used to visualize different types of anisotropy that are common in geological formations. This discussion is limited to the 2D case as the variogram function is a 3D function when 2D data are considered. The extension to 3D is conceptually trivial but difficult to visualize.

There are a number of different schemes used to categorize anisotropy, the most common being the division between zonal and geometric anisotropy (Chiles and Delfiner 1999; Deutsch 2002; Wackernagel 2003). Zonal anisotropy, also referred to as sill anisotropy (Zimmerman 1993; Eriksson and Siska 2000), occurs when the sill reaches different limiting values in different directions (Figure 2.12). Zonal anisotropy is modeled using two nested functions, one with a long range in a given direction (Figure 2.12). Geometric anisotropy is defined as any anisotropy that is not sill anisotropy. Zimmerman (1993) further separates zonal anisotropy into nugget and range anisotropy (Figure 2.12); however, it is more common to consider any type of anisotropy that is not zonal to be geometric. Nugget and range anisotropy effectively alter the shape or

geometry of the variogram before the sill is reached. The more common and less specific categories of *zonal* and *geometric* anisotropy are used here. Examples of each type of anisotropy and the geological conditions that cause them are found in Section 2.3.



C_i	Range (Red)	Range (Black)
0.8	1	1
0.2	∞	1

C_i	Range (Red)	Range (Black)
1.0	1	0.3

C_i	Range (Red)	Range (Black)
0.2	1	0
0.8	1	1

Figure 2.12: Types of anisotropy common in geological phenomena. Slices X=0 (black) and Y=0 (red) are shown. Equations for each variogram model are shown to the right, all structures are exponential models. All variograms have a sill of 1.0 and are provided for standardized distances.

2.3 Processes That Cause Anisotropy

The cause of the physical manifestation of anisotropy in a geological context requires elaboration. A number of geological processes increase or decrease continuity in a particular direction. Some of the more common processes include: fracturing; vein formation; folding; thermal gradients; fluvial and estuarine systems. The geological processes that cause LVA in these situations are described and an example deposit from the literature displaying LVA is provided.

2.3.1 Folding

Folding is the ductile deformation of a mass of rock and is most clearly illustrated when applied to sedimentary deposits. The planar structures in sedimentary deposits are

clearly visible after folding (Figure 2.13). The type and magnitude of the folding is dependent on the forces present at the time of folding and may result in symmetrical folds, asymmetrical folds, bending, buckling or parasitic folds (Leeder and Perez-Arlucea 2006).

Folding is a process that occurs after mineralization and is not limited to sedimentary deposits. It is important to understand that the process of folding can introduce non-stationary features into sedimentary, metamorphic or igneous deposits. Folding is the source of the LVA seen in the Zambian Copperbelt (Figure 2.14).

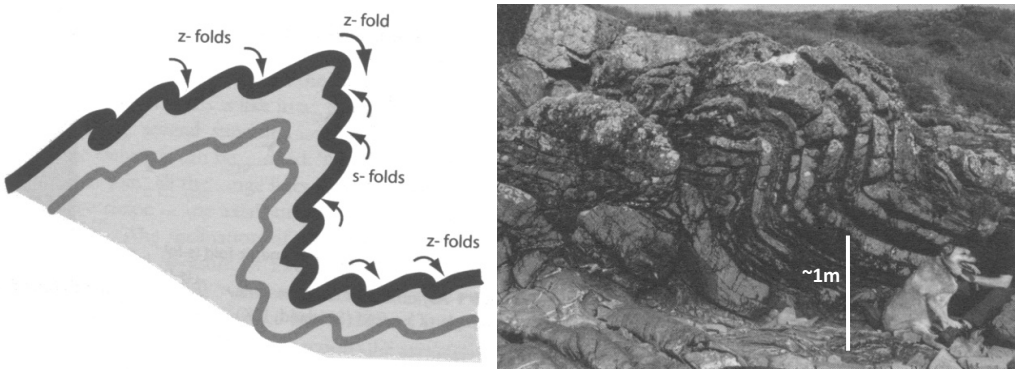


Figure 2.13: Left: Parasitic folds superimposed on a larger symmetrical fold (Leeder and Perez-Arlucea 2006). Right: Outcrop of a z-fold (Leeder and Perez-Arlucea 2006).

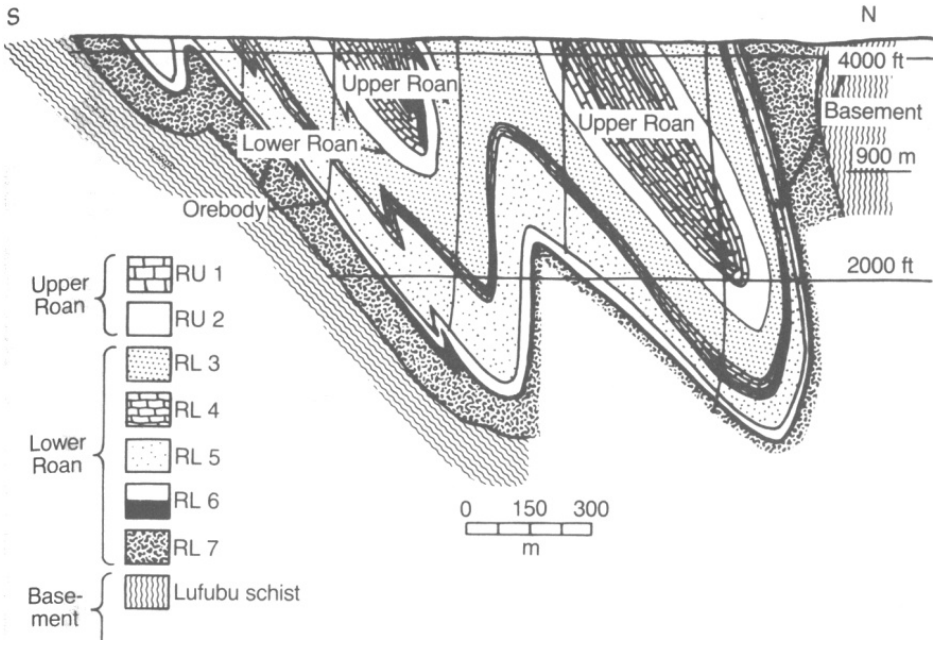


Figure 2.14: A cross section of the Zambian Copperbelt (Guilbert and Park 2007).

2.3.2 Fracturing, Faulting and Vein Formation

When the stresses on a rock mass are greater than the strength of the rock the rock fails forming fractures or faults. There is displacement of the rock mass on either side of the fracture or fault resulting in a possible conduit for fluid flow. In the presence of open fractures and faults, this fluid flow is concentrated, causing the preferential deposition of minerals, forming a mineralized vein (Guilbert and Park 2007). When the orientation of fractures or faults vary spatially, the resulting mineralization also displays locally varying characteristics (Figure 2.15).

Fractures can occur without the resulting formation of a mineralized vein. Vein formation is treated as a special case of fracturing in which mineralization is concentrated in the fracture; however, fractures can cause anisotropy in a more diffuse manner if the fractures are the dominate means of fluid flow during the formation of the deposit. The fracturing around an intrusive deposit is one example where fracturing causes the preferential deposition of minerals, as in a porphyry deposit (Section 2.3.3).

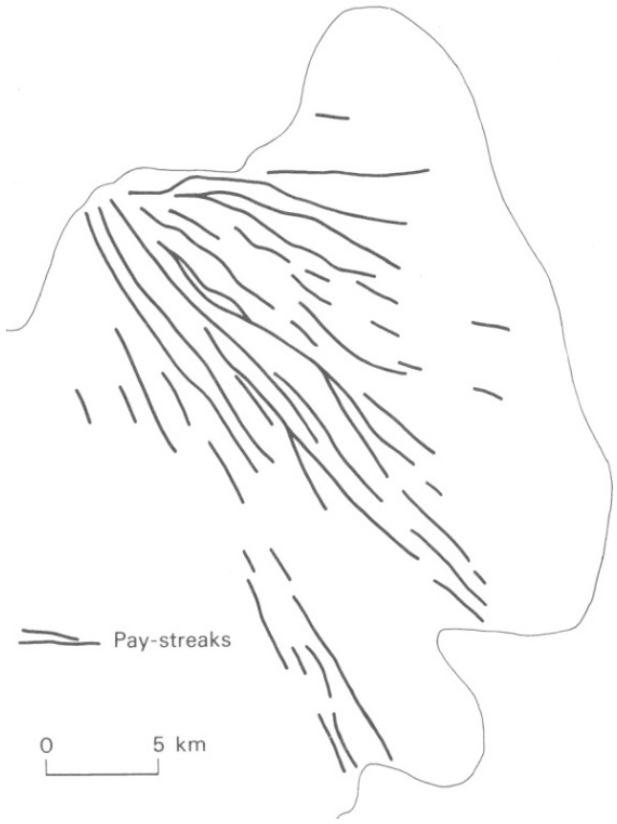


Figure 2.15: Plan view of veins in the Main Leader Reef of the East Rand Basin of the Witwatersrand Goldfield of South Africa showing locally varying directions of strike (Evans 1987).

2.3.3 Porphyry Deposit

While the exact definition of this class of deposits is in question and the formation of porphyries is a complex multistage process (Guilbert and Park 2007), typical porphyries display very interesting non-stationary characteristics. Consider the porphyry deposit at San Manuel-Kalamazoo (Figure 2.16). This intrusive deposit was formed by the convection of water followed by the precipitation and concentration of copper in the surrounding host rock. Convection is possible as the surrounding rock is highly fractured because of the intrusion (Pirajno 2000; Guilbert and Park 2007). Porphyries often display a barren core of material surrounded by a mineralized zone (Figure 2.16) resulting in an aerially radial pattern of continuity in the expected mineral grade (Figure 2.17).

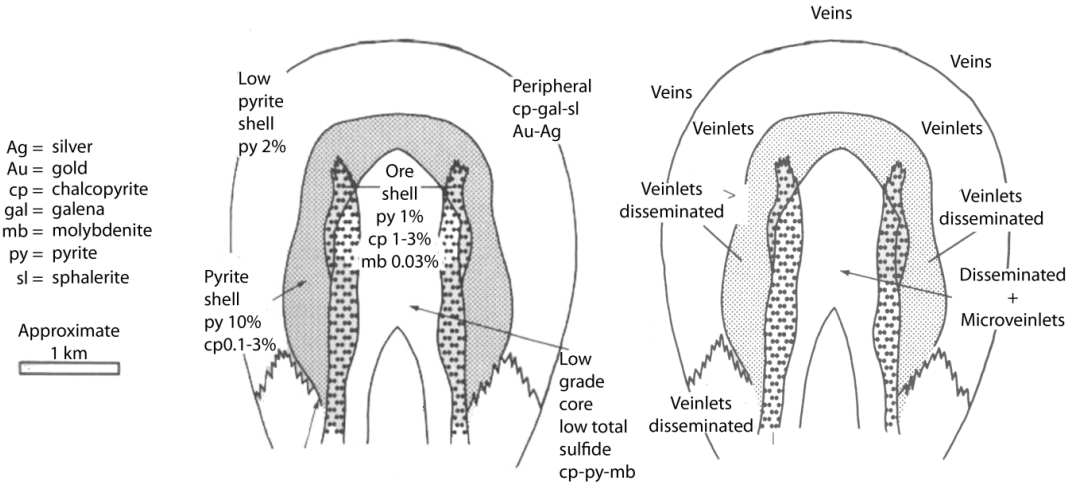


Figure 2.16: Vertical cross section of a Porphyry copper deposit at San Manuel-Kalamazoo (after Guilbert and Park 2007). Copper is concentrated in a circular region around a low grade core. Left: Alteration zones. Right: Mineralization zones.

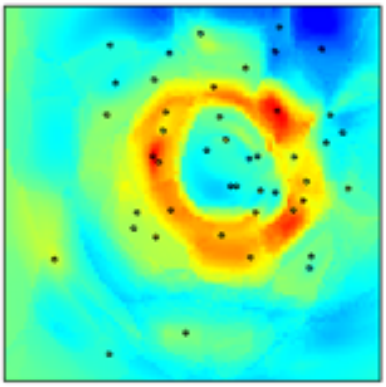


Figure 2.17: Typical cross section of the circular pattern of gold grades in the porphyry deposit modeled in Chapter 6. Mineralization is shown as hot colors (reds and yellows) with the barren core as cool colors (blues and greens). Plot dimensions are 500m in the X and Y directions.

2.3.4 Fluvial, Estuarine and Magmatic Systems

Fluvial deposits are formed from the deposition of sediments transported by rivers. Because rivers display sudden changes of direction due to varying ground conditions, hydrocarbon saturations found in fluvial deposits display similar locally varying characteristics. Slatt (2006) presents five types of fluvial deposits: alluvial fans; fan deltas; braided-river deposits; meandering-river deposits; and incised-valley-fill deposits. Each type of fluvial deposit gives rise to characteristic locally varying features, such as curvilinear muddy channels in a meandering river (Figure 2.19) or the point bars of the Rulison field (Figure 2.20). Slatt (2006) provides additional examples of many fluvial deposits, all containing significant locally varying features due to the unique evolution of different river systems.

Estuarine environments also have the potential to create hydrocarbon deposits that display highly nonlinear features. Large portions of the oil sands contained in the McMurray formation of Northern Alberta are found in sand originally deposited in an estuarine environment (Ranger and Gingras 2003). The branches of an estuary can result in clearly nonstationary directions of continuity (Figure 2.21).

Interestingly, magmatic deposits may also show similar locally varying characteristics as seen in estuarine systems. Consider an idealized model of a komatiite lava field (Figure 2.22). The branching and locally varying features are similar to the estuarine system because the basic mechanism of fluid flow is similar. Fluvial, estuarine and komatiite deposits all display locally varying features because of the complex dynamics of fluid flow involved in their formation.

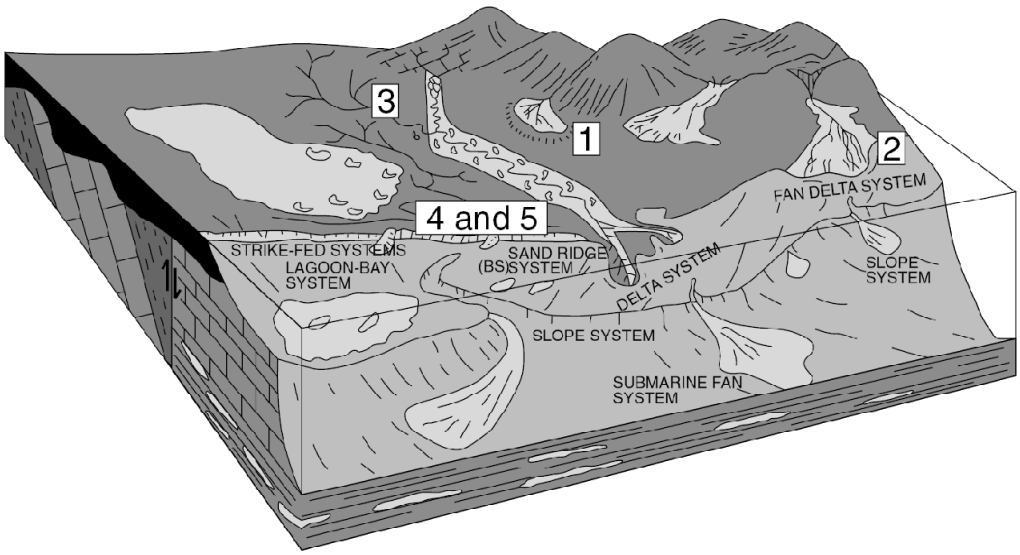


Figure 2.18: Block diagram showing five types of fluvial depositional environments (1) alluvial fans (2) fan deltas (3) braided-river deposits and (4 and 5) incised or nonincised meandering rivers (Slatt 2006). Not to scale.

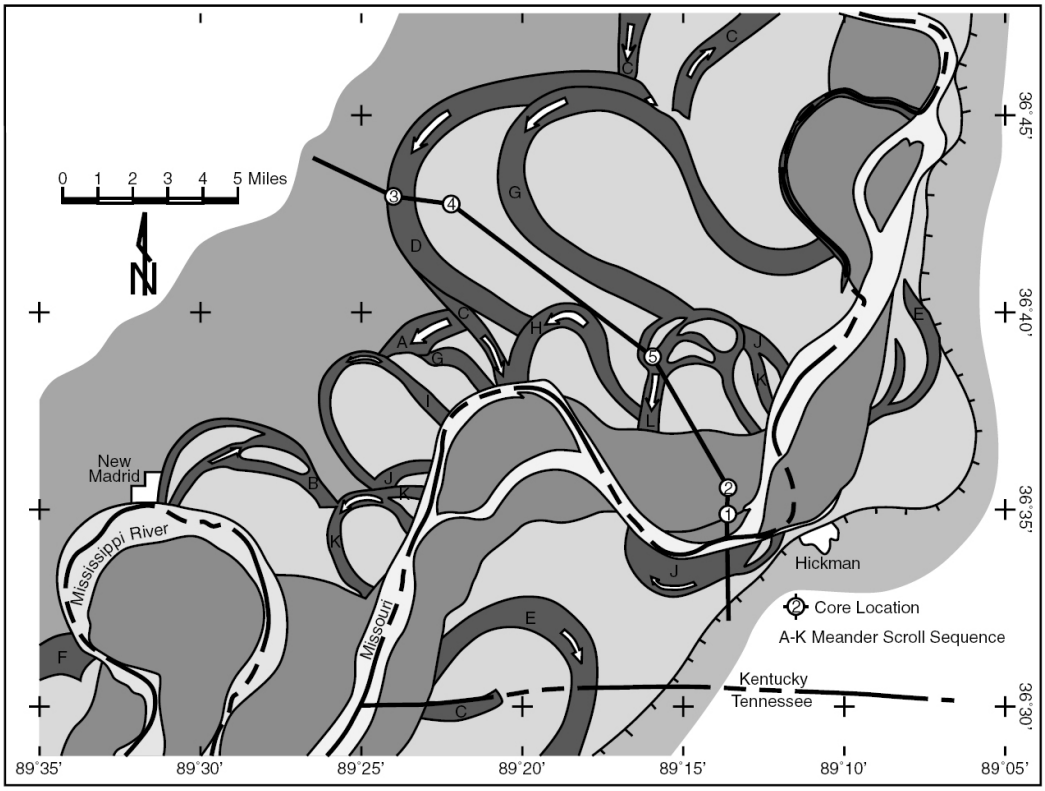


Figure 2.19: Upstream Mississippi River fluvial system. Point-bar sand deposits shown in light gray and muddy channel-fill plugs shown in dark gray (Slatt 2006).

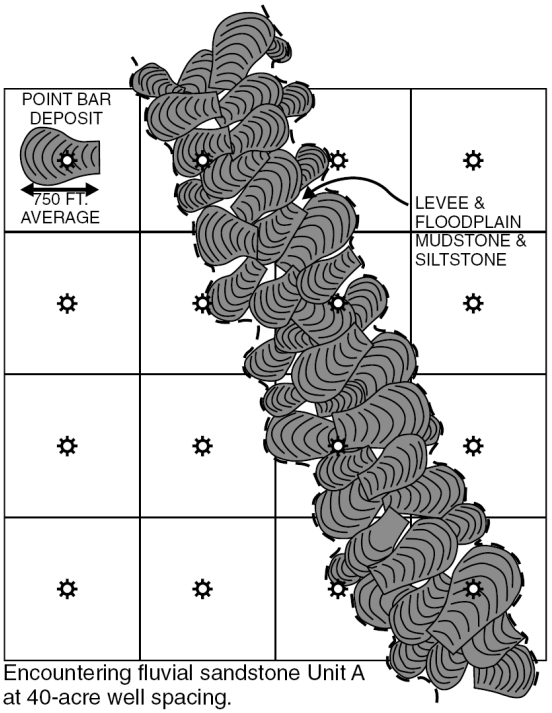


Figure 2.20: Interpreted point-bar reservoir in the Rulison field in the Piceance Basin of Colorado (Slatt 2006). Note the curvilinear point bar deposits as well as the LVA found within each point bar.

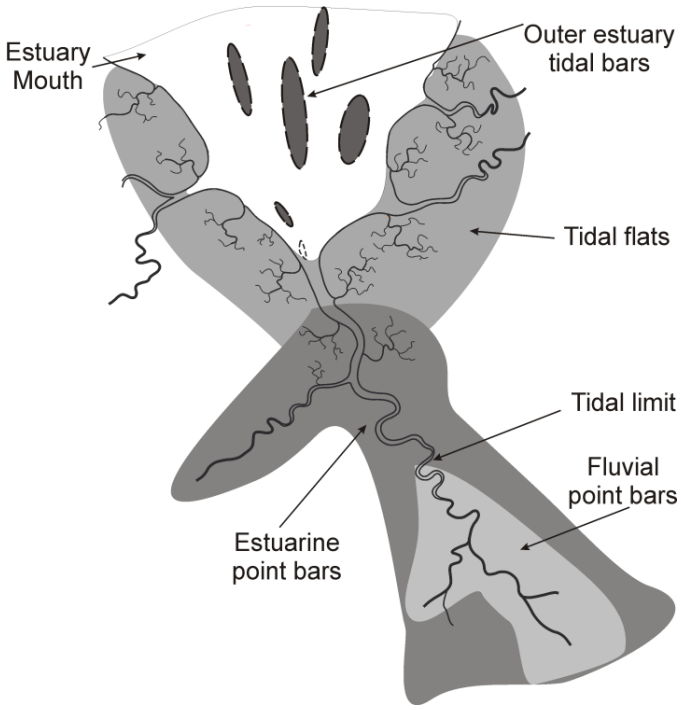


Figure 2.21: Idealized schematic of an estuarine environment (modified from Lettley 2004 and Hassanpour 2009). Not to scale.

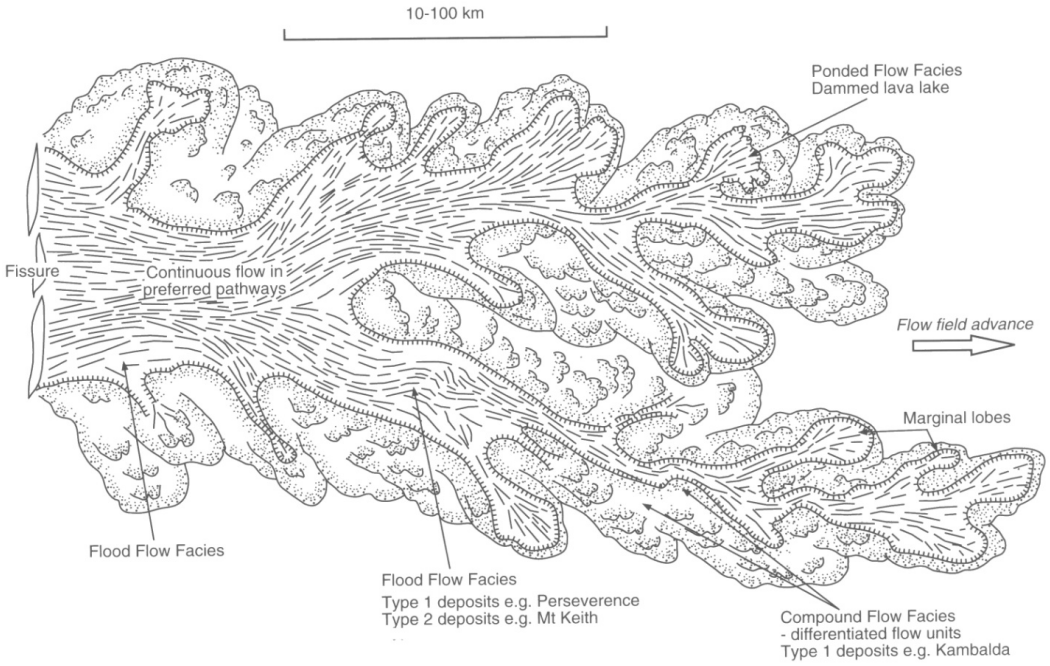


Figure 2.22: Idealized komatiite lava flow highlighting characteristic nonlinear features showing type 1 and 2 Ni sulphide ore deposits (Pirajno 2000).

2.4 The Scale of Anisotropy

The anisotropy present in any deposit is most likely a unique combination of a number of geological processes. The examples presented in Section 2.3 show clear LVA features and provide justification for the consideration of LVA in geostatistical modeling. There are valid geological reasons that the assumption of second order stationarity is often violated by geological deposits because the manner in which deposits are formed naturally results in nonstationary features. Often times these geological processes produce nonstationary features that vary according to the scale of interest. This relationship can be fractal in nature with the smaller scales mimicking the larger scale features, such as the meandering rivers in an estuarine system (Figure 2.21); however, the features seen at various scales can be different and the types of features that should be incorporated into geostatistical modeling using LVA requires clarification.

Typically there are three scales are of interest in geostatistical modeling (1) the point scale (2) the block scale and (3) the domain scale. Scale is a continuous concept and the categorization of three scales of interest is a simplistic discretization of a continuum (Figure 2.23). The appropriate features to model with LVA are features that have a scale larger than the block scale but smaller than the domain scale. Features smaller than the block scale cannot be incorporated due to block discretization limitations and features larger than the domain scale do not impact the models. This scale range of important features (Figure 2.23) is narrowed below.

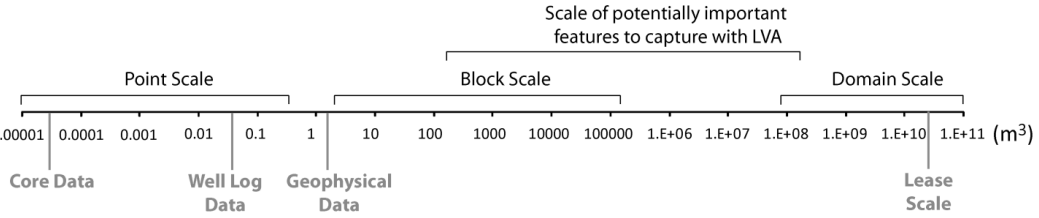


Figure 2.23: Point, block and domain scales are typically considered in a geostatistical analysis. The scale of features that should be incorporated as LVA should be larger than the block size but smaller than features captured by trends. Figure provided is for general modeling purposes, application specific conditions may see scale ranges outside the suggests bounds.

2.4.1 Point, Block and Domain Scale

The point scale and the domain scale are fixed by available data. The point scale is defined as the data scale. Typically this scale is on the order of $0.0001m^3$-$0.2m^3$. In mining applications this is the volume of the core drill hole samples; in petroleum applications this is the effective volume of the available log or core data; and in environmental applications this is the volume of collected samples. This thesis is only concerned with geostatistical models that are built on rectilinear Cartesian grids, thus the smallest scale that can be considered becomes the scale of a single block in the model which is typically several orders of magnitude larger than the point scale. There are many application specific considerations for the selection of a block size, the interested reader is referred to the following references for more discussion on block size selection (Isaaks and Srivastava 1989; Deutsch 2002). In a typical geostatistical

model the block scale is nominally between 5 m^3-100,000m^3 in order to cover the modeling domain with a manageable number of cells; thus, any micro scale feature smaller than the selected block size cannot be incorporated with LVA.

The minimum size of a feature that should be incorporated is further refined by considering the available data spacing. There must be sufficient data density to warrant the features modeled. Without secondary information, such as geological interpretations, remote sensing data or densely sampled secondary data, the lower bound on the scale of potentially important features is based on the size of feature that can be reliably supported by the available data.

The block scale provides the lower bound on the scale of features that can be considered by LVA while the domain scale provides the upper bound. The domain scale is fixed by the data configuration and the goals of the geostatistical model. It corresponds to the limits of the area modeled. This scale is typically >10^8 m^3. The appropriate scale to assess LVA is the scale of the most *relevant* geological features within the bounds provided by the block and domain scales. In this context, relevance is defined as the impact this feature is expected to have on the end use of the models. Geostatistical models are typically used as input to complex transfer functions such as the calculation of reserves (mining), flow simulation response (petroleum), or the volume of contaminated material (environmental). The practitioner must assess the potential impact of the geological features on the transfer function and model those features that are most likely to have a large impact.

Large scale features that are either well understood or can be captured deterministically should be incorporated as trends (McLennan 2008) rather than as LVA. A trend is considered to have no spatial variability and is well supported by known physical characteristics of the variable of interest. Typically these features are seen at a resolution of approximately 0.25-1.0 times the domain scale. Such large scale characteristics are better captured by incorporating a trend model (Journel and Rossi 1989; Goovaerts 1997) or considering a locally varying mean (Wackernagel 2003).

Consider the fluvial example from Figure 2.20 (reprinted as Figure 2.24). The North-Westerly striking orientation of the sediments is clear when examining the deposit at a large scale (~5000 ft in Figure 2.24). As smaller scales are considered, (~750ft in Figure 2.24) the locally varying features become apparent. In this example the North-Western anisotropy would be considered deterministically with trend modeling while the curvilinear locally varying features resulting from the deposition of sediments in the point bars would be considered the locally varying features of interest to be modeled. As a second example, consider the locally varying features in a gold deposit (Figure 2.25). There is a clear vertical trend in the features that should be accounted for with trend modeling. LVA should be used to incorporate the specific features of the individual veins (Figure 2.25 left).

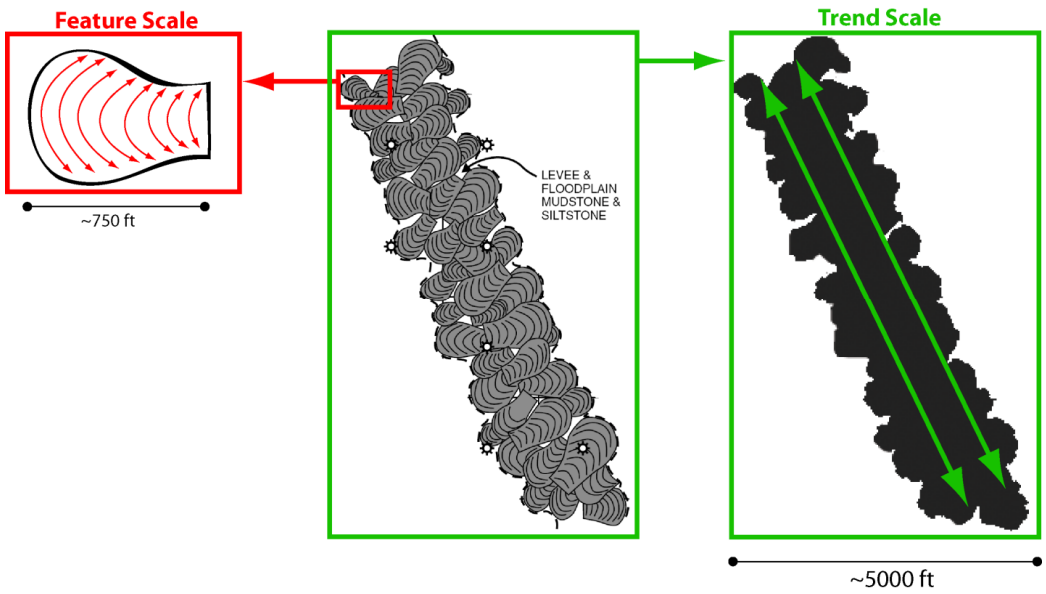


Figure 2.24: Interpreted point-bar reservoir in the Rulison field in the Piceance Basin of Colorado (Slatt 2006). The LVA at a reasonable scale for modeling is shown to the left while the larger scale anisotropy in the North West direction (right) should be accounted for with trend modeling.

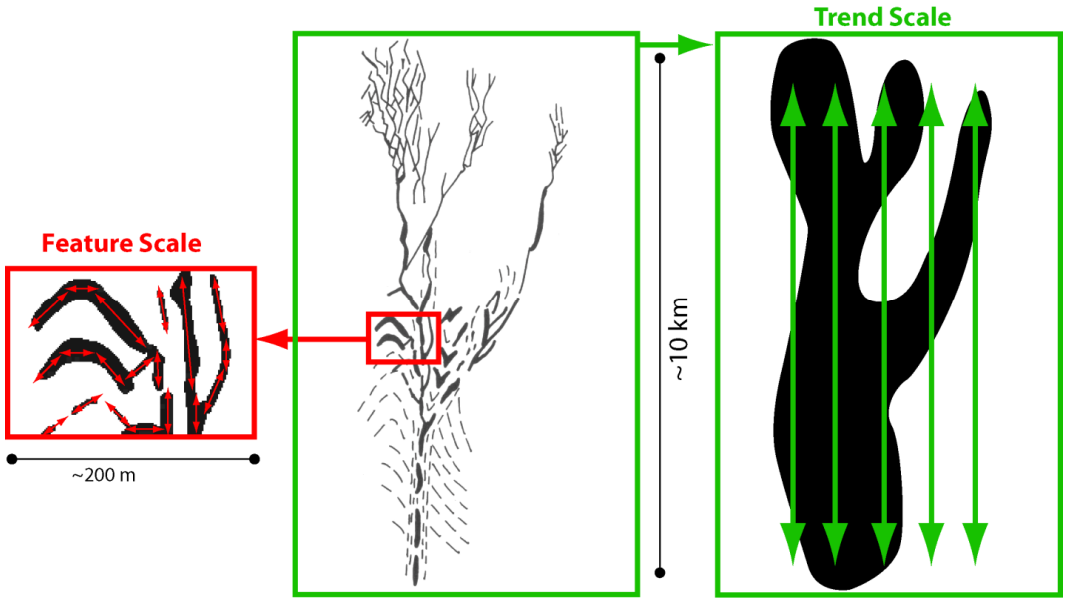


Figure 2.25: Cross section of an Archean mesothermal Au lodes (Pirajno 2000). The LVA at a reasonable scale for modeling is shown to the left while the larger scale anisotropy in the North direction (right) should be accounted for with trend modeling.

It is important to clearly identify which types of features should be modeled with LVA; to summarize the above discussion:

LVA **should not** be used to model:

- features that are smaller than the block size of the geostatistical model
- features that are too small to be reliably captured by the given data spacing
- features that are known deterministically and can be effectively modeled with a trend

LVA **should** be used to model:

- features that display locally varying directions of continuity
- features that are expected to have a significant effect on the transfer function used

2.5 Remarks

Section 2.1 clarified what anisotropy is and how it can be quantified with the variogram. The variogram is modeled based on the available sample data and is required for the implementation of traditional geostatistical techniques such as kriging and SGS. Section 2.2 quantified anisotropy and described how coordinate transformations are used to incorporate a single global direction of anisotropy into geostatistical modeling. Such rotations allow for the calculation of a scalar distance, h, between locations in the transformed space.

Section 2.3 is critical to this thesis and provides evidence that LVA is present in the types of deposits typically modeled with geostatistics. The intention was not to provide an exhaustive index of deposits that display LVA, rather, sufficient examples were presented to justify the development of a methodology for incorporating LVA into modeling.

An exploration of the nature of anisotropy necessarily discusses its dependency on scale. Geological features with LVA can exist within other locally varying geological features. Section 2.4 made clear the range of features that can be appropriately modeled with LVA. The features should be sufficiently large so that they are evident within the constraints of the selected block size and should be supported by the available data. Small features cannot be accurately described by widely spaced data without significant additional justification. Moreover, the features should be smaller than regional trends that can be effectively modeled with traditional techniques such as trend modeling.

The focus of this chapter was largely on stationary anisotropy (Sections 2.1 and 2.2) with the motivation of LVA in Section 2.3. Chapter 3 further discusses the quantification of anisotropy as a locally varying phenomenon.

Chapter 3: LVA Field Generation

This chapter is organized as follows. Section 3.1 discusses the parameterization of the LVA field. Section 3.2 explores the various sources of data that can be used to estimate the parameters that define the LVA field; however, the details of how to generate the LVA field are left to Section 3.3. Section 3.4 provides practical suggestions for selecting an LVA field generation methodology from the methods presented in Section 3.3. A number of examples of LVA fields are provided throughout this chapter for both continuous/categorical data and for exhaustively/sparsely sampled data. This chapter provides practical methodologies to generate the LVA field. Future chapters assume that the LVA field is known.

Real data sources are used in this chapter to highlight the proposed LVA field generation methodologies. Section 3.4 provides a brief summary of the data sets used throughout this chapter.

3.1 LVA Field Parameterization

Chapter 1 motivated the use of anisotropy by introducing deposits that displayed non-stationary features, some figures displaying LVA have been repeated in Figure 3.1. Further, Chapter 1 motivated the use of LVA to consider different directions or magnitudes of anisotropy within a single deposit.

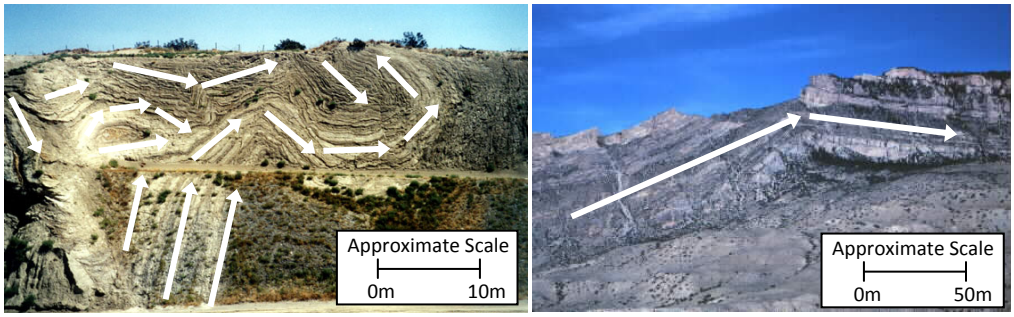


Figure 3.1: Deposits showing anisotropy from Chapter 1.

The LVA field delineates the direction and magnitude of anisotropy in a modeling domain. The direction of anisotropy is defined by three angles, strike, dip and plunge as discussed in Chapter 2. The magnitude of anisotropy is taken relative to the major direction of continuity, thus, two ratios are necessary to define the anisotropy, minor/major and vertical/major. The LVA field is used to calculate the anisotropic distance between two points (Appendix A) and can be defined by a rotation matrix (Equation 3.1).

$$R = \begin{bmatrix} \cos\alpha\cos\varphi - \sin\alpha\sin\beta\sin\varphi & -\sin\alpha\cos\varphi - \cos\alpha\sin\beta\sin\varphi & \cos\beta\sin\varphi \\ \frac{1}{r_1}\sin\alpha\cos\beta & \frac{1}{r_1}\cos\alpha\cos\beta & \frac{1}{r_1}\sin\beta \\ \frac{1}{r_2}(-\cos\alpha\sin\varphi - \sin\alpha\sin\beta\cos\varphi) & \frac{1}{r_2}(\sin\alpha\sin\varphi - \cos\alpha\sin\beta\cos\varphi) & \frac{1}{r_2}\cos\beta\cos\varphi \end{bmatrix} \quad \textbf{3.1}$$

The direction of anisotropy is defined in 3D by three angles, strike (α) dip (β) and plunge (φ), see Figure 2.1. The magnitude of anisotropy is defined by two ratios: r_1 is the ratio between the minor and major directions and r_2 is the ratio between the vertical and major direction. The calculation of the anisotropic distance between two points becomes:

$$d^2(\mathbf{h}) = \mathbf{h}^T R^T R \ \mathbf{h} \quad \textbf{3.2}$$

where two points are separated by the vector (**h**). Thus, the LVA field is an exhaustive set of five parameters (α, β, φ, r_1 and r_2) that vary locally. Calculating the anisotropic distance between points separated by **h** is accomplished using the local rotation matrix. Note that in a 2D case only α and r_1 are required. The remainder of this chapter explores sources of data and methodologies that are available to infer the necessary LVA parameters.

3.2 Sources of Data for LVA Field Inference

LVA is often ignored because there is a lack of available data to infer the LVA field. This section explores the diverse range of sources commonly available for LVA field generation (Table 3.1) but the details of how the LVA fields can be generated from these data sources is deferred to Section 3.3; rather, in this section each data source is described and an example data set with a calculated LVA field is provided. Real data is shown where available, otherwise synthetic data is generated. Section 3.3 presents techniques to generate the LVA field from the data types presented.

Table 3.1: Possible sources of information that can be exploited when generating the LVA field. 'Availability' represents this author's subjective appraisal of the potential availability of each data source.

Source	Brief Description	Availability
Drill hole data	Samples of the variable of interest.	High
Direct angle measurement	Direct measurements of the strike, dip or plunge at drill hole locations using dipmeters, logs or fullbore formation micro imagers.	Low
Remote sensing	Remote methods including seismic, magnetic, electromagnetic, ground penetrating radar and other geophysical measurements.	Medium
Structural models	Stratigraphic deposits are often modeled by layer. Layer orientation can indicate the local direction of anisotropy.	High
Facies or rock type models	Categorical variable models generated by any technique can be used to generate models of anisotropy.	High
Analogue	Features contained in analogue deposits can be borrowed to infer LVA. Features must be conditioned to the deposit of interest.	Medium

Normally, it is necessary to merge multiple sources of information from Table 3.1 to generate a single LVA field. Some areas of the deposit may be extensively sampled and the drill hole data can reliably characterize the LVA field. Sparsely sampled areas may require geological interpretation, inference from analogue deposits, exploitation of remote sensing information or even a stochastic description of the anisotropy field.

3.2.1 Drill Hole Data

If a deposit has been sufficiently sampled, the drill hole data of the variable of interest can be used to provide a reliable inference of the LVA field. Available sample data can be either continuous or categorical and are either (1) primary if it is a direct measurement of the variable for which the LVA field is being inferred or (2) secondary if it is a measurement of a different variable that has the same LVA characteristics as the primary variable.

Assayed drill hole data is available in a mining context and provides the basis for many geostatistical models. This data does not give a direct measurement of the LVA field. A similar situation exists in the petroleum industry where the available data come from wells. Such data does not provide a direct measurement of the LVA field. Finally, in the environmental industry data may be 2D samples collected at the surface, inference of a 3D LVA field is more difficult with this type of data.

The available sample data must be interpreted to extract information with respect to the underlying LVA field. In the case of densely sampled areas, such inference is simplified. In the presence of sparse data a more densely sampled secondary variable may help infer the LVA field (Figure 3.2). The primary and the secondary variable must be correlated before using the secondary variable for LVA field inference. If the spatial structure of the secondary variable is not related to the spatial structure of the primary variable it would be inappropriate to use the secondary information for LVA inference of the primary variable.

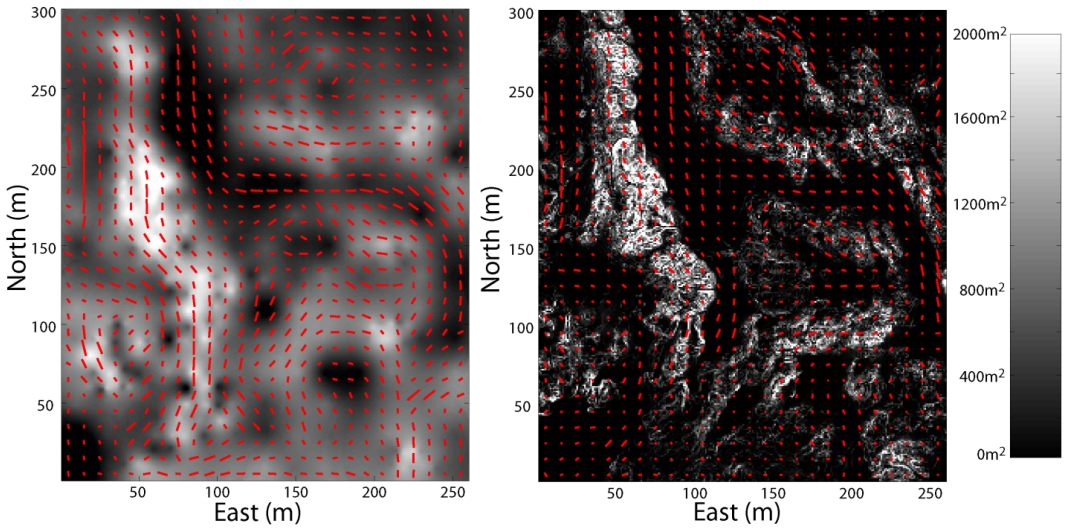


Figure 3.2: Left: Using the more extensively sampled secondary V data to generate the LVA field for U. The V data (Figure 3.3) were kriged using block kriging to obtain a smooth map. Right: Using the exhaustive secondary V data to generate the LVA field for U. The length of the line is proportional to the magnitude of the anisotropy (i.e. the anisotropy ratio).

The following example is based on the Walker Lake Data set (Isaaks and Srivastava 1989). In this data set there are two variables U and V, refer to Isaaks and Srivastava (1989) for a complete description of these variables which relate to elevation. Consider the U variable to be the variable of interest and the more extensively sampled V variable is used to generate the LVA field for U. This is possible under the assumption that U displays the same spatial features as V. This assumption is reinforced by the strong correlation between U and V (Figure 3.3) as well as the knowledge that both variables were similarly constructed (Isaaks and Srivastava 1989). The LVA field generated from the V sample data is shown in Figure 3.2. This example presents two different situations that are quite common in practice: (1) when a secondary variable, V, has been densely sampled and is correlated to the variable of interest, U; and (2) when an exhaustive secondary variable is available to generate the LVA field.

This example explored the situation when a densely sampled secondary variable is available to infer the LVA field for a primary variable. It should be noted that if there are sufficient samples of the primary variable it could be used directly for LVA field inference.

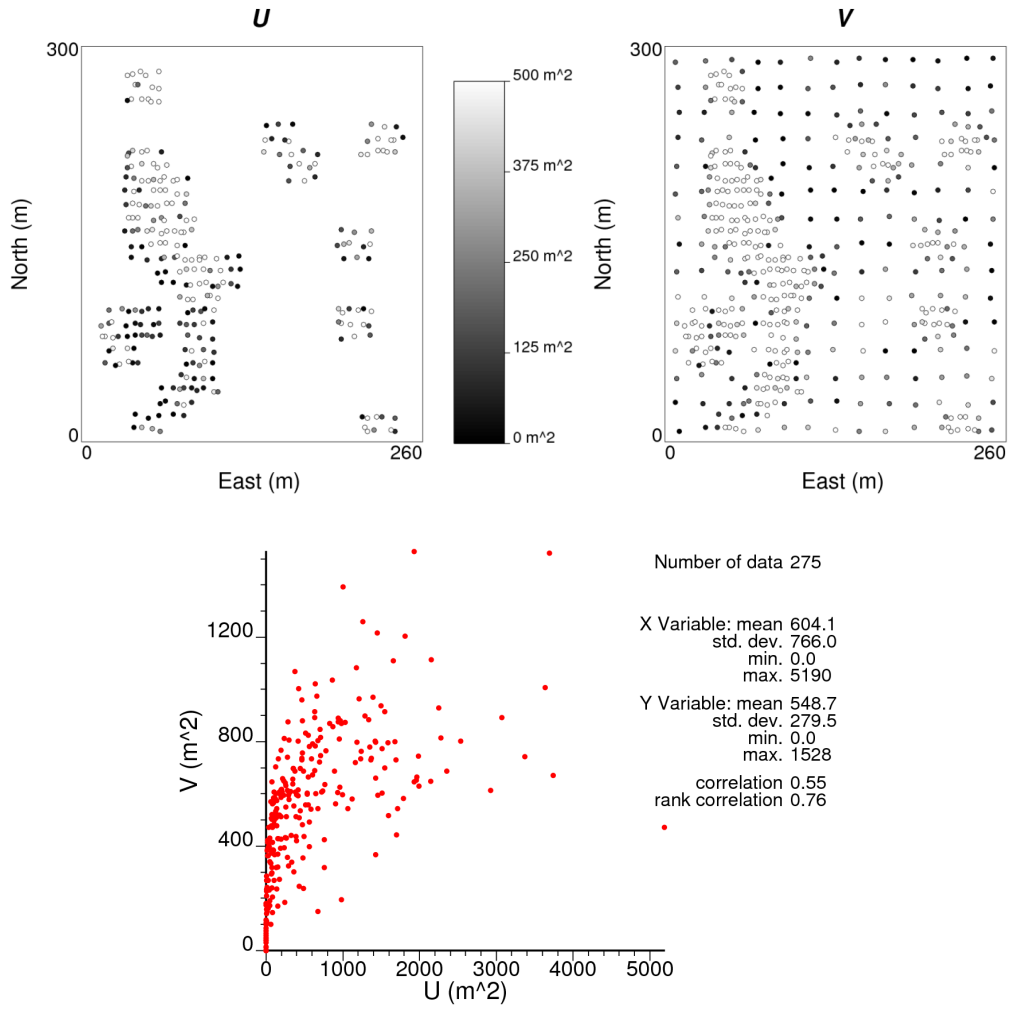


Figure 3.3: Above: Location and values of the available sample data. Below: Relationship between U and V for the Walker Lake samples.

3.2.2 Direct Angle Measurements

It may be possible to measure the orientation of a deposit at a sample location. This type of data is a direct measurement of the LVA field. It is uncommon to have an extensive data base of these types of measurements.

Direct angle measurements can be obtained from outcrops, exposed underground workings, down the hole cameras, fullbore formation micro imagers (FMI) or the formation dip can be measured directly with a dipmeter (Renard and Ruffo 1992). Typically, only the local orientation of the LVA field is measured by these devices. The magnitude of the LVA field (i.e. the anisotropy ratios) must be obtained from one of the alternative data sources presented in Section 3.2 or it can be assumed constant for the deposit. Direct measurements of the orientation and magnitude of the LVA field may be possible when considering numerical models that depend on natural transport such as wind, currents or streams where velocity may provide direction and magnitude.

When direct angle measurements are available they are often sparsely sampled. For the proposed methodology (Chapter 4-6) an exhaustive LVA field is required. This can be obtained by modeling the LVA field between the sparse samples with an interpolation technique such as kriging or inverse distance (Section 3.3.1).

Consider Figure 3.4, where there are 22 direct measurements of the orientation of the LVA field in a 2D domain. The azimuth cannot be directly modeled because of the continuous nature of angles (0° = 360°). Each angle is decomposed into its X and Y components and these components are modeled independently then recombined, see Section 3.3.1 for more details.

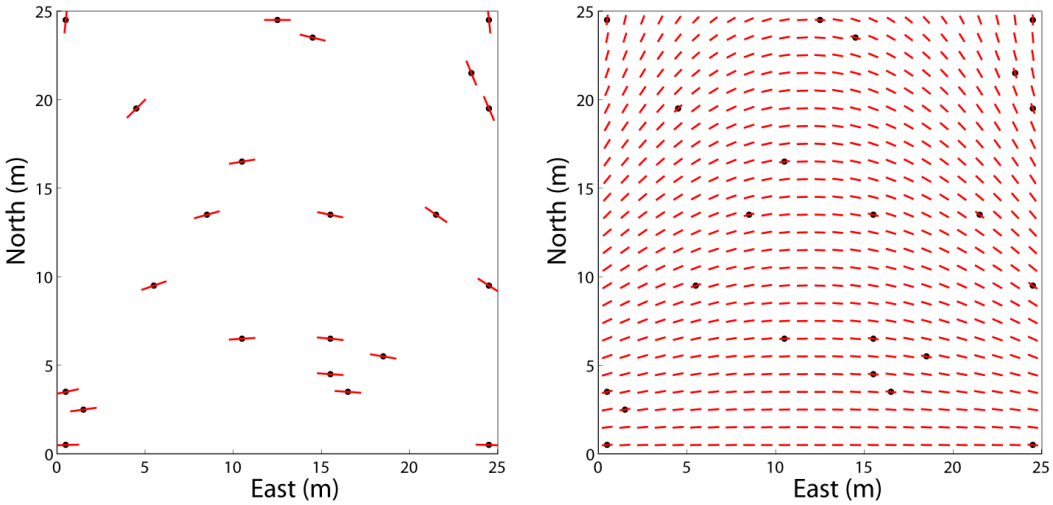


Figure 3.4: Left: Synthetic data. Direction at the locations (circles) is indicated by the orientation of the line through each circle. Right: LVA field after inverse distance interpolation of the X and Y components of each angle measurement.

3.2.3 Remote Sensing

Geophysical remote sensing data can take the form of seismic, magnetic, electromagnetic, gamma-ray or gravitational surveys. The reader is referred to Moon, Whateley and Evans (2007) for a more detailed discussion of each of these remote sensing techniques. For the purposes of this thesis, it is sufficient to understand that remote sensing surveys provide a low resolution measurement of a bulk property, such as density for gravitation surveys, resistivity for electrical surveys or wave velocities for seismic surveys.

When the bulk property measured by the remote sensing technique is related to the variable of interest, it can be used to infer the LVA field. Often this type of data is referred to as soft data. For example, in petroleum applications seismic surveys are often available and are sensitive to porosity changes in the reservoir. It is desirable to first reduce the seismic survey down to a single attribute, such as acoustic impedance, that is calibrated to porosity (Deutsch 2002). From this single attribute, the LVA field for the variable of interest can be inferred. Remote sensing surveys are often available

because of their low cost in comparison to exploration data and are attractive because of their exhaustive nature. They provide much needed information where direct sampling of the variable is sparse.

The bulk property measured by the remote sensing technique must be related to the variable of interest. If the correlation is not great the appropriateness of using the remote sensing survey should be evaluated. It may only be appropriate to use the survey as correlated secondary information when inferring the LVA field.

Consider the 62 wells in Figure 3.5 with porosity measurements. An exhaustive 2D map of an interpreted seismic attribute is also available. There is a high correlation between porosity and the seismic attribute, ρ=0.62, which is often the case as seismic data is sensitive to local porosity variations (Deutsch 2002). The seismic attribute and the inferred LVA field for porosity are shown in Figure 3.6.

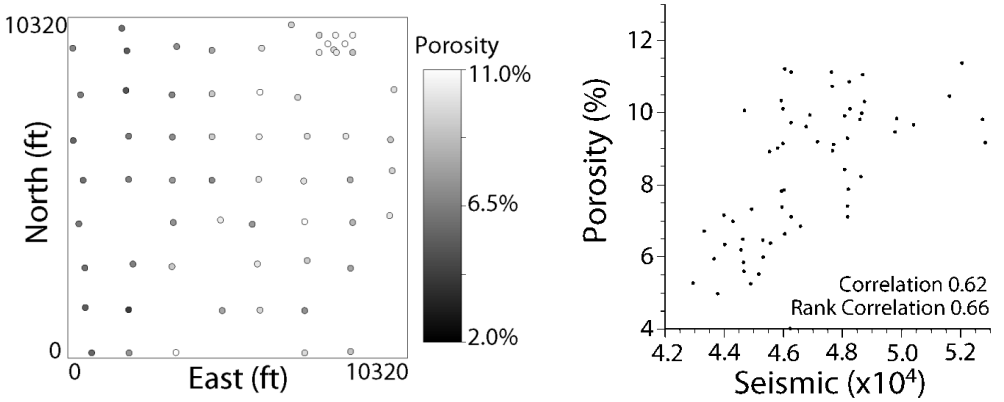


Figure 3.5: Left: location of the 62 porosity well data. Right: relationship between porosity and the seismic attribute.

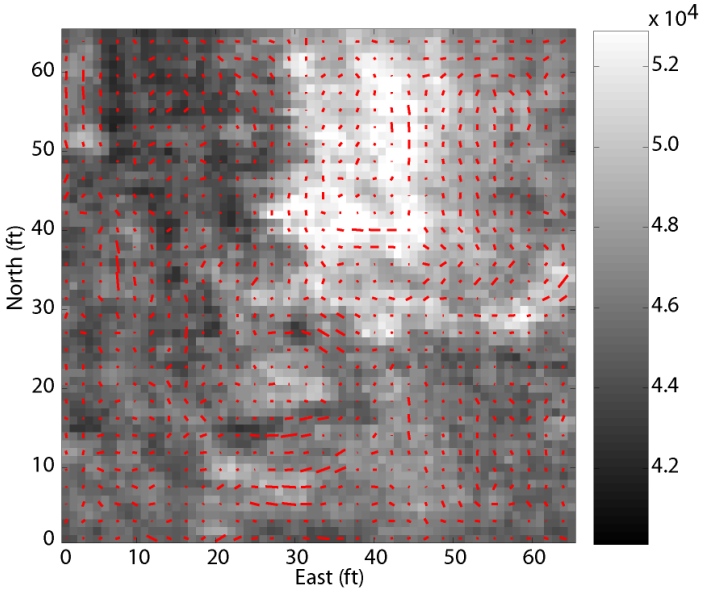


Figure 3.6: LVA field built from an exhaustive seismic survey. The gray scale image is the filtered seismic attribute. The moment of inertia method (Section 3.3.3) is used to generate the LVA field. The length of the line is proportional to the magnitude of the anisotropy (i.e. the anisotropy ratio).

3.2.4 Structural Models

Continuity within a stratigraphic deposit often follows the form of the original deposition of the layer, that is, the direction of continuity is parallel to the deposited layer. The structural model for a stratigraphic deposit can provide insight into the LVA field. Consider a reservoir with multiple stratigraphic layers (Figure 3.7). If the continuity of properties within each layer follows the stratigraphy, the surfaces of the interface between each layer provides the direction of the LVA field. When such stratigraphic deposits are geostatistically modeled, each layer is identified and considered independently (Deutsch 2002). LVA orientation can be extracted directly from the structural model; in this example the direction is obtained by averaging the slope of the top and bottom surfaces defining the layer. This would be repeated for each layer to generate the exhaustive LVA field. A constant magnitude or anisotropy ratio could be determined for each layer or a different method (Section 3.2) could be used to determine the magnitude of the LVA field.

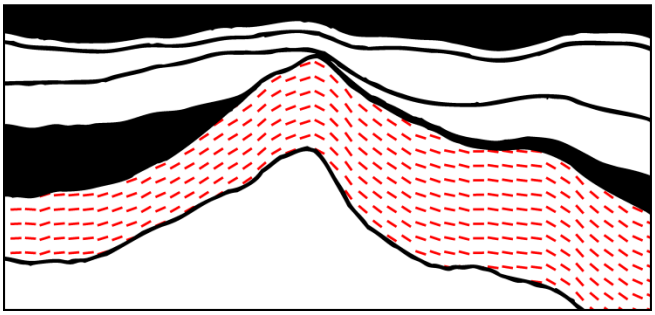


Figure 3.7: A cross section through a reservoir with 6 different stratigraphic layers (Deutsch 2002). A potential LVA field for the lower layer is shown. Scale not available.

3.2.5 Facies or Rock Type Models

The generation of facies and rock type models that respect known geological features is important. Traditional facies modeling has been based on a statistical model (SIS, truncated Gaussian, etc.) or a subjective geological interpretation generated by the geologist. In the mid 1990's two alternative modeling approaches emerged to link geological interpretations with statistical modeling. The first technique, multiple point geostatistics, uses analogue information in the form of a training image to reproduce the desired geological features in a statistical model (Deutsch 1992; Guardiano and Srivastava 1993; Caers 2001; Strebelle 2002; Ortiz and Deutsch 2004; Lyster and Deutsch 2006, Keogh *et al.* 2007). The second technique, object based modeling, parameterizes the desired geological objects and places an entire object into the modeling domain (Mackey and Bridge 1992; Deutsch and Tran 2002; Pyrcz *et al. 2008*). The end goal of both of these techniques is a facies model that is conditioned to known well/drill hole data and contains the desired geological features.

MPS and object based modeling are presented here only as methodologies to generate desired geological features in a facies model with the end goal of generating the LVA field. These facies models can be used to infer the underlying LVA field if the variable of

interest follows the geological objects in the facies models. Consider a simple channel facies model (Figure 3.8). If the property of interest follows the modeled channel feature, this can provide valuable information when inferring the LVA field.

The facies model contains uncertainty. Any LVA field generated based on facies models must consider that the facies models are themselves derived from a stochastic technique. Uncertainty in the LVA field must be carried though the modeling work flow by generating multiple LVA fields (Figure 3.9).

The following example (Figure 3.9) shows the generation of an LVA field from 3D channel objects. In this case it is assumed that the property of interest is continuous along the channel. There is no unique facies model due to the lack of available well data, thus, a different LVA field is used for each geostatistical realization. All facies models are conditioned to the available well data and represent different interpretations of the geology. Further modeling would carry forward each LVA field.

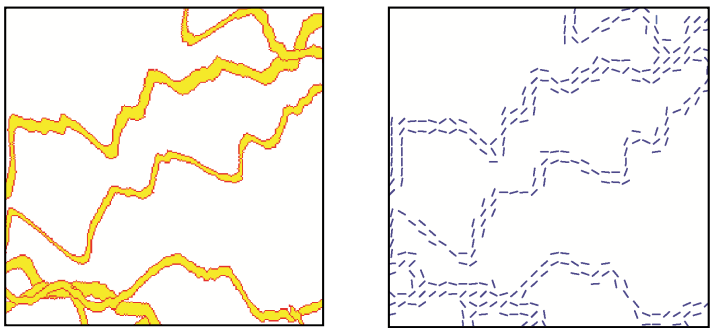


Figure 3.8: LVA field built from a channel (Chapter 1). The grid is 256 by 256 cells. The moment of inertia method (Section 3.3.3) was used to generate the LVA field.

3.2.6 Analogue Deposits

It is common in the petroleum industry to borrow the statistics of a mature reservoir and apply the statistics to a less developed reservoir. This is done because there is a lack of data available to infer the desired statistics. An analogue mature deposit could be used to infer the LVA field for a less developed deposit but the LVA field must be conditioned to any hard data available. Some geostatistical techniques, such as multiple point statistics, can employ analogue statistics directly; however, when using the LVA field of an analogue deposit the geomodeler must maintain consistency between the developed LVA field and any other data available for the less developed deposit.

The need to condition the LVA field to data makes it difficult to infer the LVA field from analogue deposits. However, if there is very little hard data available to generate the LVA field and a nearby developed deposit exists that contains similar geology, the LVA field of the mature deposit could be imposed. The LVA field would be generated for the mature deposit using one of the available techniques (Section 3.3) and manually modified to fit any of the available data for the deposit of interest.

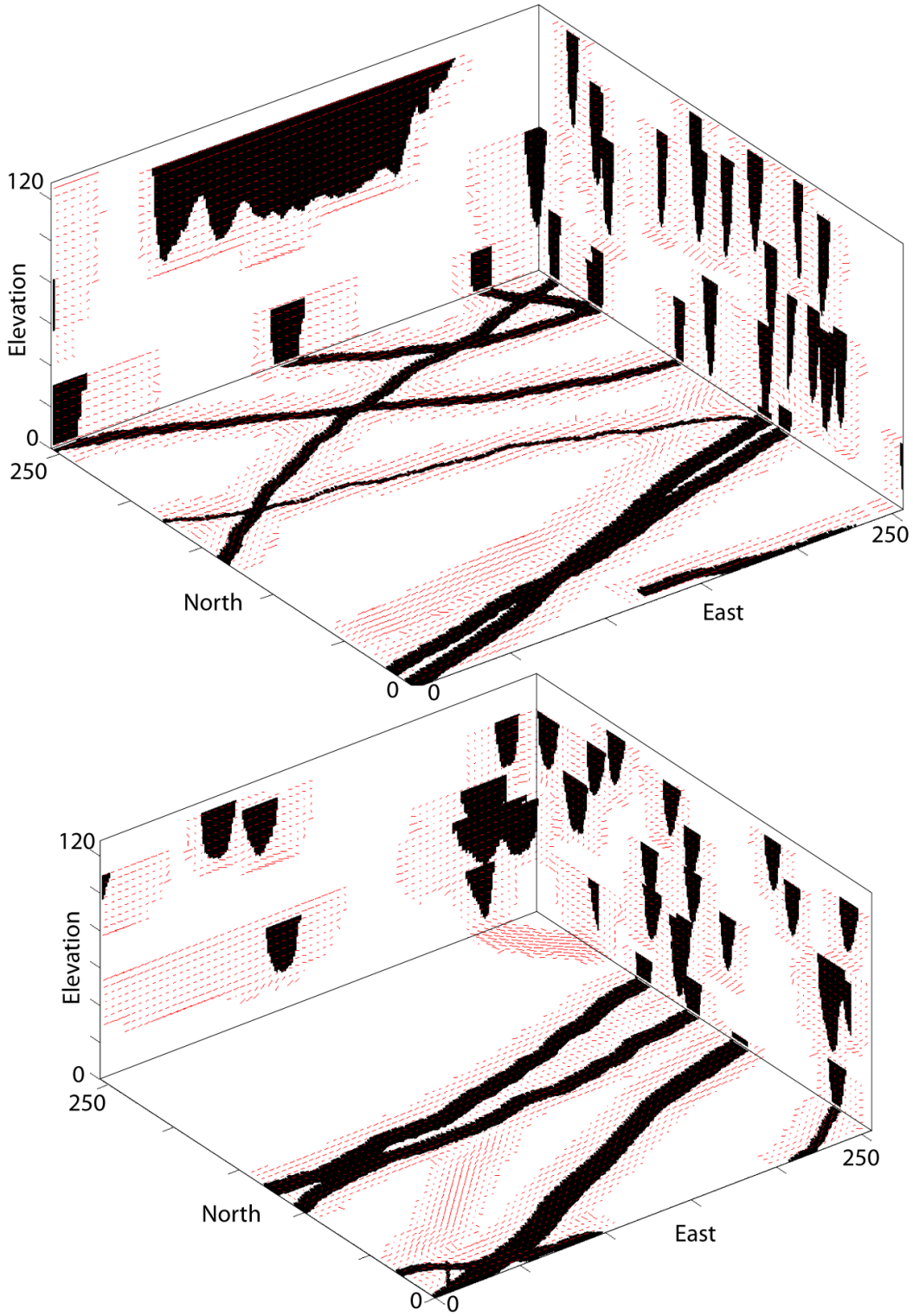


Figure 3.9: Two LVA fields built with the moment of inertia method (Section 3.3.3). The length of the line is proportional to the magnitude of the anisotropy. Models are conditioned to the available data. Where the LVA field is not indicated (white space), the anisotropy ratio is 1:1. Models taken from Pyrcz *et al.* (2008). Model scale is typically varied to fit site specific conditions such as channel width and length.

3.3 LVA Field Inference from Available Data

Section 3.2 discussed different sources of information that can be used to help infer the LVA field. The following section introduces methodologies for generating the LVA field from the data once it has been identified and obtained. All techniques work with either continuous or categorical data (Table 3.2). There may be situations that require a combination of two or more of these techniques; however, they are presented independently.

Table 3.2: Possible methods for LVA field generation. *Difficulty of application* represents this authors subjective appraisal of the difficulty involved when implementing each technique. Some techniques can only be used to determine LVA orientation.

Technique	Orientation/ Magnitude	Brief Description	Difficulty of Application
Simulation and estimation of components	Orientation and Magnitude	If direct measurements of the LVA field are available at sparse locations, the exhaustive field can be inferred using traditional geostatistical techniques.	Low
Manual LVA inference	Orientation only	A geologist can often generate a single, subjective LVA field. Such methods have the benefit of incorporating expert knowledge and experience.	Medium
Moment of inertia	Orientation and Magnitude	A moving window moment of inertia calculation can give the direction and magnitude of the LVA field based on measurements of the variable of interest.	Medium
Automatic feature interpolation	Orientation only	Data values that are similar in magnitude are automatically joined to generate samples of the LVA field orientation.	Low

3.3.1 Simulation and Estimation of Components

When direct measurements of the LVA field are available, traditional estimation techniques can be used to generate the exhaustive field (Fisher 1996). There are many such techniques available, inverse distance and kriging (as introduced in Chapter 1) are used here to demonstrate the methodology but any estimation technique could be used.

The nature of the angular data requires some unique preprocessing before the LVA field can be directly estimated. Because angles are continuous [360˚ = 0˚] the angles cannot be directly estimated from the available data. Consider this example where there are two angle measurements of 0˚ and 359˚ located at (0,0) and (25,0) respectively. Simply applying inverse distance or kriging to these values generates a nonsensical LVA field (Figure 3.10). The angles range smoothly from 0˚ to 360˚. The nature of how angles *wrap* between 360˚ – 0˚ must be explicitly accounted for.

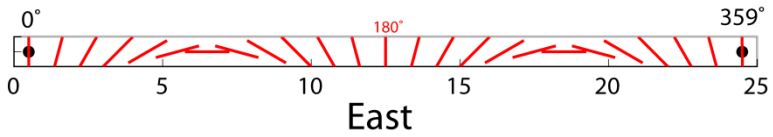


Figure 3.10: Estimation of the angles from two measurements, 0˚ and 359˚ located at 0m and 25m.

The unrealistic LVA field in Figure 3.10 can be avoided if the angles are decomposed into their X, Y and Z components. In this methodology the X, Y and Z lengths are estimated and then recombined to generate the resulting LVA field. Consider the synthetic dipmeter data from Figure 3.4 reprinted in Figure 3.11. The X and Y components of the angles can be calculated from Equations 3.3 and 3.4. The X and Y components (Figure 3.11 right) can be estimated and then recombined to generate the LVA field (Figure 3.12 and Figure 3.13). In this example inverse distance was used to interpolate the X and Y components, but kriging could also have been implemented.

$$X = \sin \alpha \quad \textbf{3.3}$$

$$Y = \cos \alpha \quad \textbf{3.4}$$

where α is the strike angle measured clockwise from North.

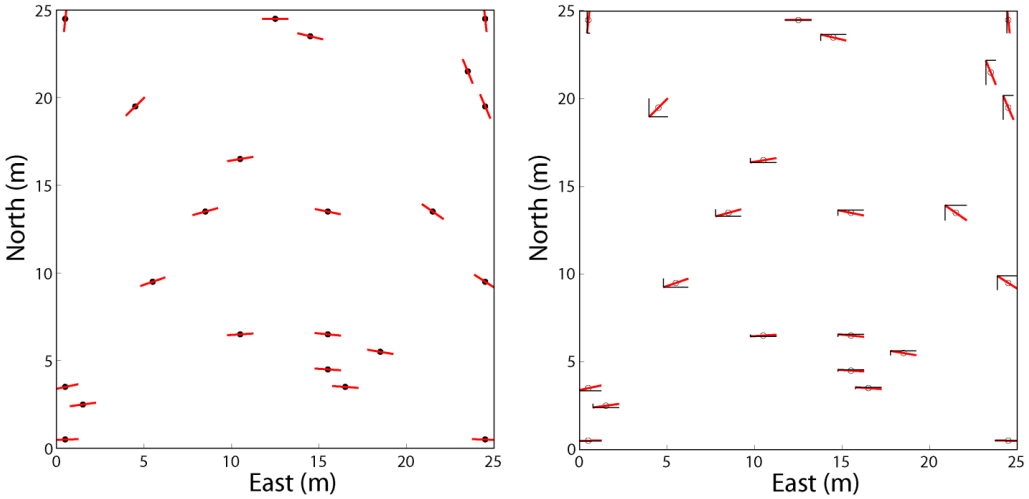


Figure 3.11: Left: Synthetic dipmeter data. Right: Each measurement has been decomposed into X and Y components.

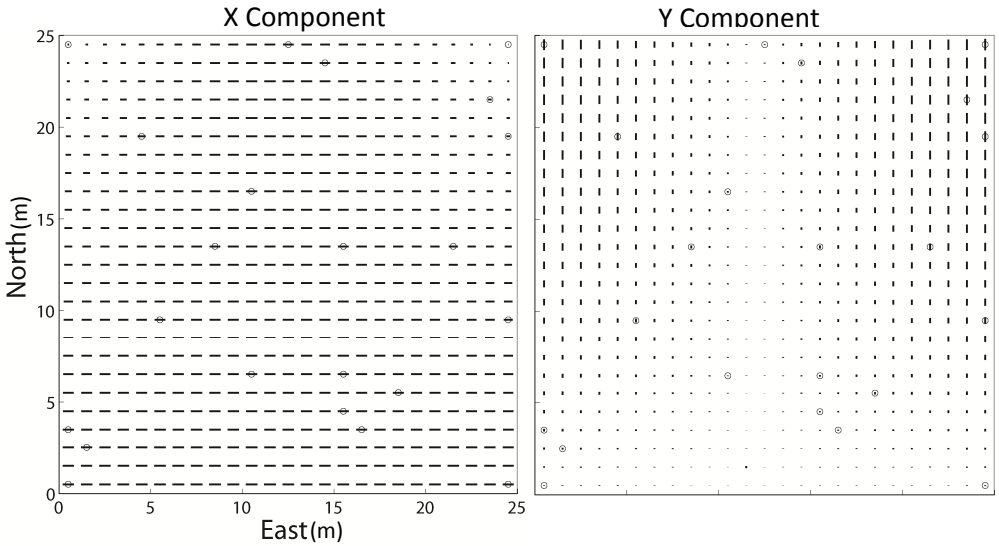


Figure 3.12: Estimation of the X and Y components.

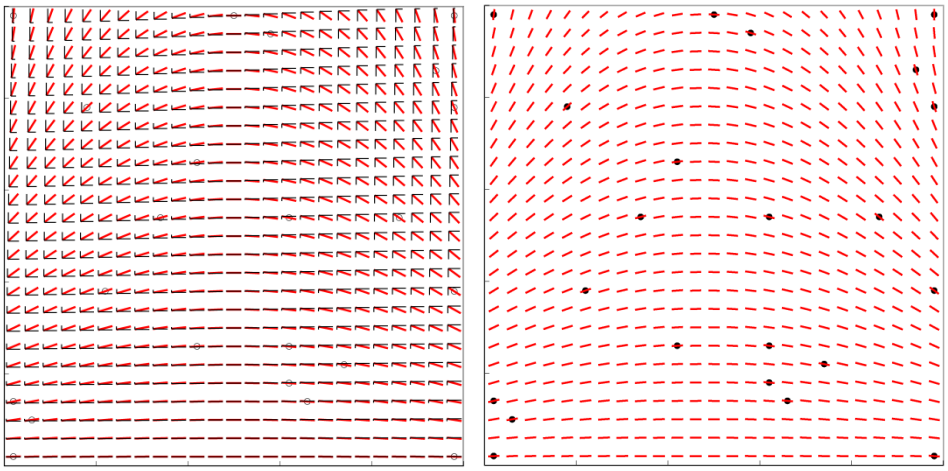


Figure 3.13: Recombining the components to create the final LVA field.

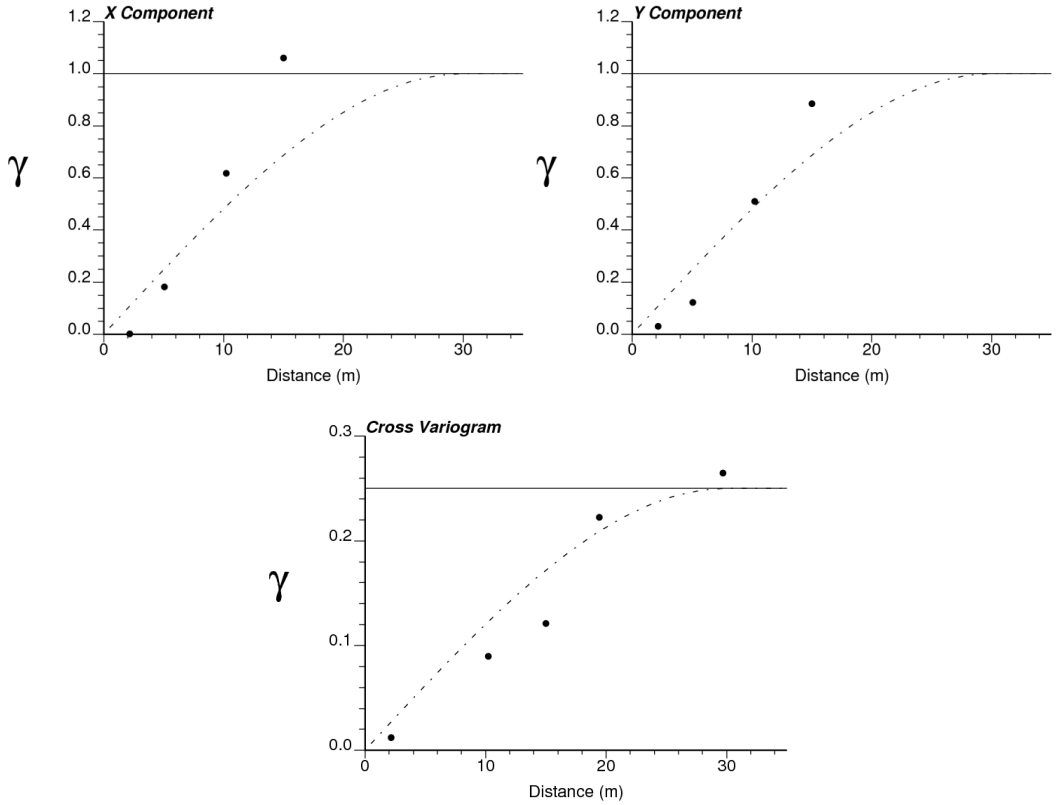


Figure 3.14: Variograms of the X and Y components using the 22 synthetic data. Variograms are fit with a single spherical structures with no nugget. X component range = 30. Y component range = 30. The cross variogram has a sill of 0.25 and is required to jointly simulate the X and Y components.

Estimation of the LVA field from the available azimuth measurements generates a single LVA field that can be used to capture the desired geological features. However, in most practical applications it is important to quantify the uncertainty in the LVA field and carry that uncertainty though the geostatistical analysis into flow simulation (petroleum), reserve calculation (mining) or contaminate classification (environmental).

This can be accomplished by considering multiple realizations of the variable of interest where each realization would be built using a different LVA field. Simulation is considered in more detail in Chapter 5; therefore, in this section the goal is to generate multiple realizations of the LVA field to quantify potential uncertainty. This can be accomplished by simulating the X and Y components. The variograms of the X and Y components from the synthetic dipmeter data are shown in Figure 3.14. Sequential Gaussian simulation is used to generate multiple realizations of the LVA field for the variable of interest (Figure 3.15).

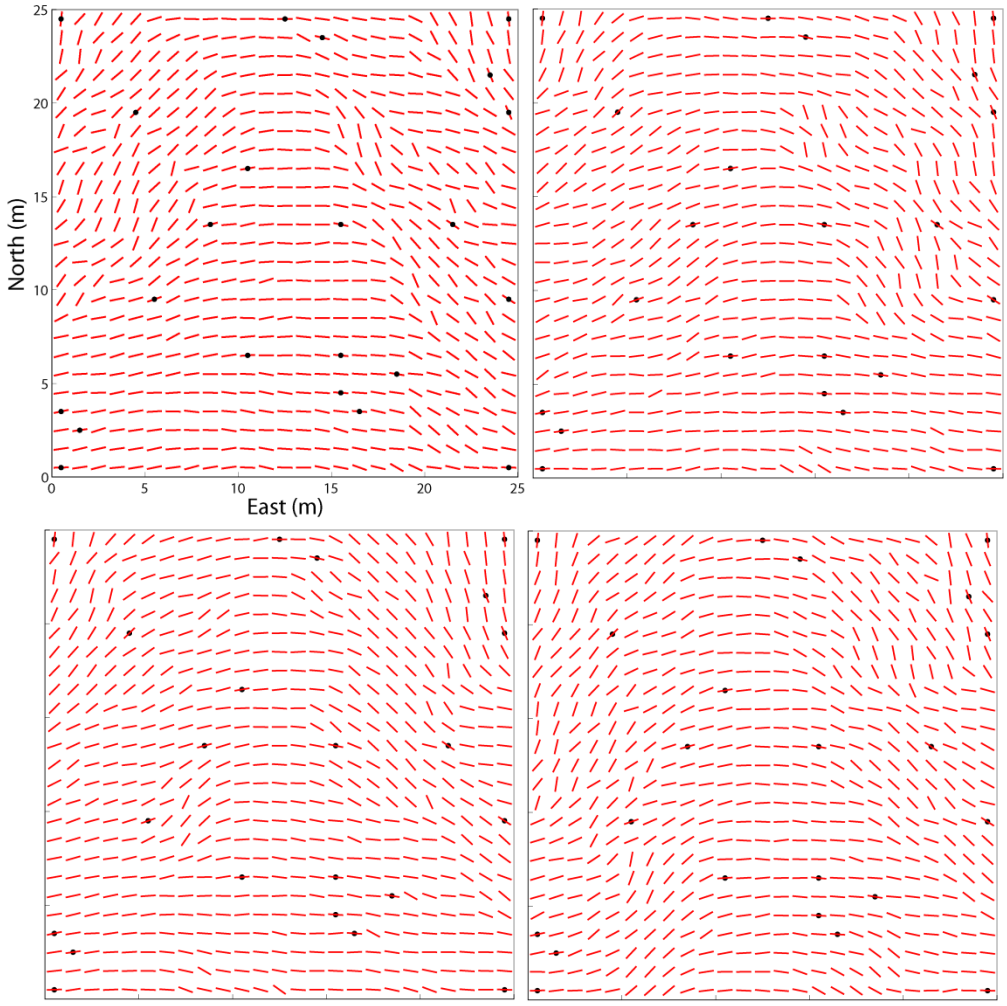


Figure 3.15: Four LVA fields from sequential Gaussian simulation.

3.3.2 Manual LVA Inference

Perhaps the most straightforward methodology to generate the LVA field is to manually assign the direction and magnitude of the LVA field based on professional expertise. Most likely this would be done by an experienced geologist with knowledge of the deposit of interest. Taking into consideration all available types of data from Section

3.2, an expert could generate a subjective LVA field that considers many different sources of information such as the depositional environment, available sample data, available dipmeter data and personal experience with similar deposits.

The expert could assess the LVA field at a number of discrete locations and then fill in the exhaustive field automatically with estimation or simulation. This would reduce the professional time required to determine the LVA field and also reduce the difficulty inferring an exhaustive field. Consider the data in Figure 3.16. The LVA field has been determined by visual assessment at various locations. The exhaustive field is then generated using kriging (Section 3.3.1). Two different interpretations are shown on (Figure 3.17) and are quite similar with the exception of the difference in the south-western portion of the LVA field.

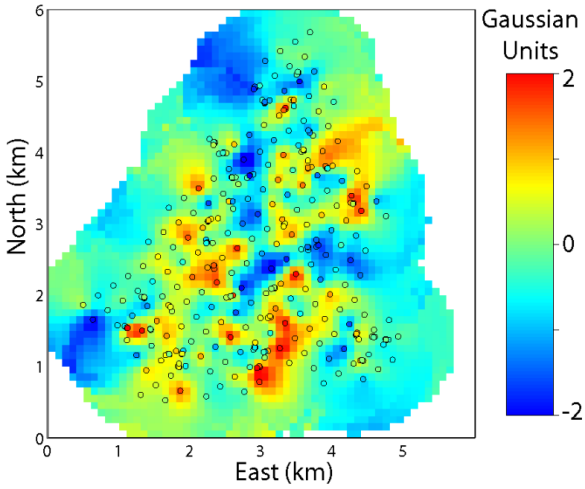


Figure 3.16: Data used for LVA field generation. Original data have been normal scored and kriged to produce the underlying smooth map.

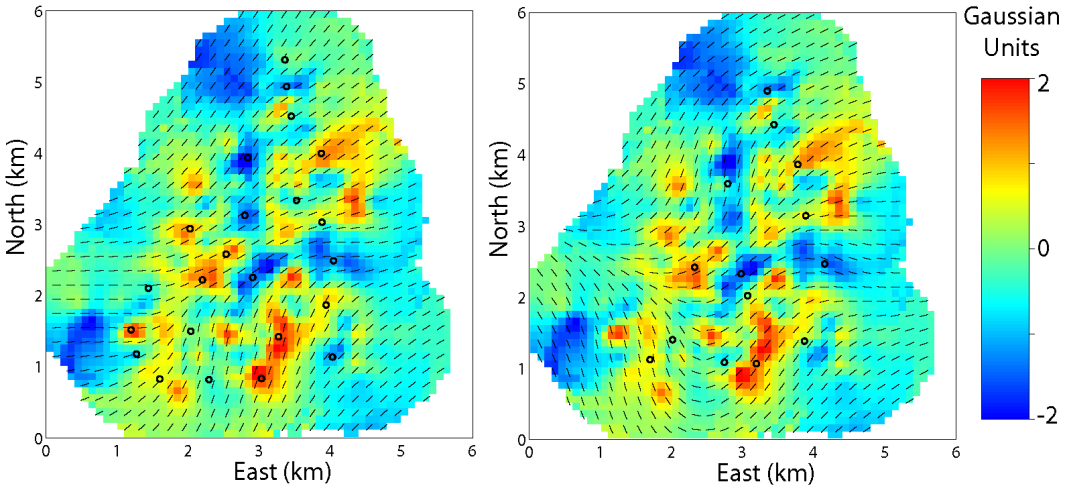


Figure 3.17: Points indicate locations were the LVA field orientation is manually assessed. Left: A deterministic interpretation based on defining LVA orientation at 14 locations. Right: A second interpretation based on defining LVA orientation at 23 locations. The exhaustive field was generated by kriging the X and Y components of the measurements.

A drawback of manually estimating the LVA field is the difficulty inferring a large three dimensional field. Likely, a 2D field would be generated by level and the third dimension extrapolated. Moreover, often there is only a single, subjective LVA field generated by the expert. Assessing and accounting for the uncertainty in the LVA field is difficult. Care should be exercised when attempting to infer the magnitude of the LVA field. The magnitude of continuity is often more difficult to assess than the orientation of the LVA field. LVA field orientation can often be seen in geological interpretations as well as the sample data available; however, it is often much more difficult to visually assess the magnitude of the anisotropy. A global anisotropy ratio can be determined and assumed to be stationary if locally varying anisotropy ratios are difficult to infer.

3.3.3 Moment of Inertia

Manual LVA field inference (Section 3.3.2) of many locally varying directions of continuity for potentially many variables of interest is time consuming and subjective. An automatic method to generate the LVA field directly from available data is required. The moment of inertia method (Mohammadhassanpour 2007) is presented here as an automatic technique for LVA field generation.

Consider the problem of determining a globally stationary direction of continuity for the data set in Figure 3.18. A standard approach is to examine the covariance map of the available data as in Figure 3.19. This is a plot of the covariance between any two locations in space separated by a lag vector (**h**). The plot is generated by scanning the data and locating pairs of points separated by lag vectors (i.e. $\mathbf{h}_1$ or $\mathbf{h}_2$ in Figure 3.19). All data separated by the appropriate lag are paired and the covariance calculated as per Equation 3.5. This is repeated for all lag distances and directions to build the covariance map. In this example, the data are more continuous in the NE direction which appears as high covariance values along the NE direction in the covariance map (Figure 3.19).

$$Cov(XY) = E\{(X - m_x) \cdot (Y - m_y)\} \quad \textbf{3.5}$$

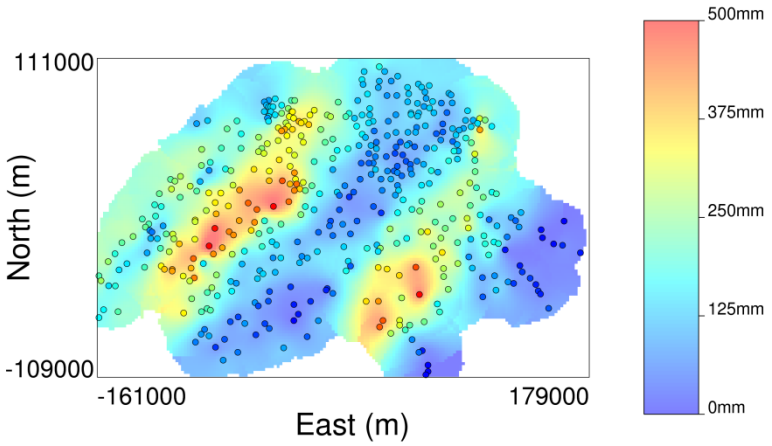


Figure 3.18: Data locations. There are 467 data with an average of 18.4mm of rainfall. Data are in units of a tenth of a millimeter of rain fall. A map of the block kriging of the data is also shown to highlight trends and local anisotropic directions.

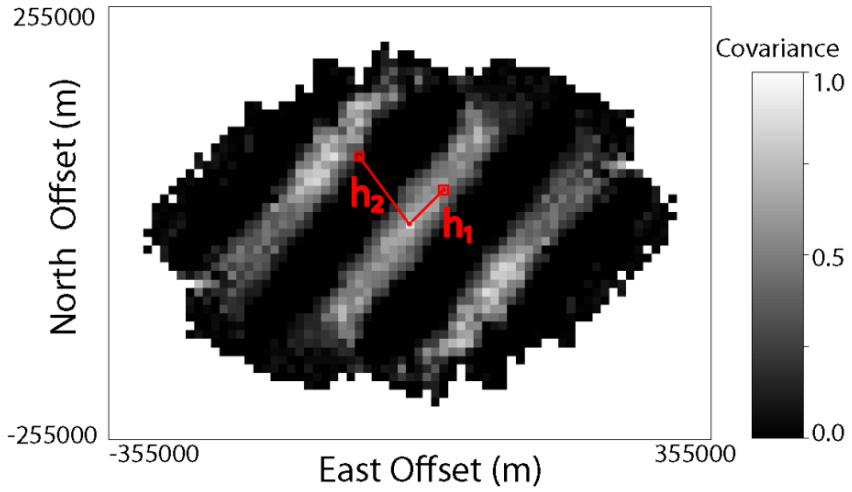


Figure 3.19: Covariance map of Sic data. Each cell in the model represents the covariance between points separated by a lag **h**. Two lags are shown: $\mathbf{h_1}$ = 56,570m in the N45°E direction; $\mathbf{h_2}$ =100,000 m in the N37°W direction. The lag can also be defined by the East and North offsets. Note that the plot is symmetric.

The global direction of continuity can often be derived from a covariance map (Figure 3.19). This concept can be extended to consider locally varying directions by restricting the calculation of the covariance map to local moving windows. For each location a local covariance map is calculated. The problem now reduces to automatically determining the direction of continuity from a covariance map. This is accomplished using the moment of inertia method (Mohammadhassanpour 2007). In this method the covariance values in the covariance map are considered as *mass*. The moment of inertia tensor of this mass can be determined and corresponds to the major direction and magnitude of continuity. The moment of inertia (*I*) of a mass (*m*) describes the rotational inertia of a rigid body (Equation 3.6).

$$I = \int_V mr^2 dm \qquad \textbf{3.6}$$

where *r* is the perpendicular radius from the axis of rotation. Each covariance value in the covariance map (Figure 3.19) is considered as a mass with distance from the origin of the plot (0,0) on Figure 3.19.

The moment of inertia tensor is:

$$\mathbf{I} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \quad \textbf{3.7}$$

where each component of $\mathbf{I}$ represents the moment of inerta of the mass about a different axis. The components of $\mathbf{I}$ can be calculated as the sum over *N* blocks in the covariance map:

$$I_{xx} = \sum_{i=1}^{N} m_i\left(y_i^2 + z_i^2\right) \qquad I_{yy} = \sum_{i=1}^{N} m_i\left(x_i^2 + z_i^2\right)$$

$$I_{zz} = \sum_{i=1}^{N} m_i\left(x_i^2 + y_i^2\right) \qquad I_{xy} = I_{yx} = \sum_{i=1}^{N} m_i x_i \ y_i$$

$$I_{xz} = I_{zx} = \sum_{i=1}^{N} m_i x_i \ z_i \qquad I_{yz} = I_{yz} = \sum_{i=1}^{N} m_i y_i \ z_i$$

The eigenvectors of the matrix in Equation 3.7 correspond to the major and minor directions of continuity while the major direction of continuity corresponds to the eigenvector with the largest eigenvalue (Beer and Johnston 1988). Once these three directions are determined, say directions A, B and C, the magnitude of the anisotropy can also be determined with respect to the three directions (Mohammadhassanpour and Deutsch 2008):

$$I_a = \frac{1}{5}m(d_b^2 + d_c^2) \qquad I_b = \frac{1}{5}m(d_a^2 + d_c^2) \qquad I_c = \frac{1}{5}m(d_a^2 + d_b^2) \quad \textbf{3.8}$$

I_a, I_b and I_c represent the magnitude of the moment of inertia of the mass about each of the 3 principal directions, A, B and C. The anisotropy ratios necessary to define the LVA field are then determined as the ratio of these magnitudes.

The eigen-decomposition of the moment of inertia tensor for a covariance map provides a methodology to automatically generate the LVA field of a variable of interest given available sample data. Specifically, the steps are:

Step 1: Generate the local covariance maps by considering data within a local window.
Step 2: Calculate the moment of inertia tensor (Equation 3.7) for the local covariance map.
Step 3: Determine the eigenvalues and eigenvectors. The eigenvector with the largest eigenvalue corresponds to the major direction of continuity. Anisotropy magnitude is calculated based on the magnitude of the moment of inertia in each direction.

Rather than determine a global direction of continuity for the data in Figure 3.19, consider determining locally varying directions with the suggested methodology. The kriged map is used to generate the locally varying directions as it is smoother than the actual data and results in an LVA field that does not contain abrupt changes due to sparse or noisy data. First, locally varying covariance maps are generated at the desired

LVA field grid resolution (Figure 3.20). In this example the LVA field is generated on a 34 x 22 grid. The moment of inertia tensor is calculated for the covariance map at each grid location and the LVA field is determined as described above.

For each local covariance map the moment of inertia tensor can be calculated, thus, the major direction and magnitude of continuity can be determined. Increasing the size of the moving window smoothes the LVA field, while applying a small window generates an LVA field with more local variations.

Many sources of data for the LVA field (Section 3.2) are not direct measures of the LVA field. The moment of inertia method provides a valuable technique for inferring the LVA field from measurements of intrinsic properties of the deposit such as grade, concentrations, porosity, seismic attributes, etc.

The implementation decisions during estimation of the kriged map have an impact on the resulting LVA field. The goal of mapping is to generate a smooth map that results in a smooth LVA field that captures relevant features. When these features are better characterized by the sample data, the data should be used directly for LVA inference. Care should be taken in sparsely sampled areas where overfitting is an issue. The *apparent* features in the estimated map should be carefully judged and disregarded if they are the result of sparse data showing geologically unjustified features.

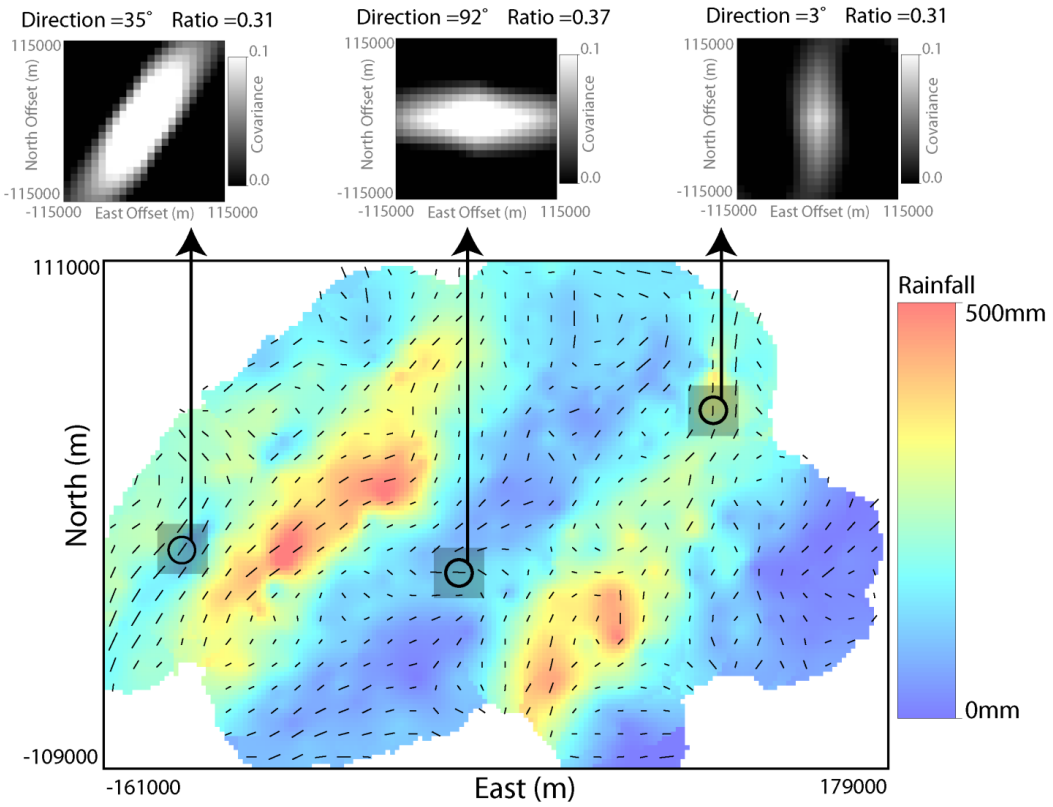


Figure 3.20: LVA field generated by considering a square moving window of +-10 cells. The size of the moving window is indicated by the shaded regions at the three highlighted locations.

3.3.4 Automatic Feature Interpolation

An alternative semi-automatic LVA field generation methodology is based on connecting nearby sample values that have similar values. The orientation of the LVA field is estimated as the azimuth of the line connecting a sample location with a nearby sample location of similar magnitude. Consider generating an LVA field for the data set presented in Figure 3.16. Data values have been joined to nearby data of similar magnitudes (Figure 3.21). The methodology for connecting data is expanded upon; first, it is important to note that there is no unique way to connect the data (Figure 3.21). Many different possible connections are generated and averaged to determine the orientation of the LVA field at each sample location.

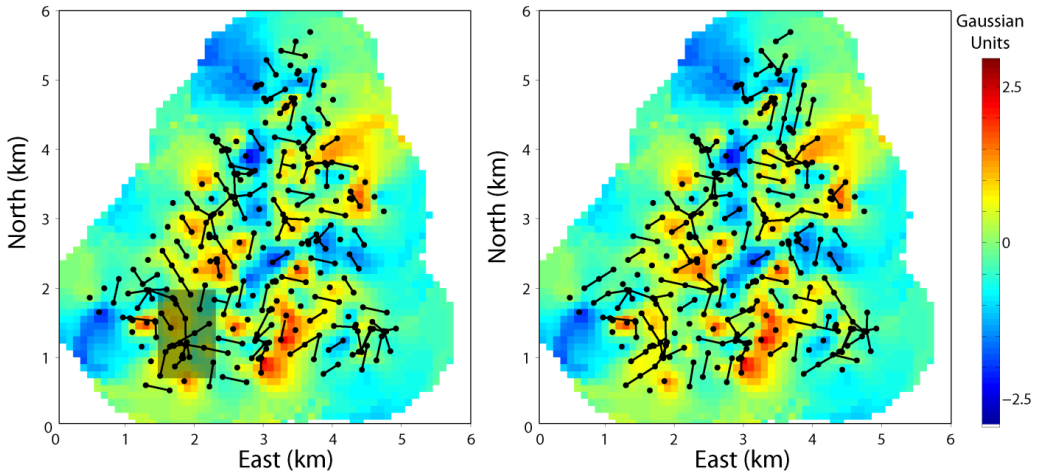


Figure 3.21: Data points of similar magnitude (+-0.5) are joined. There is no unique way to join data points. Two possible configurations are shown. The shaded region on the left plot is used to highlight the connection methodology in Figure 3.22. Variable plotted is Cu concentration after a Gaussian transformation.

The methodology used to connect data points in Figure 3.21 requires three tolerance parameters (1) a distance parameter, Δ, (2) an azimuth parameter, δ, and (3) a magnitude parameter, ξ. Each sample location is visited in a random order and the nearby data within a circular search radius, Δ, are found. The sample data that is most similar in magnitude to the current location is connected to the original location. The polyline direction is controlled by the azimuth tolerance; a change in direction greater than δ is not permitted. This process is repeated to generate pseudo-stream lines connecting the data. The polygon is terminated when either (1) there are no data nearby within ξ to connect to or (2) the polyline intersects an existing polygon, in which case the proposed polyline is clipped. This procedure is highlighted in Figure 3.22. Specifically, the steps to generate a single realization of the connections are:

Step 1: A location is randomly selected to begin the polyline (point A). The location with the most similar data value within a search radius, Δ, is selected to be the next point in the polyline (point B). 'Similar magnitude' is the difference between the sample value at A and the value at B and must be smaller than ξ or the polyline is terminated

Step 2: The connection to point B is made.

Step 3: The azimuth tolerance, δ, and the search radius, Δ, are used to determine potential points to extend the polyline. Within the search area, point C is most similar in magnitude to point B.

Step 4: The polyline is extend to point C.

Step 5: The polyline crosses an existing polyline. The polyline (ABC) is clipped to the existing polyline (123).

Step 6: Return to step 1 until all locations have been visited. The azimuth of the polyline at the data locations provides a sample of the LVA field orientation.

The above steps are repeated with a different random number seed to generate multiple realizations of the connections. These realizations are averaged to provide the LVA field orientation at the sample locations.

Once the data have been connected, the orientation of the LVA field can be determined at each data location. The azimuth of the polyline at each data location (Figure 3.21) is taken to be the orientation of the LVA field at the data location. This procedure is repeated many times (say 1000 times) and the average azimuth value at each data location can be determined (Figure 3.23). Using the kriging of components methodology presented in Section 3.3.1, the exhaustive LVA field can be obtained from the LVA field at the sample locations (Figure 3.23). The LVA field generated with the automatic feature interpolation is compared to the orientation determined manually (Section 3.3.1) to assess any inconsistencies. The comparison (Figure 3.23) shows that in general the automatic interpolation is consistent with *either* manual LVA field 1 or manual LVA field 2 with the exception of two small areas, A and B, that are consistently different. Area A is a location that shows little anisotropy and would receive an anisotropy ratio close to 1.0, thus, the difference in orientation would not be an issue. Area B is located in a transition zone where the surrounding anisotropy is highly erratic; to the north of B anisotropy is East to West, to the south it is North-West to south-East and to the East the anisotropy is highly variable. The difference seen at B is due to the highly erratic nature of the LVA field in this area. Generally, the automatic interpolation method closely matches the manual fields. Manual post processing could be applied to correct any inconsistencies deemed significant based on geological understanding of the deposit (such as area B on Figure 3.23).

The major drawback with the automatic interpolation method is the difficulty extending the technique to 3D. In 3D, the appropriate tolerance parameters become difficult to visualize and select. Moreover, inconsistencies in 3D are more difficult to evaluate in the final LVA field. For automatic 3D LVA field generation the moment of inertia methodology is recommended (Section 3.3.3).

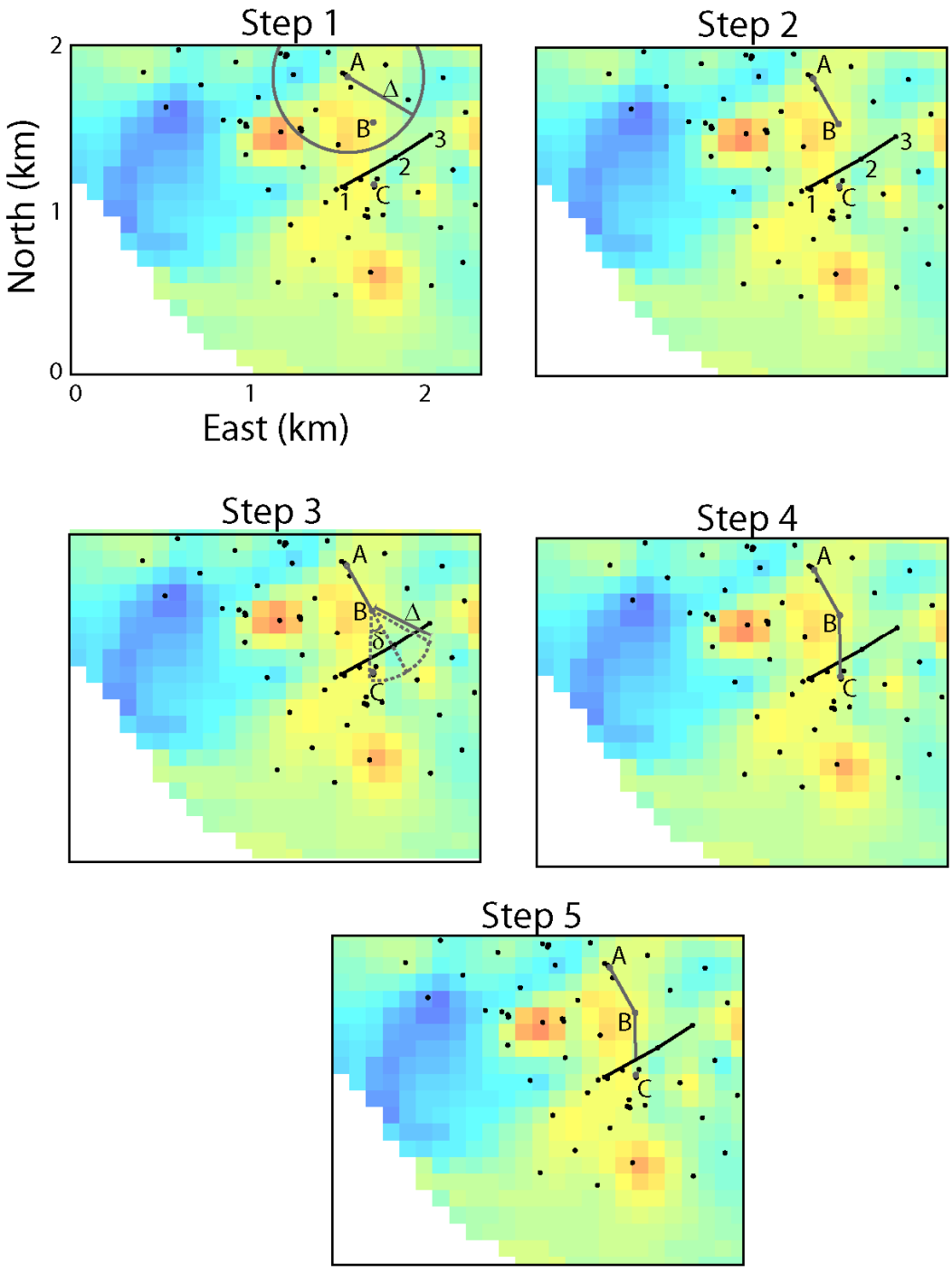


Figure 3.22: Methodology to connect sample data. The region shown is the highlighted area on Figure 3.21. Polygon 1-2-3 was constructed before starting polygon A-B-C.

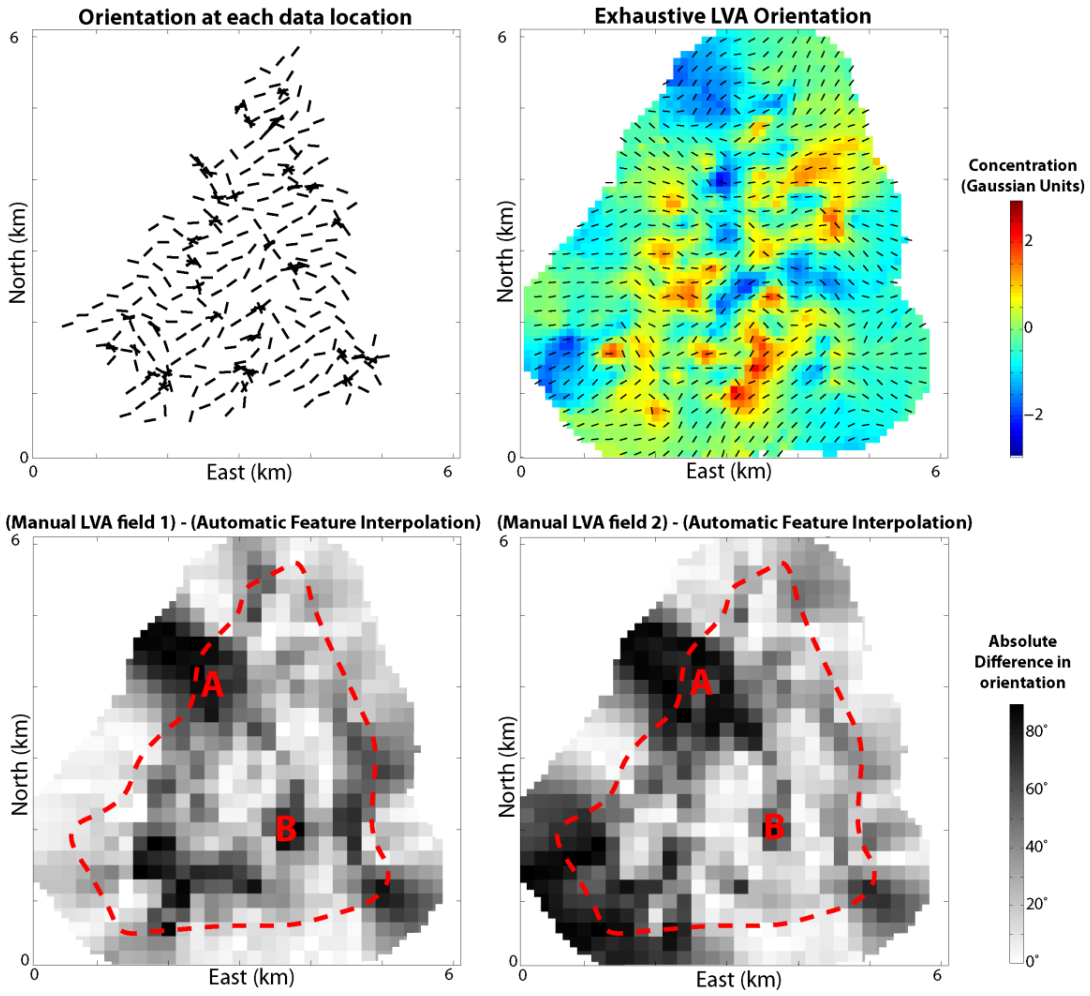


Figure 3.23: Above left: The orientation of the LVA field at each data location obtained by averaging 1000 realizations. Above right: The exhaustive LVA field using the automatic feature interpolation and kriging the components of the orientation (Section 3.3.1). Below: Difference between the orientation of anisotropy using the manual method (Figure 3.17) and the automatic feature interpolation. Dashed line indicates data extents.

3.4 Selecting an LVA Field Generation Method

A number of methodologies to generate the LVA field were presented in Section 3.3. This section presents recommendations to help determine when to apply each method. There are four available techniques:

1) Manual LVA Inference
2) Moment of Inertia
3) Estimation or Simulation
4) Automatic Feature Interpolation

The choice of method influences the resulting LVA field. Consider applying all of the available methods to the Jura data set (Figure 3.24).

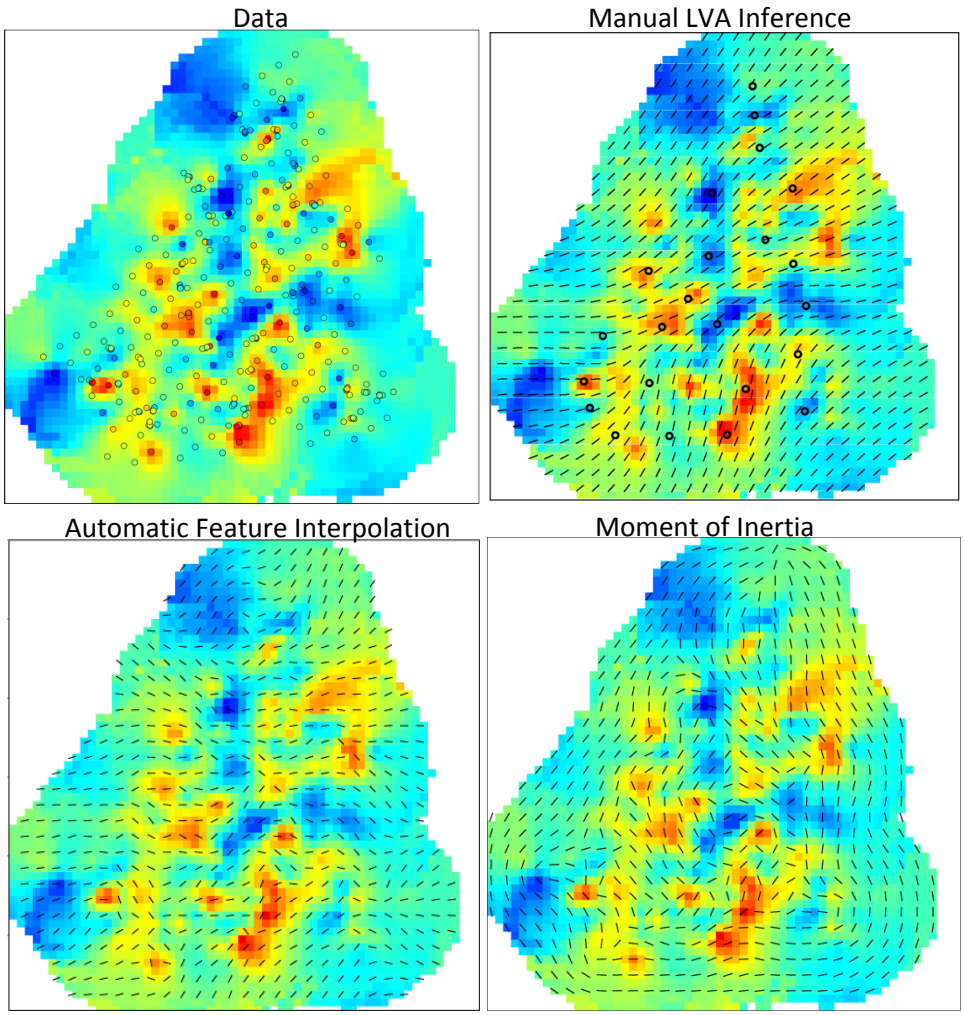


Figure 3.24: Above Left: Data used for LVA field generation, reproduced from Figure 3.16. Above Right: Manual LVA inference at 22 locations, exhaustive field determined from inverse distance estimation. Below Left: Applying the automatic feature interpolation method. Below Right: Applying the moment of inertia method. Color scale indicates Gaussian values from -2 (blue) to +2 (red). Plots cover 6km by 6km.

Of the methods proposed for LVA field generation in Figure 3.24, all seem plausible and consistent with the data with the exception of the moment of inertia method. In this case the data are too erratic for the local covariance maps to provide meaningful information about the LVA field. Visual assessment and a comparison with the available data should always be performed on the potential LVA field.

Beyond the visual inspection of the LVA field, the nature of the available data is the driving factor in determining which method to select. The following discussion considers typical situations and provides guidelines for determining which method to select once the decision to model LVA has been made. A decision tree is presented to summarize the discussion (Figure 3.25).

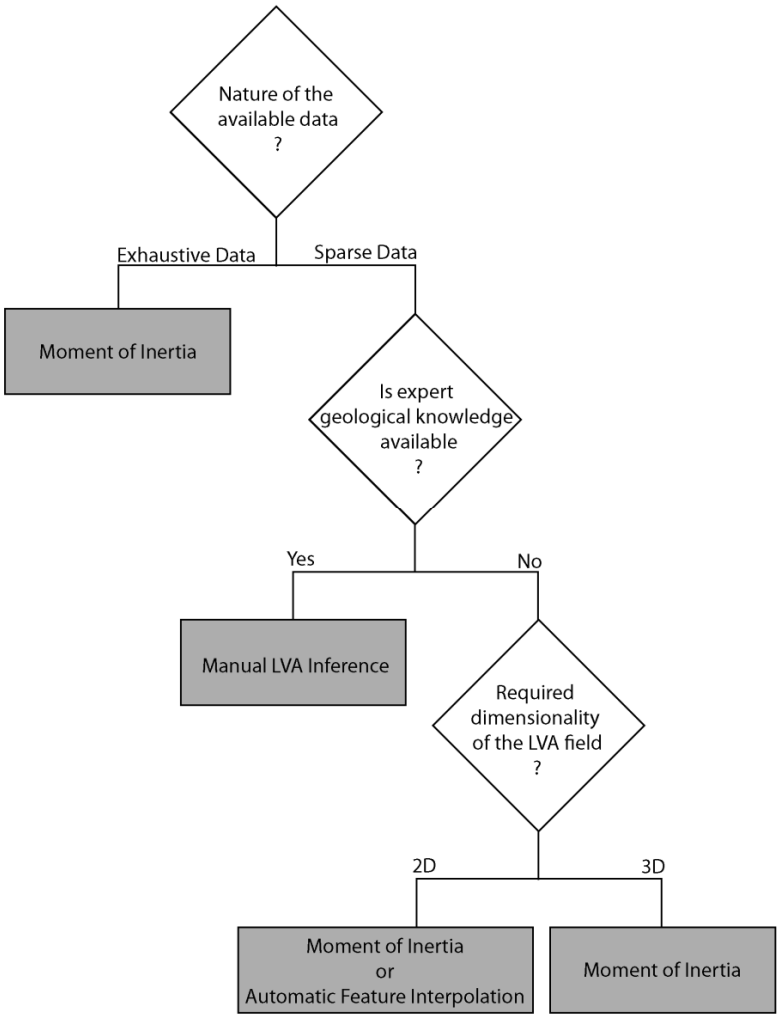


Figure 3.25: Decision tree for the selection of an appropriate LVA field generation methodology. If there is insufficient data to effectively implement one of the given methods an alternative methodology may be required to incorporate LVA.

When direct measurements of the LVA field orientation or magnitude are available, an estimation or simulation technique is used to generate the exhaustive LVA field. Such discrete measurements are available if the LVA field has been interpreted by an expert, if the LVA field has been measured from an outcrop or if the LVA field has been measured by a dipmeter.

If the available data is exhaustive the only automatic technique available is the moment of inertia method. In this situation the moving window size can be varied to obtain the desired features in the LVA field; a large moving window results in a smooth LVA field with few abrupt changes while a small moving window displays more of the local features of the exhaustive data.

If a 3D LVA field is required the manual inference of the LVA field can be difficult, although inferences could be made on a number of horizontal plans and vertical

sections. This results in a number of discreet measurements which can be estimated or simulated to generate the exhaustive 3D LVA field.

The final situation deals with the availability of sample data. If direct measurements of the variable of interest are the only available data source (as in Figure 3.24) the choice is to use either (1) the automatic feature interpolation method or (2) the moment of inertia method. From experience, the automatic feature interpolation method works well when the data are not highly clustered in areas of similar magnitude (i.e. high or low values). Where data are clustered, the method may produce unreliable and erratic LVA fields. The automatic feature interpolation method works well for 2D sample data that are not highly clustered.

3.5 Data Sources

Amoco reservoir data: This data set was used in Chu et al. (1994) and was used for Figure 3.5 and Figure 3.6. The 2D Amoco reservoir data covers an area of about 2 miles square and is the vertical averaging of the porosity of 62 wells. The reservoir is a mixed siltstone-carbonate located in West Texas.

Jura data: This data was used in Goovaerts (1997). There are 359 samples of cadmium, copper, lead, cobalt, chromium, nickel and zinc. The copper data was used in Figure 3.16, Figure 3.17, Figure 3.21, Figure 3.22 and Figure 3.23.

Spatial interpolation contest (SIC) rainfall data: This data set was used in the 1997 spatial interpolation contest and consists of 467 rainfall measurements made on May 8th, 1986 in Switzerland. The SIC data was used in Figure 3.18, Figure 3.19 and Figure 3.20.

Walker Lake data: This data was used in Isaaks and Srivastava (1989). Two variables, U and V, are provided which are functions of the mean and variance of digital elevation data from the Walker Lake region near the California/Nevada border. The exhaustive data set for both the U and V variables is available. There is also a smaller data set consisting of samples of the exhaustive field. The Walker Lake data set was used for Figure 3.3 and Figure 3.2.

3.6 Remarks

This chapter provides a critical step in the incorporation of LVA into geostatistical modeling. The difficulty inferring the LVA field is often the largest impediment when incorporating LVA in numerical models. This chapter provided a number of techniques that can be used to generate the LVA field from a number of different data sources. Moreover, situations exist when it is difficult to accurately infer the LVA field because of sparse samples. In such cases, it is necessary to generate multiple LVA fields with sequential Gaussian simulation to carry forward into the overall modeling workflow.

The remainder of this thesis assumes that the LVA field for the variable of interest is known.

Chapter 4: Calculating the Shortest Anisotropic Distance

This chapter is organized as follows. Section 4.1 discusses the transition from straight line Euclidean distances to minimized anisotropic distances in the presence of LVA. Section 4.2 presents two methodologies to calculate the shortest path between locations. The first method is based on convex optimization and is an iterative algorithm. The method is found to be impractical for large geostatistical models because of high CPU requirements. A second methodology based on graph theory is presented that proves to be faster and able to determine globally optimal paths; although, the paths are geometrically restricted to following discrete edges as defined by the graph.

The methodology proposed in this dissertation is to incorporate LVA into geostatistical models by calculating the shortest path between points in an LVA field. This calculation is the main contribution of this chapter. Chapter 5 utilizes this shortest path calculation in numerical estimation and simulation.

4.1 Moving Away from the Straight Line Path Between Points

A number of practitioners (Deutsch and Lewis 1992; Xu 1996; Sullivan *et al.* 2007) incorporate LVA into geostatistical modeling by considering the variogram to be locally variable; that is, the anisotropy specification for the location being estimated is selected to be locally representative and the straight line path between the estimate and surrounding data is used (Case *i* in Figure 4.1). Partitioning the path into piecewise linear segments, each with a constant anisotropy specification and calculating the anisotropic distance for every cell that the path intersects gives a more accurate measure of the covariance between points because this considers how the LVA field changes between two locations (Case *ii* in Figure 4.1). The methodology in this dissertation takes this further and suggests that a nonlinear path between points results in a shorter distance when anisotropy is considered (Case *iii* in Figure 4.1). There are three options to consider when determining the path between two locations:

Case i: Straight line path with the anisotropy specification obtained from the estimation location.
Case ii: Straight line path with the anisotropy specification obtained from all cells traversed.
Case iii: Nonlinear path with the anisotropy specification obtained from all cells traversed.

The distance between locations is calculated as the sum over all intersected blocks of a piecewise linear path between locations. When anisotropy is globally constant the shortest anisotropic distance between points is obtained with the straight line path. This is not true with LVA. The three cases are applied to a small example to show experimentally that nonlinear paths can be shorter than straight line paths in the presence of LVA (Figure 4.1).

Many authors have employed Case *i* (Figure 4.1) for the incorporation of LVA into geostatistical analysis; however, this does not consider that the LVA field changes between data locations. If the LVA changes smoothly over the field and changes are observed beyond the range of the variogram, considering a constant local anisotropy in each kriging neighborhood may be reasonable.

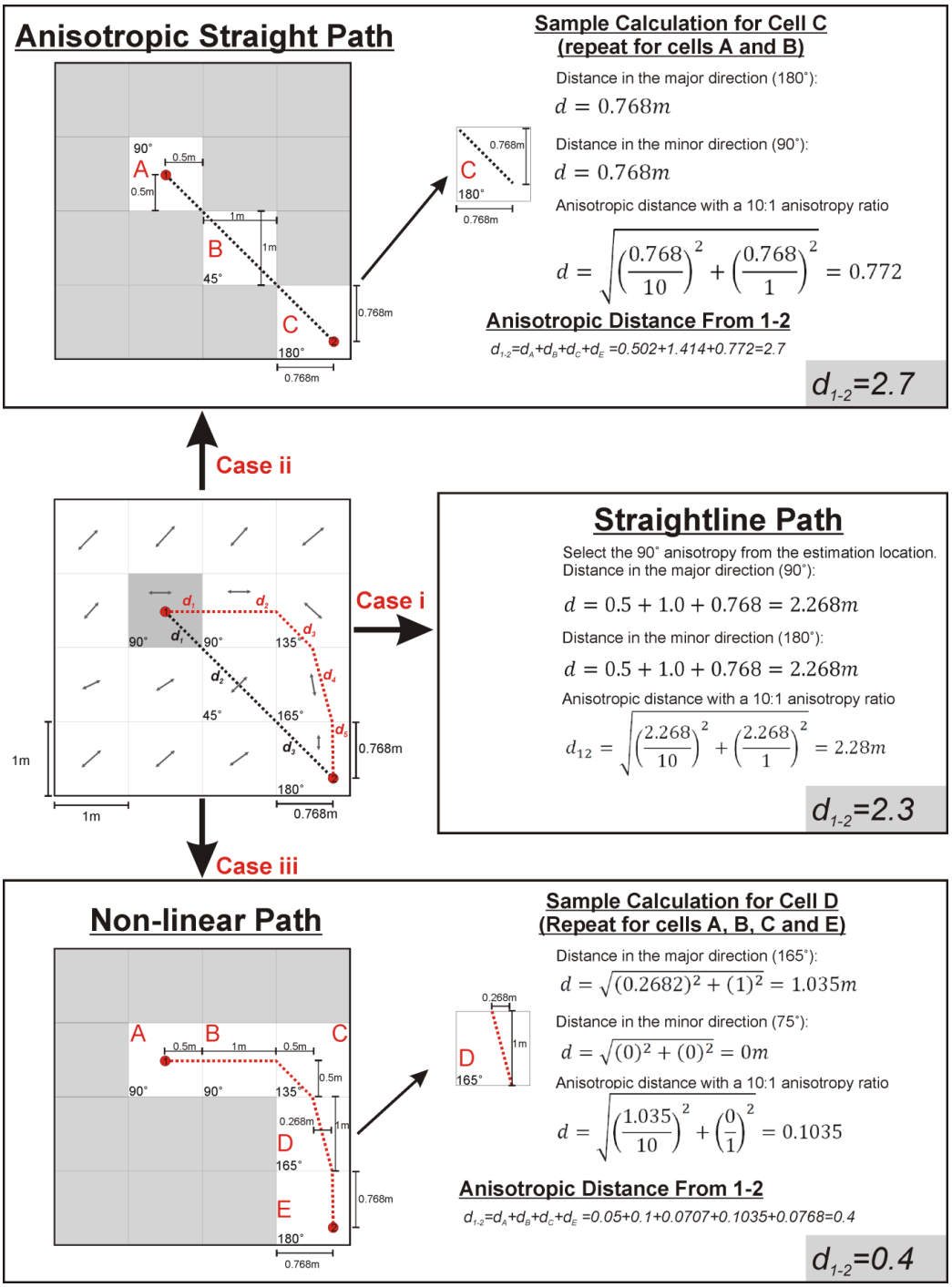


Figure 4.1: Paths between an unsampled location (p_1) and a sampled location (p_2). Case i) using a straight line path and a single direction of anisotropy. Case ii) straight line path with anisotropy defined for each cell. Case iii) nonlinear path with anisotropy defined for each cell intersected.

Selecting the shortest path between two locations results in the highest covariance between points when the covariance is a monotonically decreasing function. Thus, the problem addressed in this chapter is to determine the path between two points that results in the minimum anisotropic distance. Once this optimum path has been found, the resulting distance can be used in kriging and simulation.

4.2 Shortest Anisotropic Distance Algorithms

Optimization techniques aim to find the parameters of a function that lead to the optimal solution. Often times, this function is complex which limits the available algorithms. With the shortest path problem, the function to be minimized is the anisotropic distance between two points (Equation 4.8). Some optimization methods require the first and second derivatives of the objective function; therefore, if the objective function is twice differentiable a larger range of algorithms become available. Section 4.2.1 develops the objective function, its first and second derivatives and applies an iterative convex optimization algorithm, the guarded Newton method (Boyd and Vandenberghe 2004), to determine the anisotropic distance.

A sub-branch of numerical optimization, routing, is of particular interest (Toth and Vigo 2002; Azaron and Kianfar 2003). Ship or plane routing in changing wind/current patterns and calculating the shortest distance in the presence of LVA are similar problems; however, one major difference is the number of calculations that must be accomplished. Millions of paths must be optimized in a typical geostatistical modeling application rather than a relatively small number of vehicle routes. Despite these differences, the field of routing provides useful insight into how to solve the shortest path problem. There are many different approaches applied to routing but most often a graph is created for potential routes and the shortest path is solved using a graph theory algorithm (Laporte and Osman 1995; Anel 2005; Padhy *et al.* 2008). Section 4.2.3 reformulates the shortest path problem as a graph and applies the Dijkstra algorithm (Dijkstra 1959) to calculate the anisotropic distance.

4.2.1 Optimization: The Guarded Newton Method

Optimization techniques are applicable if an appropriate objective function is available. The objective function for this problem is the anisotropic distance between two locations (Equation 4.7). The difficulty with formulating the shortest path problem as an objective function to be minimized is that the objective function is not convex. Determining the global solution to such problems is difficult and CPU intensive (Boyd and Vandenberghe 2004). Because of the large number of paths that require optimization, CPU intensive global optimization schemes are not considered. The initial path is the straight line path between the two locations and successive iterations attempt to optimize the path until the local minima is found. The limitation of this assumption is that there could exist an alternative shorter path that is not discovered because the local minimum is not necessarily the global minimum. Moreover, it will be

shown that the CPU demands for the local solution are too high for local optimization to be considered a practical solution to the shortest path problem (Figure 4.9). The global solution would require additional CPU time and is thus impractical as well.

The Newton method is an optimization algorithm applicable to functions that are convex and twice differentiable (Boyd and Vandenberghe 2004; Sun and Yuan 2006). Although the shortest path problem is not convex, it is locally convex and the Newton method discovers a local minimum. The Newton method minimizes a function, *f*, using its second order Taylor approximation, $\hat{f}$, with ∇ the first derivative or Jacobian of *f* and ∇^2 the second derivative or Hessian of *f*.

$$\hat{f}(x+u) = f(x) + \nabla f(x)^T u + \frac{1}{2} u^T \nabla^2 f(x) u \qquad \textbf{4.1}$$

This approximation is a second order function of *u*. The minimum of this approximation occurs when $u = \Delta x_n$, the Newton step:

$$\Delta x_n = -\left[\nabla^2 f(x)\right]^{-1} \nabla f(x) \qquad \textbf{4.2}$$

When *f* is a quadratic function, the Newton step exactly minimizes it; otherwise it is a good approximation to the minimum of a non-convex function such as the anisotropic distance. In the case of a non-convex function, the Newton method is developed as an iterative algorithm starting with iteration x_{k-1}:

$$\begin{aligned} \hat{f}(x_k) &= f(x_{k-1}) + \nabla f(x_{k-1})^T u_k + \frac{1}{2} u_k^T \nabla^2 f(x_{k-1}) u_k \\ u_k &= \Delta x_{nk} = -\left[\nabla^2 f(x_{k-1})\right]^{-1} \nabla f(x_{k-1}) \\ x_k &= x_{k-1} + u_k \end{aligned} \qquad \textbf{4.3}$$

This is the pure Newton method; there is no control on where the step Δx_{nk} positions x_k for the next iteration. For strongly convex functions when x_{k-1} is near the minimum this is usually not a problem. For non-convex functions when it is unknown that x_{k-1} is near the minimum, the Newton step can position x_k further from the minimum than x_{k-1}. It becomes a divergent process, which is corrected by adding step control (*t*) to Δx_{nk} :

$$x_k = x_{k-1} + t u_k = x_{k-1} + t \Delta x_{nk} \qquad \textbf{4.4}$$

This is the guarded or dampened Newton method (Boyd and Vandenberghe 2004). Determination of *t* is accomplished with a line search algorithm in the direction of the Newton step. There are various line search algorithms available such as the golden section search, successive parabolic interpolation or the backtracking line search. These

and other line search methods are discussed in Chapter 2 in Sun and Yuan (2006). The line search in the guarded Newton method finds the step size *t* in the direction of Δx_{nk} such that the 1D function $f(x_{k-1}+t\Delta x_{nk})$ is a minimum (Figure 4.2). For this implementation the backtracking line search was used.

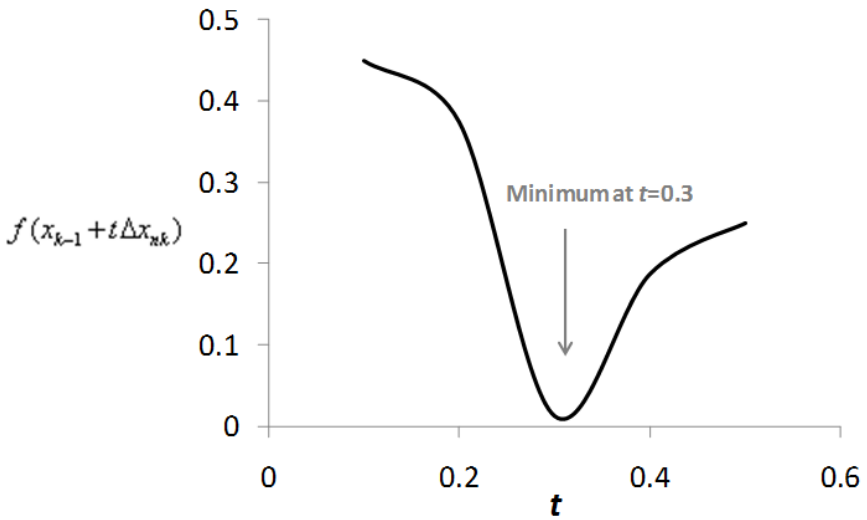


Figure 4.2: A line search algorithm discovers the step size, *t*=0.3, that results in the minimization of the function along the direction, Δx_{nk}.

The algorithm for the guarded Newton method is:

Step 1: Newton step: solve $\Delta x_{nk} = -\left[\nabla^2 f(x_{k-1})\right]^{-1} \nabla f(x_{k-1})$

Step 2: Line search: determine *t* with the backtracking line search

Step 3: Update: $x_k = x_{k-1} + t\Delta x_{nk}; k = k+1$

This optimization technique is applied to the problem of shortest anisotropic distance with LVA. In the above functions, *x* is interpreted as the path between two locations and *f* is the length of the path. Successive iterations, x_k, represent different paths with decreasing anisotropic distances.

The distance between locations depends on the LVA field. The field's complexity may lead to a distance function that has numerous local minima – there is no guarantee that the guarded Newton method determines the global minimum but it does find a local minimum. Visually, the shortest path problem is to determine the optimal red path in Figure 4.3.

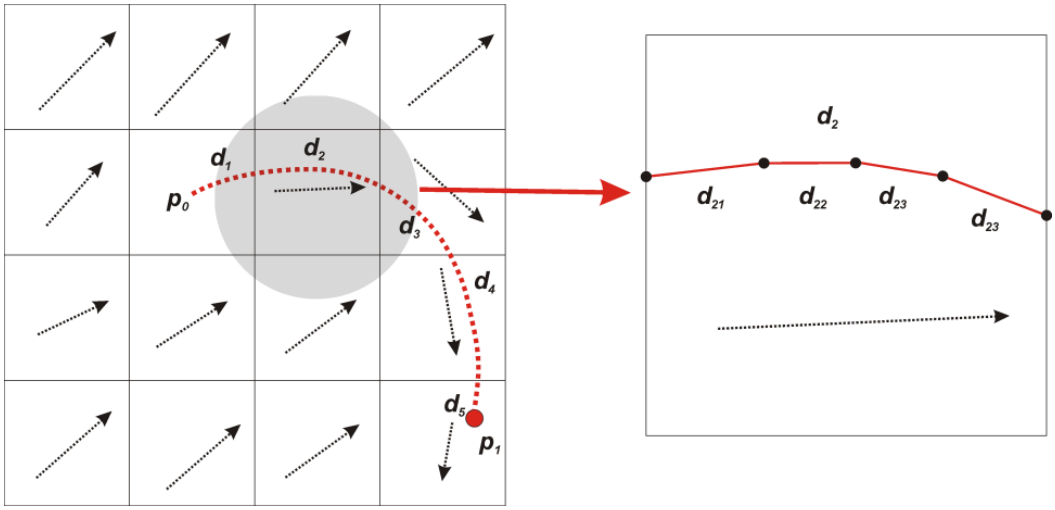


Figure 4.3: Optimized path in the presence of LVA. In a single block (right) where the anisotropy is constant, the path can be considered a series of linear segments. Calculating the distance of a single segment and summing all segments gives the distance d_2.

The equation for the distance between two locations is required; the distance is the summation of the linear segments making up the shortest path, $d_1+d_2+d_3+d_4+d_5$ (Figure 4.3). Equation 4.5 gives the squared distance of a segment when the anisotropy is constant (as in a single grid cell).

$$d_{ij}^{\,2}\left(\mathbf{p}_0,\mathbf{p}_1\right)=\left(\mathbf{p}_1-\mathbf{p}_0\right)^T R^T R\left(\mathbf{p}_1-\mathbf{p}_0\right) \qquad \textbf{4.5}$$

where the rotation matrix R is,

$$R=\begin{bmatrix} \cos\beta\cos\alpha & \cos\beta\sin\alpha & -\sin\beta \\ \frac{1}{r^1}\left(-\cos\gamma\sin\alpha+\sin\gamma\sin\beta\cos\alpha\right) & \frac{1}{r^1}\left(\cos\gamma\cos\alpha+\sin\gamma\sin\beta\sin\alpha\right) & \frac{1}{r^1}\left(\sin\gamma\cos\beta\right) \\ \frac{1}{r^2}\left(\sin\gamma\sin\alpha+\cos\gamma\cos\beta\cos\alpha\right) & \frac{1}{r^2}\left(-\sin\gamma\cos\alpha+\cos\gamma\sin\beta\sin\alpha\right) & \frac{1}{r^2}\left(\cos\gamma\cos\beta\right) \end{bmatrix} \qquad \textbf{4.6}$$

$$D^2\left(\mathbf{p}_0,\mathbf{p}_1\right)=\sum_{j=1}^{m}\left(t_j-t_{j-1}\right)^2\left(\mathbf{p}_1-\mathbf{p}_0\right)^T R_j^T R_j\left(\mathbf{p}_1-\mathbf{p}_0\right)=\sum_{j=1}^{m}\Delta t_j^2\mathbf{v}^T A_j\mathbf{v} \qquad \textbf{4.7}$$

$$\text{where } \Delta t_j^2=\left(t_j-t_{j-1}\right)^2,\ \mathbf{v}=\mathbf{p}_1-\mathbf{p}_0,\ A_j=R_j^T R_j,\ \textit{and}\ \ \mathbf{p}=\mathbf{p}_0+t\left(\mathbf{p}_1-\mathbf{p}_0\right)=\begin{bmatrix} x_0+t(x_1-x_0) \\ y_0+t(y_1-y_0) \\ z_0+t(z_1-z_0) \end{bmatrix}$$

$\mathbf{p}_0$ and $\mathbf{p}_1$ are the xyz coordinates defining a segment and d_{ij} is a single segment as shown in Figure 4.3. t is a control parameter that varies between 0 at $\mathbf{p}_0$ and 1 at $\mathbf{p}_1$. Equation 4.7 is obtained by rewriting Equation 4.5 as a summation over all m line segments.

To find the minimum squared distance between $\mathbf{p}_0$ and $\mathbf{p}_1$, an initial straight line path is assumed and the path is iteratively perturbed by adding and adjusting control points until the optimal path is found (Figure 4.4).

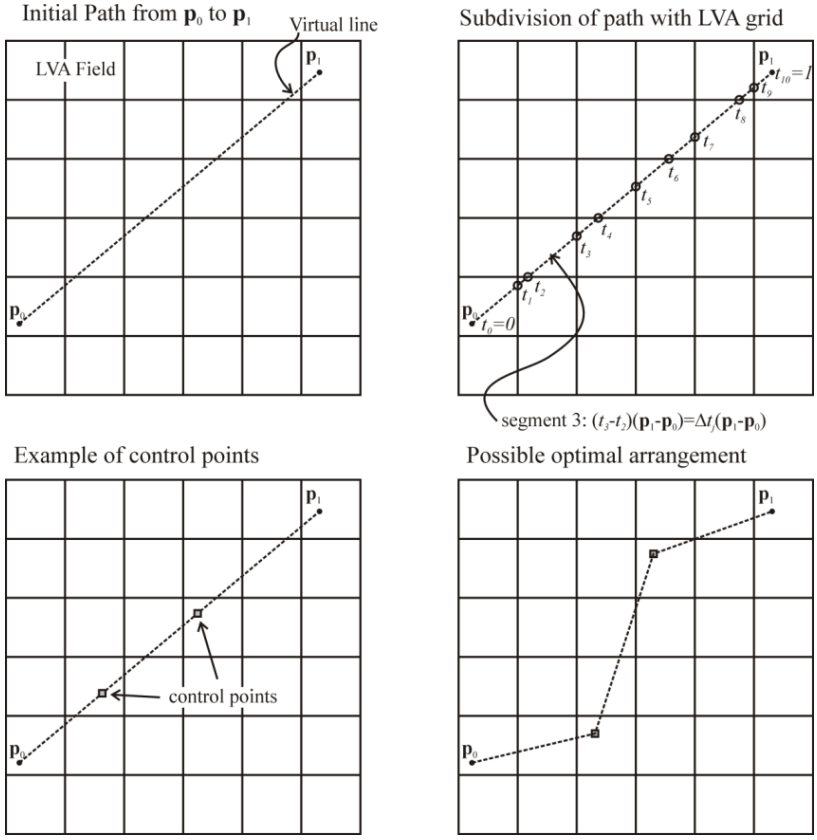


Figure 4.4: Components of minimizing the anisotropic distance.

For *n* control-points, Equation 4.7 is re-written as the final objective function to be used in optimization (Equation 4.8). The goal is to find the positions for the control-points such that the objective function (Equation 4.8) is minimal. This is accomplished with the guarded Newton method, which requires the Jacobian (J) and Hessian (H) matrices of the objective function, defined by Equations 4.9 and 4.10 respectively. Note that the objective function, $D^2(\mathbf{p_0},\mathbf{p_1})$, is non-convex because the rotation matrix, $A_{j(i)}$ in Equation 4.8 varies in every cell of the model.

$$D^2\left(\mathbf{p}_0,\mathbf{p}_1\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}\Delta t_{j(i)}^2\mathbf{v}_{(i)}^T A_{j(i)}\mathbf{v}_{(i)} \qquad \textbf{4.8}$$

$$J=\frac{\partial D}{\partial \mathbf{v}_{(i)}}=2\sum_{j=1}^{m_{(i)}}\Delta t_{j(i)}^2 A_{j(i)}\mathbf{v}_{(i)}-2\sum_{j=1}^{m_{(i+1)}}\Delta t_{j(i+1)}^2 A_{j(i+1)}\mathbf{v}_{(i+1)} \quad i=1,\ldots,n \qquad \textbf{4.9}$$

$$H = \frac{\partial^2 D}{\partial \mathbf{v}_{(i)}{}^2} = 2\sum_{j=1}^{m_{(i)}} \Delta t_{j(i)}^2 A_{j(i)} + 2\sum_{j=1}^{m_{(i+1)}} \Delta t_{j(i+1)}^2 A_{j(i+1)}$$

$$\frac{\partial^2 D}{\partial \mathbf{v}_{(i)} \partial \mathbf{v}_{(i+1)}} = -2\sum_{j=1}^{m_{(i+1)}} \Delta t_{j(i+1)}^2 A_{j(i+1)} \qquad i = 1,\ldots,n \qquad \textbf{4.10}$$

$$\frac{\partial^2 D}{\partial \mathbf{v}_{(i)} \partial \mathbf{v}_{(i-1)}} = -2\sum_{j=1}^{m_{(i)}} \Delta t_{j(i)}^2 A_{j(i)}$$

In the guarded Newton method, the Jacobian and Hessian matrices are used to calculate a Newton step (Δ**v**), which is a vector that describes the spatial translation to be applied to the control-points. Calculation of Δ**v** for iteration *k* is done by solving the system in Equation 4.11. The backtracking line search (Sun and Yuan 2006) is used to calculate the magnitude (α) of the step taken in Equation 4.12. Implementing a line search prevents translation of control-points into undesirable configurations, such as intersecting loops or into positions beyond the LVA field. More control points are added and translated according to the Newton step and the process repeats until no improvement in distance can be obtained.

$$H \cdot \Delta \mathbf{v} = -J \qquad \textbf{4.11}$$

$$\mathbf{v}^{k+1} = \mathbf{v}^k + \alpha \cdot \Delta \mathbf{v}^k \qquad \textbf{4.12}$$

Depending on the nature of the LVA field the necessary number of control points can vary; two extreme cases may arise (1) anisotropy is nearly constant everywhere in the field or (2) anisotropy is highly variable in orientation and range. In the first case, the minimum anisotropic distance can be found with few control-points, possibly with only one. However, the second scenario may demand a control-point spacing equivalent to the LVA grid spacing.

To account for this variability, control-points are inserted up to a maximum of the number of LVA blocks intersected between any two locations of interest. After adding additional control-points, a new solution is found. Optimization stops and control-points are no longer added if the decrease in distance is less than ε (~square-root of machine epsilon).

Minimizing the anisotropic distance between any two locations **u** and **v** in an LVA field is accomplished with the following algorithm:

Step 1: Calculate initial distance between **u** and **v**, $D(\mathbf{u},\mathbf{v})$.
Step 2: Set iteration k=0, change in distance $\Delta D = D^k - D^{k-1}$ and the number of control-points N_{cp}=0.
Step 3: While $\Delta D(\mathbf{u},\mathbf{v}) > \varepsilon$ and $N_{cp} < max_{cp}$:

a. Insert 2^k control-points, where k is the current iteration, $N_{cp} = N_{cp} + 2^k$.
b. Calculate Jacobian, Hessian and Newton step for all control-points.
c. Calculate step parameter α using backtracking (Sun and Yuan 2006).
d. Apply the Newton step.
e. Increment k.

A visual representation of how the gradient is used to determine the shortest path is illustrated in Figure 4.5.

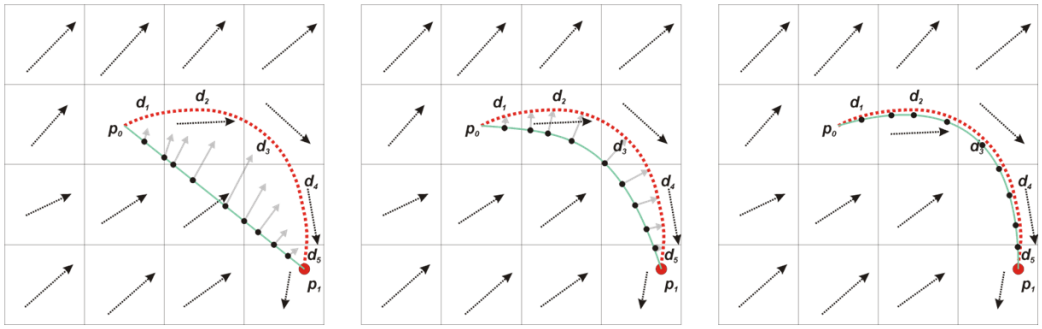


Figure 4.5: Optimizing the path between p_0 and p_1. The red path is optimal. Left: Initial path (green) is straight. The gradient (gray arrows) indicates which direction to translate the nodes in order to obtain a solution with a smaller anisotropic distance. Middle: The control points are moved in the direction of the gradient and the resulting green path represents a shorter anisotropic distance. The gradient is recalculated for the new green path Right: The final path is close to optimal.

4.2.2 Guarded Newton Method Examples

Example 1: Consider two points and the LVA field in Figure 4.6. The LVA field is defined by the arrows shown on the plot and the different paths correspond to the optimal paths considering different anisotropy ratios. If the anisotropy ratio is 1:1 there is no benefit in altering the path and the straight-line path results in the maximum covariance. As the anisotropy ratio increases it becomes progressively more beneficial to consider a nonlinear path (Table 4.1). Higher anisotropy ratios result in paths that deviate further from the straight-line path (Figure 4.6).

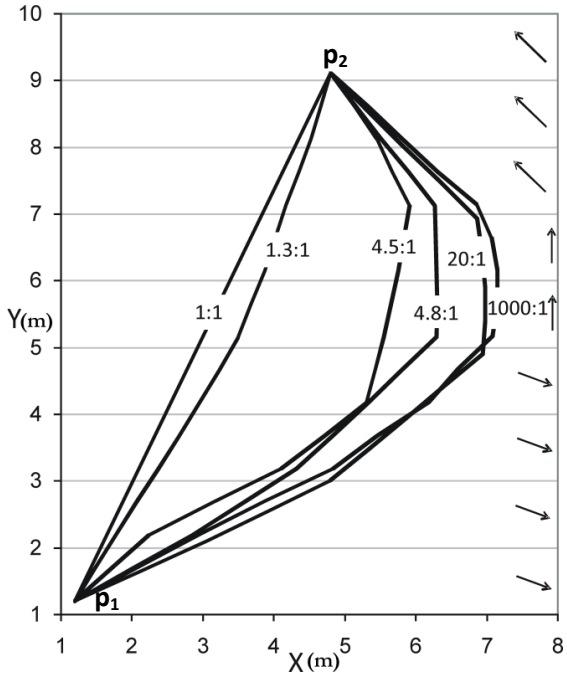


Figure 4.6: Optimized path between p_1 and p_2 with various anisotropy ratios. The major direction of continuity for each row is constant and indicated by the arrows.

Example 2: Consider the four point data configuration and the LVA field in Figure 4.7. Two cases are considered. First, Case 1 considers that the straight line path is used between the points and the LVA field is used to calculate the distances of each segment (d_1+d_2+d_3). Case 2 considers optimizing the path between points. Table 4.2 shows all distances between points for both cases; there is a decrease in the anisotropic distance when the nonlinear path is used. If an exponential variogram with no nugget effect and a range of 50m is used, the associated covariances are calculated (Table 4.2) for estimating at location 1 with data at locations 2, 3 and 4. Solving the resulting system of equations gives the kriging weights in Table 4.3. As expected, using the optimized path increases covariance between points and has a large effect on the kriging weights.

Table 4.1: Optimized and straight-line distances between data in Figure 4.6.

Anisotropy	$h_{straightline}$	$h_{optimized}$	Difference
1:1	8.7	8.7	0.0%
2:1	15.34	15.31	-0.2%
4:1	29.25	29.21	-0.1%
5:1	36.32	31.42	-13.5%
10:1	71.94	59.66	-17.1%
1000:1	7170	6147	-14.3%

Table 4.2: Anisotropic distances between locations for Example 2. Location 1 is to be estimated with data at locations 2, 3 and 4. Covariance is calculated using an exponential variogram with a range of 50m.

Locations	$h_{straight-line}$	$h_{optimized}$	$C_{straightline}$	$C_{optimized}$
$p_1 - p_2$	36.8	30.3	0.11	0.16
$p_1 - p_3$	18.1	15.2	0.34	0.40
$p_1 - p_4$	40.1	33.5	0.09	0.13
$p_2 - p_3$	29.6	29.1	0.17	0.17
$p_2 - p_4$	48.0	48.0	0.056	0.056
$p_3 - p_4$	18.6	18.5	0.33	0.33

Table 4.3: Weights (λ) obtained for Example 2 using kriging. Covariance values determined in Table 4.2.

Location	$\lambda_{straight-line}$	$\lambda_{optimized}$	Difference
p_2	0.0543	0.0952	43%
p_3	0.336	0.384	14%
p_4	-0.0229	0.00193	108%
sum	*0.37*	*0.48*	*30%*

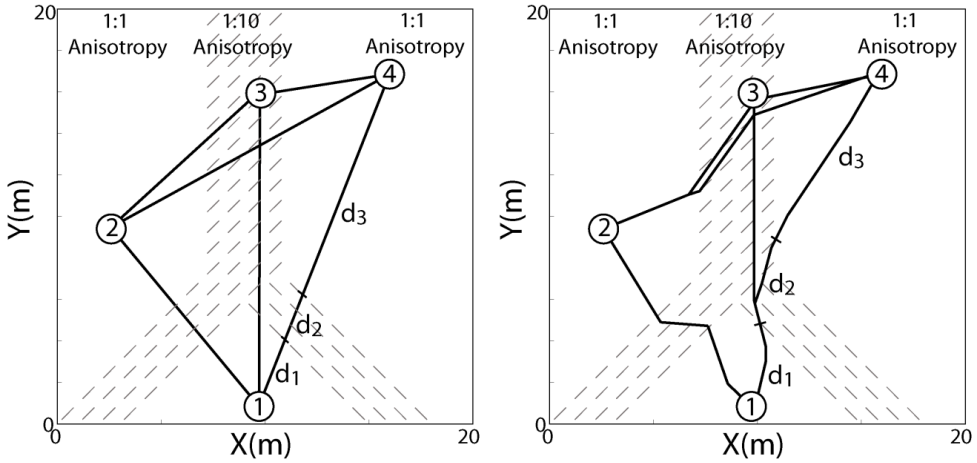


Figure 4.7: Locations 2,3 and 4 are data. Sample calculation: distance between points 1 and 4 is the sum of the segments d_1+d_2+d_3. Left: Case 1, straight line paths. Right: Case 2, optimized paths.

While the guarded Newton method succeeds at finding a locally minimum nonlinear path between points that is shorter than the straight line path, the method is not practical for typical geostatistical modeling purposes because of CPU requirements. The number of distance calculations in a typical geostatistical model is large; consider a moderate 1,000,000 cell model where the nearest 20 data are used for local estimation at each cell. If kriging is applied, the distances between the 20 local data are required (20·19/2 distances) in addition to the 20 distances between the data and the estimation location in question, for a total of 210 distances per grid cell. In total, 1,000,000·210 = 210M distances are necessary for this example. 200M distance calculations is used as a benchmark for the following CPU time study; to be practical, the distance calculations should take less than a few days.

For this time trial 10 random smooth LVA fields are generated by kriging the components of random control points that consist of an LVA angle and anisotropy ratio value. Two of the 10 LVA fields are shown in Figure 4.8. The underlying LVA field does have some effect on speed but even the most optimistic case requires 69 days to calculate the necessary distances. Moreover, considering a 3D example adds additional CPU time not considered in Figure 4.9.

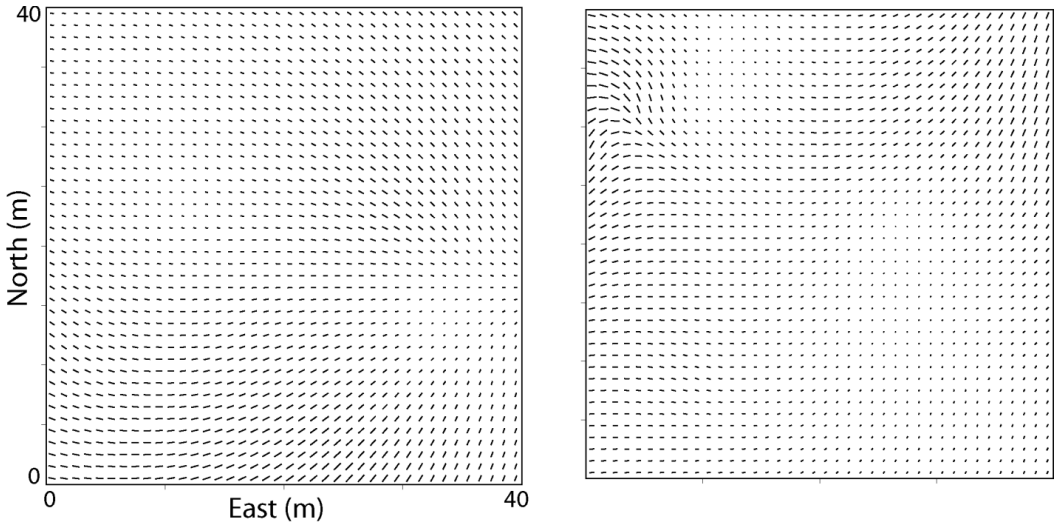


Figure 4.8: 2 of the 10 randomly generated LVA fields. The length of the line is proportional to the magnitude of the anisotropy (i.e. the anisotropy ratio).

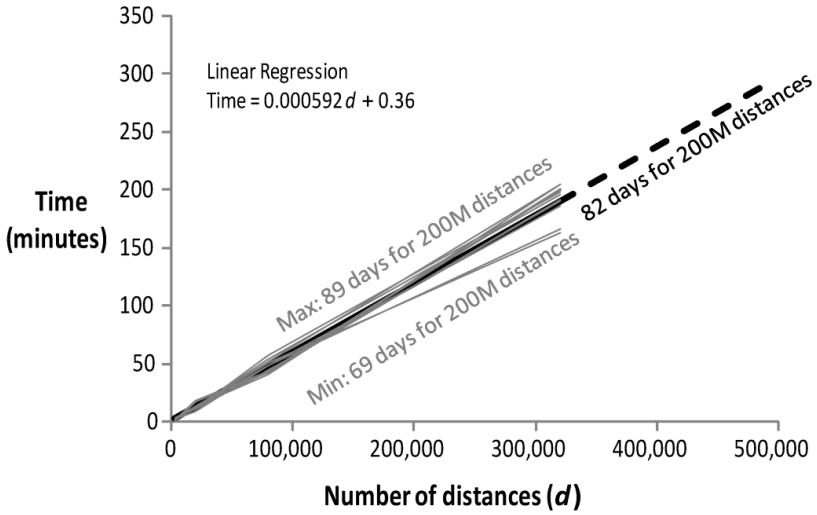


Figure 4.9: Time required to calculate the required number of distances (*d*). Each gray line represents a randomly generated LVA field. The thick black line is a linear fit based on the average time.

In addition to the excessive CPU requirements of the guarded Newton algorithm, this methodology is also unable to determine a globally optimal path; recall the initial assumption behind this application is that the locally optimal path, rather than the globally optimal path, produces sufficient results. Consider the LVA field and associated optimal paths found with the guarded Newton method in Figure 4.10. The optimal paths found are local minima which are far from the globally optimal paths found with the graph theory approach (Section 4.2.3). Global optimization, such as random restart or simulated annealing, could be applied to find better global solutions but this would come at the cost of increased CPU time, which is already an issue when determining the local solution. The drawbacks of the guarded Newton method necessitate a departure from convex optimization techniques.

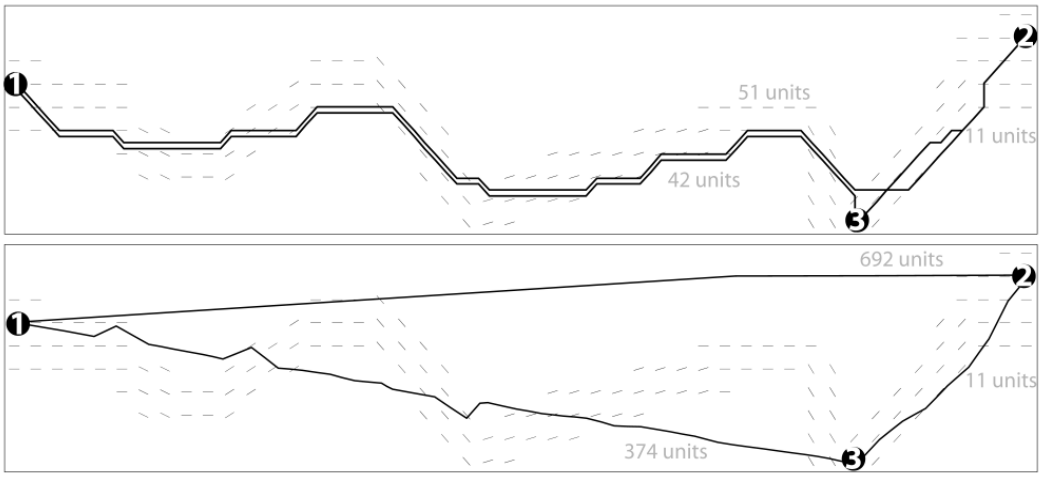


Figure 4.10: Example optimal paths using a channel LVA field indicated by the gray lines. Anisotropy ratio is 10:1. Above – Using the Dijkstra algorithm (Section 4.2.3) Below – Using the guarded Newton method the global solution is not discovered. Models are nominally 50 cells by 10 cells.

4.2.3 Graph Theory

Graph theory originated in 1736 when Euler proposed the Seven Bridges of Konigsberg problem (Euler 1736). The shortest path problem is well researched in the context of graph theory. A graph connects vertices (grid cell centers) through edges which are traversed to determine the cost between any two vertices (Figure 4.11).

Rather than consider the shortest path problem as an optimization issue with an objective function that requires minimization (Section 4.2.1), the problem is reformulated as a graph problem. Consider the graph in Figure 4.11 with the shortest path indicated in blue. The edges between vertices each have an associated distance. The distance between vertices is the summation of the individual edge distances.

A disadvantage of using this graph formulation is the limitation on the shape of the path between two locations. Because vertices must be linked with individual edges, there is a geometric restriction placed on permissible paths (Figure 4.11 middle). This limitation can be mitigated if additional vertices or edges are added to the graph, increasing the density of the vertices and increasing the flexibility of the paths (Figure 4.11 right).

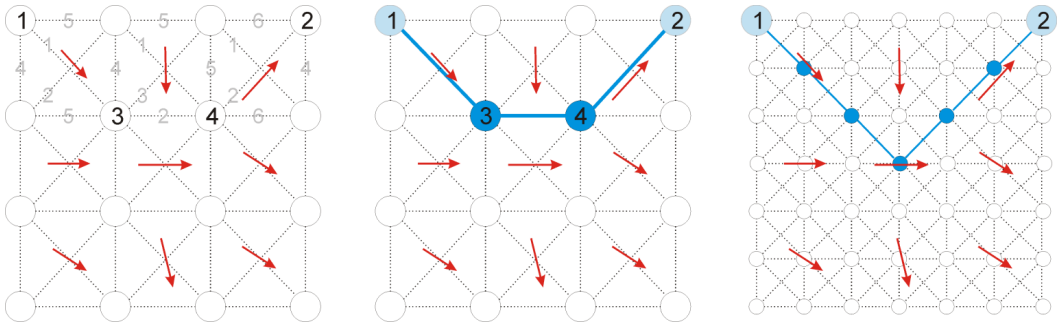


Figure 4.11: A graph representation of the shortest path problem in the presence of LVA. Left: Rather than representing a model as a sugar cube grid, a model is represented as a graph with edges between vertices. Some relevant unit distances are indicated. Middle: A permissible path between vertices must traverse along edges. The shortest distance between vertex 1 and 2 would be through vertices 3 and 4. Right: Refining the graph allows for more flexibility in the shape of the path.

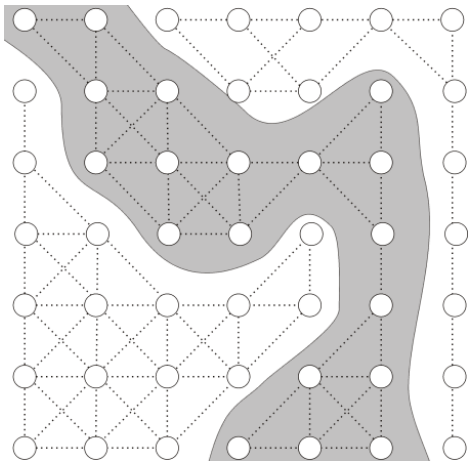


Figure 4.12: Representation as a graph makes complicated domaining possible without explicitly generating multiple models (one inside the channel and one outside).

One advantage of the graph reformulation is the ease of integrating constraints into the problem; for example, if a channel system exists and is treated as a separate domain, the vertices can be disconnected to prevent paths from crossing domains (Figure 4.12). A second advantage is the large number of algorithms available for solving the shortest path including: Dijkstra algorithm (Dijkstra 1959; Cormen *et al*. 2003); Bellman-Ford algorithm (Cormen *et al*. 2003; Medhi and Ramasamy 2007); A* search algorithm (Hart *et al.* 1968); use of a genetic algorithm (Ahn and Ramakrishna 2002 which contains references to many other genetic algorithm approaches to the shortest path routing problem); use of neural networks (Ali and Kamoun 1993; Bu and Chiueh 1999); Johnson algorithm (Johnson 1977). This list of algorithms is not exhaustive. Cherkassky *et al*. (1996) present a good comparison of algorithms for solving the shortest path problem.

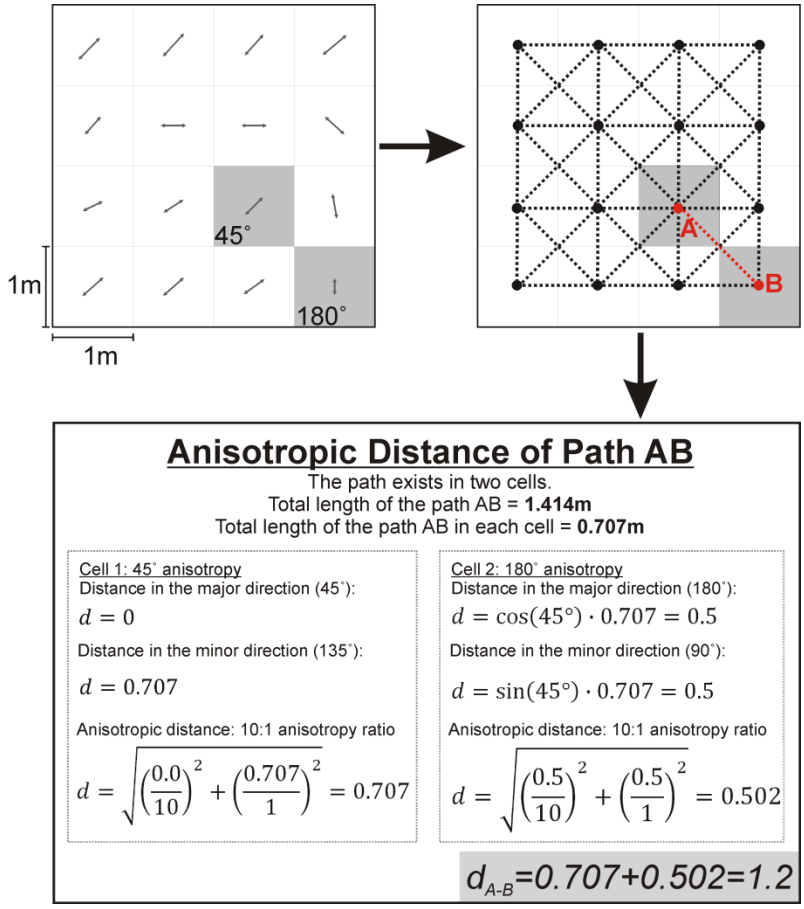


Figure 4.13: Upper Left: The LVA field from Figure 4.1. Anisotropy ratio is a constant 10:1. Upper Right: The grid is redefined as a graph with vertices (black circles) at the grid cell centers with edges (dashed lines) connecting the vertices. Below: Methodology for calculating the distance of edge AB.

In order to utilize a graph algorithm for solving the shortest path problem, the graph must first be constructed which involves defining the vertices and calculating edge lengths. Consider the LVA field from Figure 4.1 with vertices located at grid cell centers, a constant anisotropy ratio of 10:1 and vertices linked to adjacent vertices only (Figure 4.13). In this example there are 16 vertices and 42 edges. The length of each edge must

be calculated; note that each edge exists in two grid cells and its length is the summation of the portions of the edge in each cell. Figure 4.13 explains the calculation of one edge length in a graph. The final graph is constructed by repeating this process for each edge.

Once graph construction is completed the problem is to calculate the shortest path between any two vertices. In this dissertation, the Dijkstra algorithm (Dijkstra 1959) is implemented because it is very common and efficient public implementations are available (Siek *et al.* 2001). The Dijkstra algorithm determines the shortest path from A to B by spiraling out from A until B is reached. All possible paths are stored in memory and the first path to reach B is the shortest path. Figure 4.14 shows a small example with the optimal distances and paths shown between three points in a channel LVA field. The restriction of the path to following edges is apparent. The Dijkstra algorithm is explained in detail in Figure 4.15.

One significant advantage of the Dijkstra algorithm is the ability to simultaneously solve for multiple shortest paths in a single pass. Consider the final path in Figure 4.15. In addition to solving for the optimal path between vertices A and B, the optimal paths have also been determined between vertex A and all vertices highlighted blue. Thus, in a single pass with the Dijkstra algorithm the shortest distances from A to all 15 vertices are determined.

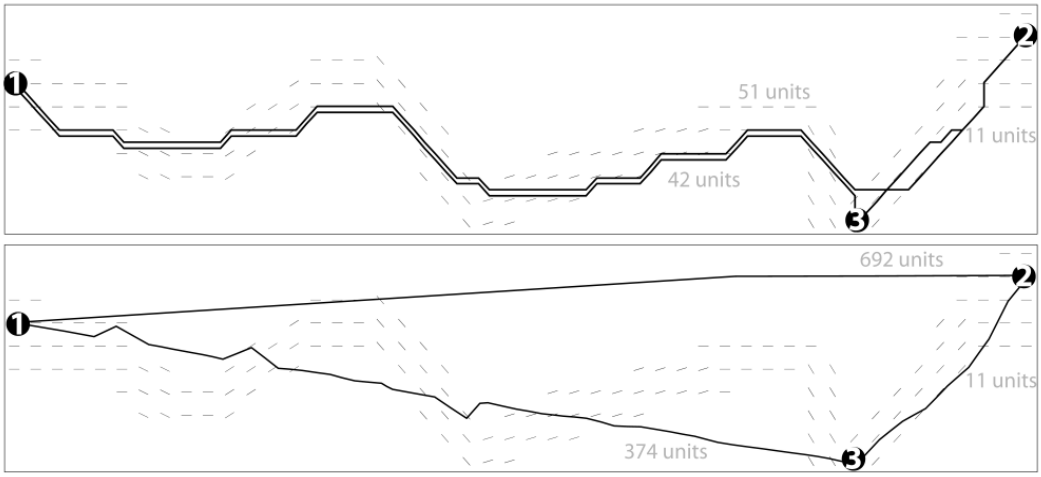


Figure 4.14: Figure 4.10 reprinted. Example optimal paths using a channel LVA field indicated by the gray lines. Anisotropy ratio is 10:1. Above – Using the Dijkstra algorithm the global solution is discovered but the path is restricted to the edges between vertices. Below – Using the Newton method the global solution is not discovered but the path is unrestricted in shape. Models are nominally 50 cells by 10 cells.

The implementation of graph theory to the shortest path problem is beneficial for two main reasons. First, within the initial assumption that potential paths are limited to following edges, the global solution is guaranteed. Regardless of the LVA field, point configuration, etc. the Dijkstra algorithm determines the globally shortest path. Secondly, there are many efficient publically available implementations of the Dijkstra

algorithm. The publicly available Boost Graph Library (Siek *et al.* 2001) is used. The Dijkstra algorithm is also parallelizable, further increasing its suitability for the proposed problem. (*NOTE: recently an experimental release of a parallel implementation of the Boost Graph Library has become available that would significantly improve CPU times. See http://www.osl.iu.edu/research/pbgl/ for more details*).

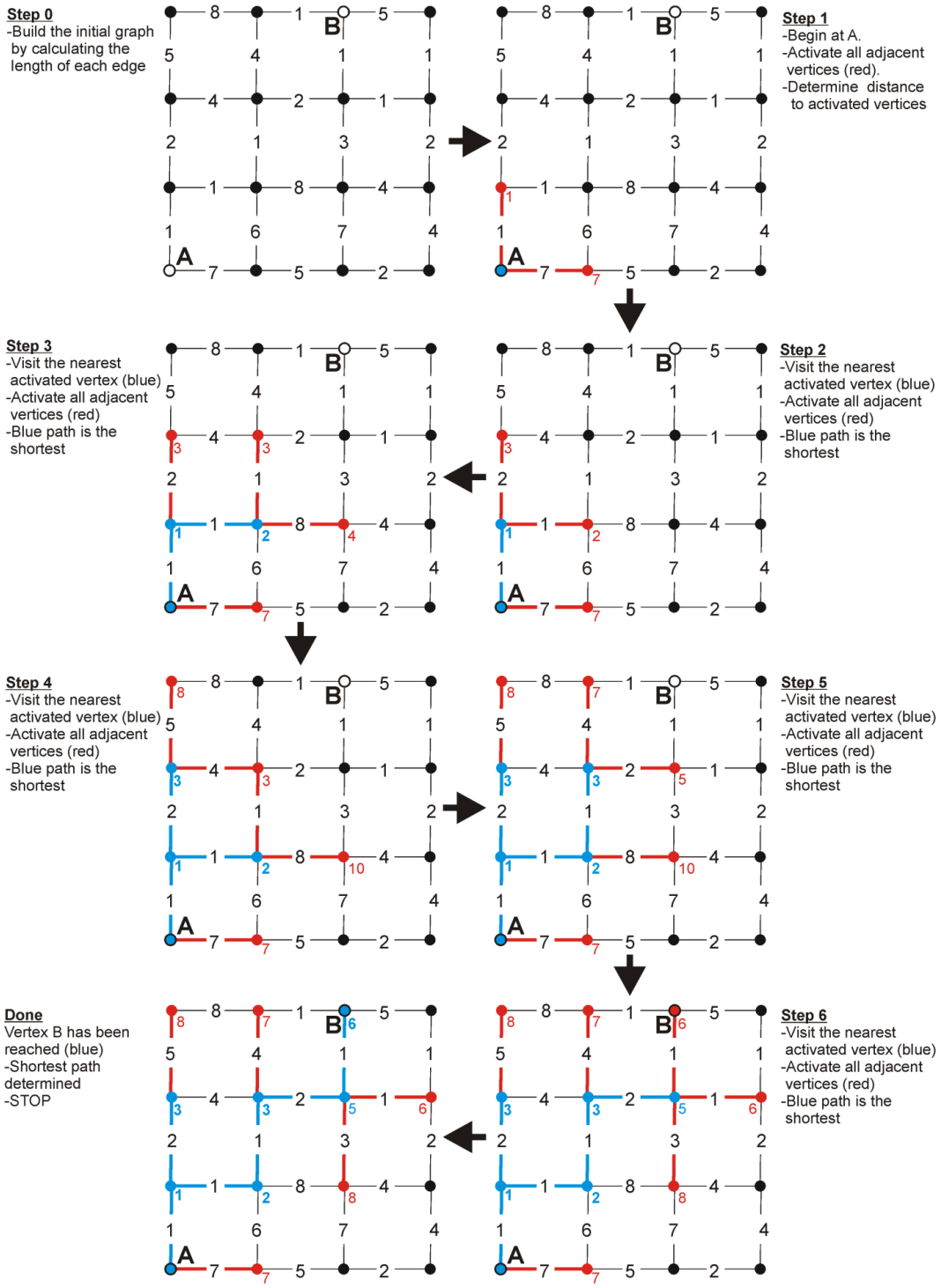


Figure 4.15: Dijkstra algorithm steps.

The Dijkstra algorithm is sufficiently fast to allow for the calculation of the 200M+ distances required in a moderately sized geostatistical model. The same 10 random LVA fields used in Figure 4.9 to test the CPU speed of the guarded Newton method are used to assess the speed of the Dijkstra algorithm (Figure 4.16). Linear regression on the average time required for all 10 LVA fields suggests that 200M distances can be calculated in under 6 hours; the range in time for the 10 different LVA fields is from 5.2 to 6.7 hours. Appendix B presents more detailed time trials for larger grids.

There are a number of factors that affect the CPU requirements of the shortest path optimization using a graph. Parameters that influence run time include (1) dimensionality (2) the number of vertices in the graph (3) the number of edges in the graph (4) the LVA field itself and (5) data spacing. Chapter 5 introduces techniques that will significantly reduce the number of distance calculations, as such, further discussion on CPU requirements are deferred to Chapter 5 when all elements of the proposed methodology are developed and can be jointly assessed.

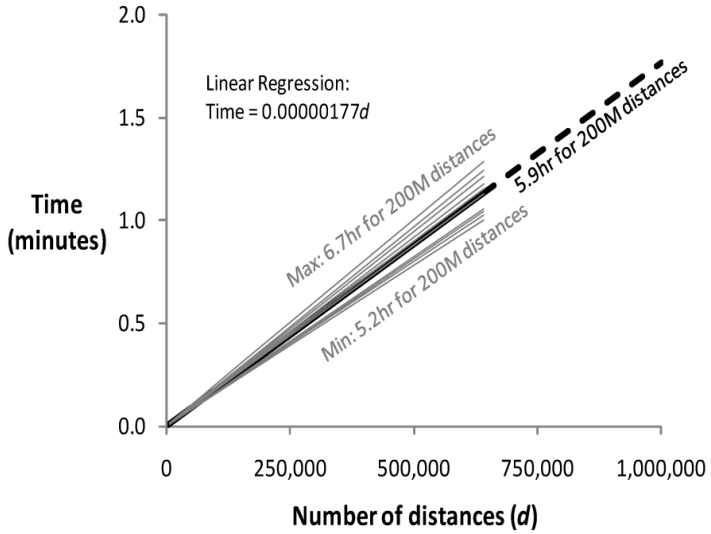


Figure 4.16: Time required to calculate the required distances (*d*). Each line represents one randomly generated LVA field. The thick black line is the linear regression based on the average time.

4.2.4 Path Flexibility with the Dijkstra Algorithm

A drawback of the Dijkstra algorithm is the restriction of the path between locations to the geometrical constraints of the graph formulation. The following discusses increasing the number of offsets considered in the graph to generate paths that are not as constrained.

Ideally, an infinite number of vertices and edges would populate the modeling domain, resulting in completely arbitrary and flexible paths. However, considering too many vertices and edges results in intractable CPU times, while considering too few vertices and edges may result in paths that are less than optimal because of the geometrical constraints on the paths. For geostatistical applications, it is recommended that the

vertices be assigned to grid cell centers to allow for direct transferability between the graph and the model (i.e. the typical sugar cube model). With vertices fixed to model cell centers, only the edges can be manipulated to obtain more accurate paths.

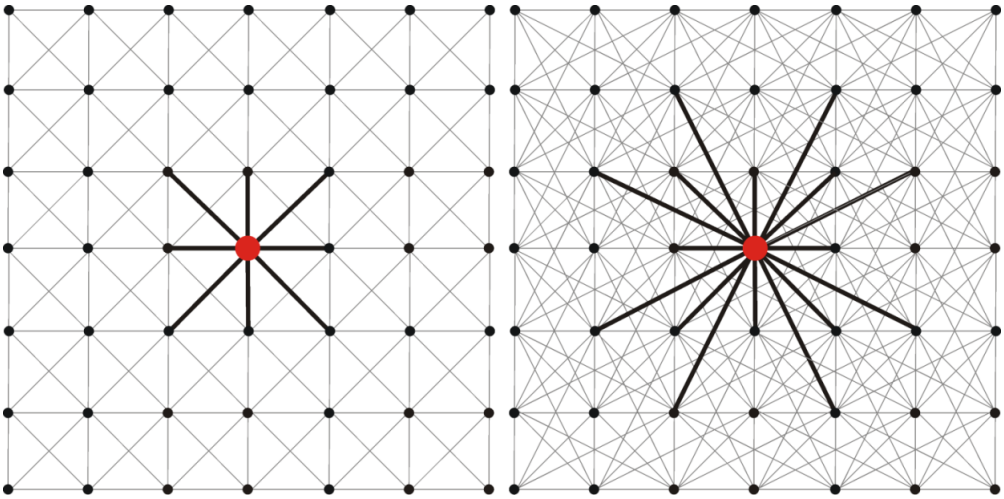


Figure 4.17: Left: Graph with vertices connected by 1 offset. Right: Graph with vertices connected by 2 offsets. Thicker lines indicate which vertices are connected to the central vertex.

Consider defining the edges by the number of grid offsets connected (Figure 4.17). Linking vertices separated by larger offsets increases flexibility in the path separating two locations. When 1 offset is considered paths are limited to following 0˚, 45˚, 90˚ or 135˚ directions with an average angle between permissible paths of 45˚ (Table 4.4). This implies that the azimuth of the piecewise linear paths is restricted to 45˚ increments. Increasing the number of offsets to 2 allows for paths that follow 0˚, 26˚, 45˚, 64˚, 90˚, 116˚, 135˚, 154˚ or 180˚ directions with an average angle between paths of 22.5˚. As more offsets are considered the range of possible azimuth directions widens and increases the flexibility of the path (Table 4.4).

While adding edges generates paths with increased flexibility there is an associated CPU cost. The complexity of the Dijkstra algorithm is O(e+$n$log$n$), where e is the number of edges in the graph and n is the number of vertices; increasing the number of edges has a linear effect on the CPU time. The number of edges required in a graph increases quickly with the number of offsets (Table 4.4).

Table 4.4: Permissible angles for increasing offsets with the Dijkstra algorithm.

Number of offsets	Number of edges (2D)	Angles available	Average angle spacing	%decrease wrt 1 offset
1	8	0˚, 45˚, 90˚, 135˚	180˚/4 = **45.0˚**	-
2	24	0˚, 27˚, 45˚, 63˚, 90˚, 116˚, 135˚, 154˚	180˚/8 = **22.5˚**	**50%**
3	48	0˚, 18˚, 27˚, 34˚, 45˚, 56˚, 63˚, 72˚, 90˚, 108˚,116˚, 124˚, 135˚, 146˚, 153˚, 162˚	180˚/16 = **11.3˚**	**75%**
4	80	0˚, 14˚, 18˚, 27˚, 34˚, 36˚, 45˚, 53˚, 56˚, 63˚, 72˚, 75˚, 90˚, 104˚, 108˚, 116˚, 124˚, 127˚, 135˚, 143˚, 146˚, 153˚, 162˚, 166˚	180˚/24 = **7.5˚**	**83%**

The effect of the number of offsets on the optimal paths is shown with two examples, an anticline (Figure 4.18) and a channel (Figure 4.19). Increasing the number of offsets beyond 3 does not visually increase the accuracy of the paths and has little effect on the average distance of the paths (Figure 4.20). From these examples and the authors past experience with larger models, 3 offsets are sufficient to determine reasonable paths. The absolute maximum number of offsets that should be considered is given in Equation 4.13. Geostatistical estimation and simulation require the distance between grid cells and data, considering more offsets than Equation 4.13 suggests would connect vertices separated by more than the data spacing. Often the number of offsets calculated from Equation 4.13 is too large to be practical and using 3 offsets as a starting point in a sensitivity analysis is highly recommended.

$$n_{off} = \frac{1}{2} \cdot \frac{DataSpacing}{LVA\ Grid\ Size} \qquad \textbf{4.13}$$

It is clear that the number of offsets has an effect on the shortest path taken. Ideally, more offsets would be considered to determine optimal paths, but CPU time and memory requirements demand fewer offsets for large examples.

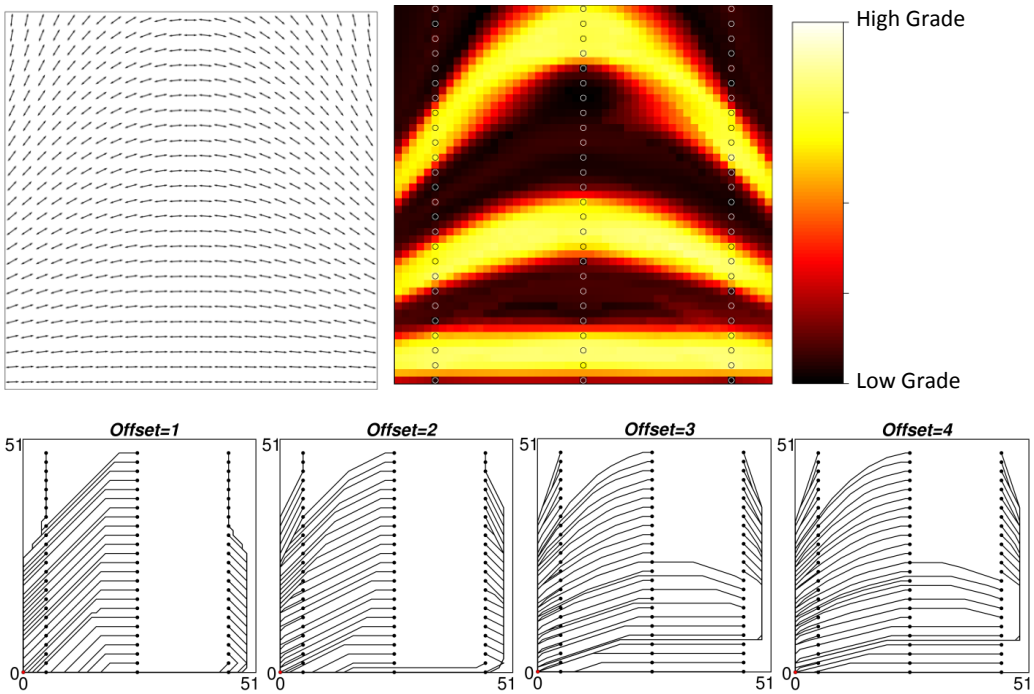


Figure 4.18: Above Left: 51x51 anticline LVA field. Anisotropy ratio = 10:1. Above Right: Kriging result with LVA (Chapter 5). Below: Shortest paths from the three drill holes to the lower left vertex. Models are 51x51.

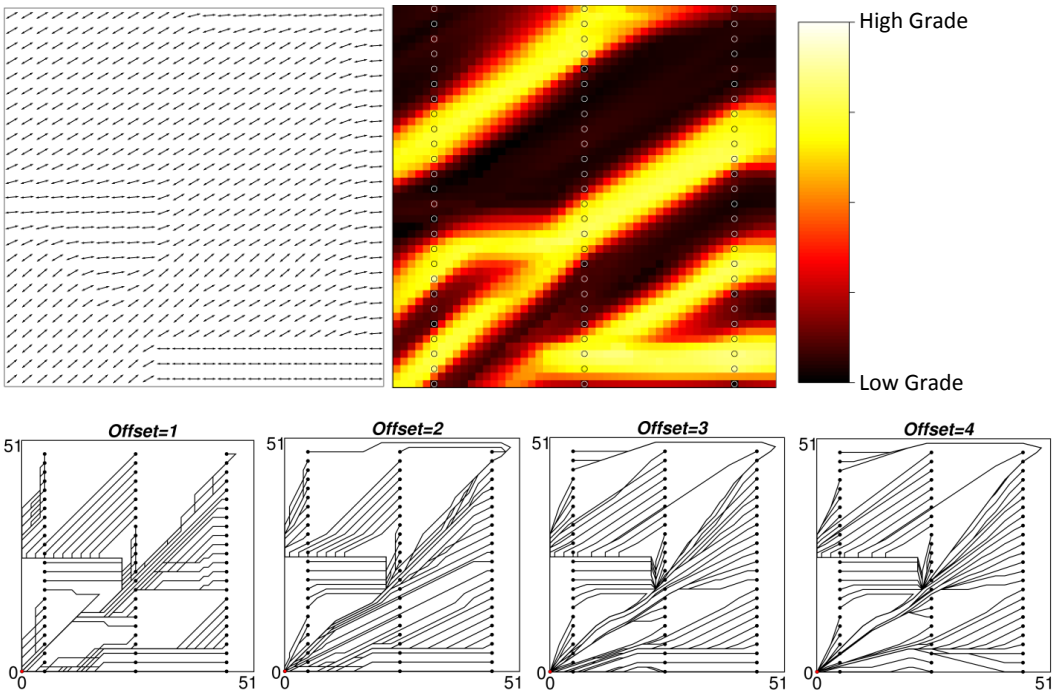


Figure 4.19: Above Left: 51x51 channel LVA field. Anisotropy ratio = 10:1. Above Right: Kriging result with LVA (Chapter 5). Below: Shortest paths from the three drill holes to the lower left vertex. Models are 51x51.

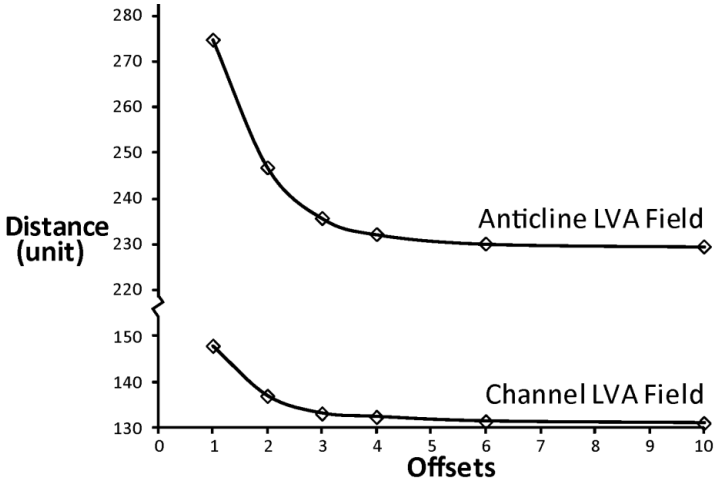


Figure 4.20: Average distance between vertices when considering more offsets for the channel and anticline examples.

4.3 Remarks

This chapter presented two methodologies for calculating the shortest path between points in the presence of LVA. The convex optimization technique, the guarded Newton method, is extremely CPU intensive because of its iterative nature and is unable to determine the globally minimal solution. Alternative global optimization techniques, such as random restart or simulated annealing, could be used to obtain better global solutions; however, these techniques require additional CPU time and are deemed

impractical. The guarded Newton method is not carried forward because of these limitations.

A second technique based on graph theory proved to be more efficient in calculating the necessary number of distances in a reasonable period of time (200M distances in ~6 hours). Only the Dijkstra algorithm is used in future chapters to calculate the shortest path between locations.

This chapter was limited to discussions of 2D graphs because the extension to 3D is not difficult, the only difference is in the construction of the graph (Figure 4.21) and the associated increase in the number of edges (Table 4.5). Recall that the complexity of the Dijkstra algorithm is linear with respect to the number of edges in the graph.

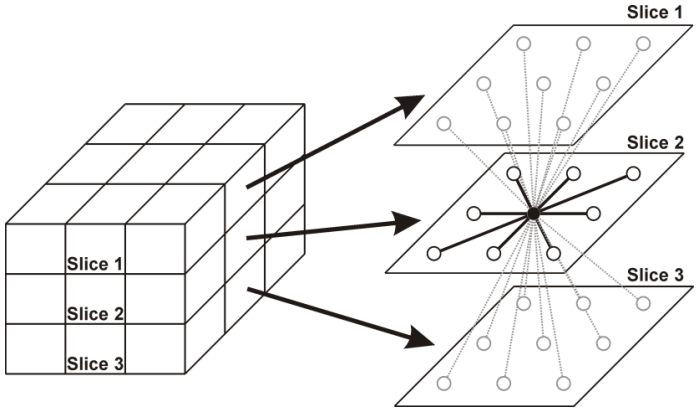


Figure 4.21: Left: 3x3x3 block model. Right: 26 edges connected to the central vertex (black) in a 3x3x3 graph.

The main limitation with the Dijkstra algorithm is the restriction of the optimal path to following the edges. It has been shown that this limitation can be mitigated by adding additional edges to vertices separated by larger offsets. As more edges are added the range of angles the path can follow increases but the number of edges also increases (Table 4.5). The addition of edges to obtain path flexibility must be balanced against the CPU requirements of the Dijkstra algorithm. For smaller models more offsets can be considered but larger models are restricted to a smaller number of offsets (Chapter 6).

Table 4.5: Number of edges per vertex in a graph.

n_{off}	2D	3D
1	8	26
2	24	124
3	48	342
4	80	728
5	120	1330
10	440	9260

Chapter 5: Estimation and Simulation with Anisotropy

Geostatistics is focused on the spatial prediction of variables from sparse sample data. From these spatial predications of grade, porosity, saturation, concentration, etc. resource calculations can be made (Chapter 1). This chapter incorporates LVA into the spatial prediction of variables by using the shortest path distance (SPD) developed in Chapter 4. Section 5.1 presents a modified version of inverse distance weighted interpolation (IDW) to introduce the idea of spatial predictions as well as demonstrate the effects of using the SPD in geostatistical modeling.

Kriging and sequential Gaussian simulation (SGS) are geostatistical tools often applied in resource evaluation studies. These techniques use the best linear unbiased estimator (BLUE) to obtain optimal spatial predictions. This requires the solution to a positive definite system of equations. Section 5.2.2 introduces dimensionality reduction techniques that are implemented to guarantee the positive definiteness of this system of equations. Sections 5.3 and 5.4 explore kriging and SGS with the SPD to incorporate LVA into resource estimation.

Only small illustrative examples using simple LVA fields and relatively few cells (<3000) are discussed in this chapter. A larger case study involving 1M cells and many data is presented in Chapter 6.

5.1 Inverse Distance Weighted Interpolation (IDW)

The most common form of spatial interpolation is the prediction of estimates from a weighted average of nearby data (Babak and Deutsch 2008). This concept was introduced in Chapter 1 and is repeated here. Weights, λ_α, (Equation 5.1) are assigned to surrounding data, z_α, to generate an estimate, z^*, at an unsampled location, $\mathbf{u}$. With IDW, the weights are inversely proportional to the distance between the unsampled location and the nearby samples (Equation 5.2). The only parameter that is required for IDW is ω. The purpose of this dissertation is not to discuss the parameterization of IDW, as such, a constant ω=2 is used throughout this chapter. Interested readers are referred to Mueller *et al.* (2005) for further discussion on the optimization of ω.

$$z^*(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha \, z_{(u_\alpha)} \qquad \textbf{5.1}$$

where n is the number of data available for estimation.

$$\lambda_\alpha = \frac{\frac{1}{d_\alpha^\omega}}{\sum_{\alpha=1}^{n} \frac{1}{d_\alpha^\omega}} \qquad \textbf{5.2}$$

The incorporation of the SPD into IDW is straightforward, the distance in Equation 5.2 is replaced with the SPD:

$$\lambda_\alpha = \frac{\frac{1}{SPD_\alpha^\omega}}{\sum_{\alpha=1}^{n} \frac{1}{SPD_\alpha^\omega}} \qquad \textbf{5.3}$$

Using the weights in Equation 5.3 the IDW estimate maps are constructed. Consider the anticline LVA field with three drill holes in Figure 5.1. There are a total of 78 sample data, thus the distance matrix (**D**) between the 2601 grid nodes and the 78 sample data is required (Figure 5.2) to calculate the weights in Equation 5.3. The resulting estimated map is shown in Figure 5.1.

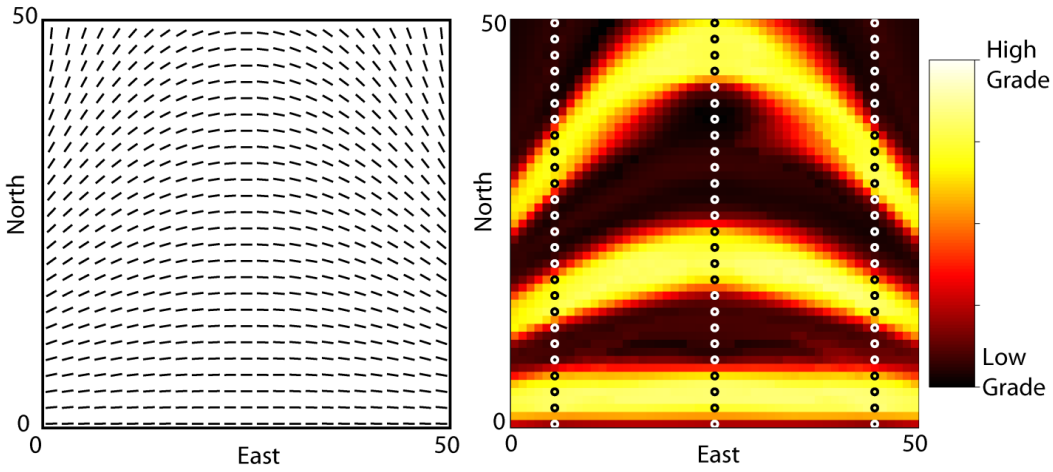


Figure 5.1: Left: Anticline LVA field with a constant anisotropy ratio of 10:1. Right: Spatial estimates using IDW estimation with the three strings of data shown as hollow points. Distances provided as unit distance.

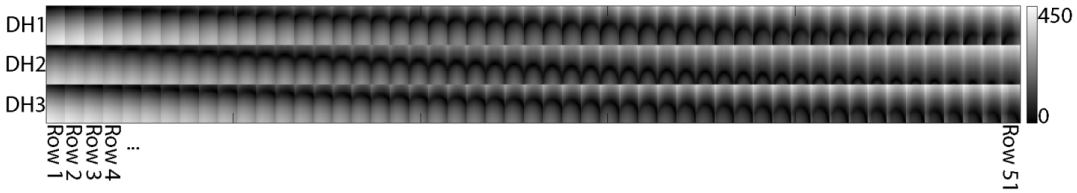


Figure 5.2: Unit distance matrix (**D**) between 78 data (3 drill holes with 26 data per drill hole). Each row of the distance matrix represents the distance from one of the 78 samples to the 2601 cells in the model. 3 offsets were used with the Dijkstra algorithm to determine the distance between locations. Distances provided as unit distance.

The locally varying dips of the anticline are well reproduced in Figure 5.1. Often IDW is used as an exploratory technique in the early stages of a geostatistical project; however, there are situations when it is preferred, such as when the variogram is difficult to infer for a particular data set. Notwithstanding the simplicity of IDW with LVA, if an estimated map is required it is more common to apply kriging. Kriging estimates are optimal in a least squared error sense when considering all possible linear unbiased estimates (recall the description of kriging in Chapter 1). Moreover, there are practical issues with applying IDW such as, sensitivity to the characteristics of the available data distributions (skewness, clustered data) and the exponent parameter (Babak and Deutsch 2008). Kriging with LVA is presented as an alternative to IDW. IDW estimates are compared to kriging in Chapter 6.

5.2 Embedding in *q*-Dimensional Euclidean Space

Incorporating the SPD into kriging and SGS is not straightforward. The mathematical foundation of kriging and SGS is based on the solution of a positive definite system of equations known as the kriging equations (Equation 5.4).

$$\sum_{\beta=1}^{n} \lambda_\beta \mathrm{C}(\mathbf{u_\alpha}, \mathbf{u_\beta}) = \mathrm{C}(\mathbf{u}_?, \mathbf{u_\alpha}) \quad \alpha = 1, \dots, n \qquad \textbf{5.4}$$

If the covariances in Equation 5.4 form a positive definite covariance matrix, the solution to the kriging system of equations is unique and can be determined (Christakos 1984; Cressie 1993). Normally, positive definiteness is ensured by using an Euclidean distance metric and a positive definite covariance function (Christakos 1984); however, positive definite covariance functions do not guarantee positive definiteness when used with the SPD metric (Curriero 1996 and 2005). The SPD metric is mapped to a high dimensional Euclidean space using dimensionality reduction algorithms. In the high dimensional Euclidean space it is relatively straightforward to find valid covariance models that are positive definite.

5.2.1 SPD Metric

The SPD is a valid distance metric (Curriero 2005). A distance function, $d(\mathbf{p}_i,\mathbf{p}_j)$, is a valid distance metric if it meets the following conditions for all locations *i* and *j*:

Condition 1: $d(\mathbf{p}_i,\mathbf{p}_j)\geq 0$ and $d(\mathbf{p}_i,\mathbf{p}_j)=0$ iff $\mathbf{p}_i=\mathbf{p}_j$
Condition 2: $d(\mathbf{p}_i,\mathbf{p}_j)= d(\mathbf{p}_j,\mathbf{p}_i)$
Condition 3: The triangle inequality: $d(\mathbf{p}_i,\mathbf{p}_j)\leq d(\mathbf{p}_i,\mathbf{p}_k) + d(\mathbf{p}_k,\mathbf{p}_j)$

The Euclidean distance metric, $d(\mathbf{p}_1, \mathbf{p}_2) = \|\mathbf{p}_1 - \mathbf{p}_2\|_2$, is a valid distance metric as it meets all three conditions; constraints on the SPD can ensure that it is also a valid distance metric. If the graph is bidirectional and all edges have a distance greater than 0 (condition 1) the SPD function is a metric. A graph is bidirectional when the cost to travel from node *i* to node *j* is equivalent to the cost to travel from node *j* to node *i* (condition 2). Because the Dijkstra algorithm always discovers the shortest path between locations, the triangle inequality is also satisfied. Cases exist when the SPD is not a valid distance metric, such as when there is zero cost between nodes (rare in a geostatistical application) or when the graph is directional ($d_{ij} \neq d_{ji}$). The assumption that the graph is bidirectional ($d_{ij}=d_{ji}$) is often valid for typical earth science applications but can be violated when considering time series or unidirectional phenomenon such as wind patterns or stream currents.

The SPD is a valid distance metric but this is not a sufficient condition to guarantee positive definiteness of the resulting covariance matrix in kriging or SGS (Curriero 2005). One solution to this problem is to embed the geostatistical grid into a high dimensional Euclidean space where valid covariance functions can be applied to ensure positive definiteness. This embedding can be accomplished using a dimensionality reduction technique. The main goal of these techniques is to reduce the dimensionality of data such that it can be analyzed and meaningful relationships determined. One common technique is principal component analysis (PCA) (Pearson 1901; Hotelling 1933). The

following discussion explores various dimensionality reduction techniques that can be applied to the SPD with the end goal of obtaining a representation of the input distance matrix in a high dimensional Euclidean space such that the interpoint SPD is reproduced and positive definiteness is guaranteed.

Dimensionality reduction algorithms take data that are of a high dimensionality and attempt to find a lower *q*-dimensional representation of the data such that the original features of the data are optimally preserved. The original coordinates of earth science data are either in one, two or three dimensional space. It should be noted that in this dissertation dimensionality *reduction* techniques are used to *increase* the dimensionality of the data. The reduction in dimensionality is not a reference to the coordinates of the locations (2D or 3D); rather, consider the input distance matrix, **D**, between all grid locations in a model to be the original space and every location, defined by **D**, is mapped to a lower *q*-dimensional space. The goal is to discover the optimal configuration of points in an Euclidean *q*-dimensional space such that **D** is best preserved.

The majority of dimensionality reduction algorithms take the *coordinates* of high dimensional data as input and find a lower dimensional representation of the coordinates; however, a sub class of dimensionality reduction algorithms specifically attempt to preserve the interpoint distances between locations rather than preserving the data configuration as defined by a set of coordinates.

Two algorithms, isometric feature mapping (ISOMAP, Tenenbaum *et al.* 2000) and local linear embedding (LLE, Roweis and Saul 2000), are carried forward in this chapter. These algorithms are selected to be representative of the two categories of available techniques, global and local. This selection, and subsequent rejection of algorithms presented in the extensive review of Maaten *et al.* (2008), is further justified below:

1) PCA cannot be applied directly to a distance matrix and is very similar to ISOMAP, thus, PCA is not considered.
2) All techniques optimize the dimensionality reduction in some manner; however, Autoencoders, LLC and Manifold charting (Maaten *et al.* 2008) require the optimization of a non-convex problem that is not considered because of the additional CPU requirements.
3) As noted by Maaten *et al.* (2008), Kernel PCA, ISOMAP and Laplacian Eigenmaps are highly similar techniques and are special cases of Kernel PCA with different kernel functions. As such ISOMAP is selected to represent these techniques as it is widely used in the literature.
4) LTSA is similar to the Hessian LLE technique which is a variant of LLE (Maaten *et al.* 2008). LLE is carried forward to represent these local techniques.
5) Diffusion maps require multiple random walks through the original graph and are not considered because of the CPU time required by the random walks used in these methods.

The following discussion introduces ISOMAP as an algorithm to embed points defined by a distance matrix into a *q*-dimensional Euclidean space. The computational requirements of ISOMAP are intractable for the size of problems normally encountered in geostatistics; therefore two variants on ISOMAP are explored.

5.2.2 ISOMAP

Multidimensional scaling (MDS) (Mardia *et al.* 1976) is a statistical technique used to explain/map any type of similarities (i.e. covariances) or dissimilarities (i.e. distances or travel times) between data observations. When these dissimilarities are obtained from a graph rather than from an Euclidean measure, MDS is referred to as ISOMAP (Tenenbaum *et al.* 2000; Maaten *et al.* 2008). As an instructional example consider a road map, the problem is to spatially represent the travel time between cities on a 2D map (Figure 5.3). The map can be considered a graph with roads (edges) connecting cities (vertices). Travel time depends on a number of factors: distances between cities; speed limits; condition of roads; traffic; and road construction. Two geographically close cities may require a large travel time, whereas two distant cities separated by a freeway may have a short travel time. To visualize these travel times on a 2D map, the cities must be rearranged until the Euclidean distance between them is roughly equivalent to the appropriate travel time. Because the proposed dissimilarity measure (travel time) is unlikely to be Euclidean, it is usually impossible to exactly reproduce the dissimilarities even in a higher *q*-dimensional space; however, MDS can be applied to accomplish the rearrangement in a way that minimizes stress (Equation 5.5) after mapping the cities to a new coordinate system (Figure 5.3). The dimensionality of this Euclidean space can be as high as the input number of points less one. Retaining more dimensions lowers the stress measure as there is more flexibility for locating the cities; alternatively, two dimensions could be retained for visualization but the resulting stress measure increases (Table 5.1).

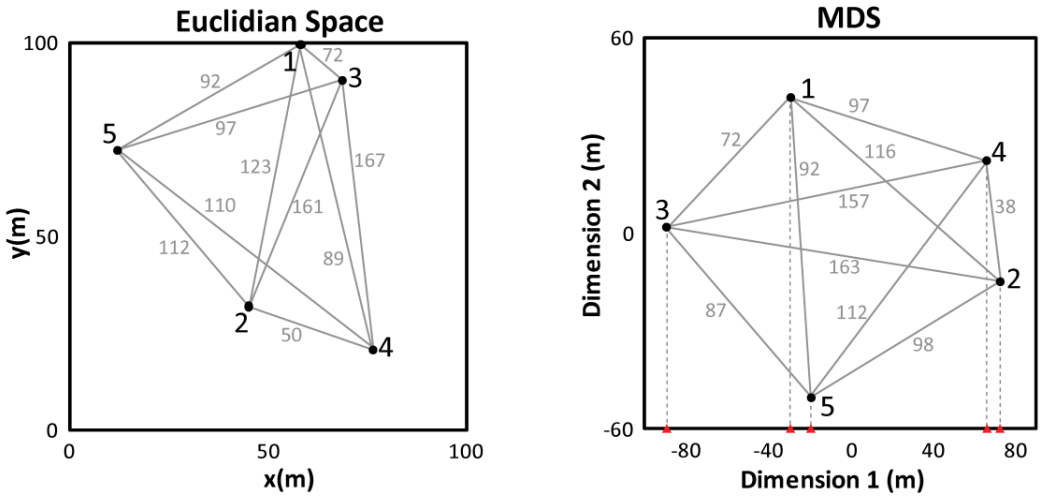


Figure 5.3: Above Left: Locations of 5 cities in Euclidean space with travel times indicated (hours). Above Right: Cities in the first two dimensions (also an Euclidean space). Note that the cities can also be represented in a 1D Euclidean space as a projection onto the first dimension (triangles on the x-axis).

Table 5.1: Table of distances/times between cities.

Cities	Euclidian (m)	Time (hr)	(MDS)/(% difference to time) 1D MDS (hr)	2D MDS (hr)	3D MDS (hr)
1-2	69	123	102 / 17%	116 / 6%	123 / 0%
1-3	14	72	60 / 51%	72 / 41%	79 / 36%
1-4	81	89	95 / 22%	97 / 21%	97 / 21%
1-5	53	92	10 / 92%	92 / 25%	93 / 24%
2-3	63	161	162 / 32%	163 / 33%	163 / 32%
2-4	33	50	6 / 95%	38 / 69%	59 / 52%
2-5	52	112	92 / 25%	98 / 20%	112 / 9%
3-4	70	167	155 / 26%	157 / 28%	161 / 31%
3-5	59	97	70 / 43%	87 / 29%	99 / 20%
4-5	82	110	86 / 30%	112 / 9%	113 / 8%
Average % Difference			**43%**	**28%**	**23%**

In the context of LVA, ISOMAP is applied with the SPD. Rather than travel times between locations, there are now optimized paths to provide the interpoint dissimilarity measure to be used in ISOMAP.

Consider *N* observations with *M=N(N-1)/2* (dis)similarities between grid cells (nodes). These (dis)similarities are often the covariance between points or the inter-point distances. The inter-point distances from the Dijkstra algorithm (Chapter 4) provide the dissimilarity measure to be used in ISOMAP; therefore, the inputs to ISOMAP are *M* distances in a symmetric matrix (**D**). ISOMAP maps the input dissimilarities, **D**, to a *q*-dimensional Euclidean space where the straight line distances between locations approximately match the input dissimilarities. There is some mismatch in the distances in the embedded space because of the impossible geometric configuration of the input distance matrix. ISOMAP minimizes the mismatch between the original distances (d_{ij}^q) and the mapped distances ($\hat{d}_{ij}^q$) in *q* dimensions as defined by a stress measure:

$$Stress(q) = \left\{ \frac{\sum_{j=1}^{N} \sum_{i=1}^{j-1} (d_{ij}^q - \hat{d}_{ij}^q)^2}{\sum_{j=1}^{N} \sum_{i=1}^{j-1} (d_{ij}^q)^2} \right\}^{1/2} \quad \textbf{5.5}$$

where *N* is the number of grid nodes in a geostatistical model.

Applying ISOMAP to a distance matrix involves the following steps (Mardia *et al.* 1979; Tenenbaum *et al.* 2000; Loland and Host 2003). Note that in this application of ISOMAP, the distance matrix, d_{ij}^2, is obtained from calculating the SPD between nodes *i* and *j* with the Dijkstra algorithm:

Step 1: Construct matrix $\mathbf{A} = (-0.5 d_{ij}^2)$. Note that in **D**, $d_{ii} = 0$ and $d_{ij} \geq 0$

Step 2: Construct matrix **B**. Elements of **B** are $b_{ij} = a_{ij} - \overline{a}_{i\bullet} - \overline{a}_{\bullet j} + \overline{a}_{\bullet\bullet}$, $\overline{a}_{i\bullet} = \sum_{j=1}^{N} \frac{a_{ij}}{N}$, $\overline{a}_{\bullet j} = \sum_{i=1}^{N} \frac{a_{ij}}{N}$,

$\overline{a}_{\bullet\bullet} = \sum_{j=1}^{N} \sum_{i=1}^{N} \frac{a_{ij}}{N^2}$.

Step 3: Find *q* normalized eigenvectors (**V**) of **B** that correspond to positive eigenvalues:

$$\mathbf{V} = (v_1, \ldots, v_q) \qquad \textbf{5.6}$$

Step 4: The coordinates of the input data are the *q* rows of **V**.

The number of dimensions to retain, *q*, is a *choice* and can be selected to be 2 or 3 for visualization purposes; however, the size of *q* can be as large as the number of positive eigenvalues of **B** up to a maximum of *N*-1. Calculating the stress resulting from different values of *q* can help in the selection of the number of dimensions to retain by locating the inflection point on a scree plot; however, using more dimensions always reduces stress but increases the CPU requirements of kriging and SGS. Using the maximum number of dimensions results in the best reproduction of the input distances and is recommended. It is also important to note that the dimensions can be ordered by eigenvalue and if *q* is selected to be less than the maximum, then the *q* dimensions with the largest eigenvalues should be selected to minimize stress.

ISOMAP is a coordinate transformation where initially the grid cells are defined by the interpoint distance matrix, **D**. After applying ISOMAP, each cell has a new set of coordinates in *q* dimensions defined by the rows of **V**. Performing ISOMAP for the anticline example (Figure 5.1) results in the embedding shown in Figure 5.4. The volume of each cell in the embedded Euclidean space may not be identical; however, volume variance relationships need not be considered as the embedded space is used only as a higher dimensional representation of the relationship between locations in the original space.

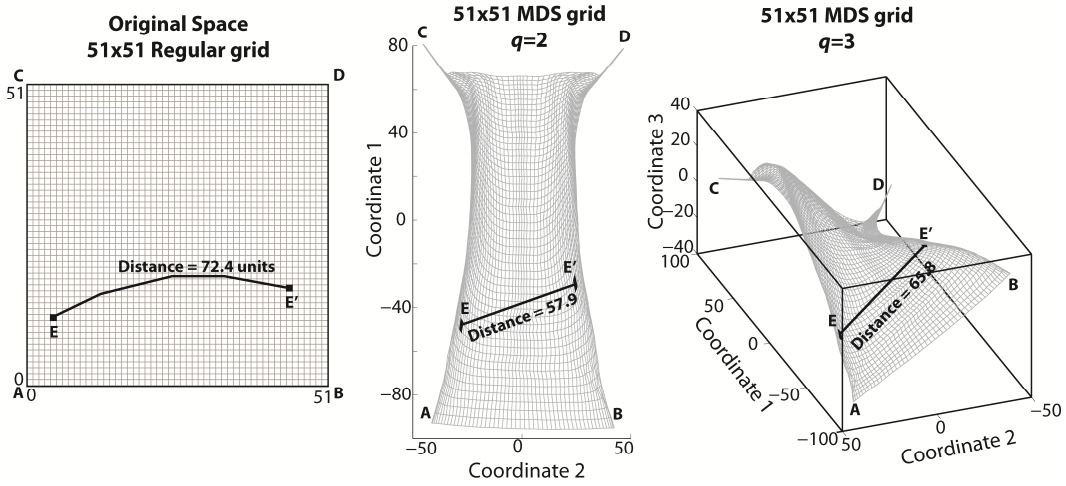


Figure 5.4: ISOMAP transformation of a 51x51 grid using the anticline LVA field, distances are considered as unit distance. The length of the path E-E' is shown in both coordinate systems. Note that in the ISOMAP coordinate system (left) there are 2600 dimensions (2601– 1) but visualization is only possible in the first 2 or 3 dimensions. Considering all 99 dimensions the distance E-E' is 72.9units.

The benefit of applying ISOMAP is illustrated in Figure 5.4. All paths between locations are considered as straight line paths in the embedded space. Because the embedded space is Euclidean, the resulting covariance matrix, when applying kriging or SGS, is positive definite. One drawback to this embedding is that it is not perfect, thus, the SPD between locations are only approximately honored as per the stress criterion (Equation 5.5).

ISOMAP requires the eigenvalues/vectors of the distance matrix of all points that must be transformed, which would be the distance matrix between ALL nodes in the model at the required grid resolution. This poses two problems (1) solving the eigen problem for a large *dense* system is not feasible, a million grid cell model would require the eigen solution of a 1M x 1M dense matrix and (2) generating the difference matrix between ALL nodes in a grid using the Dijkstra algorithm would be computationally infeasible (~$5x10^{11}$ distances). There are solutions to these computational problems which involve using the approximate ISOMAP of a subset of the grid. Two main types of algorithms exist, local and global. Local techniques, such as LLE, require a sparse difference matrix informed by local neighborhoods only. This reduces the *dense* eigen problem to a *sparse* eigen problem, which is solvable even for large 1M x 1M matrices. Global algorithms, such as landmark-ISOMAP (L-ISOMAP) reduce the number of nodes that are used in the distance matrix such that the eigen solution can be feasibly solved.

Both global and local dimensionality reduction algorithms are attractive for the incorporation of LVA. It is expected that global algorithms should maintain the complex nonlinear characteristics of large scale geological features while local algorithms would maintain the small scale details found within the larger scale structures. LLE is a local algorithm that uses the distance between points found within a local search neighborhood only and ignores the distances between points outside the search, generating a sparse distance matrix that can be solved (Watkins 2007). L-ISOMAP is a global algorithm that uses a series of widely spaced landmark points in calculating the distance matrix, reducing the size of the distance matrix and thus, reducing computational requirements.

5.2.3 LLE and L-ISOMAP

LLE (Roweis and Saul 2000) is a local dimension reduction technique that attempts to reduce the computational demands of dimensionality reduction by reducing the number of dissimilarities in **D**. LLE maps the input data to a set of coordinates where the neighborhood points in the original space are still nearby in the mapped space. *Nearby* is defined by a search radius in original coordinates such that all pairs separated by a distance greater than the search are ignored and thus, not explicitly reproduced in the mapped coordinate system. LLE assumes that each point and its neighbors lie on a locally linear patch of a manifold defined by the inter-point distances of the points in the

neighborhood. Points are embedded in the new coordinate space using a set of reconstruction weights which optimize the squared error between the input coordinates of the neighborhood points ($\vec{X}_i$) and the reconstructed coordinates:

$$Error(W) = \sum_i \left(\vec{X}_i - \sum W_{ij} \vec{X}_i \right)^2 \qquad \textbf{5.7}$$

The LLE procedure is summarized in Figure 5.5. Roweis and Saul (2000) give a detailed description of the LLE algorithm.

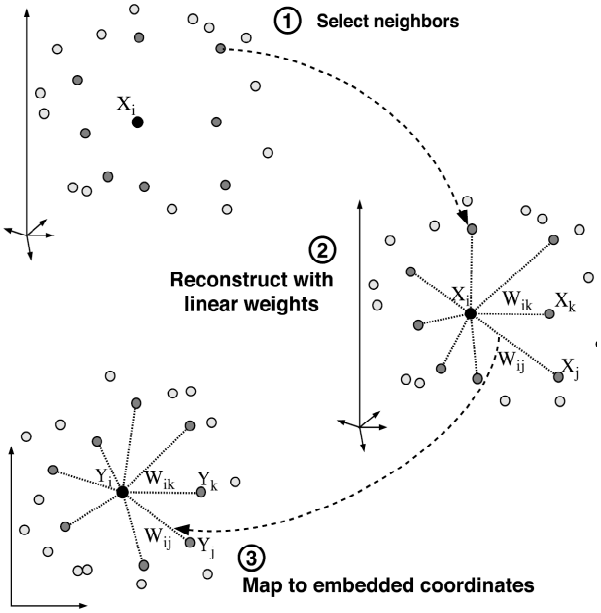


Figure 5.5: Steps of LLE: (1) Assign neighbors to each data point (for example by using the *k* nearest neighbors). (2) Compute the weights W_{ij} that best linearly reconstruct **D** from its neighbors, solving the least-squares problem in Equation 5.7. (3) Compute the low-dimensional embedding vectors best reconstructed by W_{ij}, minimizing Equation 5.7 (as shown in Roweis and Saul 2000).

L-ISOMAP is a global dimensionality reduction technique, significant current references include Silva and Tenenbaum (2003), Balasubramanian *et al.* (2002) and Tenenbaum *et al.* (2000). L-ISOMAP uses a set of *L* landmark points to reduce the computational demands of ISOMAP. Consider the grid in Figure 5.6 with 121 grid cells (nodes). Applying ISOMAP to this grid would require the eigen solution of a 121x121 matrix. With L-ISOMAP, only the coarser landmark grid points are used in the embedding. This requires the eigen solution of a 9x9 matrix (*L*=9). Moreover, there are fewer dissimilarities to calculate with the Dijkstra algorithm, only the *L(L*-1*)/*2 SPD between landmark points are needed, rather than the full *N(N*-1*)/*2. Note that this is equivalent to performing ISOMAP on the landmark points only. To locate the remaining points in the grid their distances to the landmark points are also required. Calculating the coordinates of all grid cells not considered as landmark point is analogous to

trilateration (for example, determining location based on distance from a known set of satellites as done with GPS).

There is an interesting synergy between L-ISOMAP and the use of graph theory to determine the distance between locations. Recall that when applying the Dijkstra algorithm the distance from one node to all other nodes is calculated in a single execution of the algorithm (Figure 4.14). As such, the Dijkstra algorithm need only be applied *L* times to obtain all necessary distances. Further, because the user is free to select the value of *L*, the computational requirements of the proposed methodology can be mitigated by selecting a small value of *L*. The consequence of a small value of *L* is an increase in the stress. Chapter 6 shows that a very small value of *L* can result in reasonable results and considerable CPU savings. This synergy does not exist when applying LLE; with LLE the Dijkstra algorithm is applied at each node in the grid to find the distances to the nearest *k* points.

Many dimension reduction algorithms exist that could be applied to this problem, Maaten (2007) provides Matlab code to implement 27 techniques. Examining LLE and L-ISOMAP are representative of local and global dimensionality reduction techniques and are selected because they are often used in the literature as benchmarks. Further implementation details of LLE and L-ISOMAP can be found in Maaten (2007).

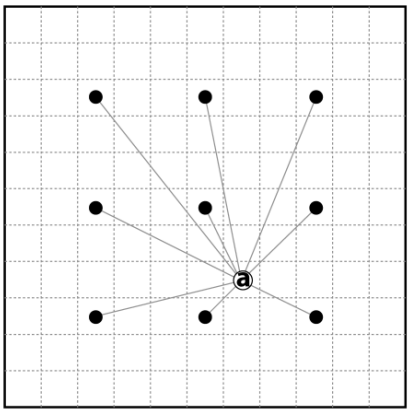


Figure 5.6: Locating point *a* using distance to nine landmark points (black).

5.2.4 Comparing LLE and L-ISOMAP

There are two criteria that must be considered when selecting a dimensionality reduction algorithm (1) minimization of the stress and (2) CPU time considerations for scalability to practical geostatistical model sizes. The ideal algorithm provides the minimum stress with reasonable effort. Locations are only embedded once, which implies that some extra CPU time can be incurred to minimize stress. The CPU requirements and minimization of the stress for L-ISOMAP and LLE both depend on their respective input parameters, number of landmark points (L) and the size of the search

radius to define the nearest neighbors (*k*). Table 5.2 provides a comparison between these input parameters.

It is expected that both LLE and L-ISOMAP should perform better as *k* or L increases. Consider applying the algorithms to a channel example (Figure 5.7) with increasing values of *k* and L. Stress is used to measure the performance of the algorithms; the lowest theoretical stress is obtained with classical ISOMAP and is 0.04. Erratic results are obtained for some values of *k* when implementing LLE (Figure 5.7) and in some cases the stress measure even increases as *k* increases (Figure 5.8). Conversely, L-ISOMAP behaves as expected with improving visual results as L increases (Figure 5.7) and a consistent decrease in stress (Figure 5.8). Figure 5.7 also shows the effect of considering a limit on the number of dimensions, *q,* retained.

LLE is unstable and provides a larger value of stress when compared to L-ISOMAP (Figure 5.7 and Figure 5.8); therefore, it is no longer considered. The potential danger in selecting a single dimensionality reduction technique, L-ISOMAP, to carry forward is that it may be unable to capture the small scale variations seen in natural deposits. Local techniques such as LLE should better reproduce short scale structures; however, in practice the unpredictability of LLE is a significant disadvantage (Figure 5.7 and Figure 5.8). Both LLE and L-ISOMAP were implemented but only L-ISOMAP will be discussed in the remainder of this dissertation.

Table 5.2: Comparison of dimensionality reduction algorithms, CPU requirements and the impact on stress.

Algorithm	Control Parameter	Distances to calculate with Dijkstra (N=nodes in grid)	Matrix size for Eigensolution	Expected effect on stress
L-ISOMAP	Number of landmark points (*L*)	L·N	Dense L x L	↑L = ↓Stress & ↑Time
LLE	Number of nearest neighbors in search (*k*)	N·*k*	Sparse N x N	↑*k* = ↓Stress & ↑Time
Classical ISOMAP	None	N·N	Dense N x N	Stress minimized

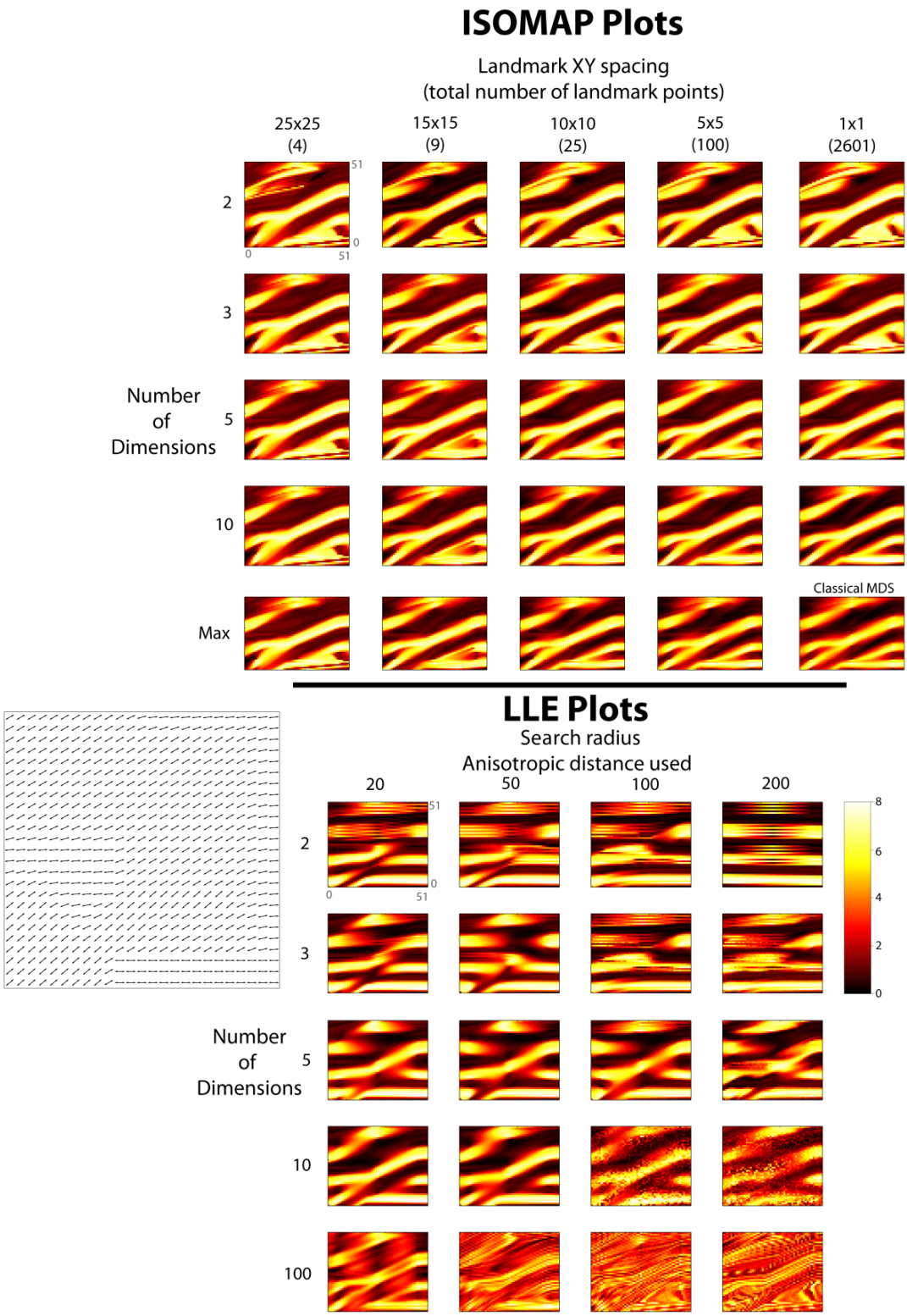


Figure 5.7: Left: Channel LVA field used. Above plots: Kriging with L-ISOMAP. Number of dimensions and landmark points are varied (i.e. 25x25 = a landmark point every 25 blocks in the x and y directions). Lower plots: Kriging with LLE. Neighborhoods are defined by a search radius. Plot dimensions are shown as unit distances with the hot colors (white and yellow) as high grades and black as low grade.

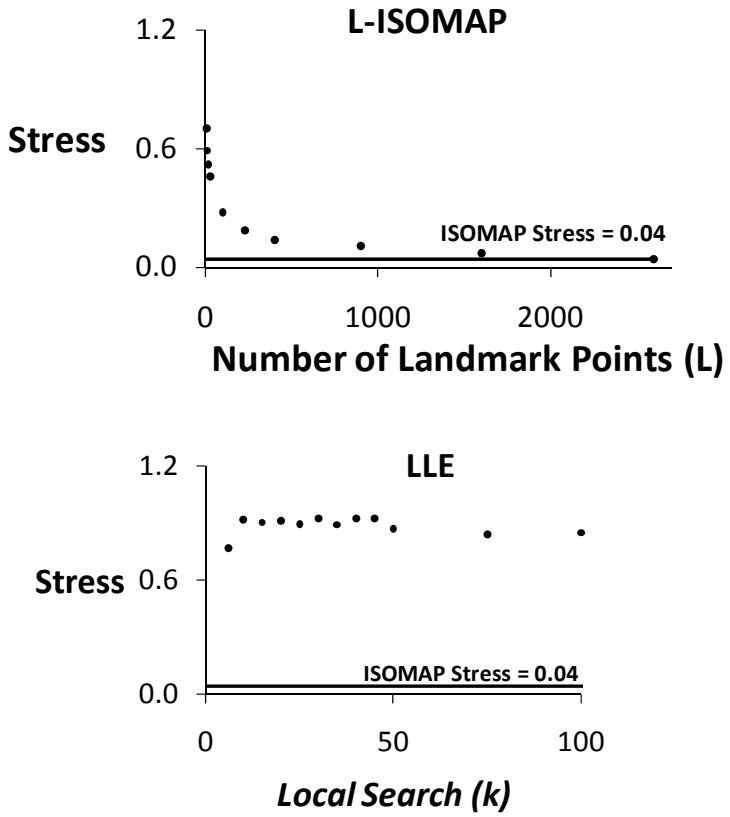


Figure 5.8: Above: Stress using L-ISOMAP. Below: Stress using LLE. Points indicate the stress using various k or L parameters. The solid line represents the lowest possible stress using complete ISOMAP with 2600 dimensions, note this is equivalent to using L-ISOMAP with L=2601.

5.3 Kriging

The only input required to kriging is a covariance function used to calculate the covariance between locations separated by a lag vector (**h**). Once the covariance function is known the kriging weights are solved for all unsampled locations in the geostatistical model and a smooth estimated map is generated. Kriging is implemented by the following algorithm. Steps 4 and 5 have not been discussed and are expanded upon below.

Step 1: Generate the LVA field (Chapter 3).

Step 2: Calculate initial distance matrix between the landmark vertices (L-ISOMAP) and all cells in the model with the Dijkstra algorithm (Chapter 4).

Step 3: Perform L-ISOMAP to embed all cells in a high dimensional Euclidean space (Section 5.2.3)

Step 4: Model an isotropic variogram. This variogram is used to obtain the covariance between locations given the SPD (Section 5.3.2).

Step 5: For every grid cell:

a. Determine the nearest *n* neighbors (Section 5.3.1)
b. Calculate the required *n* by *n* distance matrix
c. From the *n* by *n* distance matrix, calculate the covariance matrix using the modeled variogram
d. Solve the resulting system of equations to determine weights for each datum
e. Calculate the kriging mean and error variance

Kriging for a small example is shown in Figure 5.9. Steps 1, 2 and 3 have been discussed previously. Step 3 is necessary to ensure positive definiteness (Curriero 2005). The required input parameters to kriging are:

1) The LVA field (Chapter 3).
2) The value of *n*. In practice this value is usually set between 20-50 depending on the application.
3) The isotropic variogram. This variogram is modeled from the available sample data.
4) The necessary parameters for dimensionality reduction. These parameters, introduced in Section 5.2, are discussed more thoroughly in relation to a case study (Chapter 6).

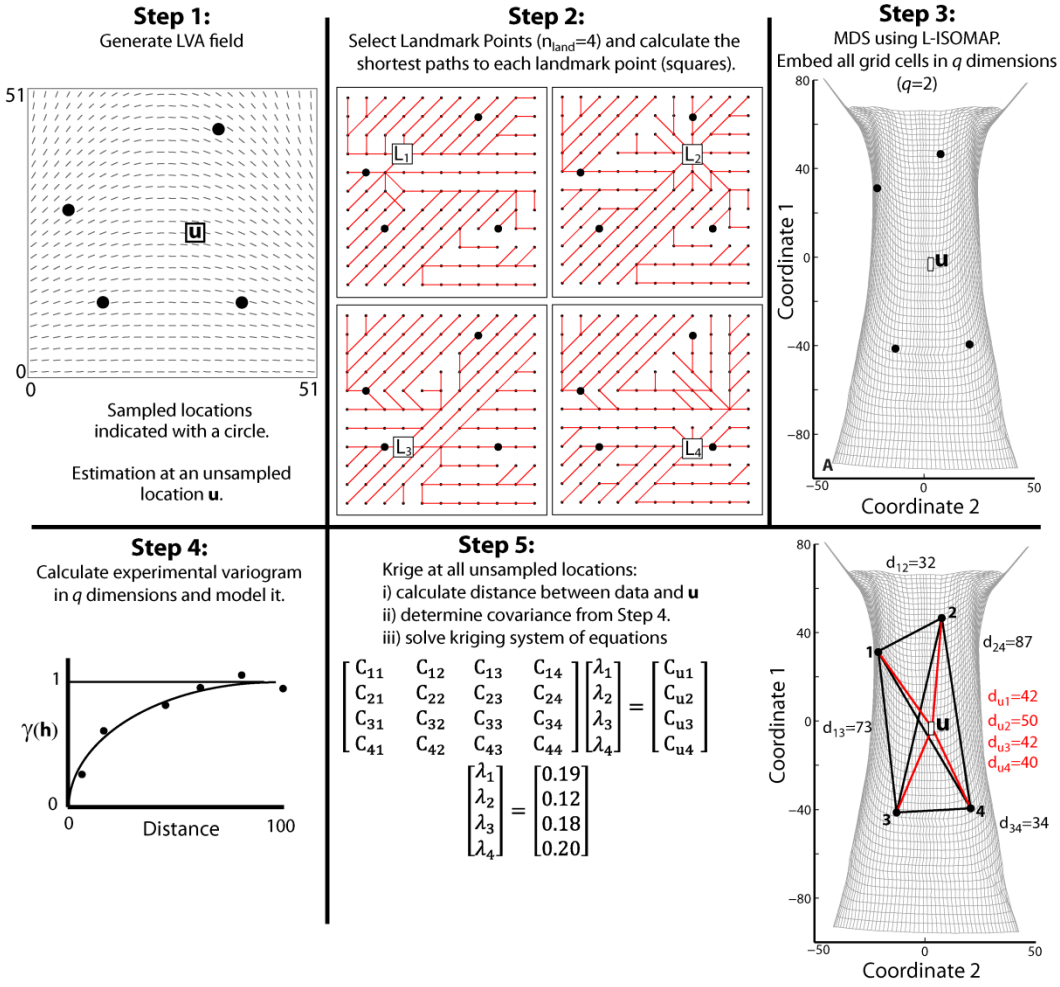


Figure 5.9: Kriging with LVA explained. Note in Step 5, the distance calculated is the Euclidean distance between points in *q* dimensions, which approximates the SPD.

5.3.1 Searching for Nearby Data

For computational considerations, it is common to restrict the number of data used to estimate a given location to the nearest *n* data in the local neighborhood (Deutsch and Journel 1998). This step in kriging requires searching for the *n* nearest neighbors to the estimation location. One search that has been implemented effectively is the superblock search (Deutsch and Journel 1998). This search strategy indexes the available data based on a coarse regular grid and allows for fast retrieval of the nearest neighbors. This indexing is done for each dimension considered, for example, data located in three dimensions would require n^3 superblocks where each dimension is discretized into *n* superblocks (a detailed description of this search can be found in Deutsch and Journel 1998, page 33). In the proposed methodology the available data are embedded in a *q*-dimensional space, use of the superblock search strategy would require indexing n^q superblocks, which is prohibitive in terms of memory requirements. Consider a very small number of discretizations, $n = 10$, while retaining 15 dimensions; this would require the storage of 10^{15} indices. Assuming 4 bytes per number for storage, this would require 3730 TB of storage, exceeding current RAM limits.

An alternative search strategy is to use a *k*-dimensional tree (*k*d tree). Note the change of notation, when discussing dimensionality reduction algorithms *q* denotes the number of dimensions for embedding, while *k* denotes the number of dimensions searched in the tree ($k \leq q$). The *k*d tree is a binary search tree specifically designed for searching in high dimensional space (Kennel 2004). The superblock and *k*d tree search strategies are compared in 3D using a 51x51x26 grid (67,626 cells). Data sets of 1,000, 10,000, 100,000 and 1,000,000 are randomly generated in the grid and the time required to search for *n* nearest neighbors to each grid cell is calculated for each method (Figure 5.10). The large data sets were selected because the *k*d tree is also used in SGS (Section 5.4.1) where all grid cells are included requiring searching through many locations.

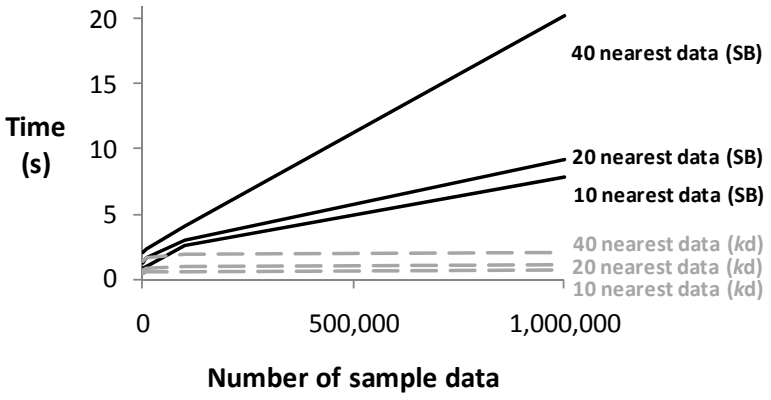


Figure 5.10: Comparing search times for the *k*d tree (*k*d) and the superblock (SB).

Even in low dimensional 3D space the *k*d tree outperforms the superblock search (Figure 5.10). For this study the search radius of the superblock search was optimized for each run such that 10, 20 or 40 nearest data are found at each estimation location with the smallest search radius possible; this represents the minimal time for the superblock search, in practice the search radius is rarely optimized and would likely require additional time. The *k*d tree is more efficient in both CPU time (Figure 5.10) and memory requirements (allocation of a n^q superblock grid is not required). Moreover, in higher dimensions the superblock search becomes intractable while the *k*d tree simply takes additional time and memory (Figure 5.11). Thus, the superblock search (Deutsch and Journel 1998) is replaced with a *k*d tree (Kennel 2004) to allow for searching in the *q*-dimensional embedded space.

Using all available dimensions for searching requires additional CPU time (Figure 5.11). In cases of extremely large models it may be necessary to limit the number of dimensions such that $k<q$ to decrease run times; however, large models require hours of processing time while using 100+ dimensions requires only an additional ~30 minutes for a 2M cell model. The limitation of reducing the number of dimensions searched is that data that are in the search neighborhood of the estimation location may not be discovered. Because the additional CPU time for searching is not large, it is recommended that all *q* dimensions be used for searching while kriging. Searching for neighbors is made more difficult with SGS (Section 5.4.1) and considering $k<<q$ may be necessary for large models due to memory requirements.

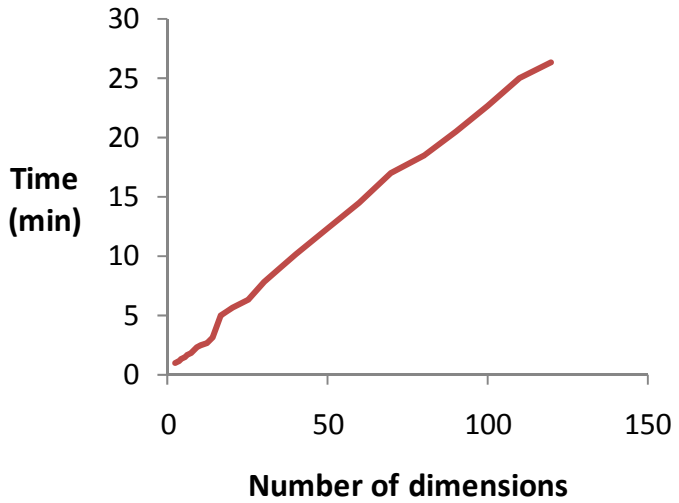


Figure 5.11: Effect of increasing dimensionality and run time of the *k*d tree. The 30 nearest neighbors of 2M grid cells are searched.

5.3.2 An Isotropic Covariance Function

Kriging requires a function to calculate the covariance between locations, such as the variogram, $\gamma(\mathbf{h})$. The variogram represents the variance between two locations in space separated by a lag vector, **h** (Equation 5.8). The variogram is often modeled by first calculating experimental variogram points from the method of moments technique (Equation 5.9) and fitting them with an analytical function.

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \quad \textbf{5.8}$$

$$2\gamma(\mathbf{h}) = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}(x_i - y_i)^2 \quad \textbf{5.9}$$

The experimental variogram can be calculated for discrete values of **h** by pairing sampled data values separated by an appropriate lag, **h**, and calculating the variance. The covariance can depend on the direction of **h** as well as the magnitude; however, an isotropic covariance function is applied in LVA kriging. The magnitude of **h** separating two locations in space, say point *a* and point *b*, can be determined by calculating the Euclidean distance between locations once they have been embedded in *q* dimensions (Equation 5.10).

$$d_{a-b}^2 = \sum_{i=1}^{q}\left(l_a^i - l_b^i\right)^2 \quad \textbf{5.10}$$

where l_a^i is the coordinate of point *a* in the i^{th} dimension.

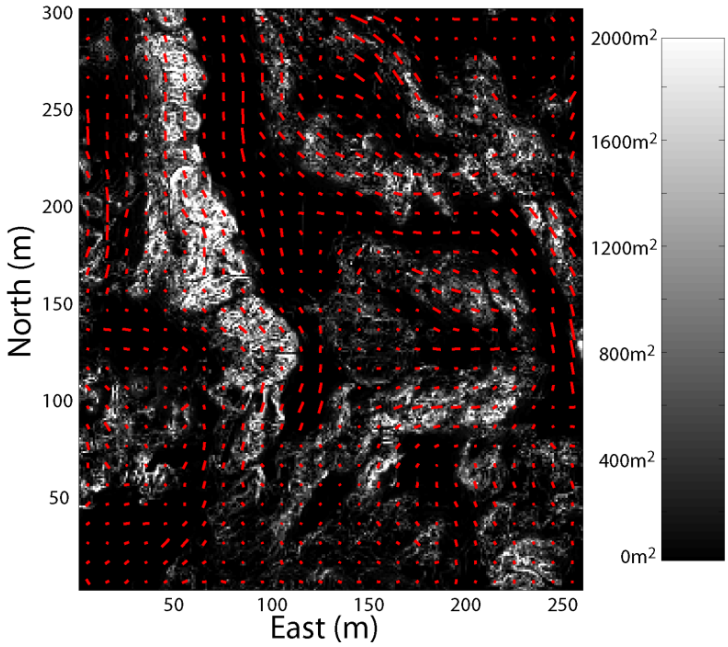


Figure 5.12: Using the exhaustive secondary V data to generate the LVA field for U (from Chapter 3).

Consider the Walker Lake data set introduced in Chapter 3 (Figure 5.12). Using the LVA field developed in Chapter 3 the experimental variogram can be generated after applying L- ISOMAP with L=49 landmark points (7x7 pattern). The experimental

variogram is calculated for the first 3 dimensions (Figure 5.13) and displays very little anisotropy, all three directions are virtually identical with respect to the variogram. Of concern is the reduction in the range of the variogram, initially the range of correlation is between 20m-50m while after applying L- ISOMAP the range is reduced to ~10m. This is because of the imperfect embedding in the new space. This can be improved upon if all 48 possible dimensions are considered (Figure 5.13). Moreover, increasing the number of landmark points, say to 100, better reproduces the variogram range in the original space (Figure 5.14 and Figure 5.15). It is not critical to reproduce the apparent range of the variogram in the embedded space. Consider three cases:

Case 1) Retaining 3 dimensions using 49 landmark points (apparent range = 11m)
Case 2) Retaining 48 dimensions using 49 landmark points (apparent range = 18m)
Case 3) Retaining 99 dimensions using 100 landmark points (apparent range = 20m)

Kriging with each case (Figure 5.15) reveals that the variogram ranges in the 76˚ direction and the 166˚ are actually increased, even though in the embedded space the ranges appeared lower. There is no explicit use of the experimental variograms in the 76˚ and 166˚directions, thus it is not expected that the experimental ranges of 20m and 50m respectively be reproduced. The increase in the variogram ranges (Figure 5.15) is due to the incorporation of the LVA field into the modeling process.

The adoption of an isotropic covariance function is justified as the embedding of the grid removes all anisotropy from the data (Sampson and Guttorp 1992), this is the effect seen in Figure 5.13 where the variogram is nearly identical in the first three embedded dimensions. An isotropic variogram in *q* dimensions can reproduce LVA because these features are captured by the inter-point distance matrix generated from the Dijkstra algorithm, rather than in an anisotropic variogram. Considering anisotropy in the higher dimensional space would be redundant as the anisotropy should be captured in the underlying LVA field used to determine the shortest path between locations.

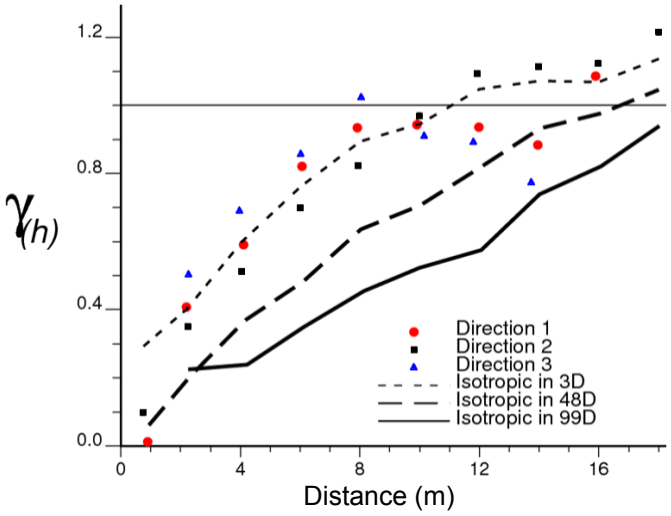


Figure 5.13: Experimental variogram considering data embedded in *q* dimensions using L-ISOMAP.

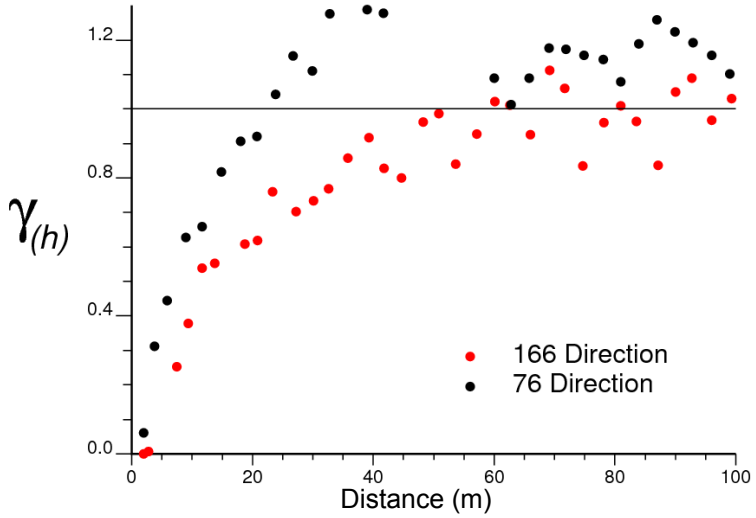


Figure 5.14: Traditional experimental variogram using the Euclidean metric for the V variable in the 166˚ and 76˚ directions.

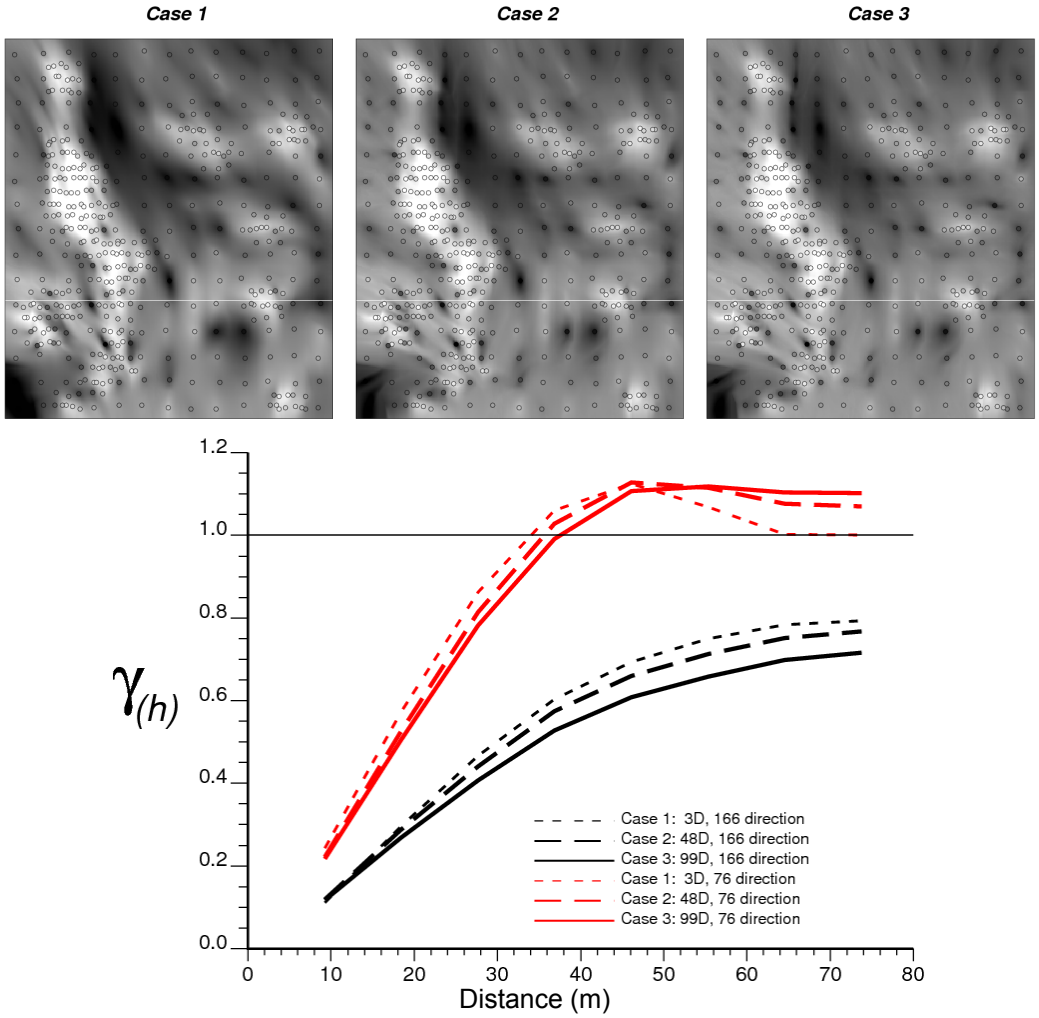


Figure 5.15: Above: Kriging with different cases, color scale ranges from 0 (black) to 1700 (white), plots cover 260m by 300m. Below: Variograms for each case.

Kriging requires the variogram so that a model of the relationship between the SPD and covariance can be determined. Normally, the experimental variogram is calculated from Equation 5.9. A positive definite function must be fit to the available experimental variogram to facilitate modeling. Table 5.3 lists the common variogram models used in geostatistics. It should be noted that any linear combination of these different functions also guarantees positive definiteness; this fact can be used to increase flexibility when modeling the variogram.

Not all functions in Table 5.3 are positive definite for all dimensions. Extensive literature exists that deals with the mathematical validity of covariance functions as well as testing potential functions for positive definiteness, the interested reader is referred to the following references for more information (Matheron 1973; Journel and Huijbregts 1978; Christakos 1984; Armstrong and Diamond 1984; Myers 1991; Cressie 1993; Curriero 2005). Because the available sample data are located in a *q* dimensional Euclidean space, the selected variogram function must also be positive definite in *q* dimensions (Curriero 2005). If *q* is selected to be 3, then all variograms commonly used in practice are permissible. However, to reduce stress (Section 5.2) it is recommended that *q* be as large as possible, thus limiting the available variogram models.

Table 5.3: Some known positive definite variance functions.

Function Name	Equation	Dimensions for positive definiteness	Comments	Relevant References
Spherical	$\gamma(\mathbf{h}) = 1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3$	(see comment)	One of the most common variogram types. Could be made positive definite by considering the volume of intersecting hyper spheres.	
Exponential	$\gamma(\mathbf{h}) = 1 - exp\left(-\frac{(3h)^\omega}{a^\omega}\right)$ $0 < \omega < 2$	Any Dimension	Positive definite in any dimension. Useful as a default variogram type.	
Nugget Effect	$\gamma(\mathbf{h}) = c$	Any Dimension		
Power	$\gamma(\mathbf{h}) = ch^\omega \;\; 0 < \omega < 2$	Any Dimension		
Matern Class	$\gamma(\mathbf{h}) = 1 - \tau^2 + \sigma^2\left\{\left(2^{\kappa-1}\Gamma(\kappa)\right)^{-1}\left(\frac{\lVert\mathbf{h}\rVert_\alpha^{\beta/2}}{\phi}\right)^\kappa K_\kappa\left(\frac{\lVert\mathbf{h}\rVert_\alpha^{\beta/2}}{\phi}\right)\right\}$	Any Dimension	See references for more details.	Curriero, 2006 or Paciorek and Schervish 2006

5.3.3 The Kriging Variance

An interesting result of considering LVA is that the kriging variance is bounded between traditional kriging with the maximum range and the minimum range used in the LVA field. Consider a simple channel example where the anisotropy is 10:1 inside the channel and 1:1 outside the channel. The range of the variogram used in LVA kriging is 8 units with an exponential function (Table 5.3). The resulting kriging variance is bounded by simple kriging with an isotropic range of 8units and simple kriging with an isotropic range of 80units (Figure 5.16). Simple kriging with an isotropic range of 80units

represents a best case, minimum variance scenario for LVA kriging; the shortest path distance cannot be smaller than considering the straight line path between points with a range equivalent to 80units. Similarly, simple kriging with an isotropic range of 8units represents a worst case, maximum variance scenario for LVA kriging; the SPD is smaller than or equal to the straight line path between points with a range equivalent to 8units. As such, the variance for LVA kriging (considering a monotonically decreasing covariance function) lies between the two extremes.

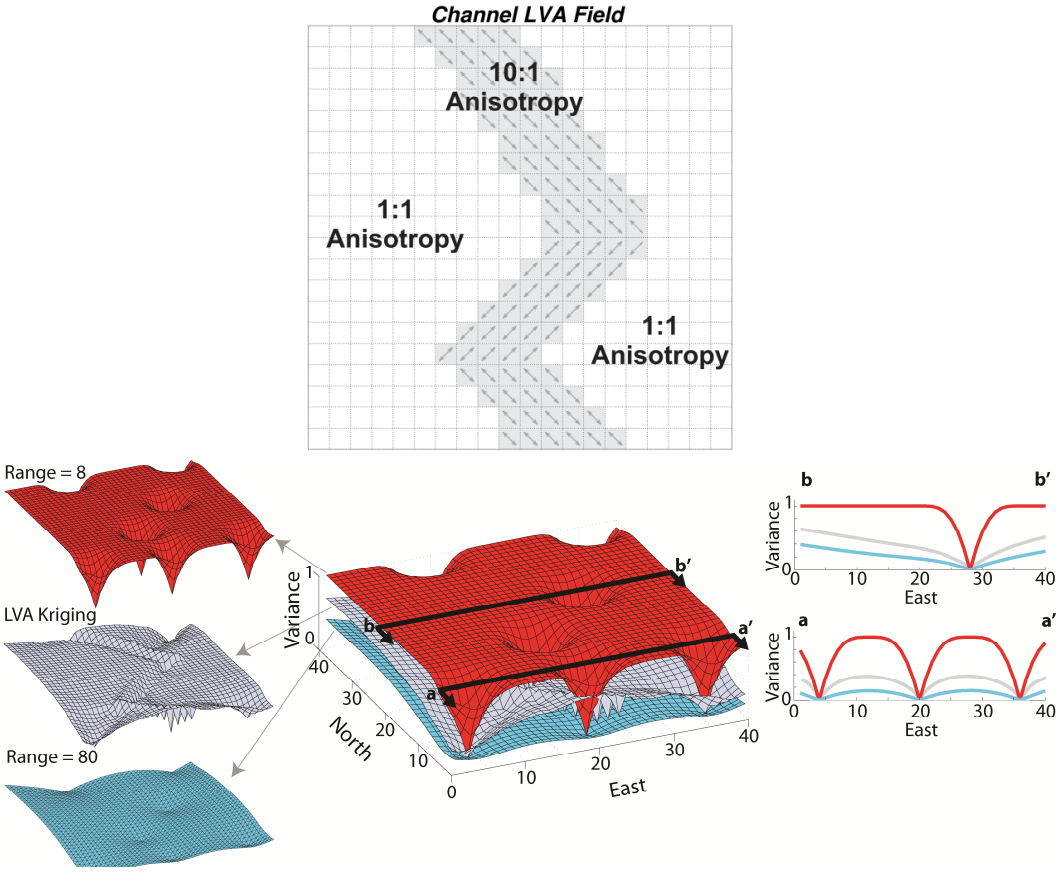


Figure 5.16: Above: LVA field, modeling area is 40units by 40units. Below: Variance with simple kriging (range = 8units and range = 80units) and variance with LVA kriging. Section aa' and bb' are shown to the right.

5.3.4 Synthetic Examples of Estimation with LVA

A single specification of anisotropy is often insufficient to fully describe natural phenomenon. Figure 5.17 contains a number of synthetic examples that highlight some of the types of geological features that can be reproduced with LVA including (1) folding (2) nonlinear channel/vein deposits or (3) smoothly changing local directions of anisotropy.

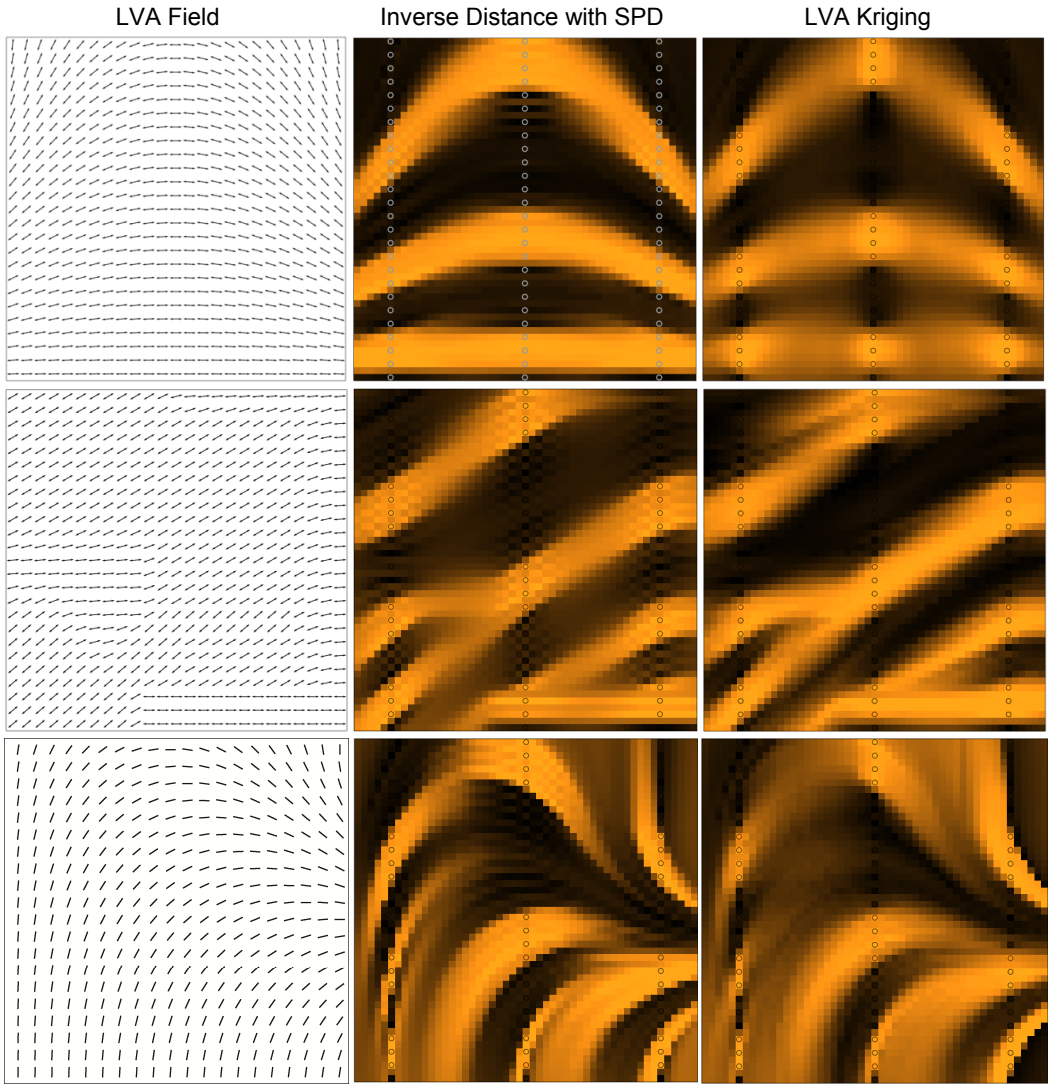


Figure 5.17: Examples of kriging with LVA. Left: LVA field direction, anisotropy ratio is constant 10:1. Middle: Estimates with IDW. Right: Estimates with kriging with an exponential variogram with a range of 200 units. Dimensions of all plots are 51x51 units. The same three data strings with 78 total data were used in kriging. The scale ranges from Black = 0 orange/gray=6.

5.4 Sequential Gaussian Simulation (SGS)

Recall that the motivation for considering SGS is the reproduction of the correct variability seen in the original data (Chapter 1). This is accomplished by generating a number of different equiprobable realizations of the desired random function {Z(**u**) **u** ∈ A}. SGS uses kriging to calculate local conditional distributions; first, a location, **u**, is selected and kriging is performed using the nearby data to determine the parameters of the local Gaussian distribution, $N(m_u, \sigma_u)$. A simulated value is then drawn from the $N(m_u, \sigma_u)$ distribution. This simulated value is added to the growing list of simulated nodes and is used when simulating all successive locations. SGS is described in detail in Figure 5.18. Multiple realizations are accomplished by repeating the algorithm with a new set of random numbers.

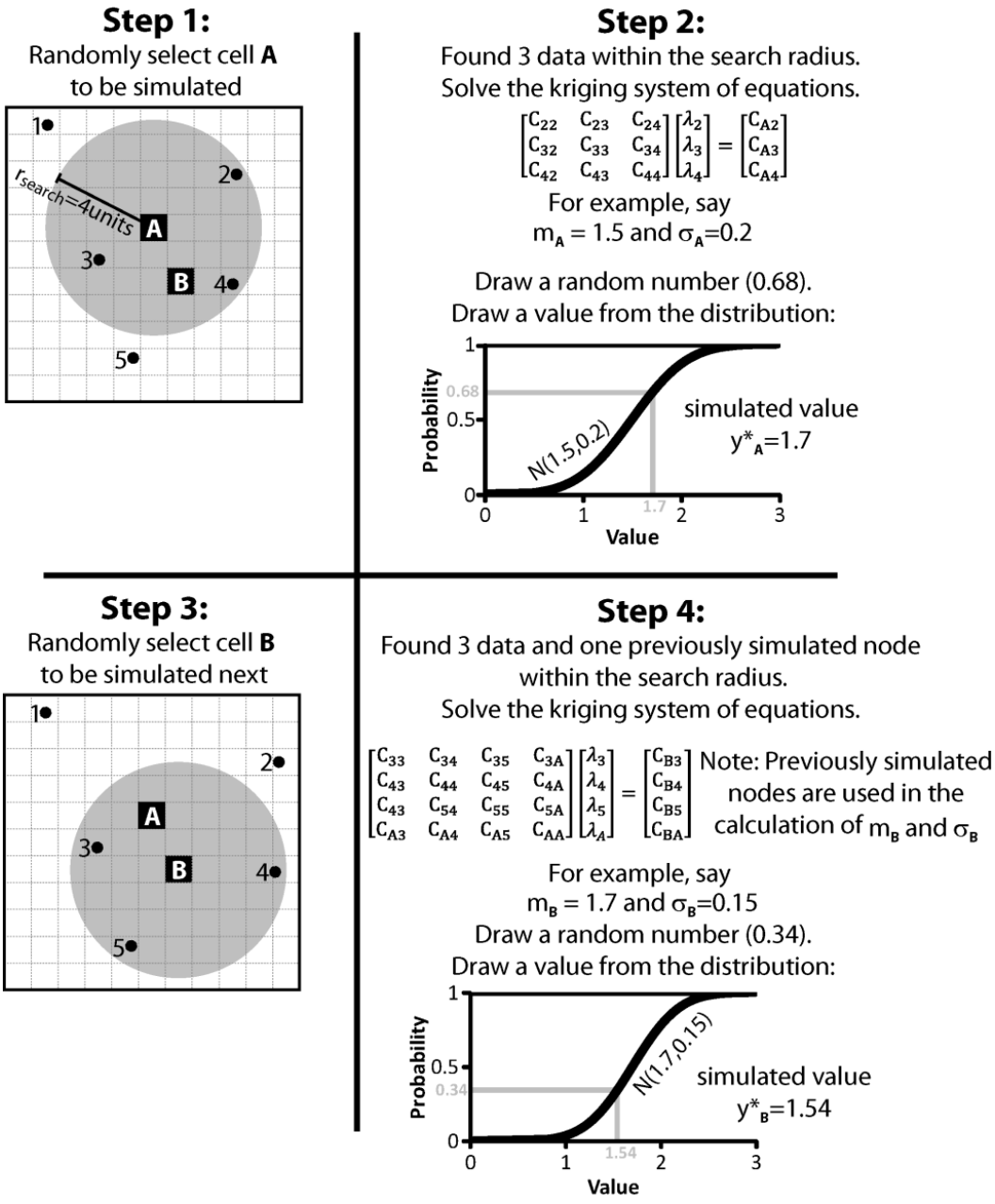


Figure 5.18: SGS explained. Simulation of node A is followed by simulation of node B. This process is sequentially repeated for all locations in the model. Previously simulated nodes are included in the building of the kriging system of equations for each location.

The necessary inputs to SGS are identical to those discussed in kriging: an LVA field; the number of nearby data to consider; an isotropic variogram; and the necessary parameters for L-ISOMAP. There are two aspects of SGS that significantly increase computational requirements and require practical solutions. The first aspect is the use of previously simulated locations when performing SGS. As the sequential process proceeds, more simulated nodes become available. When simulating at **u**, the growing list of previously simulated nodes must be searched to find the nodes in the

neighborhood of **u**, which is more computationally demanding than simply searching for nearby sample data as with kriging (Section 5.3.1). The second aspect of SGS that increases computational requirements is the common practice of generating multiple realizations. While the CPU time required to generate a single realization is reasonable (see Chapter 6) the time required to generate many realizations would not be practical. This section presents solutions to these two issues.

5.4.1 Searching for Nearby Previously Simulated Nodes

Implementation of the SGS algorithm requires the user to limit the number of data used to improve CPU performance. Not all data can be considered as the solution to the kriging system of equations would require excessive CPU time; therefore, it is common to use *m* nearest data to the simulation location and ignore all other data (Deutsch and Journel 1998).

All available conditioning data, such as drill holes or wells, are assigned to grid nodes. This is required for the usage of the Dijkstra algorithm when determining the SPD between locations. As such, no distinction is made between previously simulated nodes and the initial conditioning sample data; data behave as previously simulated nodes that are fixed for each realization. The term *informed nodes* is used hereafter to refer to grid cells that have been previously simulated or have been assigned data values. The problem then reduces to finding the *m* nearest informed nodes to a given location, **u,** in a model with *N* total nodes. This can be accomplished with the *k*d tree presented in Section 5.3.1 but become more difficult as more dimensions are considered (Figure 5.19). All *N* grid locations are loaded into the search tree and the nearest informed nodes are returned.

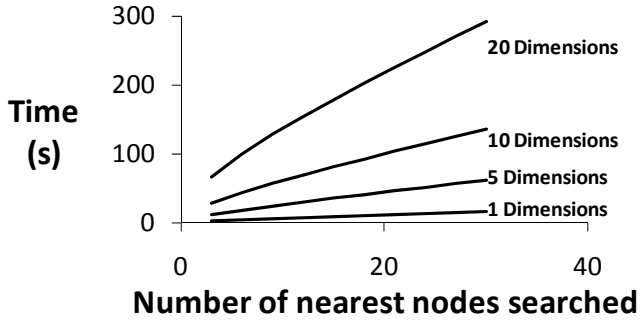


Figure 5.19: Time required to search for the nearest informed nodes for a single realization. A 100x100x50 model (500,000 cells) was sequentially simulated. Retaining more dimensions increases CPU times.

The *k*d tree built for simulation requires more memory than the tree used for kriging because all *N* grid locations must be added to the tree rather than only the initial sample data, often *N* > 2M. Memory limitations may require a truncation of the number of dimensions used for searching (q_{srch}) such that $q_{srch}<q$. The drawback of considering fewer dimensions in searching is the associated error in incorrectly identifying informed nodes that are near to **u**. This is highlighted with the channel example (Figure 5.20). Consider the two cases where (1) all dimensions are used for searching and (2) only the first three dimensions are used for the search. When all dimensions are used the ordering of nodes by distance is shown in Figure 5.21 (above), when only the first three dimensions are used the order is different Figure 5.21 (below). The cross plot of these two orderings reveals that there is some difference in the orderings, but there is a high correlation (0.91) between them (Figure 5.22) and a very good match when the nearest 50 data are considered. The correlation between the ordering using an exhaustive search verses the limited search provides a measure of the error in limiting the search to only three dimensions. This correlation would be 1 if the same ordering is found. Repeating this exercise for each node in the grid and plotting the correlation coefficient between the exhaustive search and the limited search reveals that the error is small when retaining only three dimensions for this example (Figure 5.23). The error depends on the LVA field and is higher in areas of complex anisotropy (Figure 5.23).

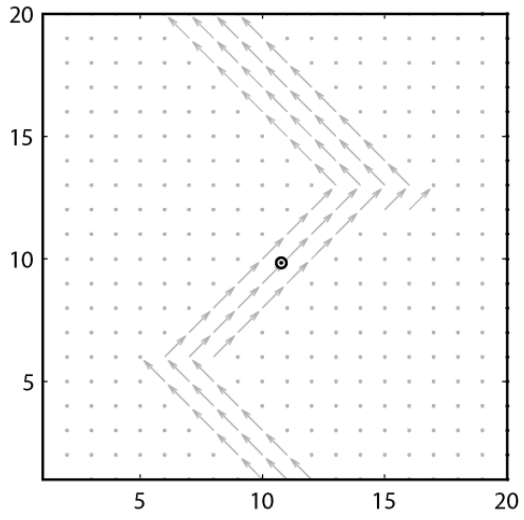


Figure 5.20: Channel LVA field showing central node (circle) used for searching. Inside channel, anisotropy = 10:1.

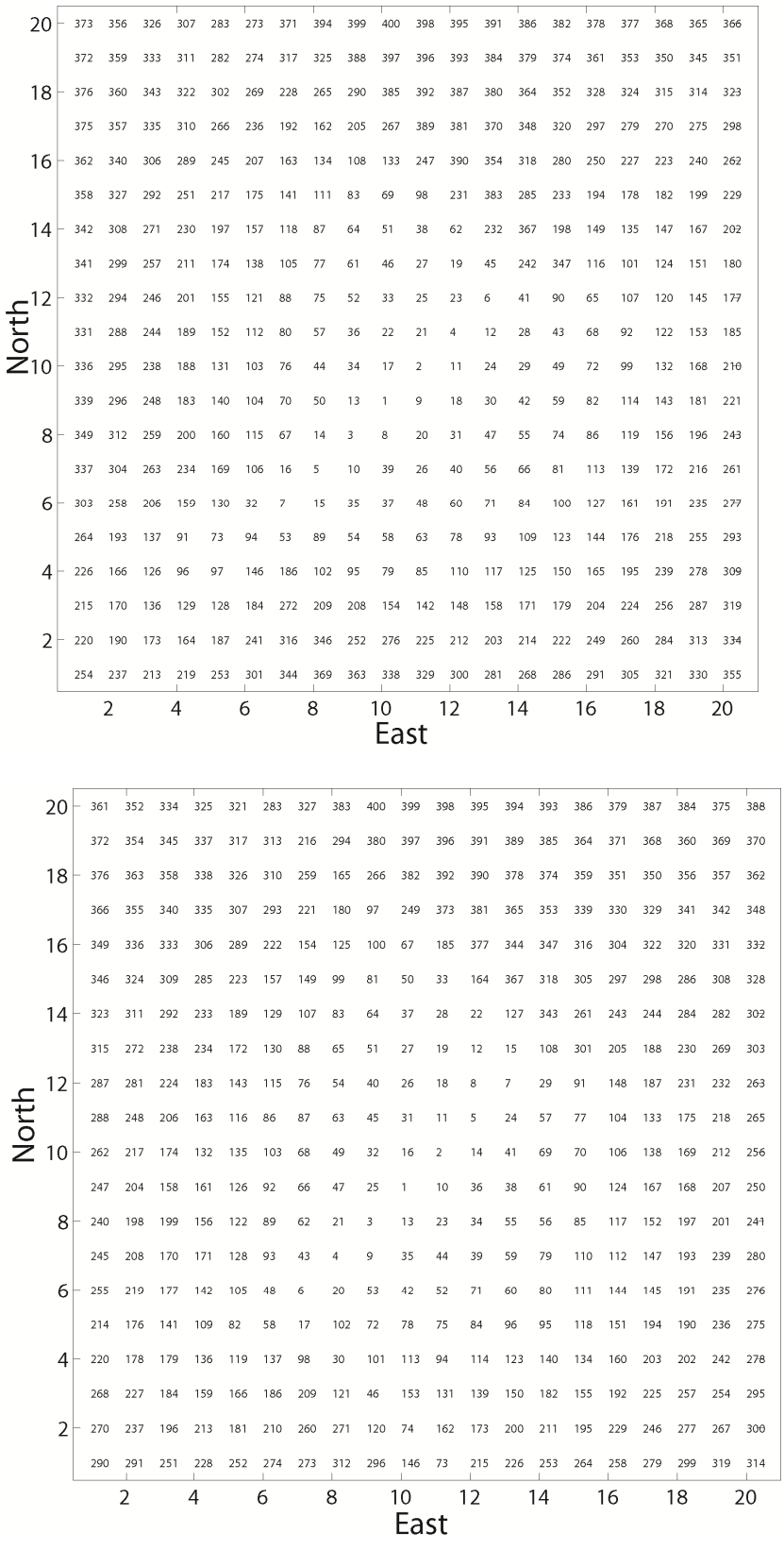


Figure 5.21: Above: Searching using an exhaustive search. Below: Searching using only the first three dimension. Grid cells are numbered by the order found when searching for neighbors of cell 1.

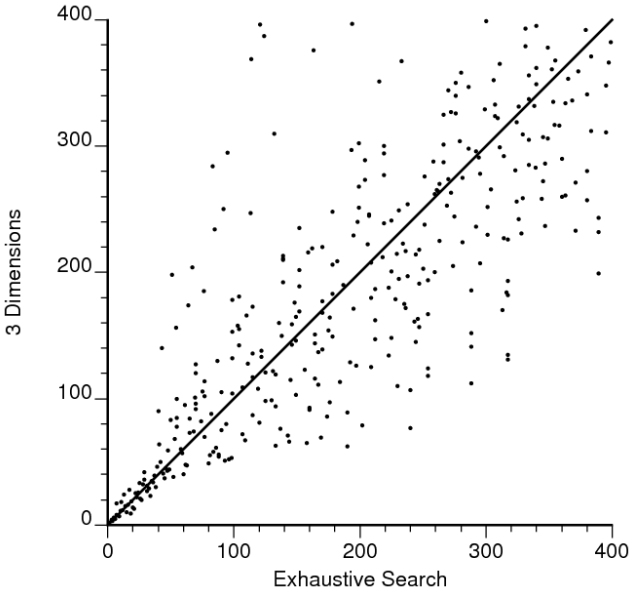


Figure 5.22: Scatterplot between the rank of each cell for the two searches, ρ=0.91.

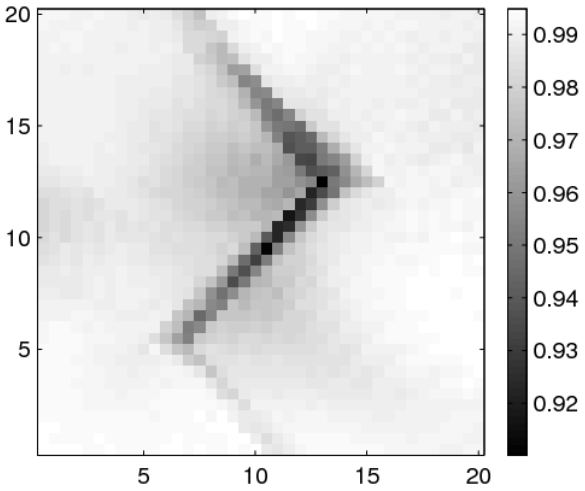


Figure 5.23: Each grid cell is visited and the 400 nearby cells are searched using the exhaustive search and the search limited to 3 dimensions. The rank correlation between these two searches is plotted at every location. The minimum correlation is 0.909.

5.4.2 Considering Multiple Realizations

The CPU requirements of SGS in this implementation are larger than considering constant anisotropy because of the calculation of covariance between locations. When anisotropy is constant, a covariance lookup table can be created to determine the covariance between grid cells given their xyz index (Deutsch and Journel 1998). This is not possible with LVA for two reasons (1) the data are no longer on a regular grid (recall Figure 5.4) and (2) the data are in *q* dimensions and would require a *q* dimensional lookup table that would be too large to store in RAM. Regardless of the memory requirements, a covariance lookup table is only possible with regularly gridded data.

Thus, when simulating, the distances between all informed nodes within the local neighborhood must be calculated as in Equation 5.10.

Multiple realizations, each with a different random path, are typically considered in a geostatistical work flow (Deutsch 2002). Generating realizations with the same random path produces kriging matrices that are identical for all realizations. The kriging matrices can be stored and used for all realizations to improve CPU time by a factor of *r*, where *r* is the number of realizations considered. The calculation of the necessary distances between locations (Equation 5.10) cannot be avoided but if the same random path is used for all realizations, these distance calculations are identical for all realizations.

A different random path is used for each realization (Figure 5.18) because some practitioners believe artifacts may arise otherwise. Artifacts do occur if an axis aligned sequential path is selected (Isaaks 1991); however, there is no evidence that generating realizations using the same *random* path produces artifacts. The concern is that considering only a single random path may produce realizations that are *too similar,* that is, realizations that to do not properly span the space of uncertainty in the random variable modeled, Z(**u**). This can be assessed by comparing the statistics of realizations produced with the same random path to the statistics of realizations produced with different random paths.

Consider the Walker Lake data set, the same data set used in Isaaks (1991) to show that an axis aligned path produces artifacts. 50 realizations are generated with a single random path and compared to 50 realizations generated with different random paths, no visual artifacts are noted (Figure 5.24). Two statistics are assessed (1) the histogram of the realizations and (2) the variogram of the realizations. If use of a single random path does not adequately span the space of uncertainty, the variations in these measured statistics would be reduced. Figure 5.25 indicates that there is no reduction in the space of uncertainty for either the histogram or variogram when using a single random path for each realization; however, this result cannot be generalized to all models and models generated with a single random path explore a more limited region of all possible models. Realizations with different random paths could be generated at an added CPU cost.

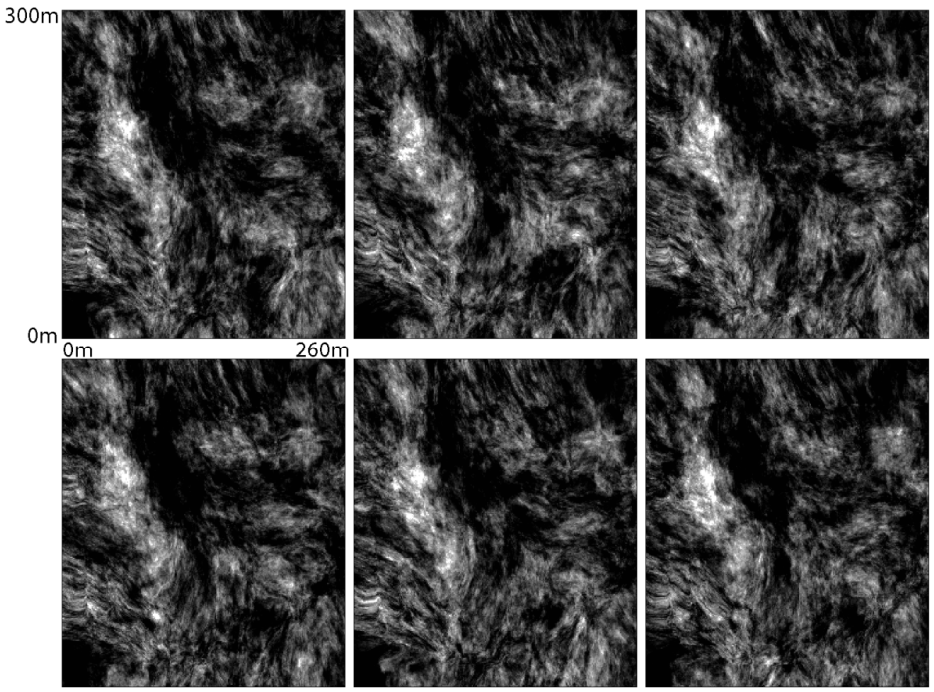


Figure 5.24: Above: Three realizations using different random paths. Below: three realizations using the same random path. The scale ranges from black = 0m^2, gray=1400m^2.

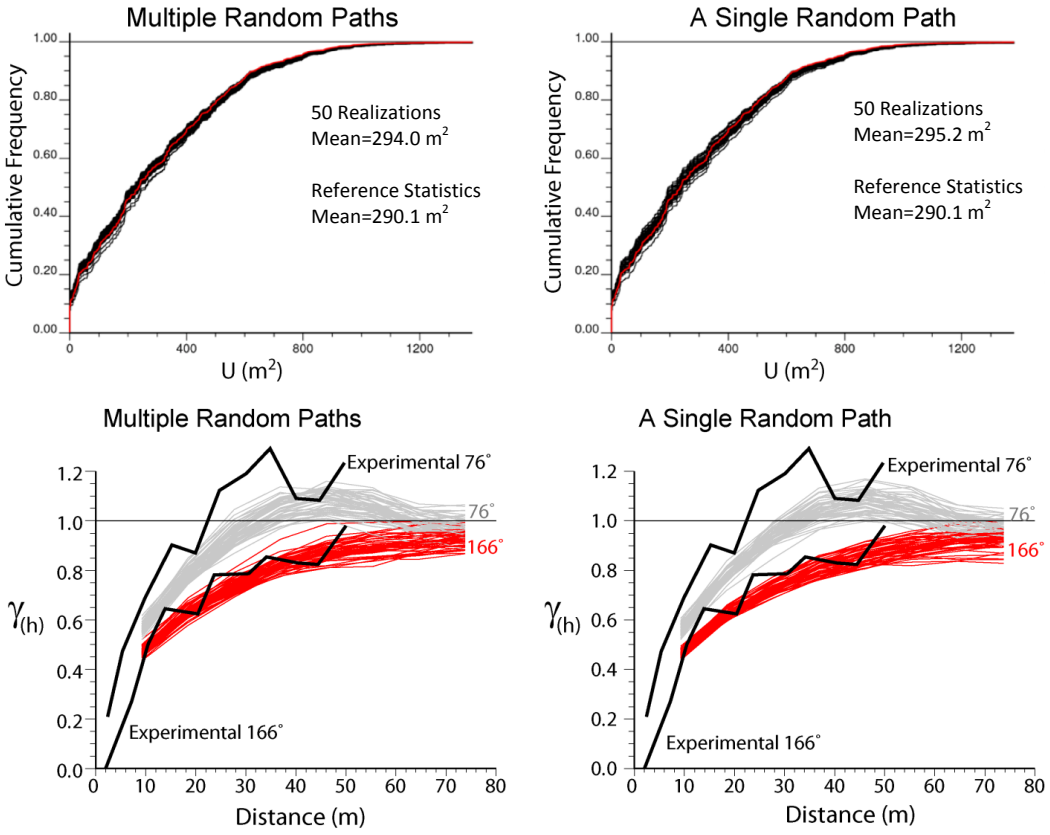


Figure 5.25: Above: Histograms of 50 realizations. Below: Standardized variograms of 50 realizations. SGS with LVA does not explicitly reproduce the experimental variograms in the 76˚ and 166˚ directions. Of interest here is only the variability of the resulting statistics using a single random path.

5.5 Remarks

Estimation and simulation algorithms form the core of geostatistical resource calculations. This chapter presented IDW, kriging and SGS implementations that incorporate LVA. Resource estimation methodologies, such as presented in Journel and Kyriakidis (2004), can be used with the modified algorithms to incorporate LVA.

IDW techniques have the advantage of simplicity. Inference is only required for a single power parameter, ω. There is no need for a positive definite covariance function because a system of equations is not solved; the weights are calculated directly from the SPD. This also implies that dimensionality reduction algorithms are not required to perform IDW. The only computational requirement is the calculation of the necessary distances with the Dijkstra algorithm.

Much of this chapter was driven by practical considerations of implementing kriging and SGS. Ideally, a covariance function that ensures positive definiteness with the SPD metric would be available; however, no such function is known to exist. Dimensionality reduction techniques, such as ISOMAP, provide a mathematical framework within which kriging and SGS can be applied and the original SPD matrix is approximately reproduced. Moreover, L-ISOMAP can be applied without the full SPD matrix between all grid locations providing a significant reduction in computational requirements.

A number of approximations were made in this chapter to reduce the computational requirements of the methodology: (1) embedding of the grid nodes with an approximate dimensionality reduction technique (L-ISOMAP) rather than ISOMAP; (2) reducing the number of dimensions retained for kriging/SGS when searching for nearby informed nodes; (3) considering only a single random path with SGS. These approximations increase the potential for error in the final estimates but have the advantage of reducing CPU run times. Chapter 6 examines the errors resulting from these approximations in the context of resource estimation (i.e. grade/tonnage above a cutoff). The effect of these approximations on the stress criterion (Equation 5.5) is of little interest in a practical application, rather the effect on resource estimation provides a more meaningful measure of the impact of the errors. Thus, the focus of this chapter was on methodology while Chapter 6 demonstrates the practical implications and benefits of considering LVA.

Chapter 6: Case Studies

The methodology for performing geostatistics with LVA was developed in Chapters 4 and 5 with the generation of the LVA field discussed in Chapter 3. This chapter presents the application of this methodology to model two data sets (1) the McMurray formation thickness in the Alberta oil sands and (2) gold grade in a copper/gold porphyry deposit. The first case study examines the application of geomodeling with LVA to a 2D thickness data set. 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 parameters, variogram modeling, LVA field resolution, and sensitivity to other model inputs are examined. Resulting errors from approximations such as reducing the number of landmark points in the application of L-ISOMAP are quantified. The second case study involves modeling gold grade within a porphyry deposits and demonstrates the methodology in three dimensions to provide guidance on practical issues when extending to the third dimension.

6.1 Alberta Oil Sands Thickness

The data used in this case study can be obtained from www.ercb.ca. The data was used by the Energy Resources Conservation Board (ERCB) 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 layer is the McMurray Channel (Figure 6.1) which contains a significant proportion of the hydrocarbon resource in the region. The ERCB study area consisted of two distinct regions, the northern and southern study areas (Figure 6.2). Only the southern region, where the majority of the wells are located, is considered for this case study. Of the 3265 wells evaluated for bitumen thickness, 2342 are located in the southern region.

Further information on the depositional environment and stratigraphy of the Athabasca oil sands can be found in Ranger and Gingras (2003), Hein *et al.* (2000) or Carrigy (1959).

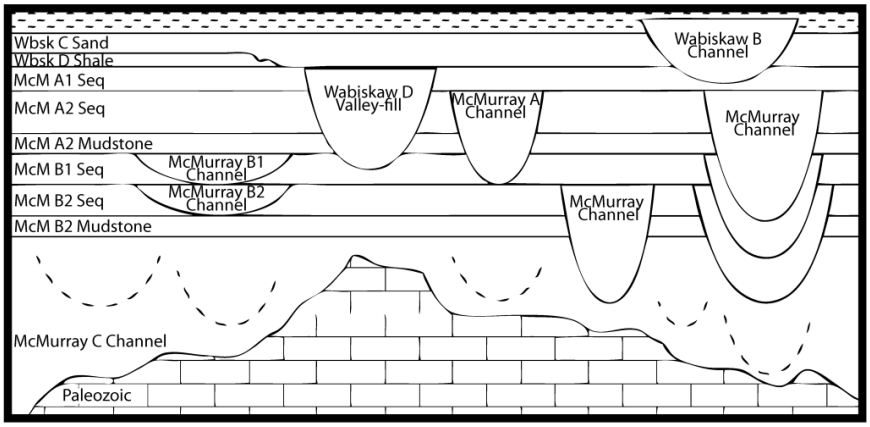


Figure 6.1: Stratigraphic model of the area (Warren 2003). Average thickness of the McMurray channel is 17m.

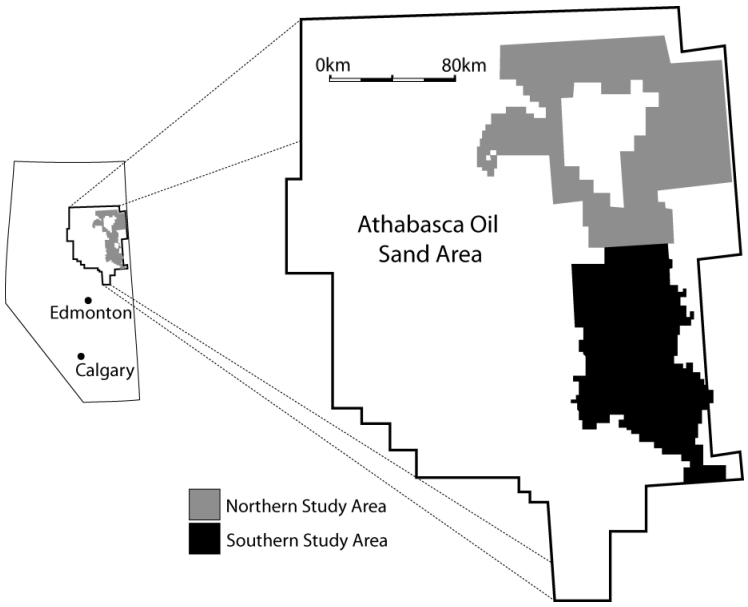


Figure 6.2: Location of the ERCB study area (Warren 2003).

6.1.1 The Data

The available data extend 111km in the x-direction and 170km in the y-direction. In the southern area there are a total of 2342 wells where the net continuous bitumen (NCB) thickness has been determined (Figure 6.3). The data is partitioned into four subsets based on drill date (Table 6.1). The base data set contains all wells drilled prior to the year 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. This mimics the realistic evolution of a deposit as past 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. The cell size is selected to generate a model size that is small enough to assess the sensitivity of the methodology to the approximations discussed in Chapter 5; implementing a large number of offsets in the Dijkstra algorithm or a large number of dimensions with MDS is only possible on smaller models.

Table 6.1: ERCB data summary. Only data in the southern area (Figure 6.2) are considered.

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

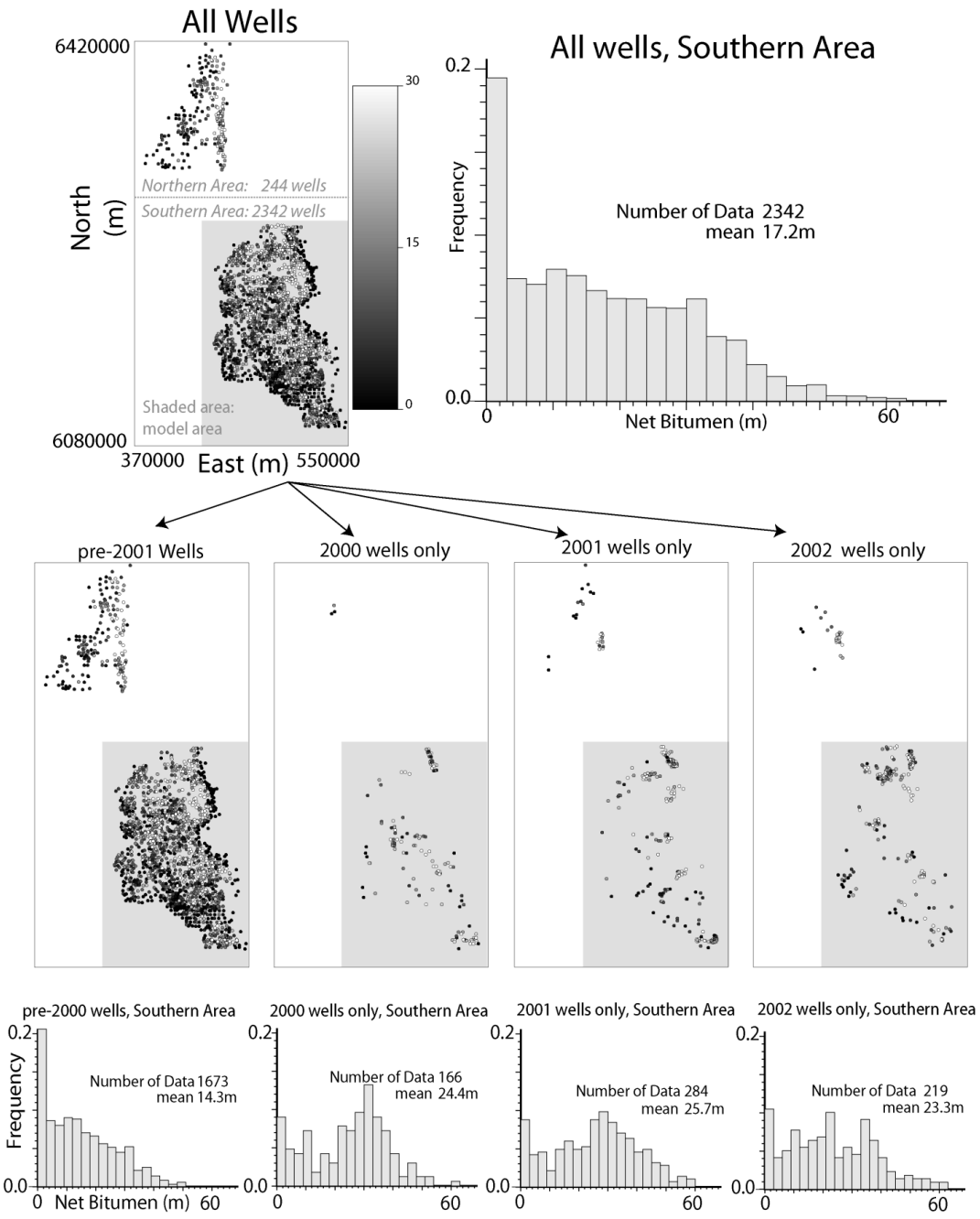


Figure 6.3: Wells in the ERCB study area. Wells are subdivided into four data sets based on their completion dates. The shaded area indicates the extents of the geostatistical model for this case study. Histograms are provided for wells in the southern area.

6.1.2 The LVA Field

In this case study it is assumed that 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 correlated 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 an LVA field is the moment of inertia method (Section 3.3.2). Recall that this methodology provides 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 covariance maps. Square windows of 10km, 12.5km and 15km are used (Figure 6.4 and Figure 6.5). The window sizes were selected based on a visual inspection of the resulting LVA fields; window sizes larger than 15km produce LVA fields that appear too smooth and do not capture the locally varying directions of continuity while window sizes smaller than 10km appear too erratic.

Two manual LVA fields (Figure 6.6) based on the 2000 data are generated in addition to generating the LVA field from the moment of inertia method. This is accomplished by block kriging (Deutsch and Journel 1998) the NCB data with a long range variogram and analyzing the resulting smoothed map. Block kriging is implemented to produce a smoother map that highlights large scale features. A normal score transform (Deutsch and Journel 1998) is used before kriging as simulation is performed in Gaussian units and requires input parameters, such as the LVA field and the variogram, in Gaussian units as well.

Directions of continuity are manually selected based on the block kriged map (Figure 6.6) and the x and y components of the directions are kriged to generate an exhaustive directional field (Section 3.3.4). 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 6.7). The procedure, also known as leave one out analysis, is to estimate each data location with LVA kriging and compare the estimated values to the known truth. This is done for each datum and 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 improvement in the cross validation results (Figure 6.7).

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 (Figure 6.8 left). The anisotropy ratio is then block kriged to generate the locally varying anisotropy ratio map for the second manual LVA field (Figure 6.8 right).

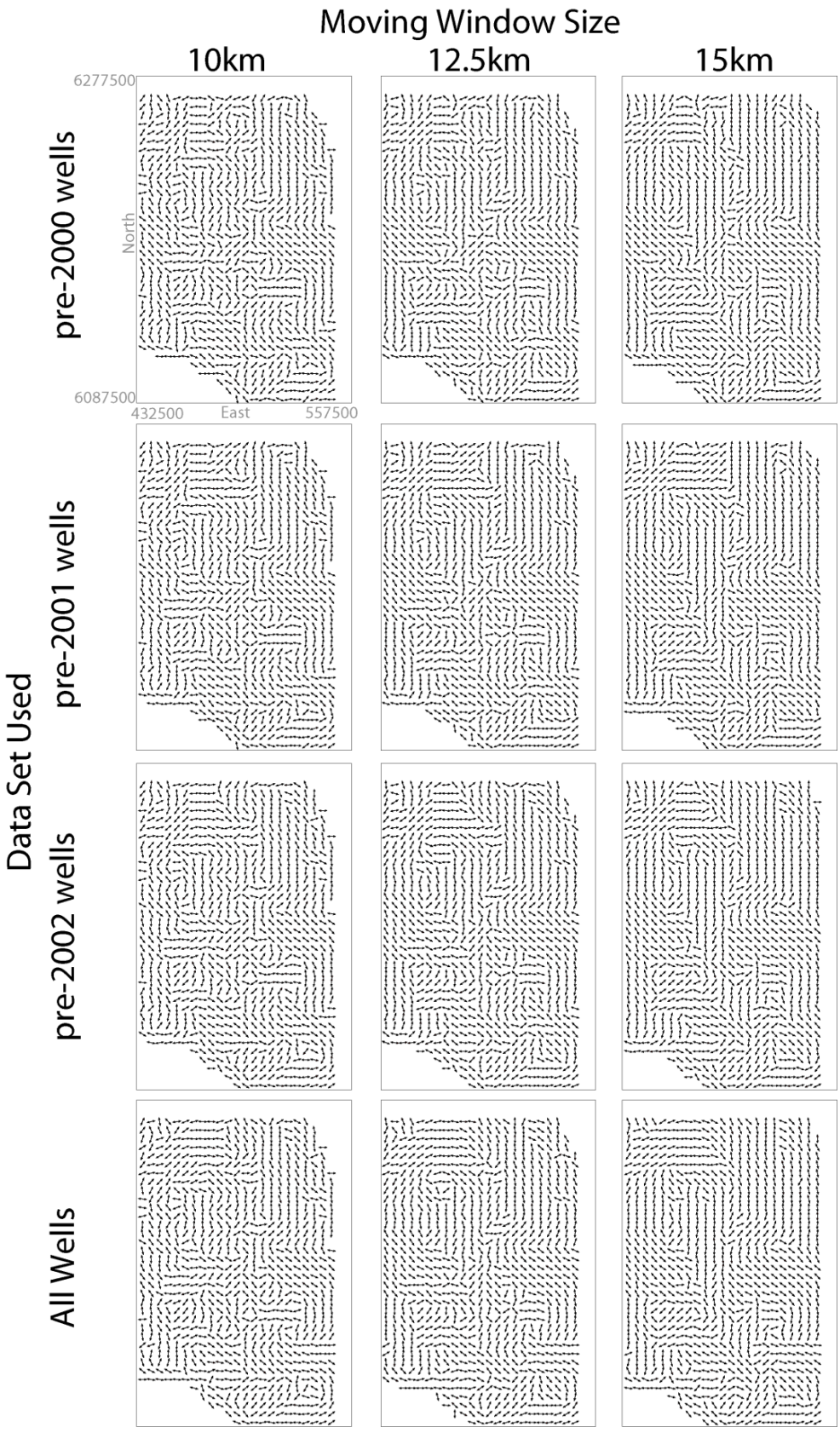


Figure 6.4: Major direction of continuity for the LVA fields from the moment of inertia method.

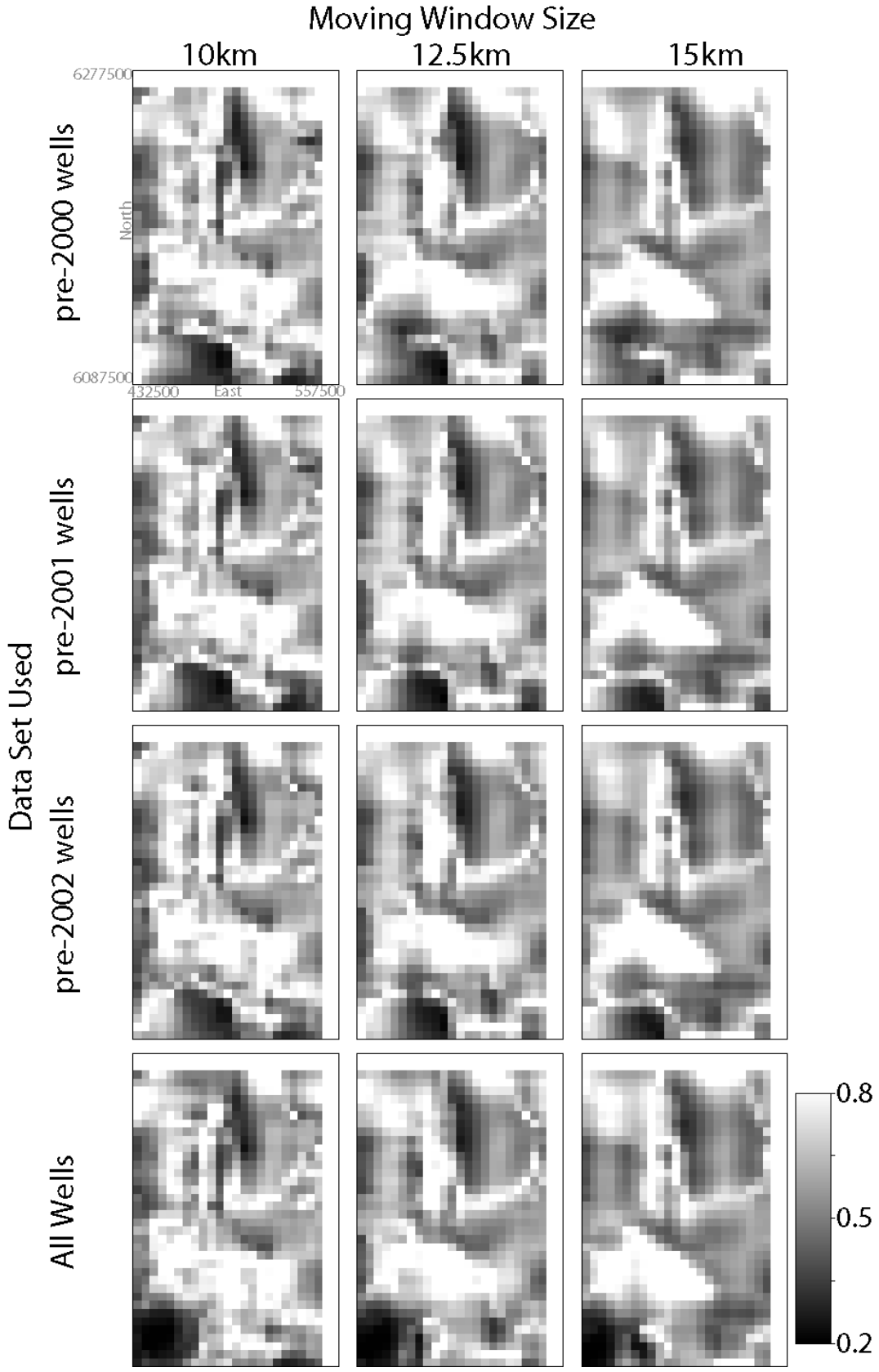


Figure 6.5: Anisotropy ratio for the LVA fields built with the moment of inertia method.

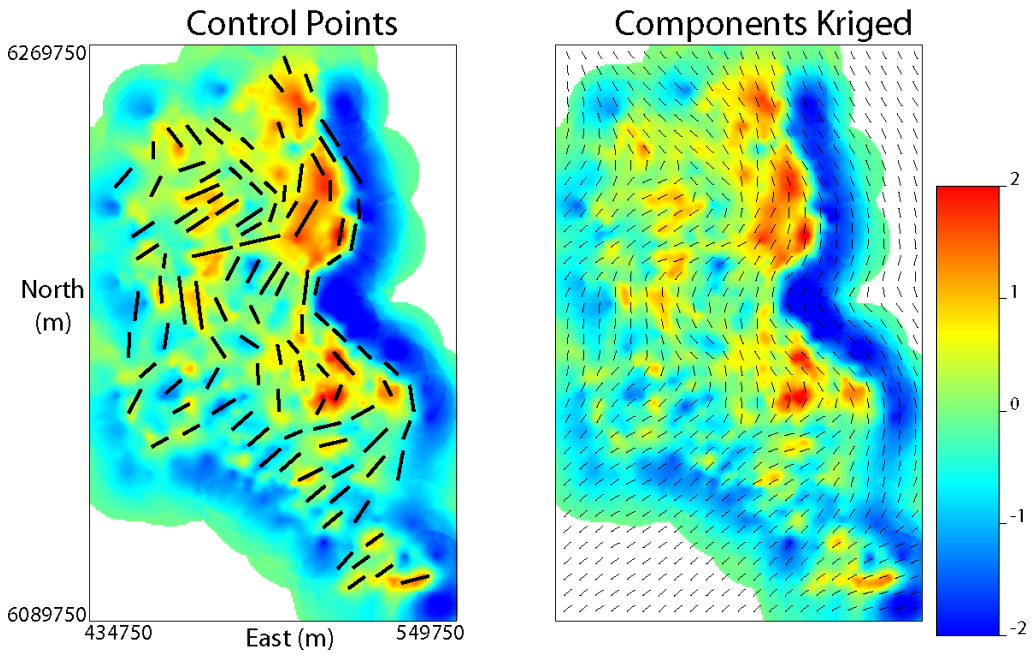


Figure 6.6: Left: Lines manually selected based on block kriged NCB. Right: Direction of anisotropy from kriging the components of the manually selected directions. NCB in Gaussian units is plotted. Plot dimensions are identical.

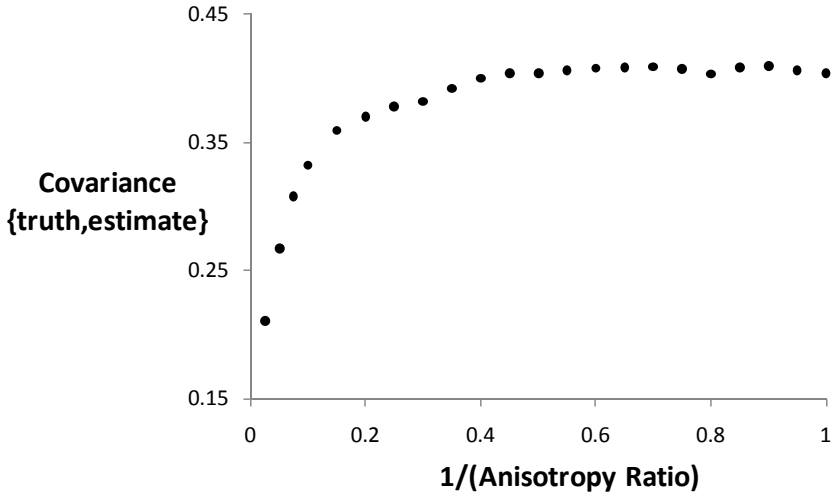


Figure 6.7: Resulting covariance between the truth and the estimate for various anisotropy ratios.

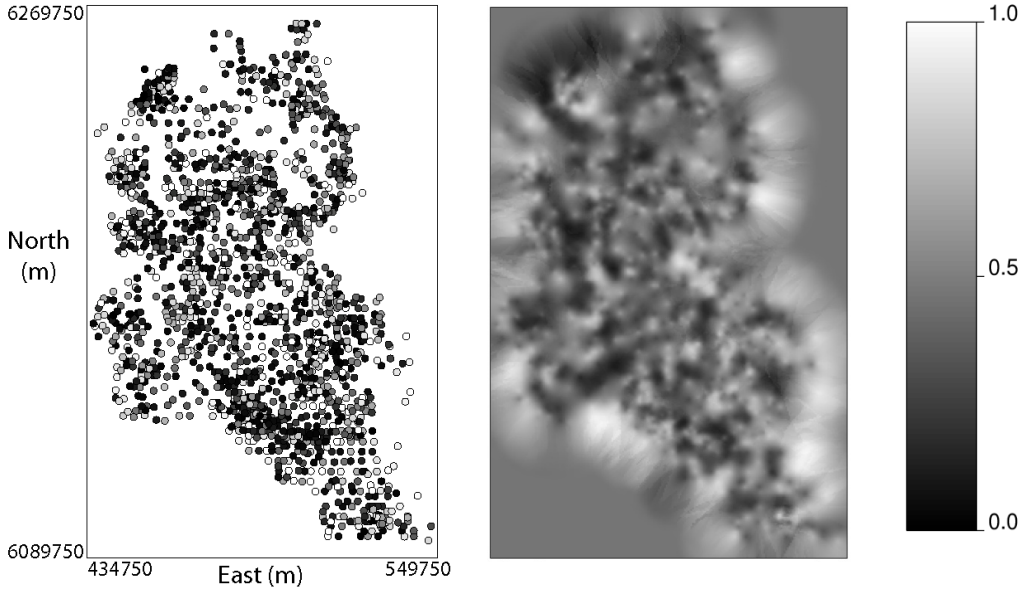


Figure 6.8: Left: Data for kriging the anisotropy ratio. Right: Resulting anisotropy ratio map. Plot dimensions are identical.

6.1.3 LVA Field Assessment

Using the moment of inertia method, an LVA field is built for each data set (Figure 6.4 and Figure 6.5). Preferably, the inclusion of additional data from 2000, 2001 and 2002 should not significantly affect the LVA field if the LVA field is consistent with the deposit. If the LVA field changes significantly with additional data, the appropriateness of the LVA field to this case study would require reassessment. Two graphical methodologies are provided to analyze the difference in LVA fields: (1) vector difference plots and; (2) maps and histograms of the errors (Figure 6.10).

To generate the vector difference plots the anisotropy specification is treated as a vector with the direction of anisotropy corresponding to its orientation and the anisotropy ratio to its magnitude. Comparing two LVA fields is accomplished by subtracting the vectors and plotting the result. This provides a graphical representation of the variation in anisotropy at a single location and is shown for the LVA fields generated with the moment of inertia method and a 12.5km moving window (Figure 6.9). The addition of new data does not significantly alter the features of the LVA field; similar results were found for 10km and 15km moving window averages.

A second measure of the difference between LVA fields is the difference between the anisotropy ratios and angles (Figure 6.9). There are very few visual differences between the LVA fields generated with additional data, indicating the LVA field generation methodology is stable for this example. The main differences in the LVA fields are concentrated at the edges of the available well data where the addition of new wells can have a large effect. This is expected as these areas have the most uncertainty in the LVA field due to limited data availability.

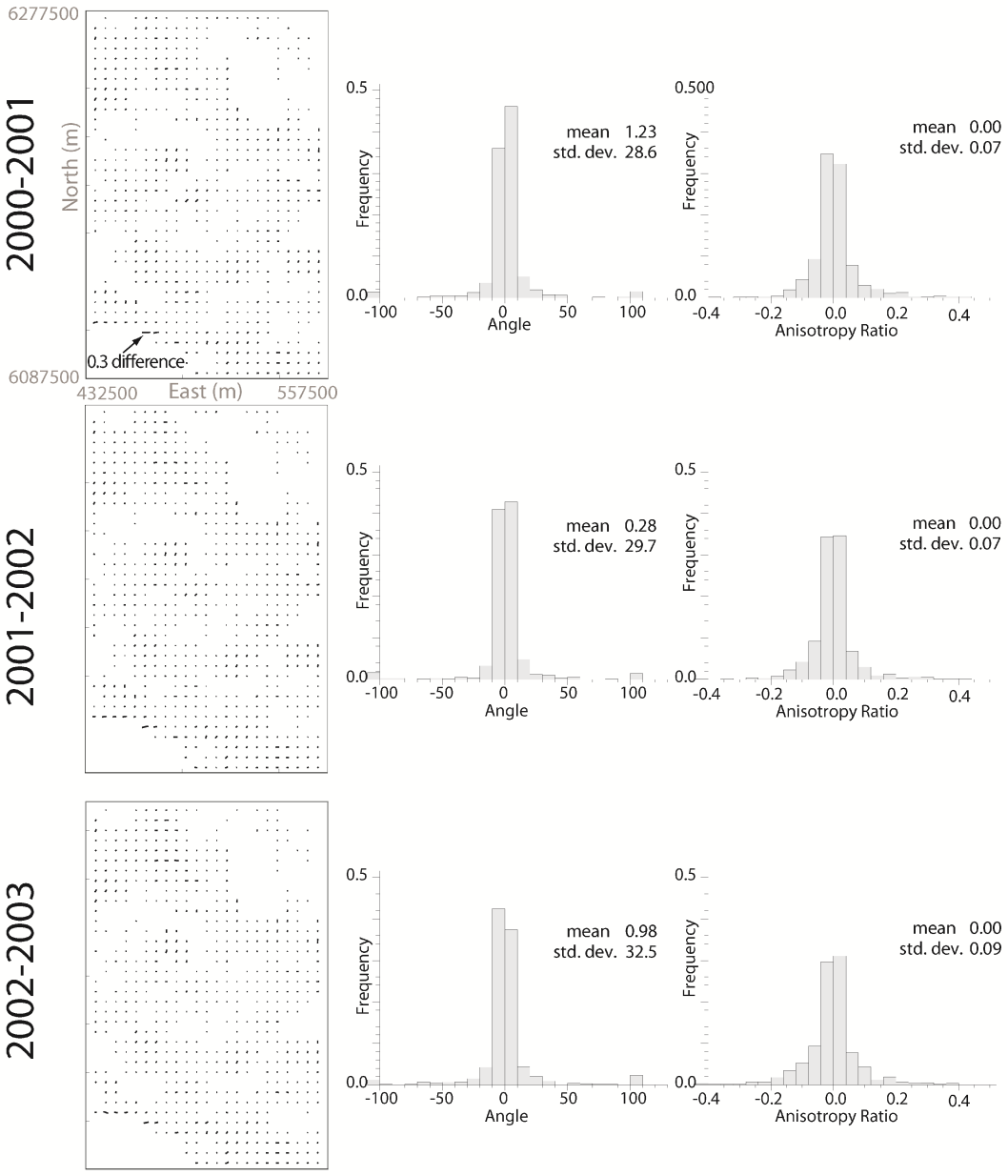


Figure 6.9: Left: vector difference between data sets (i.e. 2000-2001 shows the vector difference between the LVA field built for the 2000 data and the 2001 data). Right: Histograms show the **difference** in the angles and anisotropy ratios between the respective LVA fields. All LVA fields used a 12.5 km moving window average.

The LVA fields are similar (Figure 6.9) because all the moment of inertia fields are generated with the same methodology and additional data to not overly alter the LVA field. Of greater interest is the comparison between the LVA field manually generated and the moment of inertia method. Because the LVA fields are similar for each year considered, only the LVA field generated from the pre-2000 data set and a 12.5km moving window is shown (Figure 6.10). There are four highlighted regions within the data extents where the LVA fields vary significantly. The data in these regions do not display clear anisotropy, thus, the difference is attributed to the uncertainty in the LVA field. The two LVA fields show similar anisotropy outside the highlighted areas where the anisotropy is well defined by denser data.

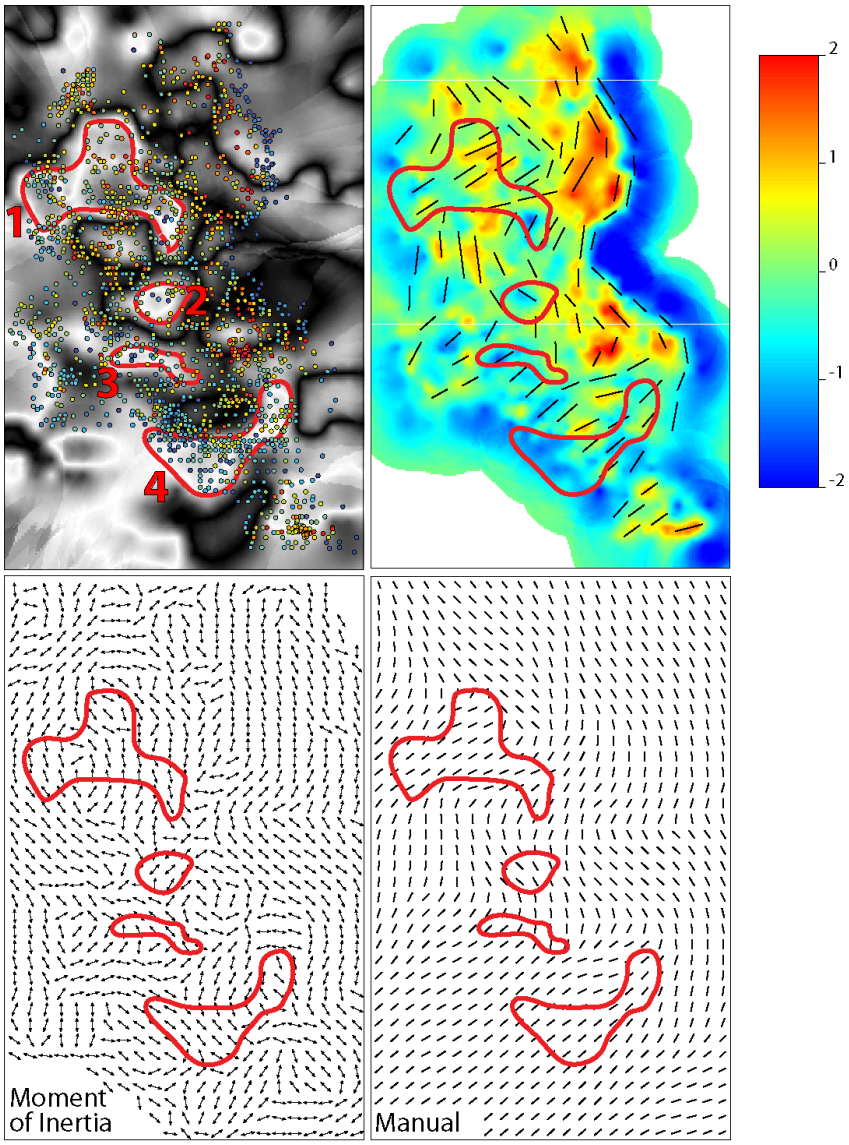


Figure 6.10: Comparison between the moment of inertia LVA field (below left) and the manual LVA field (below right). Above Left: Difference in the angle between the LVA fields (white = 90˚ difference, black = 0˚ difference). Above Right: block kriging, in Gaussian units, used to determine manual LVA field (Figure 6.6) provided for comparison. Both fields were built using the pre-2000 data. Plot dimensions are 434750-549750 in x and 6089750-6269750 in y.

Often the manual generation of the LVA field, when available, is preferred as it is a product of all expert knowledge available for the project; however, deterministic manual LVA fields often do not capture the uncertainty in the LVA field because generating many equiprobable fields is difficult. Close visual examination of the highlighted regions on Figure 6.10 does not suggest that one methodology is more consistent with the conditioning data. Rather, all LVA fields are carried forward to assess the impact of the LVA field on the methodology. A single LVA field may suffice in situations where the LVA field is known with more certainty.

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. The different LVA fields are summarized in Table 6.2.

Table 6.2: Summary of different LVA fields generated.

Methodology	Data set used	Number of LVA fields	Comments
Moment of Inertia	2000,2001, 2002,2003	12	3 LVA fields were generated for data belonging to each year.
Manual A	Pre-2000	1	Manual A considers a constant anisotropy ratio of 1:2.2
Manual B	Pre-2000	1	Manual B considers a variable anisotropy ratio determined from crossvalidation

6.1.4 L-ISOMAP

The mathematical requirement for positive definiteness of the kriging system of equations is addressed by L-ISOMAP (Chapter 5). The input to L-ISOMAP is the location of *L* landmark points that are used to embed the geostatistical grid into a high dimensional 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 shortest path distance (SPD, Chapter 4) from the *L* landmark points to each grid cell is required for embedding with L-ISOMAP. Ideally, every cell in the geostatistical grid would be considered a landmark point, however, this may require prohibitive CPU time to determine the required distances with the Dijkstra algorithm. Figure 6.11 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 5M 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 $L = 380$.

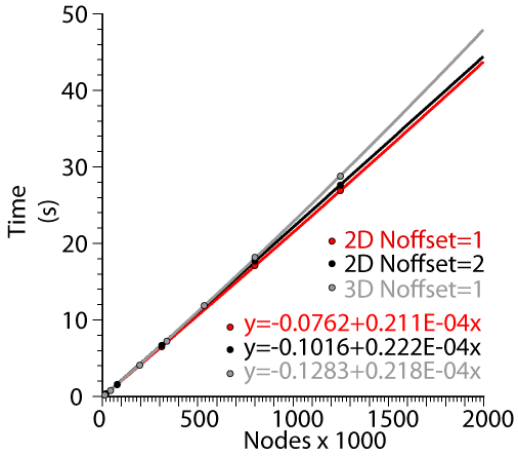


Figure 6.11: Time required to calculate the distance from *n* grid nodes to a single landmark point.

A more theoretical determination of *L* requires the use of an error measure such as the stress criterion (Equation 6.1). *L* could be selected such that the error criterion is below an acceptable threshold. A drawback to this methodology is that the threshold can vary depending on the LVA field used; it is not possible to select an acceptable stress level without knowing the lower bound on the stress. This lower bound is determined by implementing traditional classical ISOMAP, i.e. L-ISOMAP with a landmark point at each grid cell. Determining this lower bound is too CPU intensive for large models. Moreover, the stress is not necessarily a measure of importance in a geostatistical application; consider increasing every $\hat{d}_{ij}^{q}$ in Equation 6.1 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). Alternatively, a more practical solution based on the necessary CPU time can be applied.

$$Stress(q) = \left\{ \frac{\sum_{j=1}^{N} \sum_{i=1}^{j-1} (d_{ij}^{q} - \hat{d}_{ij}^{q})^2}{\sum_{j=1}^{N} \sum_{i=1}^{j-1} (d_{ij}^{q})^2} \right\}^{1/2} \qquad \textbf{6.1}$$

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 in a geostatistical model.

Once the number of landmark points has been selected their location must be determined. The most straightforward implementation would be a regularly spaced pattern of points. The alternative is to optimize the location of the landmark points so as to minimize an error function. As discussed, the stress measure may not be an ideal error measure because the lower bound on the stress is not known aprori.

The effect of L-ISOMAP on the input grid can be visualized by plotting each grid in the first 2 or 3 dimensions of the embedded space (Figure 6.12). For visualization of a 3D grid see Figure 6.24. The smoothness of the manual LVA field is evident, while the moment of inertia method generates more erratic grids. Note the length of the grid in the first three coordinates. 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.

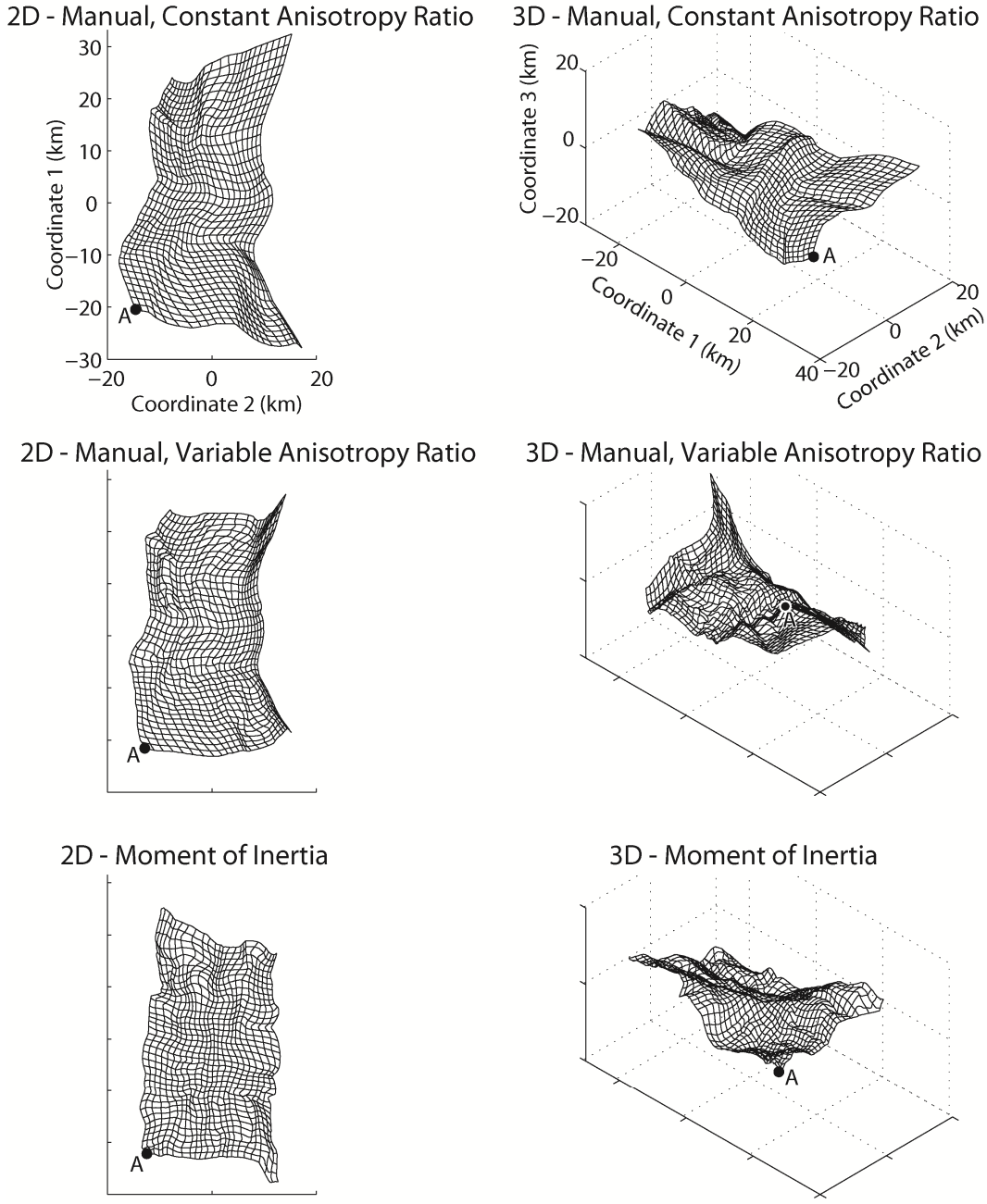


Figure 6.12: Grids embedded in 2 and 3 dimensions for select LVA fields. The moment of inertia LVA field was generated with a window size of 12.5km. All plots have identical axes.

6.1.5 Estimation

With the previous determination of the LVA field (Section 6.1.2) and the location of the landmark points for L-ISOMAP (Section 6.1.4), kriging can be applied to map the NCB. The same modeled variogram was used for all kriged maps (Figure 6.13) as the experimental variograms are similar for each LVA field considered. Thus, the only difference in the kriging maps is the LVA field (Figure 6.14). The modeled variogram consists of two exponential structures with ranges of 900m and 12,000m 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. Inverse distance maps using the SPD are also provided for comparison.

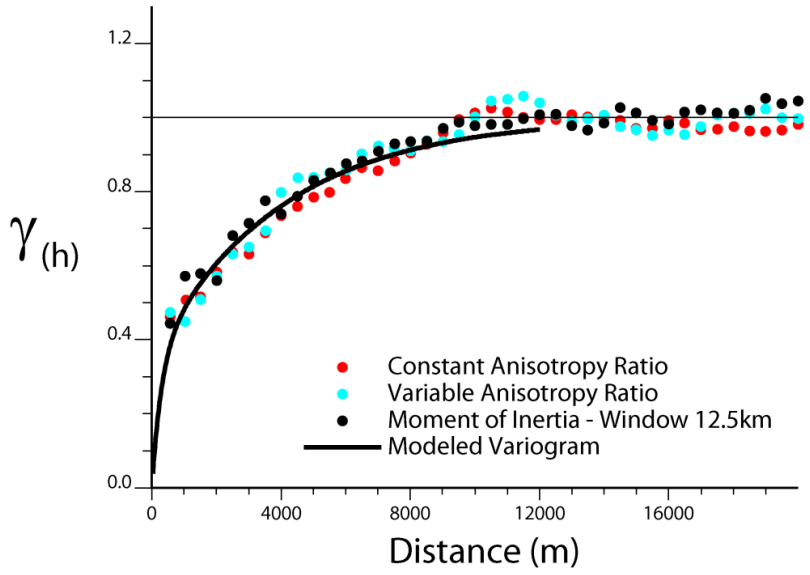


Figure 6.13: The experimental variogram for different LVA fields is different since the distance depends on the LVA field.

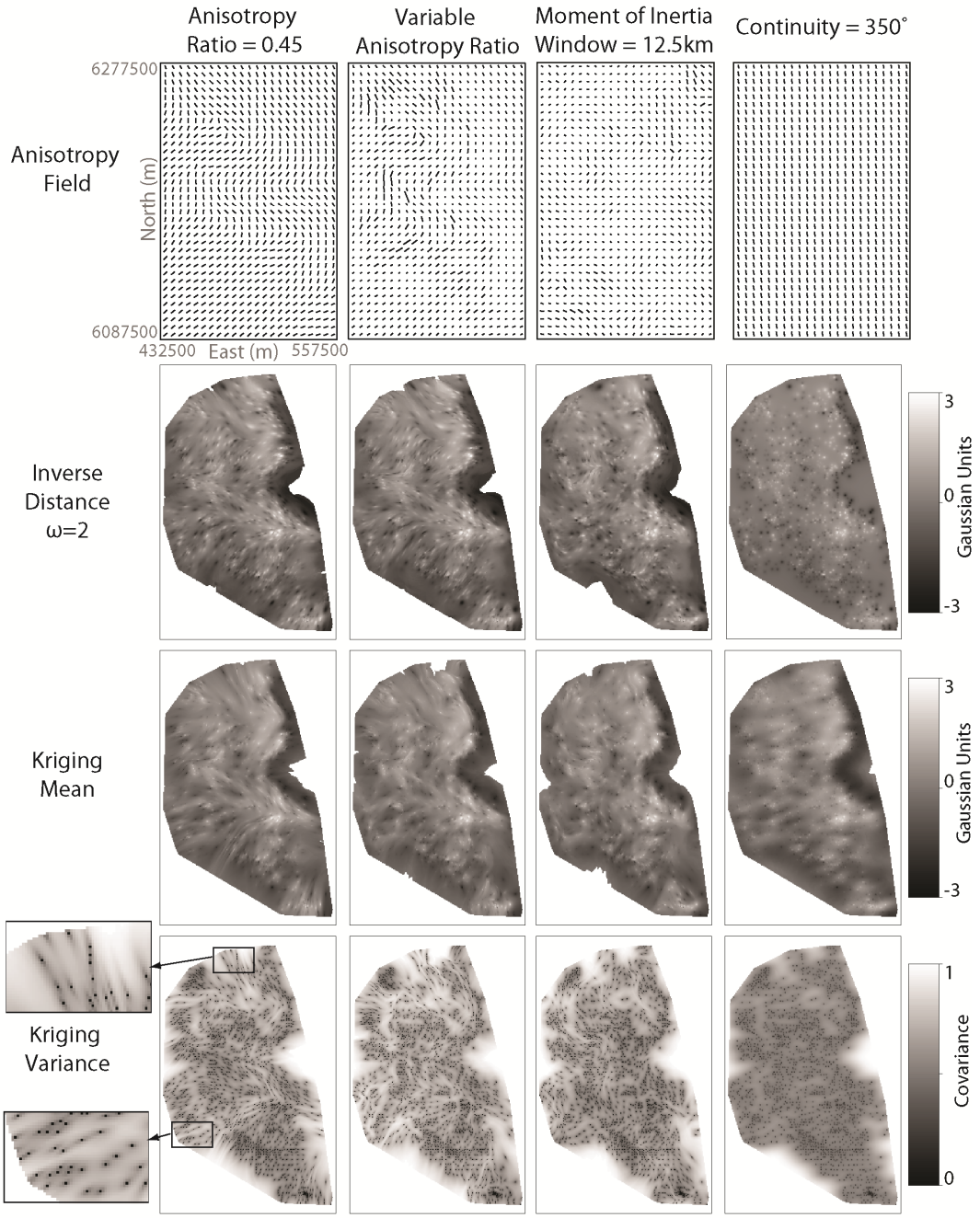


Figure 6.14: Simple kriging and inverse distance in Gaussian units using three different LVA fields. A search range of 7,000m and a maximum of 30 data were used in kriging. The pre-2000 data is used as conditioning data for these maps. Models are trimmed to the convex hull of the conditioning data.

A few comments can be made about the maps in Figure 6.14:

1) The LVA field has a large effect on the continuity 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 on the kriging variance is an elongation of lower variance parallel to the directions and an increase in variance perpendicular to the

directions (highlighted regions on Figure 6.14, lower left). As expected, with traditional kriging the variance is elliptical and aligned with 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 (Figure 6.15). This increase in distance results in a lower weight assigned to nearby 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.

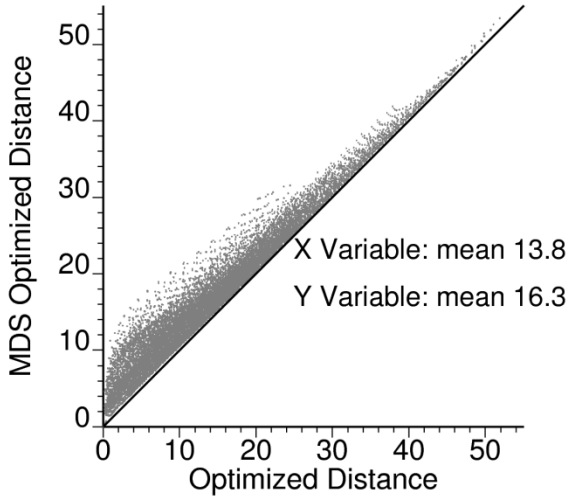


Figure 6.15: Distance reproduction (km) with L-ISOMAP.

6.1.6 Sequential Gaussian Simulation

Kriging maps are useful as they provide a local estimate that is optimal in the least squared error sense; 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) emerged as a practical solution to the biased reserves calculated with kriging. The same parameters are required for SGS as with kriging. The variogram in Figure 6.13 is also used here for SGS. Figure 6.16 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.

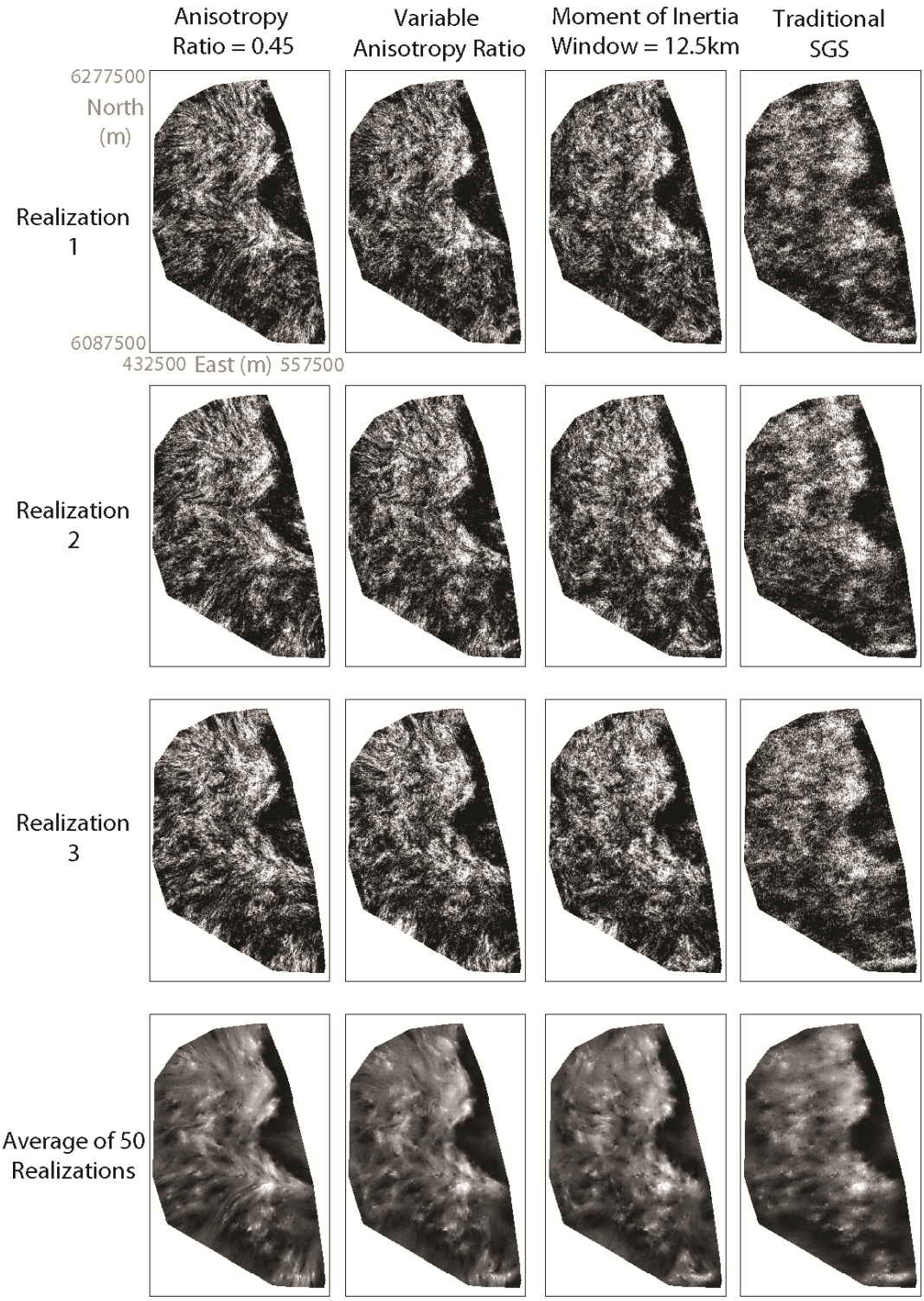


Figure 6.16: 3 SGS realizations using various LVA fields. Gray scale ranges from 0m NCB (black) to 40m (white).

At a minimum, SGS realizations must reproduce the histogram and the variogram of the input statistics (Leuangthong *et al.* 2004). These statistics are shown for the LVA realizations in Figure 6.17 and indicate good agreement with the input. Note that histogram reproduction is expected as simulations are performed in Gaussian units and back-transformed to match the input data histogram; however, the variogram of the input data is not explicitly used in the generation of realizations with LVA. The variogram shown in Figure 6.17 is the variogram of the data before coordinate transformation with L-ISOMAP, while this variogram is not used in simulation, it is well reproduced. A large discrepancy between the input data and the LVA realizations would indicate either (1) the use of an LVA field that was inconsistent with the model area or (2) poor parameter selection in L-ISOMAP such as using too few landmark points.

The LVA is not as visually dominant as with kriging (Figure 6.14) because of the addition of variability to match the input histogram and variogram. The LVA is somewhat masked; however, when the realizations are averaged (Figure 6.16) the features are more distinguishable. Of interest when applying SGS is the resulting reserves calculation (Figure 6.18). Often, steam assisted gravity drainage (SAGD) is implemented to recover the hydrocarbons present in these areas. A cutoff of 18m-20m NCB thickness is common (Shin and Polikar 2005; Ren *et al.* 2006; Jimenez 2008); thus, of interest in the study area is how reserve calculations change for cutoff values >18m. Reserve calculations are quite similar globally (Figure 6.18) but there is a difference in the local reserves (Figure 6.19) when two arbitrary areas, A and B, within the model are considered.

The reserves calculation is sensitive to the choice of LVA field. Consider the difference for area A when using the two manual LVA fields (Figure 6.19). Both LVA fields consider the same directions of anisotropy (Section 6.1.2); however, there is considerable difference in the calculated reserves, most notably for the important cutoffs above 20m. Further sensitivity to the LVA field is assessed using the various LVA fields generated with the moment of inertia method (Section 6.1.2). Recall that three LVA fields were generated with different moving window sizes. The changes in the reserve calculations due to the LVA field are shown in Figure 6.20.

The LVA field for this case study is not known with certainty and it may be inappropriate to assume a single deterministic LVA field. Three moment of inertia LVA fields and two manual LVA fields were carried through the analysis to quantify the uncertainty in reserve calculations due to the uncertainty in the LVA field. 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 accurate information is not available, the LVA field should be treated as an additional unknown variable and various LVA fields carried though the analysis, as shown here.

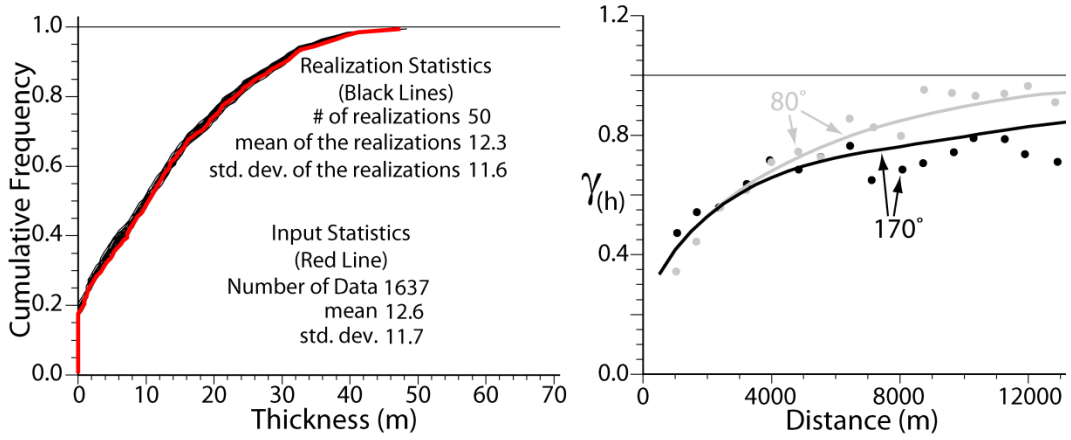


Figure 6.17: Above: Histogram reproduction with SGS-LVA. Below: Original variogram reproduction with SGS-LVA. The average variogram for the 50 realizations is shown as a solid line with points indicating the variogram of the original data.

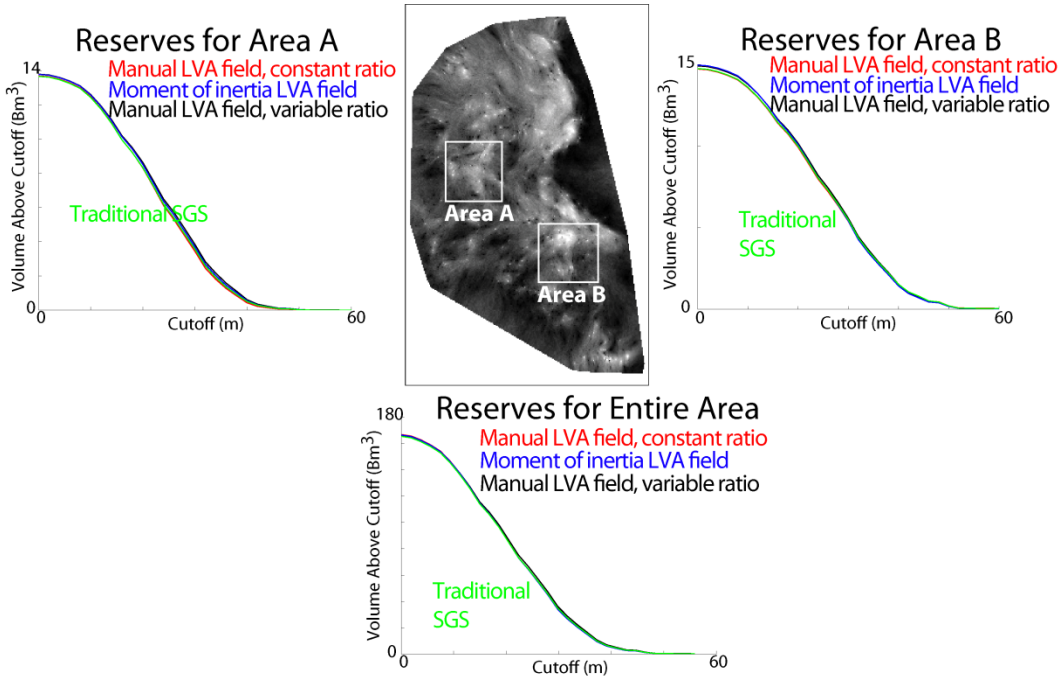


Figure 6.18: Reserve calculations using various LVA fields. Reserves were calculated for 50 realizations and averaged.

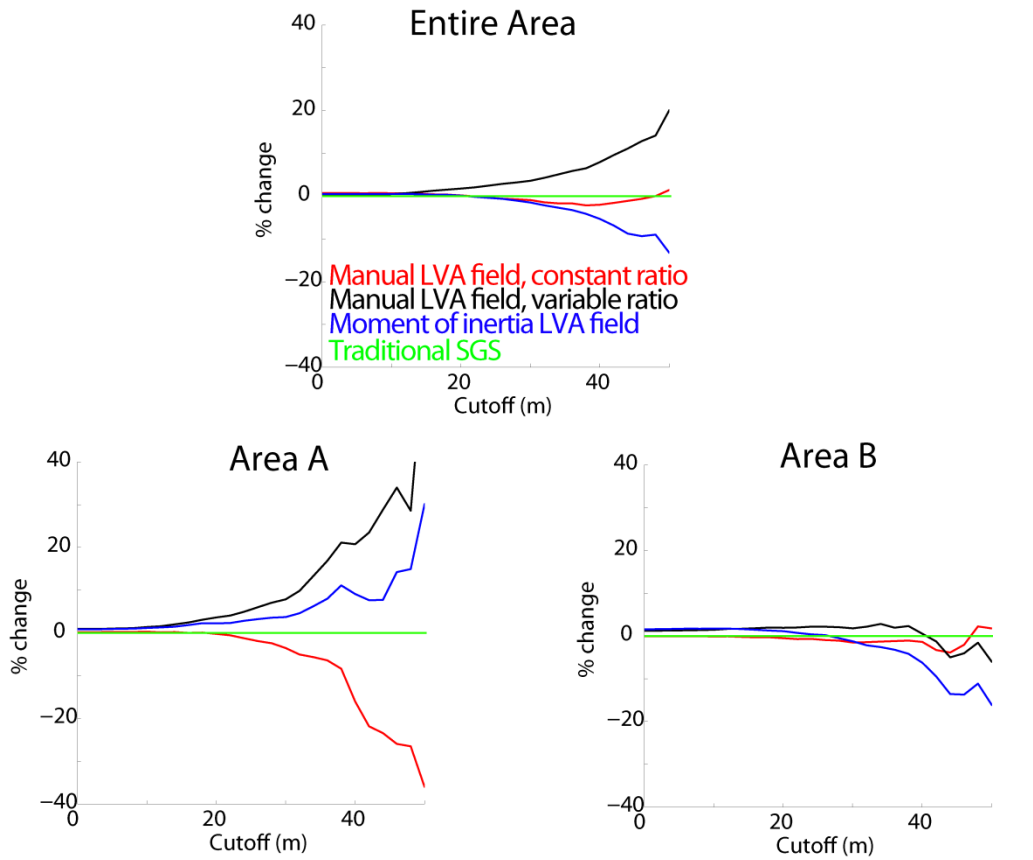


Figure 6.19: % change in reserves compared to traditional SGS. Reserves were calculated for 50 realizations and averaged.

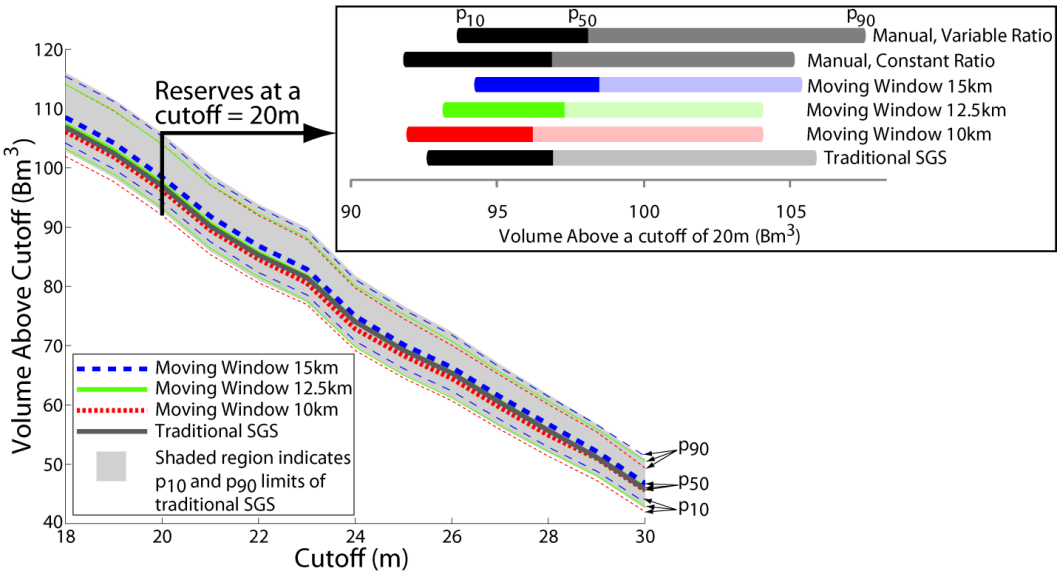


Figure 6.20: Reserves for the entire area showing different LVA fields generated with the moment of inertia method. Upper right figure shows the reserves for a NCB cutoff of 20m. The two manual LVA fields were added to the upper right plot.

6.1.7 Sensitivity Studies

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

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 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 predicted model value 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 L-ISOMAP
3) The grid resolution of the LVA field.
4) The number of offsets used in the calculation of the SPD with the Dijkstra algorithm.
5*) The input LVA field (direction of continuity and the anisotropy ratio).

* The LVA field is a significant modeling parameter that must be carefully analyzed and was the focus of Chapter 3. The sensitivity of this case study to LVA field selection is not discussed here as the effect of the LVA field on reserves for this data set was analyzed in Section 6.1.6.

Sensitivities are assessed for input parameters 1-4. Each parameter is varied within a practical range of possible values (Table 6.3). Because these parameters affect the implementation of L-ISOMAP, the variogram for each model is slightly different. Therefore, automatic variogram fitting software was applied for each variogram and visually inspected to ensure reasonable fits.

Table 6.3: Summary of first 4 parameters assessed for model sensitivity.

Parameter	Minimum Value	Maximum Value	Default Value
Number of landmark points	2	225	100
Number of dimensions to retain	15	100	99
Grid resolution of the LVA field	4 x 4	760 x 1200	230 x 360
Number of offsets used with Dijkstra	1	5	2

The grid resolution of the LVA field (Table 6.3) 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 SPDs but requires additional CPU time.

The number of dimensions to retain (Table 6.3) refers to the number of dimensions after L-ISOMAP 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 with the *k*d tree. It may not be possible to store all $(L\text{-}1)N$ coordinate locations. Selecting $d \ll L\text{-}1$ reduces memory requirements.

As seen in Figure 6.21, the methodology is insensitive to the L-ISOMAP and Dijkstra parameters. Thus, approximations can be made by significantly reducing the number of landmark points or offsets without adverse effects, so long as a minimum value is selected such that the results are stable (Figure 6.21).

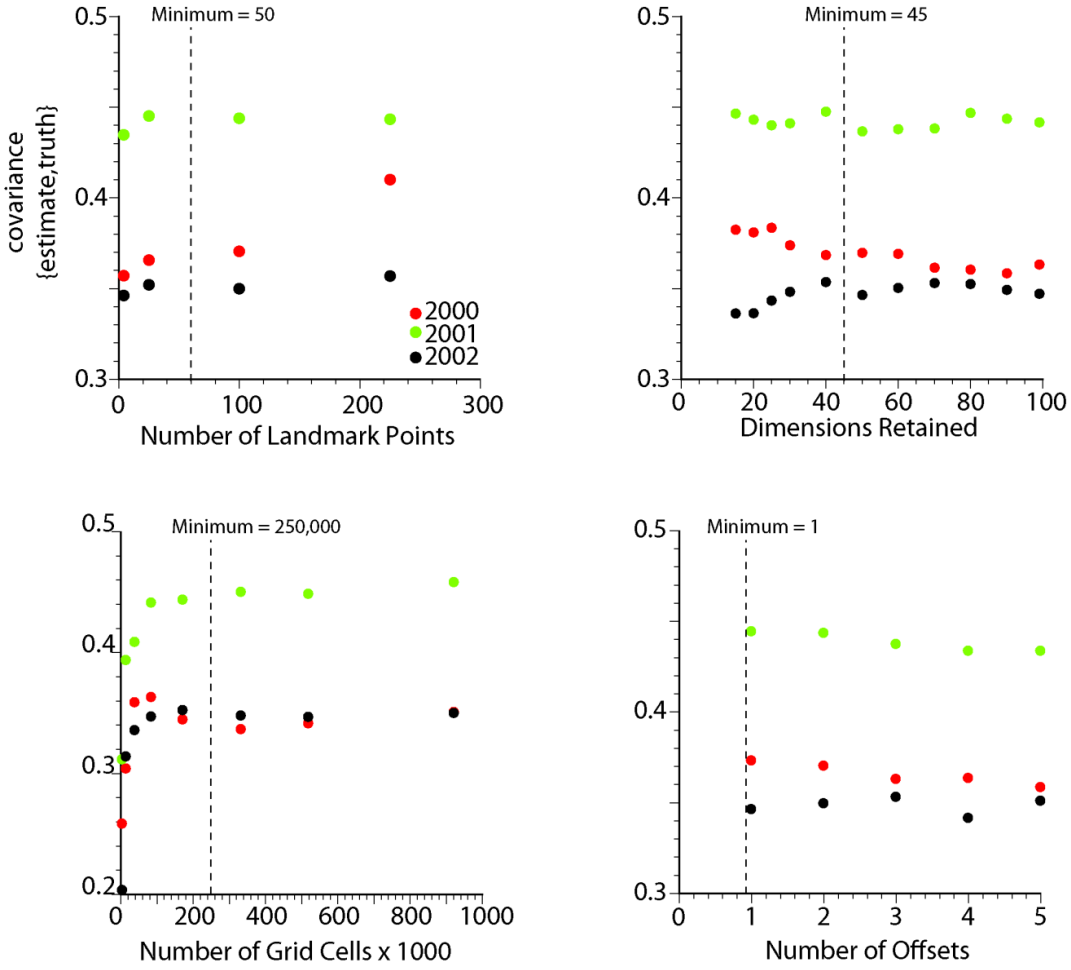


Figure 6.21: Effect of various parameters on the covariance measure. Note the different scale in the lower left figure.

6.2 Modeling Gold Grade in a Porphyry Deposit

The previous 2D case study focused on the sensitivity of the methodology to the necessary input parameters for estimation and simulation. The 2D nature of the ERCB data is convenient for a detailed analysis of the input parameters; however, the extension to 3D must be considered. As discussed in Chapter 4 the additional issue arising from considering a third dimension is the increase in the number of nodes and

the edges for calculating the SPD. This case study demonstrates the application of LVA to a 3D model.

6.2.1 The Data

The data consist of 119 drill holes for a copper/gold porphyry mine. More details on the geology of porphyry deposits can be found in (Guilbert and Park 2007). The data has been altered slightly to protect its origin; however, the spatial features remain the same. For confidentially reasons, the exact location and geology of this data cannot be discussed.

Data is available for both Cu and Au; however, only the Au grades are considered, although both variables display similar LVA. Moreover, results are only provided for Gaussian transformed values to further ensure protection of the original data. The same spatial features are present after the Gaussian transformation and are sufficient for this discussion.

The data displays very clear nonlinear anisotropy (Figure 6.22). Selecting a single direction of continuity for such a deposit is difficult. Perhaps the best results could be obtained with an isotropic variogram because any direction selected is inappropriate for some locations in the model (see Figure 6.27 and Figure 6.28). Taking advantage of the known locally varying directions of continuity can improve model performance. This type of deposit is an ideal candidate for kriging with LVA as local features are clearly shown in the data but there is not sufficient data density to fully define the complex geology and control the resulting geostatistical models.

A grid cell size of 5m in the x and y directions and 2m in the z direction is used for this model. A model of size 100x100x100 cells is used for a total of 1M cells.

6.2.2 Generating the LVA Field

Block kriging generates a smooth map that can be used to determine the orientation of the LVA (Figure 6.22 left). The LVA field is generated manually for this case study and consists of three distinct zones (1) within the circular feature, where visually there appears to be no anisotropy (2) the circular mineralized zone and (3) a transition zone between zones 1 and 2. Plan sections were examined every 20m and the core of the deposit is manually defined by the center and radius of the inner dashed circle (Figure 6.22 middle). Parameters for this isotropic core are shown in Figure 6.23. Within the isotropic core an anisotropy ratio of 1:1 is assigned to the LVA field. Outside the core, the magnitude of the anisotropy increases linearly to 10:1; this transition zone has a thickness of 20m (Figure 6.22 middle). The direction of continuity for this LVA field is tangential to a circle centered at the manually selected location (Figure 6.22 right). Such an LVA field effectively captures the radial nature of the nonlinear geological features (Figure 6.22) as well as the isotropic barren core.

The LVA field is subjective and a sensitivity analysis as discussed in Section 6.1 could be applied for a better understanding of the uncertainty resulting from the LVA field; however, the purpose of this case study is to demonstrate the methodology in 3D. As such, only this single, geologically realistic but subjective LVA field is carried forward.

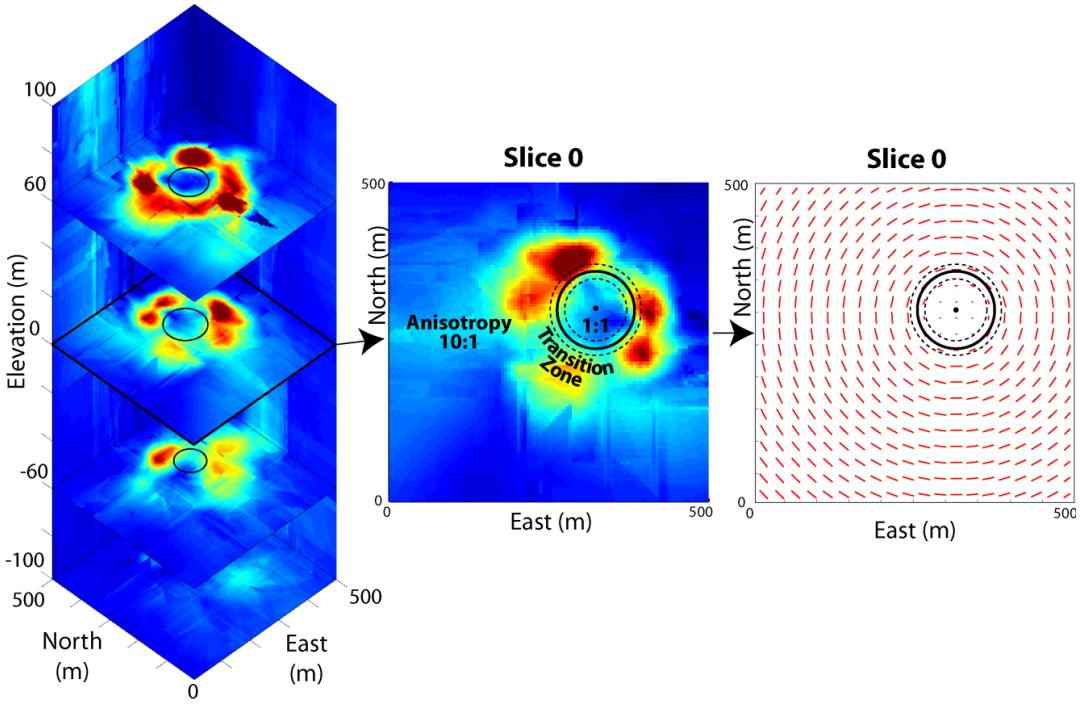


Figure 6.22: Left: Multiple slices of the block kriging map used to generate the LVA field. Grade is shown in Gaussian units with low grade (blue and green) and high grade (red and yellow). Middle: For each slice the anisotropy ratio is 1:1 inside the circle (defined manually for each slice) and 10:1 outside. Right: The LVA field. Length of the line is proportional to the anisotropy ratio.

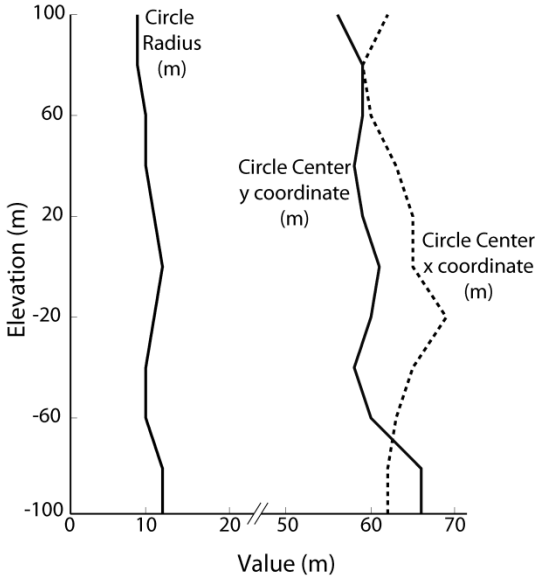


Figure 6.23: Parameters to fully define the isotropic core of the LVA field. The radius and xy coordinates of the center of the circle were manually fit every 20m and linearly interpolated between.

6.2.3 Estimation with LVA

For L-ISOMAP a total of 64 landmark points on a regular pattern of 4 x 4 x 4 are distributed evenly in the modeling area. For the calculation of the SPDs a single offset is used. After application of L-ISOMAP, the resulting 3D grid can be visualized in the first three coordinates of the embedded 63 dimensional Euclidian space (Figure 6.24). Of interest in this visualization is the apparent overlapping of multiple gird cell locations (highlighted on Figure 6.24 right). This can occur because there are no restrictions placed on the embedding; cell locations are free to be placed in any orientation such that the initial distance matrix is best reproduced (Chapter 5). However, in this example, the overlapping of the grid cells occurs because the grid has been embedded in 63 dimensions and is only displayed here in the first three. If the grid could be visualized in 63 dimensions, this overlapping would not be seen.

2D slices of the transformed model indicate that the grid for each elevation is nearly identical (Figure 6.24 below right). This is expected as the LVA field for each elevation is also nearly identical, as defined by an isotropic circular region surrounded by radial 10:1 anisotropy (Figure 6.22).

Kriging with a horizontal isotropic variogram is compared to kriging with LVA. The variograms used to model the deposit are shown in Figure 6.25 with associated cross validation results in Figure 6.26. There is a small improvement of 0.7% in the correlation when applying LVA; however, kriging with LVA realizes a greater gain of 101% when measured by the covariance. Covariance can be a better assessment of model performance as it considers the standard deviation of the resulting models.

Performing kriging with 64 landmark points, a single offset with the Dijkstra algorithm, 30 nearest data and 1M cells requires a total of 70 minutes of CPU time with Microsoft Windows Server 2003 Standard x64 Edition and a 2.41GHz AMD processor.

The goal of this case study is to generate realistic geostatistical models that incorporate LVA. The conceptual model of this deposit suggested that the mineralized zone is concentrated in a radial pattern. Further study of the kriging maps in Figure 6.27 and Figure 6.28 indicates that the assumption of a circular barren core may be too simplistic. Consider elevation -20m, in some areas an ellipsoidal core may be more consistent with the available data. In practice, such modifications can be made iteratively to the LVA field if deemed necessary.

When applying traditional methods with either a single direction of continuity or no anisotropy the high valued Au zones are often disjointed and interrupted when they are expected to be more continuous (Figure 6.27 and Figure 6.28). When the data density is sufficient to adequately describe the geology, such as at an elevation of -20m, traditional methods tend to do well; however, this is not the case when the data density decreases, such as in elevation -60m.

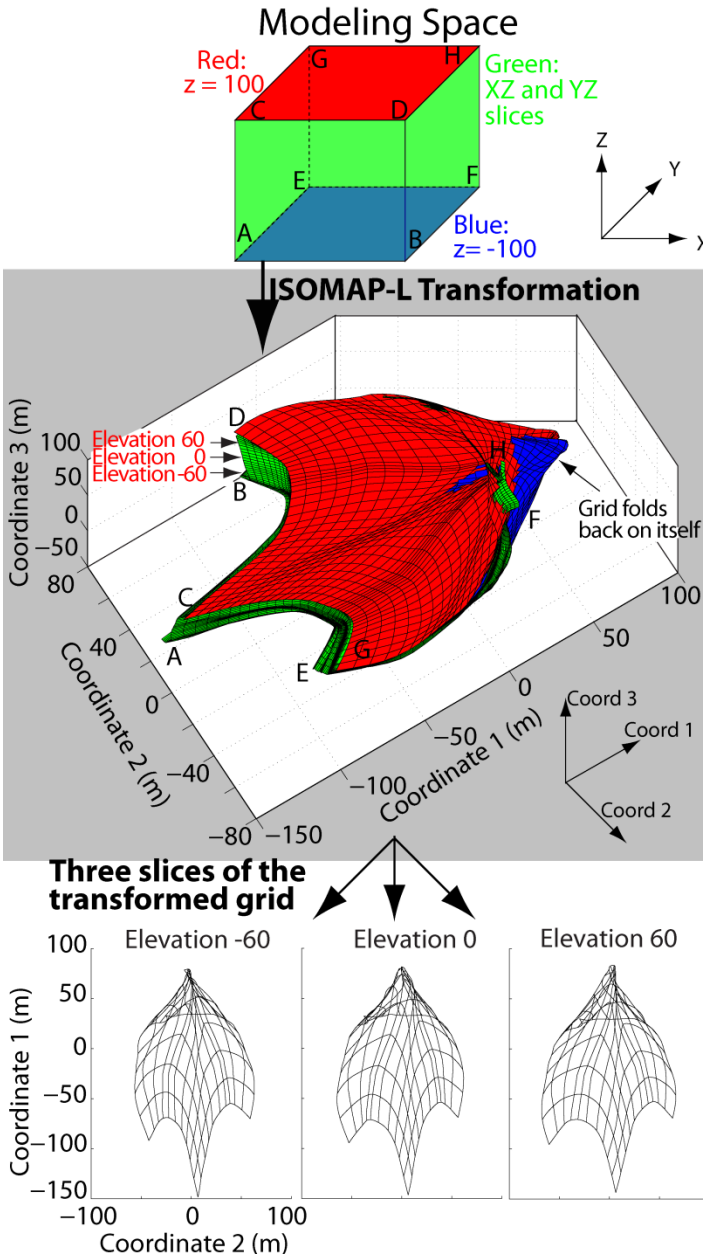


Figure 6.24: 3D visualization of the embedded grid. Above: the original grid shown in Cartesian coordinates. In both figures, the red surface is the upper slice of the model (z = 100) while the blue surface is the lower (z = -100). The green surfaces represent the XZ and YZ boundaries of the block model. Slices of the transformed grid for a constant elevation are shown below.

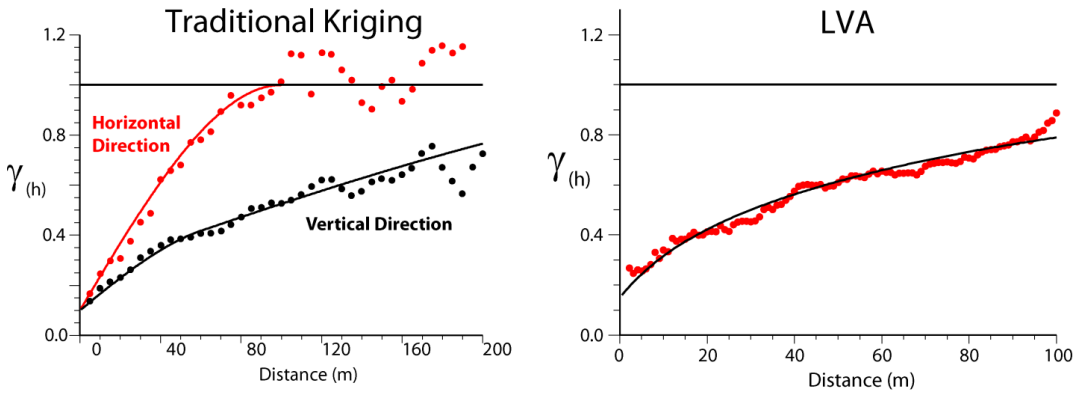


Figure 6.25: Variograms used to build models.

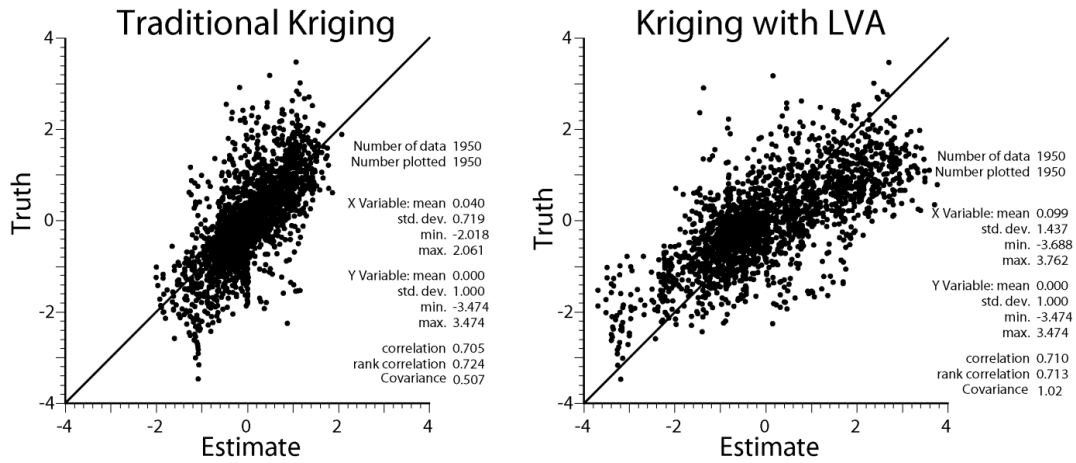


Figure 6.26: Cross validation using traditional kriging (left) and LVA kriging (right) in Gaussian units.

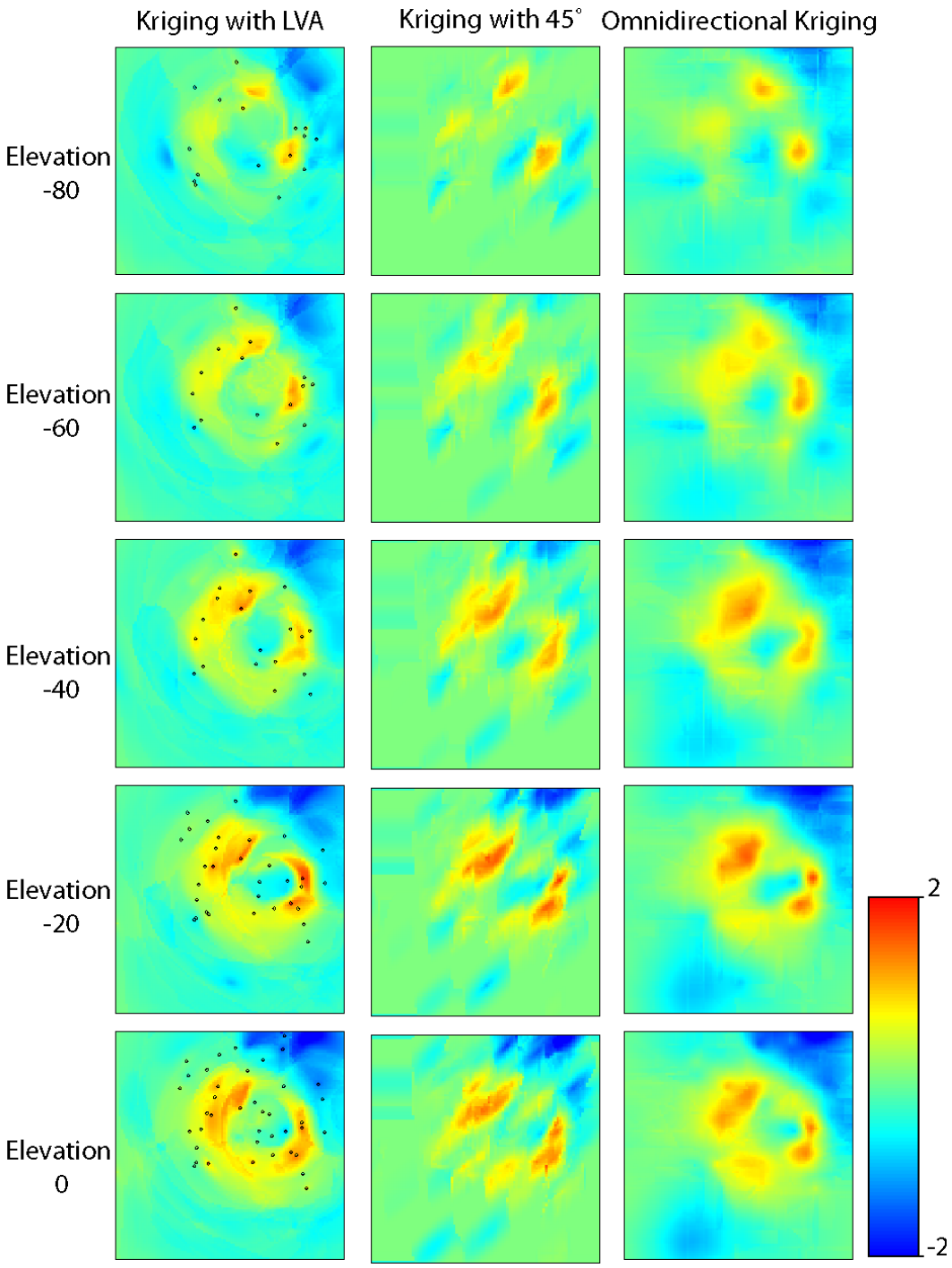


Figure 6.27: Left: Kriging with LVA in Gaussian units. Middle: Kriging with constant anisotropy in the 45° direction. Right: Kriging without horizontal anisotropy. Elevation -80m through 0m. Dimensions are 500m in the X and Y directions.

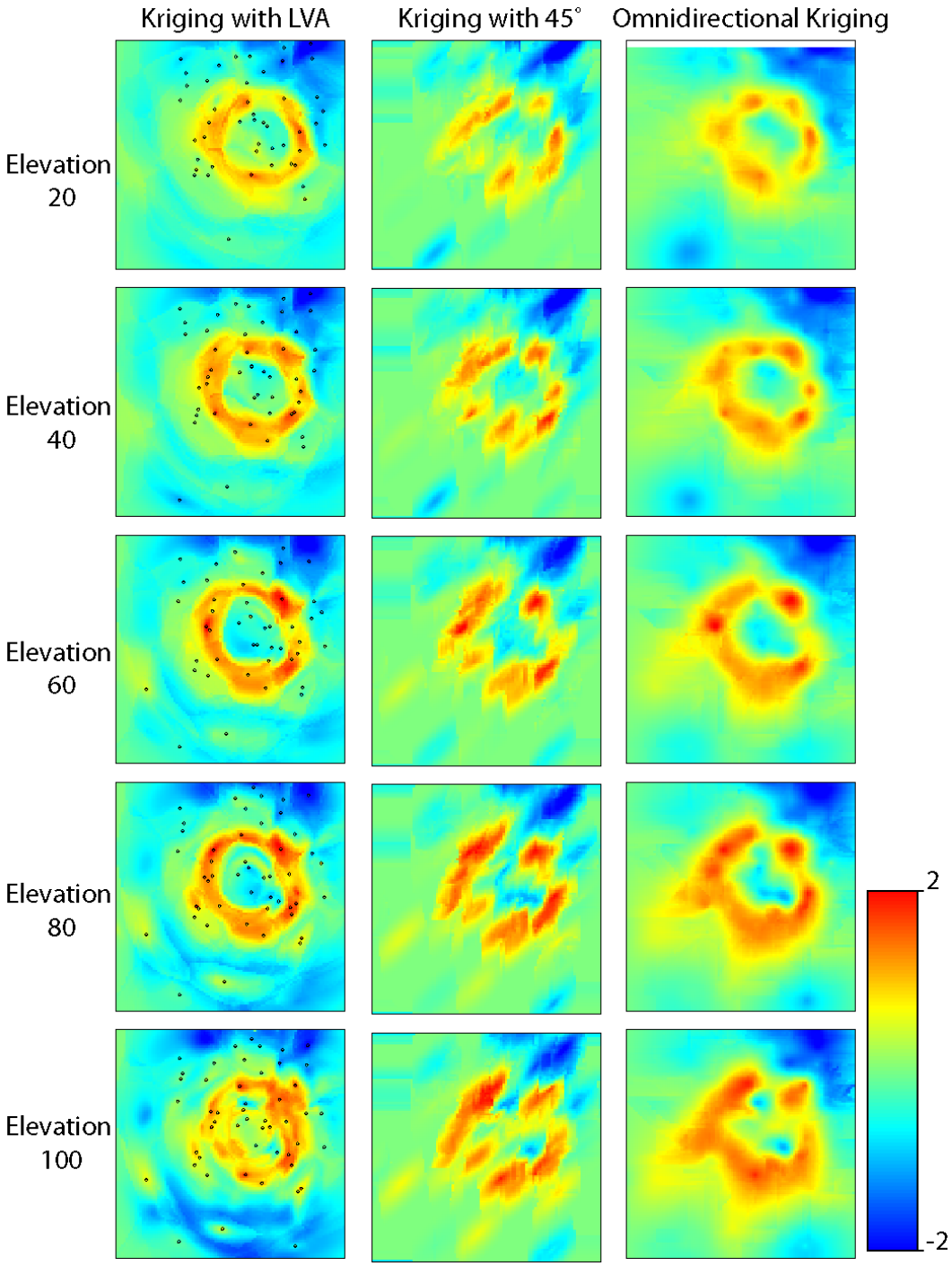


Figure 6.28: Left: Kriging with LVA in Gaussian units. Middle: Kriging with constant anisotropy in the 45° direction. Right: Kriging without horizontal anisotropy. Elevation 20m through 100m. Dimensions are 500m in the X and Y directions.

6.3 Remarks

The methodology proposed in Chapters 3, 4 and 5 was applied to two natural geologic deposits. The first case study modeled the NCB thickness for an oil sands deposit in Northern Alberta. The large extents of this deposit indicated that there were numerous directions of continuity within the modeling domain. These locally varying features were effectively incorporated into geostatistical estimation and simulation.

The main purpose of the NCB thickness 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 the Dijkstra algorithm and L-ISOMAP. To summarize, 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. A sensitivity analysis of the methodology to the selection of L-ISOMAP parameters showed that they have a small effect on the resulting model and can be lowered to a minimum threshold of 50 landmark points to improve CPU performance if necessary. In contrast, selection of the LVA field was shown to have a large effect on important reserve calculations and even more of an impact when local areas are analyzed.

The application of LVA to generating 3D geomodels was shown for a Cu/Au porphyry deposit. There are two main issues with extending the LVA modeling methodology from 2D to 3D, the first was highlighted in Chapter 3 with the discussion of the increased CPU time required for determining distances in 3D. The second issue is the increased difficulty in modeling the LVA field. Often data can be adequately visualized in 2D and a number of techniques can be used to generate the LVA field, even manually assigning the directions and ratios. Such data visualization is more difficult in 3D when the data are often aligned along a string making it difficult to simultaneously visualize all samples in the modeling area. In the case of the Cu-Au porphyry deposit, the LVA could be adequately visualized in 2D slices of the data and manually fit to generate a realistic LVA field. Such data visualization can be more difficult in 3D when the LVA is not axis aligned. Overcoming the difficultly of modeling the LVA field in 3D is case specific.

The benefits of modeling with LVA were apparent in both the reserve calculations for the net thickness (Section 6.1.6) as well as a 101% increase in cross validation covariance found in the porphyry example (Section 6.2.3).

Chapter 7: Conclusions

In the case of modeling a variable that displays a single stationary direction of continuity over the entire domain of interest, traditional estimation and simulation algorithms adequately reproduce the existing anisotropy. When the orientation or the magnitude of the anisotropy of a variable is not constant in the modeling domain, traditional methodologies fail to reproduce the LVA unless there is sufficient data density to drive the modeling and overcome the limitation of a single global direction of continuity.

The proposed methodology can be used to incorporate LVA when there is geological knowledge that a deposit is characterized by LVA and when there is sufficient quantitative or qualitative knowledge to reasonably infer the LVA. Models that are built to reflect this additional information have higher local accuracy and can be used to make better engineering decisions.

7.1 Summary of Contributions

Two main contributions emerged from this thesis. The first contribution is the generation of the LVA field from a wide variety of data sources. Typically, the exhaustive LVA field is not known for a deposit of interest and it must be inferred. The second contribution is the integration of LVA into inverse distance, kriging and sequential Gaussian simulation.

7.1.1 LVA Field Inference

Inference of the LVA field is difficult as the available data rarely measures the LVA field itself. Often, orientation and magnitude must be inferred from static point measurements of the property at widely spaced intervals. A number of novel and existing methodologies were presented to generate the LVA field:

1) **Estimation or Simulation from Point Data** (Section 3.3.1): In the desirable situation when there are point measurements of anisotropy orientation, any estimation or simulation algorithm can be used to generate an exhaustive field. This requires a decomposition of the anisotropy vector into X, Y and Z components to facilitate modeling of the anisotropy orientation. This technique is also useful when generating a manual LVA field; the practitioner can manually assign the LVA specification at discrete locations (on section in 3D) and the exhaustive field can be estimated or simulated. Simulation is used to generate multiple LVA fields when there is uncertainty in the LVA field inference.
2) **Manual LVA Inference** (Section 3.3.2): Knowledge of the depositional characteristics of the deposit of interest can be used to infer the LVA field. Such qualitative data can be incorporated into numerical modeling by imposing an LVA field with inferred characteristics from an in depth understanding of the genesis of the deposit.

Generating manual LVA fields also provides a base line with which to measure any automatic method employed.

3) **Moment of inertia** (Section 3.3.3): The orientation and magnitude of anisotropy can be obtained by using the moment of inertia of a moving window covariance map. The orientation and magnitude are obtained by considering the covariance within each cell of the covariance map to be a mass; high covariance corresponds to high mass. The moment of inertia tensor for the covariance map is used to calculate the axis direction in which the moment of inertia is minimized. This corresponds to the major direction of continuity. The magnitude of the moment of inertia can be used to determine the magnitude of anisotropy. This methodology can be applied to a densely sampled primary or secondary variable with the covariance calculated from the sample data or the methodology can be applied to an exhaustive secondary variable that is known to have the same LVA structure as the variable of interest.
4) **Automatic Feature Interpolation** (Section 3.3.4): This technique is an attempt to automatically replicate the process of a geologist generating an LVA field. Locations that have similar data values are connected by polylines. The orientation of these polylines corresponds to the orientation of the underlying LVA field. Repeating this process many times and averaging the orientations results in a discrete measurement of the LVA field orientation at each data location. These discrete angles can be estimated or simulated to generate an exhaustive LVA field orientation. The methodology is straightforward in 2D but the necessary tolerance parameters become difficult to infer in 3D.

The creation of the LVA field is a critical step in the incorporation of LVA into numerical modeling. The LVA field has a significant effect on the resulting realizations or estimates. Relevant geological knowledge should be combined with all primary and secondary data available to formulate a geologically realistic LVA field using the techniques presented. If this LVA field is uncertain, which is nearly always the case, multiple fields should be generated to span the uncertainty in the LVA field and different scenarios carried though any relevant analysis.

7.1.2 Algorithms: Estimation and Simulation with LVA

The second contribution of this thesis is the incorporation of LVA into inverse distance, kriging and sequential Gaussian simulation which facilitates the generation of spatial numerical models for earth science variables. In order to incorporate LVA into modeling, the relationship between locations must be known. This relationship, or distance, between points is determined using the Dijkstra algorithm which calculates the shortest path through the LVA field. Using this shortest path the relationship between points can be determined and numerical models can be generated that display complex nonlinear geological features.

Considering LVA with inverse distance estimation is straightforward. The weights are inversely proportional to the distance raised to a power selected by the geomodeler. The shortest path distance is used rather than the traditional Euclidian distance.

Incorporating LVA into a kriging or sequential Gaussian simulation framework is more difficult because the use of the shortest path distance between locations does not guarantee a positive definite system of equations. To solve this problem, the entire modeling domain, characterized by a finite number of geological modeling cells, is embedded in a high dimensional Euclidian space with L-ISOMAP. Appropriate variogram functions can now be applied to ensure positive definiteness because the space is Euclidian and the distances between locations are calculated along straight paths.

The nonlinear distance calculated with the Dijkstra algorithm is the means by which LVA is incorporated into modeling. The embedding of the grid using L-ISOMAP (or another MDS like technique) is simply a correction of the distances such that the kriging matrices can be solved. While this step allows for a sound theoretical framework in which valid weights can be calculated, error is introduced into the process. The distance between locations is only approximately reproduced in the embedded space, typically measured by the stress (Chapter 4). This error must be accepted until a covariance or variogram function is discovered that proves to be positive definite with the shortest path distance metric.

7.2 Context of Modeling with LVA

The modeling methodology for an earth sciences problem is highly dependent on the unique characteristics of each case, but there are a number of steps that are often followed (Figure 7.1). Many of these steps were not discussed in this thesis as the focus was on property modeling in the specific case when information is available that indicates that the assumption of second order stationarity is not valid. In such cases, the methodologies presented throughout this dissertation can be applied to the *Numerical Property Modeling* step in a geostatistical workflow (Figure 7.1).

The proposed methodology does not change the numerical modeling workflow when there is a single global direction of continuity that accurately describes the spatial structure of the variable of interest (i.e. second order stationarity). Moreover, when there is so little information available about the LVA in a domain that it is not known if the assumption of second order stationarity holds, it is unlikely that the proposed methodology would improve spatial modeling. The methodology proposed is intended for the situations when additional knowledge, such as geological surveys, extensive sampling of secondary variables, understanding of the genesis of the deposit or past experience can be relied upon to generate a reasonable assessment of local anisotropy.

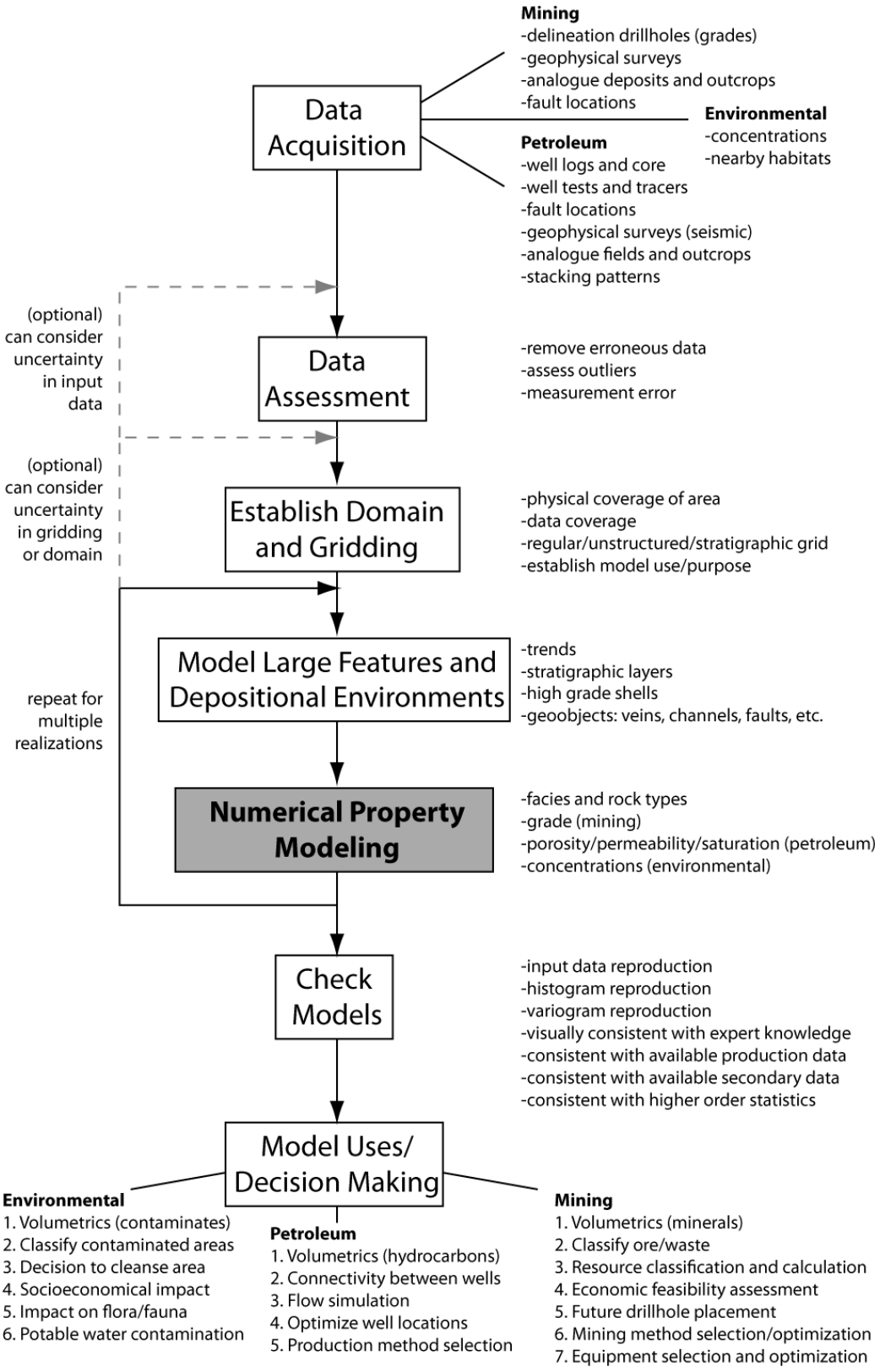


Figure 7.1: High level flow chart of typical steps in a numerical modeling workflow. This thesis is concerned with the *Numerical Property Modeling* step highlighted in gray.

7.3 Advantages and Limitations

The main advantage of incorporating LVA into numerical models is an increase in local accuracy. When complex nonlinear features are known to exist they should be modeled. Directly placing these features into numerical models can be difficult and time consuming; however, LVA can be used to model the effect these features have on the variable of interest. For example, recall the porphyry deposit from Chapter 6. Reproducing the physics of the complex geological events that created the radial pattern of grades would be difficult; however, the effect these complex geological events have on gold grade is a simple radial pattern of grades that can be modeled with LVA.

A beneficial property of the models generated with LVA is that they appear to be more geologically realistic. A criticism of conventional modeling is that it does not generate models that are visually realistic. While a numerical model is rarely created to visually match reality, there is merit to the argument that major geological features that significantly impact the variable of interest should appear in any numerical models of that variable.

Often a transfer function is applied to numerical models, such as (1) flow simulation for hydrocarbon production prediction (2) reserve calculation for mineral commodities or (3) contaminate level prediction for hazardous materials. An apriori assessment of the impact of the perceived LVA on the appropriate transfer function is required. It is not necessary to incorporate LVA if it is unlikely that the LVA will affect the outcome of the transfer function. For example, the calculation of global resources without a cutoff grade depends almost entirely on the data histogram rather than the spatial arrangement of grades, as such, models built with and without LVA return similar resource estimates; however, when models are built to calculate local reserves the effect of LVA could be significant depending on the features and the area under consideration.

There are some drawbacks to implementing the proposed methodology. The most significant limitation is based on available computer resources. Even with efficient implementations, the Dijkstra and L-ISOMAP algorithms are extremely CPU and memory intensive. This may limit the size of model that can be generated. This limit also affects the accuracy of the embedding as it is not computationally feasible to consider classical MDS (or ISOMAP) for the large grids that are required in the earth sciences.

With respect to the necessary input parameters of numerical modeling with LVA, kriging and simulation require (1) a modeled variogram (2) necessary search parameters (3) the LVA field. The modeled variogram is typically straightforward to estimate in this methodology because in most situations there is sufficient data density to inform an isotropic variogram; however, with extremely sparse data this becomes an issue. The necessary search parameters are typically straightforward to select and are often based

on the range of correlation. Where the input to kriging/SGS with LVA diverges from traditional geostatistical tools is in the inference of the LVA field. This is the most difficult, subjective and time consuming parameter that must be inferred for the application of LVA. Selecting an inappropriate LVA field will have a large impact on the resulting models. Carrying multiple LVA fields through the modeling process is recommended to account for the inherent uncertainty in the LVA field but this increases the professional time required. In the situation where there is no reliable information to inform the inference of the LVA field, the application of the proposed methodology in this thesis is not recommended.

The final drawback of the methodology is the application of L-ISOMAP. The use of L-ISOMAP can be thought of as a fix necessary to ensure that the weights in kriging and SGS can be determined. In the embedding process, some information is lost as a perfect embedding (stress=0) is rarely possible. As such, this fix introduces error in the distances. Ideally, classical MDS would be applied to minimize this error; however, the computational requirements of MDS are high and L-ISOMAP is required for large models. Note that with inverse distance this is not the case as no embedding is necessary.

7.4 Future Work

The overall methodology for numerical modeling with LVA was developed in this thesis. There are a number of details that could be further researched that are likely to generate sizeable incremental improvements over the presented methodology:

1) The main assumption inherent in the implementation of the Dijkstra algorithm is that the path between locations follows the edges that define the graph. This limits the shortest path to a restrictive piecewise linear geometry. The true minimum distance would be determined if an arbitrary geometry was allowed.
2) The development of a variogram function that is positive definite with the shortest path algorithm would eliminate the requirement of embedding with L-ISOMAP.
3) The incorporation of LVA into numerical modeling for a single variable was considered. In practical applications there are often many variables to consider. Modeling multiple correlated variables with potentially different LVA fields should be explored.
4) Ultimately, the numerical models generated with LVA are used for a purpose (Figure 7.1). Exploring the effect of LVA on these transfer functions, such as flow simulation or contaminate classification, would highlight the utility of the methodology presented.
5) Optimal sample placement in the presence of global anisotropy has been explored (McBratney *et al.* 1981). New samples are placed such that the most information is obtained (i.e. far from existing samples and in areas of high heterogeneity).

Exploration of sample design in the presence of LVA would allow for better optimization of future sampling campaigns for deposits with LVA.

6) Integration of the proposed methodology into commercial software packages to encourage usage of the methodology.
7) Creation of additional tools and methodologies to assess and quantify the uncertainty in the LVA field and its impact on resulting transfer functions.
8) Validation of the methodology on other real data sets would increase confidence in the methodology. Demonstrating improved predictions of flow response after flow simulation of realizations that display LVA would emphasize the benefits of the methodology and establish the place of these tools.

7.5 Final Remarks

Inverse distance, kriging and SGS with LVA account for second order stationarity but do not explicitly deal with issues of first order stationarity. In domains that are neither first nor second order stationary, the proposed methodology to incorporate LVA should be coupled with techniques that consider first order stationarity (McLennan 2008).

It should be noted that the proposed methodology does have some reliance on the assumption of second order stationarity. Rather than classical second order stationarity, where the variogram is constant in the domain, this methodology assumes that the isotropic variogram in the transformed q dimensional space is stationarity. This can be thought of as a relaxation of typical second order stationarity assumption.

The scope of the proposed methodology to incorporate LVA is not limited to mining and petroleum applications. Any industry that requires numerical models of spatial variables could potentially require incorporation of LVA. Chapter 1 briefly discussed the application of LVA in an environmental context with the illustrative example of pollution spread in the presence of wind. It is expected that many industries could benefit from integrating LVA into numerical modeling.

Complex nonlinear geological features often have an impact on the end use of the numerical models. Care must be taken such that the geomodeler does not *see* anisotropic features in a deposit that are simply a result of preferential, sparse or erroneous sampling. If the features have a reasonable geological explanation and are a characteristic of the modeling domain rather than simply a sampling phenomenon they should be incorporated into numerical modeling when they are deemed to have an impact on the end use of the model.

Bibliography

Ahn C and Ramakrishna R, 2002. A genetic algorithm for shortest path routing problem and the sizing of populations. IEEE Transactions. 6(6):566-579.

Ali M and Kamoun F, 1993. Neural networks for shortest path computation and routing in computer networks. IEEE Transactions on Neural Networks. 4(6):941-954.

Almendral A, Abrahamsen P and Hauge R, 2008. Multidimensional scaling and anisotropic covariance functions. In Ortiz J and Emery X (eds.) Geostats 2008: Proceedings of the eighth international geostatistics congress, Gecamin. 187-196.

Anel A, 2005. Network shortest path application for optimum track ship routing. M.Sc. Thesis, Naval Postgraduate School, Monterey, California. 93p.

Armstrong M and Diamond P, 1984. Testing variograms for positive-definiteness. Mathematical Geology. 16(4):407-421.

Arpat G and Caers J, 2007. Conditional Simulation with Patterns. Mathematical Geology. 39(2):177-203.

Azaron A and Kianfar F, 2003. Dynamic shortest path in stochastic dynamic networks: Ship routing problem. European Journal of Operational Research. 144(1):138-156.

Babak O and Deutsch C, 2008. Statistical approach to inverse distance interpolation. Stochastic Environmental Research and Risk Assessment. 23(5):543-553.

Balasubramanian M, Schwartz E, Tenenbaum J, Silva V and Langford J, 2002. The ISOMAP algorithm and topological stability. Science. 295(5552):7p.

Beer F and Johnston E, 1988. Vector Mechanics for Engineers: Statistics and Dynamics 5th edition. McGraw-Hill. 1280p.

Boisvert J, Pyrcz M and Deutsch C, 2007. Multiple point statistics for training image selection. Natural Resources Research. 16(4):313-321.

Boisvert J and Deutsch C, 2008. Shortest anisotropic path to reproduce complex geological features. In Ortiz J and Emery X (eds.) Geostats 2008: Proceedings of the eighth international geostatistics congress, Gecamin. 1041-1046.

Boisvert J, Manchuk J and Deutsch C, (2009). Kriging in the presence of locally varying anisotropy using non-Euclidean distances. Mathematical Geosciences 41(5): 585-601.

Boisvert J and Deutsch C, 2010. Programs for Kriging and Sequential Gaussian Simulation with Locally Varying Anisotropy Using Non-Euclidean Distances. Computers and Geosciences. In press, accepted March 2010.

Boisvert J, Leuangthong O, Ortiz J and Deutsch C, 2008. A methodology to construct training images for vein type deposits. Computers and Geosciences. 34(5)491-502.

Borgman L, Taheri M and Hagan R, 1983. Three-dimensional, frequency-domain simulations of geological variables. In Verly G, David M, Journel A and Marechal A, Geostatistics for Natural Resources Characterization. 517-541.

Boyd S and Vandenberghe L, 2004. Convex Optimization. Cambridge University Press, 716p.

Brown P, Nhu D and Zidek J, 1994. Multivariate spatial interpolation and exposure to air pollutants. The Canadian Journal of Statistics. 22(4)489-505.

Bu L and Chiueh T, 1999. Solving the shortest path problem using an analog network. Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions. 46(11)1360-1363.

Caers J, 2001. Geostatistical reservoir modeling using statistical pattern recognition. Journal of Petroleum Science and Engineering. 29(3):177-188.

Carrigy M, 1959. Geology of the McMurray formation: Part III, General geology of the McMurray area, Alberta Research Council, Memoir I. 130p.

Cherkassky B, Goldberg A and Radzik T, 1996. Shortest path algorithms: Theory and experimental evaluation. Mathematical Programming. 73(2):129-174.

Chiles J and Delfiner P, 1999. Geostatistics: Modeling Spatial Uncertainty, Wiley Inter-Science Publication, New York. 695p.

Christakos G, 1984. On the problem of permissible covariance and variogram models. Water Resources Research. 20(2):251-265.

Chu J, Xu W, Journel A, 1994. 3-D implementation of geostatistical analyses – the Amoco case study. In: Yarus J and Chambers R (eds.) Stochastic Modeling and Geostatistics, AAPG Computer Applications in Geology 3, 201–216.

Cormen T, Leiserson C, Rivest R and Stein C, 2003. Introduction to Algorithms Second Edition. MIT Press and McGraw-Hill. 1180p.

Cressie N, 1993. Statistics for Spatial Data Revised Edition. John Wiley and Sons, New York, 900p.

Cressie N and Majure J, 1997. Spatio-temporal statistical modeling of livestock waste in streams. Journal of Agricultural, Biological and Environmental Statistics. 2(1):24-47.

Curriero F, 1996. The use of non-Euclidean distance in geostatistics. Ph.D. Thesis, Kansas State University. 213p.

Curriero F, 2005. On the use of non-Euclidean distance measures in geostatistics. Mathematical Geology. 38(8):907-926.

Damian D, Sampson P and Guttorp P, 2000. Bayesian estimation of semi-parametric non-stationary spatial covariance structures. Environmetrics. 12(2):161-178.

Deutsch C, 1992. Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data. Ph.D. Thesis, Stanford University. 306p.

Deutsch C and Lewis R, 1992. Advances in the practical implementation of indicator geostatistics. In Proceedings of the 23rd APCOM Symposium Tucson, AZ, SME/AIME. 169-179.

Deutsch C and Journel A, 1998. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York. 369p.

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

Deutsch C and Tran T, 2002. FLUVSIM: A program for object-based stochastic modeling of fluvial depositional systems. Computers and Geosciences. 28(4):525-535.

Dijkstra E, 1959. A note on two problems in connection with graphs. In Numerische Mathematik. 1(1):269-271.

Dimitrakopoulos R, Mustapha H and Gloaguen E, 2010. High-order Statistics of Spatial Random Fields: Exploring Spatial Cumulants for Modeling Complex Non-Gaussian and Non-linear Phenomena. Mathematical Geosciences. 42(1):65-99.

Eriksson M and Siska P, 2000. Understanding anisotropy computations. Mathematical Geology. 32(6):683-700.

Euler L, 1736. Problematis ad Geometriam Situs Pertinentis Academiae Scientiarum Imperialis. Petropolitanae. 8:128-140.

Evans A, 1987. An Introduction to Ore Geology Second Edition. Blackwell Scientific Publications, Oxford. 358p.

Fisher N, 1996. Statistical Analysis of Circular Data. Cambridge University Press. 296p.

Frank H, 1969. Shortest paths in probabilistic graphs. Operations Research. 17(4):583-599.

Fuentes M, 2002a. Interpolation of nonstationary air pollution processes: A spatial spectral approach. Statistical Modeling. 2(4):281-298.

Fuentes M, 2002b. Spectral methods for nonstationary spatial processes. Biometrika 89(1):197-210.

Fuentes M, Chen L, Davis J and Lackmann G, 2005. Modeling and predicting complex space-time structures and patterns of coastal wind fields. Environmetrics. 16(5):449-464.

Ganio L, Torgersen C and Gresswell R, 2005. A geostatistical approach for describing spatial pattern in stream networks. Frontiers in Ecology and the Environment. 3(3):138-144.

Gardner B, Sullivan P and Lembo A, 2003. Predicting stream temperatures: geostatistical model comparison using alternative distance metrics. Canadian Journal of Fisheries and Aquatic Sciences. 60(3):344-351.

Goovaerts P, 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press. 496p.

Gringarten E and Deutsch C, 2001. Variogram interpretation and modeling. Mathematical Geology. 33(4):507-534.

Guardiano F and Srivastava M, 1993. Multivariate geostatistics: Beyond bivariate moments. In Soares A (ed.) Geostats Troia 1992. 133-144.

Guilbert J and Park C, 2007. The Geology of Ore Deposits. Waveland Press Inc., Illinois. 985p.

Guttorp P and Sampson P, 1994. Methods for estimating heterogeneous spatial covariance functions with environmental applications. In Patil G and Rao C (eds.) Handbook of Statistics XII: Environmental Statistics. 663-690.

Hart P, Nilsson N and Raphael B, 1968. A formal basis for the heuristic determination of minimum cost paths. IEEE Transaction of Systems Science and Cybernetics. 4(2):100-107.

Hassanpour M, 2009. A review of McMurray formation geology in Athabasca oil sands. Center for Computational Geostatistics Annual Report 11. University of Alberta. 8p.

Hein F, Cotterill D and Berhane H, 2000. An Atlas of Lithofacies of the McMurray Formation Athabasca Oil Sands Deposit, North Alberta: Surface and Subsurface. Alberta Geological Survey, Earth Sciences Report. 217p.

Higdon D, 1998. A process-convolution approach to modeling temperatures in the North Atlantic Ocean. Journal of Environmental and Ecological Statistics. 5(2):173-190.

Higdon D, Swall J and Kern J, 1998. Non-Stationary Spatial Modeling. Bayesian Statistics 6. Oxford University Press. 10p.

ver Hoef J, Peterson E and Theobald D, 2006. Spatial statistical models that use flow and stream distance. Environmental Ecological Statistics. 13(4):449-464.

Hotelling H, 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology. 24(6):417-441.

Isaaks E, 1991. The application of Monte Carlo methods to the analysis of spatially correlated data. Ph.D. Thesis, Stanford University .226p.

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

Izenman A, 2008. Modern Multivariate Statistical Techniques: Regression, Classification and Manifold Learning. Springer, New York. 731p.

Jimenez J, 2008. The field performance of SAGD projects in Canada. International Petroleum Technology Conference. 11p.

Johnson D, 1977. Efficient algorithms for shortest paths in sparse networks. Journal of the ACM. 24(1):1-13.

Journel A, 2004. Beyond covariance: The advent of multiple-point geostatistics. In Leuangthong O and Deutsch C (eds.) Geostats 2004: Proceedings of the seventh international geostatistics congress, Springer. 225-235.

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

Journel A and Rossi M, 1989. When do we need a trend model in kriging? Mathematical Geology. 21(7):715-739.

Journel A and Kyriakidis P, 2004. Evaluation of Mineral Reserves: A Simulation Approach. Oxford University Press. 232p.

Kennel M, 2004. KDTREE 2: Fortran 95 and C++ software to efficiently search for near neighbors in a multi-dimensional Euclidean space. eprint arXiv:physics/0408067
http://adsabs.harvard.edu/abs/2004physics...8067K

Keogh K, Martinius A and Osland R, 2007. The development of fluvial stochastic modelling in the Norwegian oil industry: A historical review, subsurface implementation and future directions. Sedimentary Geology. 202(1):249-268.

Kern J and Higdon D, 1999. A distance metric to account for edge effects in spatial data analysis. In American Statistical Association Proceedings of the Biometrics Section. 49-52.

Krivoruchko K and Gribov A, 2004. Geostatistical interpolation and simulation in the presence of barriers. In GEOENV IV: Geostatistics for Environmental Applications, Barcelona, Spain. 331–342.

Laporte G and Osman I, 1995. Routing problems: A bibliography. Journal Annals of Operations Research. 61(1):227-262.

Leeder M and Perez-Arlucea M, 2006. Physical Processes in Earth and Environmental Sciences. Blackwell Publishing, Oxford. 321p.

Legleiter C and Kyriakidis P, 2007. Forward and inverse transformations between Cartesian and channel-fitted coordinate systems for meandering rivers. Mathematical Geology. 38(8):927-958.

Lettley C, 2004. Elements of a genetic framework for inclined heterolithic strata of the McMurray formation, Northeast Alberta. M.Sc. Thesis. University of Alberta. 150p.

Leuangthong O, McLennan J and Deutsch C, 2004. Minimum acceptance criteria for geostatistical realizations, Natural Resources Research. 13(3):131-141.

Little L, Edwards D and Porter E, 1997. Kriging in estuaries: As the crow flies, or as the fish swims? Journal of Experimental Marine Biology and Ecology. 213(1):1-11.

Liu Y, Harding A, Gilbert R and Journel A, 2004. A workflow for multiple-point geostatistical simulation. In Leuangthong O and Deutsch C (eds.) Geostats 2004: Proceedings of the seventh international geostatistics congress, Springer. 245-255.

Loland A and Host G, 2003. Spatial covariance modelling in a complex costal domain by multidimensional scaling. Environmetrics. 14(3):307-321.

Lyster S and Deutsch C, 2006. A new MPS simulation algorithm based on Gibbs sampling. Report Eight: Center for Computational Geostatistics. 26p.

van der Maaten L, Postma E and van den Herik H, 2008. Dimensionality reduction: A comparative review. Maastricht University. Preprint submitted to Elsevier. 22p.

van der Maaten L, 2007. An introduction to dimensionality reduction using Matlab. Maastricht University. 38p.

Mackey S, Bridge J, 1992. A revised FORTRAN program to simulate alluvial stratigraphy, Computers and Geosciences. 18(2):119-181.

Mallet J, 2002. Geomodeling. Oxford University Press First Edition. 624p.

Mardia K, Kent J and Bibby J, 1979. Multivariate Analysis. Academic Press. 521p.

Matheron G, 1962. Traite de Geostatistique Appliquee. Technip, Paris.

Matheron G, 1973. The intrinsic random functions and their applications. Advances in Applied Probability. 5(3):439-468.

McBratney A, Webster R and Burgess T, 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables – I. Computers & Geosciences. 7(4):331-334.

McLennan J, 2008. The Decision of Stationarity. Ph.D. Thesis. University of Alberta. 167p.

Medhi D and Ramasamy K, 2007. Network Routing: Algorithms, protocols and architectures. Elsevier Morgan Kaufmann, Boston. 788p.

Meiring W, Monestiez P, Sampson P and Guttorp P, 1997. Developments in the modeling of nonstationary spatial covariance structure from space time monitoring data. In Baafi E and Schofield N (eds.) Geostats Wollongong 1996. 162-174.

Meiring W, Guttorp P and Sampson P, 1998. Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. Environmental and Ecological Statistics. 5(3):197-222.

Mohammadhassanpour R, 2007. Tools for multivariate modeling of permeability tensors and geometric parameters for unstructured grids. M.Sc. Thesis, University of Alberta. 61p.

Mohammadhassanpour R and Deutsch C, 2008. Fitting local anisotropy with mass moments of inertia. Center for Computational Geostatistics Annual Report 10. University of Alberta. 9p.

Moon C, Whateley M and Evans A, 2007. Introduction to Mineral Exploration. Blackwell Publishing. 481p.

Morris W, 1975. The Heritage Illustrated Dictionary of the English Language: International Edition. American Heritage Publishing Co. 1550p.

Mueller T, Dhanikonda S, Pusuluri N, Karathanasis A, Mathias K, Mijatovic B and Sears B, 2005. Optimizing inverse distance weighted interpolation with cross-validation. Soil Science. 170(7):504-515.

Myers D, 1991. Pseudo-cross variograms, positive-definiteness, and cokriging. Mathematical Geology. 23(6):805-816.

Newham R, 2005. Properties of Materials: Anisotropy, Symmetry, Structure. Oxford University Press, New York. 390p.

Nott D and Dunsmuir W, 2002. Estimation of nonstationary spatial covariance structure. Biometrika. 89(4):819-829.

Ortiz J and Deutsch C, 2004. Indicator simulation accounting for multiple-point statistics. Mathematical Geology. 36(5):545-565.

Paciorek C and Schervish M, 2006. Spatial modelling using a new class of nonstationary covariance functions. Environmetrics. 17(5):483-506

Padhy C, Sen D and Bhaskaran P, 2008. Application of wave model for weather routing of ships in the North Indian Ocean. Journal Natural Hazards. 44(3):373-385.

Pearson K, 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine. 2:559-572.

Perrin O and Meiring W, 1999. Identifiability for nonstationary spatial structure. Journal of Applied Probability. 36(4):1244-1250.

Pirajno F, 2000. Ore Deposits and Mantel Plumes. Kluwer Academic Press. 556p.

Pyrcz M, Boisvert J and Deutsch C, 2008. A library of training images for fluvial and deepwater reservoirs and associated code. Computers and Geosciences. 34(5):542-560.

Ranger M and Gingras M, 2003. Geology of the Athabasca Oil Sands – Field guide and overview. Canadian Society of Petroleum Geologists. 123p.

Rathbun S, 1998. Spatial modeling in irregularly shaped regions: kriging estuaries. Environmetrics. 9(2):109-129.

Ren W, Deutsch C, Garner D, Wheeler T, Richy J and Mus E, 2006. Quantifying resources for the Surmont lease with 2D mapping and multivariate statistics. SPE Annual Technical Conference and Exhibition. 24-27.

Renard D and Ruffo P, 1992. Depth, dip and gradient. In Soares A (ed.) Geostats Troia 1992. 167-178.

Roweis S and Saul L, 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500):2323-2326.

Sampson P and Guttorp P, 1992. Nonparametric estimation of nonstationary spatial covariance structure. Journal of the American Statistical Association. 87(417):108-119.

Sampson P, Damian D and Guttorp P, 2001. Advances in modelling and inference for environmental processes with nonstationary spatial covariance. In Monestiez P, Allard D and Froidvaux R (eds.) GeoENV 2000: Geostatistics for Environmental Applications. Kluwer Academic Publishers. 17-32.

Schmidt A and O'Hagan A, 2003. Bayesian inference for nonstationary spatial covariance structure via spatial deformations. Journal of the Royal Statistical Society. 65(3):743-758

Shin H and Polikar, 2005. Optimizing the SAGD process in three major Canadian oil-sands areas. SPE Annual Technical Conference and Exhibition. 9-12.

Siek J, Lee L and Lumsdaine A, 2001. The Boost Graph Library: User Guide and Reference Manual. Addison-Wesley. 352p.

Silva V and Tenenbaum J, 2003. Global versus local methods for nonlinear dimensionality reduction. In Becker S, Thrun S and Obermayer K (eds.) Advances in Neural Information Processing Systems 15. MIT Press, Cambridge. 721-728.

Slatt R, 2006. Stratigraphic Reservoir Characterization for Petroleum Geologists, Geophysicists, and Engineers. Elsevier, Oxford. 478p.

Strebelle S, 2002. Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology. 34(1):1-22.

Strebelle S and Remy N, 2004. Post-processing of multiple-point geostatistical models to improve reproduction of training patterns In Leuangthong O and Deutsch C (eds.) Geostats 2004: Proceedings of the seventh international geostatistics congress, Springer. 979-988.

Stroet C and Snepvangers J, 2005. Mapping curvilinear structures with local anisotropy kriging. Mathematical Geology. 37(6):635-649.

Sullivan J, Satchwell S and Ferrax G, 2007. Grade estimation in the presence of trends – the adaptive search approach applied to the Andina Copper Deposit, Chile. In Magri J (ed.) Proceedings of the 33rd International Symposium on the Application of Computers and Operations Research in the Mineral Industry. GECAMIN. 135-143.

Sun W, Yuan Y, 2006. Optimization Theory And Methods: Nonlinear Programming. Springer. 687p.

Tenenbaum J, Silva V and Langford J, 2000. A global geometric framework for nonlinear dimensionality reduction. Science. 290(550):2319–2323.

Toth P and Vigo D, 2002. The Vehicle Routing Problem. Society for industrial and applied mathematics, Philadelphia. 367p.

Wackernagel H, 2003. Multivariate Geostatistics Third Edition. Springer-Verlag Berlin. 387p.

Warren A, 2003. Report 2003-A: EUB Athabasca Wabiskaw-McMurray Regional Geological Study. Alberta Energy and Utilities Board. 187p.

Watkins D, 2007. The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods. Society for Industrial Mathematics. 452p.

Wu J, Boucher A and Zhang T, 2008. A SGeMS code for pattern simulation of continuous and categorical variables: FILTERSIM. Computers & Geosciences. 34(12):1863-1876.

Xu W, 1996. Conditional curvilinear stochastic simulation using pixel-based algorithms. Mathematical Geology. 28(7):937-949.

Yao T, Calvert C, Jones T, Foreman L and Bishop G, 2007. Conditioning geologic models to local continuity azimuth in spectral simulation. Mathematical Geology. 39(3):349-354.

Yuan L, 2004. Using spatial interpolation to estimate stressor levels in unsampled streams. Environmental Monitoring and Assessment. 94(1):23-38.

Zhang T, Switzer P and Journel A, 2004. Sequential conditional simulation using classification of local training patterns. In Leuangthong O and Deutsch C (eds.) Geostats 2004: Proceedings of the seventh international geostatistics congress, Springer. 265-275.

Zimmerman D, 1993. Another look at anisotropy in geostatistics. Mathematical Geology. 25(4):453-470.

Images used in Chapter 1:

Field image, accessed January 2008, http://www.surreyproperty.com

Gold vein core, accessed January 2008, http://jpgold.com

Coal face, accessed January 2008, http://www.scienceclarified.com

San Andreas Fault, accessed January 2008, http://strike-slip.geol.ucsb.edu

Rocky Mountains, accessed January 2008, http://mkutis.iweb.bsu.edu

Meandering River, accessed January 2008, http://www.hi.is

Anticline accessed January 2009, http://geology.about.com

Appendix A: Detailed Numerical Example: Calculating Distance and Covariance in the Presence of Anisotropy

In a traditional implementation of kriging, anisotropy is constant in the modeling domain and is used to calculate the distance between points. To estimate at an unknown grid cell location, the anisotropic distance is used to generate a covariance between points with a variogram function; the covariance between points is used to generate a weight to give to each datum with the kriging equations. Once the weights have been determined the estimate is calculated. The following numerical example demonstrates how to calculate the anisotropic distance and covariance between two points.

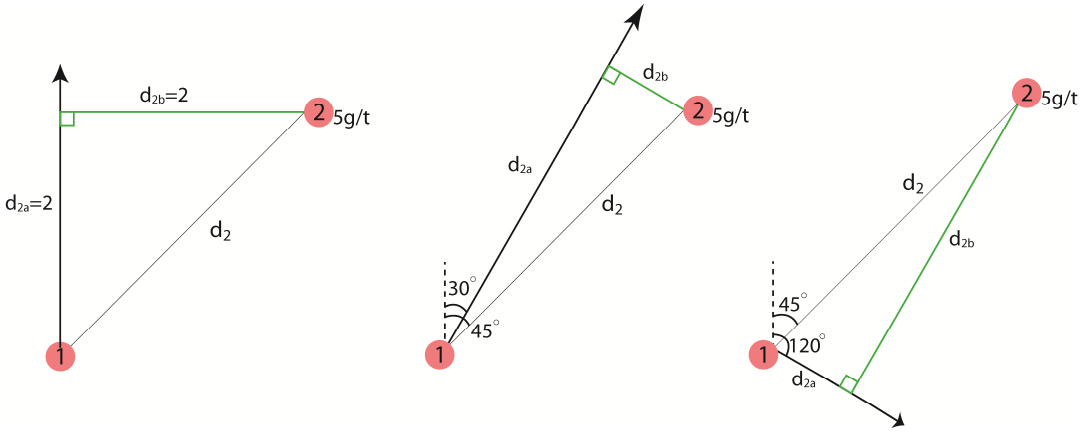


Figure A1: The Euclidian distance between the two points (distance without considering anisotropy) is 2 units in the vertical direction and 2 units in the horizontal direction. For the anisotropic calculations an anisotropy ratio of 10:1 is assumed. Left: No anisotropy. Middle: Anisotropy in the 30° direction. Right: Anisotropy in the 120° direction.

Case 1: No Anisotropy

The distances between the two points is, $d_2 = \sqrt{2^2 + 2^2} = 2.82$ units.

Case 2: Anisotropy in the 30° Direction.

First the components in the 30° (d_{2a}=2.73) and 120° (d_{2b}=0.732) directions are calculated using basic trigonometry. Once these components are known, the anisotropy ratio (which defines the magnitude of the anisotropy in the 45° direction) can be used to calculate the distance between points. For this example a 10:1 anisotropy is assumed, thus the anisotropic distance (d_2) is:

$$d_{30^\circ} = \sqrt{\left(\frac{d_a}{a_{30^\circ}}\right)^2 + \left(\frac{d_b}{a_{120^\circ}}\right)^2} = \sqrt{\left(\frac{2.73}{10}\right)^2 + \left(\frac{0.732}{1}\right)^2} = 0.781 \text{ units}$$

Case 3: Anisotropy in the 120° Direction.

This same equation can be used to calculate the distance if the anisotropy is in the 120° direction. The components are now: d_{2a}=0.732 in the 120° direction and d_{2b}=2.73 in the 30° direction:

$$d_{120^\circ} = \sqrt{\left(\frac{d_a}{a_{120^\circ}}\right)^2 + \left(\frac{d_b}{a_{30^\circ}}\right)^2} = \sqrt{\left(\frac{0.732}{10}\right)^2 + \left(\frac{2.73}{1}\right)^2} = 2.73 \text{ units}$$

The effect of anisotropy is to reduce the distance between points *along* the direction of anisotropy.

Using a Rotation Matrix to Calculate Distance

The operations shown above can be automated by introducing a rotation matrix:

$$R = \begin{bmatrix} \cos\alpha\cos\varphi - \sin\alpha\sin\beta\sin\varphi & -\sin\alpha\cos\varphi - \cos\alpha\sin\beta\sin\varphi & \cos\beta\sin\varphi \\ \frac{1}{r_1}\sin\alpha\cos\beta & \frac{1}{r_1}\cos\alpha\cos\beta & \frac{1}{r_1}\sin\beta \\ \frac{1}{r_2}(-\cos\alpha\sin\varphi - \sin\alpha\sin\beta\cos\varphi) & \frac{1}{r_2}(\sin\alpha\sin\varphi - \cos\alpha\sin\beta\cos\varphi) & \frac{1}{r_2}\cos\beta\cos\varphi \end{bmatrix} \quad \textbf{A.1}$$

where two points are separated by the vector (**h**). The direction of anisotropy is defined in 3D by three angles, strike (α) dip (β) and plunge (φ). The magnitude of anisotropy is defined by two ratios: r_1 is the ratio between the minor and major directions and r_2 is the ratio between the vertical and major direction. The calculation of the anisotropic distance becomes:

$$d^2(\mathbf{h}) = \mathbf{h}^T \; R^T R \; \mathbf{h} \quad \textbf{A.2}$$

For the example, when anisotropy is considered in the 30° direction:

$$R = \begin{bmatrix} \cos 30\cos 0 - \sin 30\sin 0\sin 0 & -\sin 30\cos 0 - \cos 30\sin 0\sin 0 & \cos 0\sin 0 \\ \frac{1}{10}\sin 30\cos 0 & \frac{1}{10}\cos 30\cos 0 & \frac{1}{10}\sin 0 \\ 0(-\cos 30\sin 0 - \sin 30\sin 0\cos 0) & 0(\sin 30\sin 0 - \cos 30\sin 0\cos 0) & 0\cos 0\cos 0 \end{bmatrix} = \begin{bmatrix} .866 & -0.5 & 0 \\ 0.05 & 0.0866 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The distance between two point separated by a lag, **h=**[2,2,0], becomes:

$$d_{30^\circ}(2,2,0) = \sqrt{\mathbf{h}^T R^T R \; \mathbf{h}} = \sqrt{\begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}^T \bullet \begin{bmatrix} .866 & -0.5 & 0 \\ 0.05 & 0.0866 & 0 \\ 0 & 0 & 0 \end{bmatrix}^T \bullet \begin{bmatrix} .866 & -0.5 & 0 \\ 0.05 & 0.0866 & 0 \\ 0 & 0 & 0 \end{bmatrix} \bullet \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}} = 0.781$$

This could be repeated for anisotropy in the 120˚ direction by substituting α=120˚ in Equations A.1 and A.2.

Calculating Covariance from the Anisotropic Distance

Once the anisotropic distance in each direction is determined, the variogram function, $\gamma(\mathbf{h})$, provides the covariance between the two points. Consider a spherical variogram with a range of 10m in the 30° direction and 1m in the 120° direction (Case 2 above, with anisotropic distance, $||\mathbf{h}||$=0.781 units). The equation for a spherical variogram is:

$$\gamma(\mathbf{h}) = Sph(\mathbf{h}) = \begin{cases} [1.5h - 0.5h^3], if\ h \leq 1 \\ 1 \qquad\qquad\qquad , if\ h > 1 \end{cases}$$

Thus the variogram value for a distance of $||\mathbf{h}||$=0.781 units is:

$$\gamma(0.781) = Sph(0.781) = 1.5(0.781) - 0.5(0.781)^3 = 0.933$$

From the variogram value the covariance between these two points, $C(\mathbf{h})$, can be determined if the variance of the variable, $C(0)$, is known. In geostatistics it is common to consider that the variable has been transformed to a N(0,1) distribution and by definition the variogram is:

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$$

Thus, if the variance of the variable is 1.0, the covariance between these two points is:

$$C(\mathbf{h}) = 1.0 - \gamma(\mathbf{h}) = 1.0 - 0.933 = 0.067$$

Appendix B: Implementation Details of Kriging and Simulation Source Code

FORTRAN based GSLIB style (Deutsch and Journel 1998) programs are provided for variogram calculation, kriging and simulation. The majority of the inputs to these programs are identical to their GSLIB equivalents and all additional parameters are discussed in detail below. Further detail on programs and access to source code can be found in Boisvert (2010).

Support Program Boost_dijkstra.exe and External Dependencies

Boost_dijkstra.exe is a C++ wrapper to run the Dijkstra algorithm using the Boost Graph Library (BGL) (Siek *et al.* 2001). The program reads in the specification of the graph from the files grid.out and nodes2cal.out and returns the necessary SPDs in a file dist_cpp.out. This program is called by gamv_LVA.exe, kt3d_LVA.exe and SGS_LVA.exe and should not be run manually by the user.

Gamv_LVA.exe

To implement kriging or SGS with LVA a variogram of the data in the embedded space after L-ISOMAP is required. Thus, the experimental variogram calculation must first embed the data into q dimensions. Embedding with L-ISOMAP depends on the input grid used to locate the landmark points. When the desired block model and the available LVA model have different model specifications, L-ISOMAP must be performed on the block model grid. Within gamv_LVA.exe, there is the ability to regrid an LVA model to the desired block model if they differ (lines 11-16, Figure B1). When these grids are identical, no regridding is required. The LVA model must fully cover the desired block model.

Lines 1-8 are identical to gamv.exe (Deutsch and Journel 1998). Lines 9-16 are the specifications of the required grids. In the LVA file the angles measure strike (measure positive clockwise from north) dip (positive downwards) and plunge. The anisotropy ratios are defined as the min/max range (i.e. a 10:1 anisotropy ratio is 0.1).

Line 17 indicates how many nodes are connected in the graph. Connecting more nodes results in more flexible, and thus shorter, paths after the implementation of the Dijkstra algorithm but requires more memory and may be infeasible for large 3D models. Line 18 allows gamv_LVA.exe to read a set of distances previously generated by the Dijkstra algorithm to reduce CPU requirements. Line 19 specifies the location of the landmark points for L-ISOMAP. Line 20 specifies the number of dimensions to retain after L-ISOMAP for the calculation of the distance between datum. Using fewer than the maximum number of dimensions can reduce run times but does not reduce memory requirements. When using fewer than the maximum number of dimensions, the most important dimensions, based on the eigenvalues of each dimension, are retained.

```
LINE 1 :  START OF PARAMETERS:
LINE 2 :  data.out           -file with data
LINE 3 :  1 2 0 3            -columns for X, Y, Z coordinates, VAR
LINE 4 :  -998 1e21          -trimming limits
LINE 5 :  LVA.out            -file for variogram output
LINE 6 :  40                 -number of lags
LINE 7 :  5                  -lag separation distance
LINE 8 :  3                  -lag tolerance
LINE 9 :  grid_LVA.out       -file containing the locally varying anisotropy grid
LINE 10:  4 5 6 7 8          -LVA columns for ang1, ang2, ang3, aniso ratio min/max, aniso ratio vert/max
LINE 11:  100 0.5 1          -nx,xmn,xsiz (ESTIMATION GRID see NOTE1)
LINE 12:  100 0.5 1          -ny,ymn,ysiz (ESTIMATION GRID see NOTE1)
LINE 13:  1   0.5 1          -nz,zmn,zsiz (ESTIMATION GRID see NOTE1)
LINE 14:  100 0.5 1          -nx,xmn,xsiz (LVA GRID see NOTE1)
LINE 15:  100 0.5 1          -ny,ymn,ysiz (LVA GRID see NOTE1)
LINE 16:  1   0.5 1          -nz,zmn,zsiz (LVA GRID see NOTE1)
LINE 17:  1                  -noffsets for graph (number of offsets described below, see NOTE2)
LINE 18:  2                  -use MDS? 2=L-ISOMAP 3=read dist from 'grid_cpp.out' and use L-ISOMAP (see NOTE3)
LINE 19:  10 10 1            -number of landmark points in x,y,z (evenly spaced grid nodes)
LINE 20:  -1                 -max number of dimensions to use (set -1 to use max posible)
LINE 21:  0                  -standardize sills? (0=no, 1=yes)
```

Figure B1: Parameters for gamv_LVA.exe.

kt3d_LVA.exe

Kriging with LVA requires (1) the embedding of the estimation grid into q dimensions with L-ISOMAP and (2) searching for n nearest data to an estimation location in q dimensions. The parameters for L-ISOMAP are identical to the gamv_LVA.exe program (lines 8-19, Figure B2). Searching for n nearest data is accomplished in *q* dimensions with a *k*d tree (Kennel 2004). Using fewer dimensions in the search (line 22, Figure B2) can reduce run times and memory requirements.

Because the data are located in q dimensions the distance calculations can be CPU intensive. The onetime calculation of the *n*x*n* distance matrix between data can reduce CPU time (Figure B3) if fewer than ~10,000 data are used. If possible, the *n*x*n* matrix is stored, otherwise the distances are calculated as needed.

KT3D_LVA.exe can be used to implement IDW. Setting line 23 <0 indicates IDW. To perform IDW with ω=2.5 set line 23 to "-2.5 0". The second value, the simple kriging mean, is ignored.

The remaining parameters (lines 1-7 and 23-25, Figure B2) are identical to the kt3d.exe GSLIB program (Deutsch and Journel 1998). Note that the specified variogram model must be positive definite in q dimensions (line 19, Figure B2), the exponential variogram model is recommended. Time trials using Microsoft Windows Server 2003 Standard x64 Edition, 2.41GHz AMD processor are provided (Figure B4).

```
START OF PARAMETERS:
LINE 1 :  data.out          -file with data
LINE 2 :  1 2 0 3           -columns for X,Y,Z,var
LINE 3 :  -998   1.0e21     -trimming limits
LINE 4 :  0 0               -option (0=grid 1=cross) , col for  DHID (kriging a grid does not use the DHID)
LINE 5 :  0                 -debugging level: 0,1,2,3
LINE 6 :  kt3d.dbg          -file for debugging output
LINE 7 :  kriging.out       -file for kriged output
LINE 8 :  100 0.5 1         -nx,xmn,xsiz (ESTIMATION GRID see NOTE1)
LINE 9 :  100 0.5 1         -ny,ymn,ysiz (ESTIMATION GRID see NOTE1)
LINE 10:  1   0.5 1         -nz,zmn,zsiz (ESTIMATION GRID see NOTE1)
LINE 11:  grid_LVA.out      -file containing the locally varying anisotropy grid
LINE 12:  4 5 6 7 8         -LVA columns for ang1, ang2, ang3, aniso ratio min/max, aniso ratio vert/max
LINE 13:  100 0.5 1         -nx,xmn,xsiz (LVA GRID see NOTE1)
LINE 14:  100 0.5 1         -ny,ymn,ysiz (LVA GRID see NOTE1)
LINE 15:  1   0.5 1         -nz,zmn,zsiz (LVA GRID see NOTE1)
LINE 16:  1                 -noffsets for graph (number of offsets described below, see NOTE2)
LINE 17:  2                 -use MDS? 2=L-ISOMAP 3=read dist from 'grid_cpp.out' and use L-ISOMAP (see NOTE3)
LINE 18:  10 10 1           -number of landmark points in x,y,z (evenly spaced grid nodes)
LINE 19:  -1                -max number of dimensions to use (set -1 to use max, see NOTE4)
LINE 20:  2 60              -min, max data for kriging
LINE 21:  1000              -maximum search radii (a 1D isotropic distance in q dimensions)
LINE 22:  -1                -maximum number of dimensions to use in search (-1 uses all dimensions see NOTE4)
LINE 23:  0     0           -0=SK,1=OK, simple kriging mean
LINE 24:  1 0.00            -# of nested structures, nugget effect (1D variogram)
LINE 25:  2 1 35            -it,cc,range
```

Figure B2: Parameters for kt3d_LVA.exe.

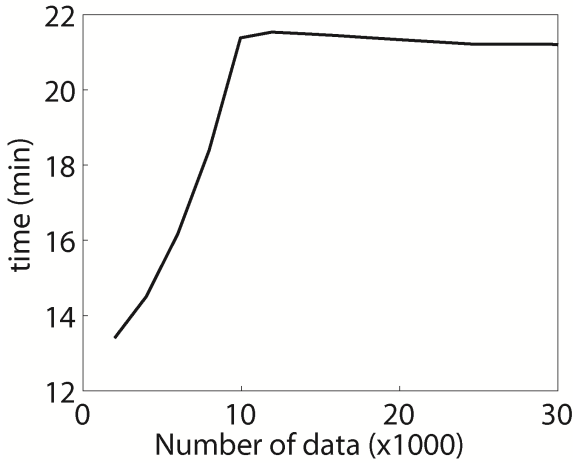


Figure B3: Time required to krige a 1M cell model using 100 landmark points. The inflection point at 10,000 data is a result of the expense of calculating the *nxn* distance matrix for a large number of data.

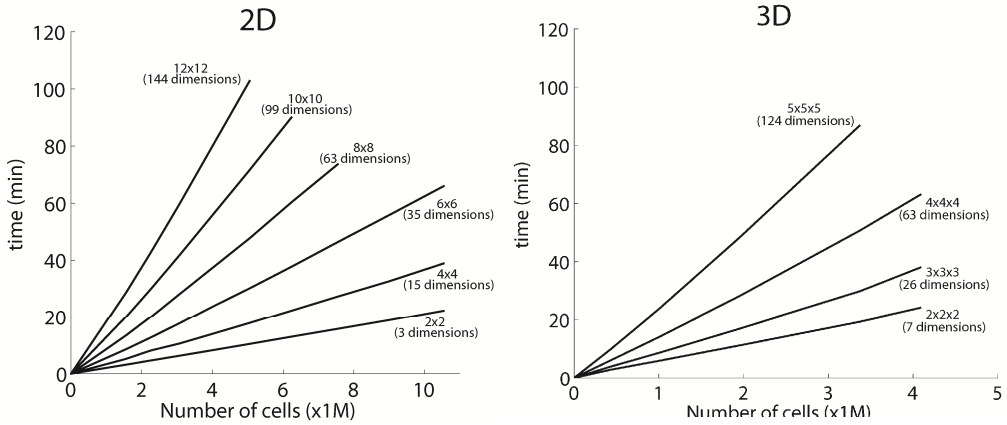


Figure B4: Left: Time required to krige a 2D grid with 3000 data using 20 nearest data. Each line represents a square pattern of landmark points with *L*-1 dimensions used in MDS. Right: Time required to krige a 3D grid with 3000 data using 20 nearest data.

SGS_LVA.exe

The parameters for SGS_LVA.exe (Figure B5) are identical to kriging (Figure B2). The only implementational difference is in the searching for previously simulated nodes. The *k*d tree is not efficient when there are many empty branches that lead to unsimulated nodes; therefore, an exhaustive search is performed until the *k*d tree is more efficient at which point the *k*d tree is used. Also note that the input data to the program must be normal scored. There is no normal score transform within the SGS_LVA.exe program. Time trials using Microsoft Windows Server 2003 Standard x64 Edition, 2.41GHz AMD processor are provided (Figure B6).

```
LINE 1 :  START OF PARAMETERS:
LINE 2 :  nsdata.out          -file with data
LINE 3 :  1 2 0 3             -columns for X,Y,Z,var,sec var
LINE 4 :  -998   1.0e21       -trimming limits
LINE 5 :  0                   -debugging level: 0,1,2,3
LINE 6 :  sgs.dbg             -file for debugging output
LINE 7 :  sgs.out             -file for output
LINE 8 :  10                  -nreal
LINE 9 :  100 0.5 1           -nx,xmn,xsiz (ESTIMATION GRID see NOTE1)
LINE 10:  100 0.5 1           -ny,ymn,ysiz (ESTIMATION GRID see NOTE1)
LINE 11:  1   0.5 1           -nz,zmn,zsiz (ESTIMATION GRID see NOTE1)
LINE 12:  32146               -random number seed
LINE 13:  grid_LVA.out        -file containing the locally varying anisotropy grid (LVAG)
LINE 14:  1 2 3 4 5           -LVA columns for ang1, ang2, ang3, aniso ratio min/max, aniso ratio vert/max
LINE 15:  100 0.5 1           -nx,xmn,xsiz (LVA GRID see NOTE1)
LINE 16:  100 0.5 1           -ny,ymn,ysiz (LVA GRID see NOTE1)
LINE 17:  1   0.5 1           -nz,zmn,zsiz (LVA GRID see NOTE1)
LINE 18:  1                   -noffsets for graph (number of offsets described below, see NOTE2)
LINE 19:  2                   -use MDS? 2=L-ISOMAP 3=read dist from 'grid_cpp.out' and use L-ISOMAP (see NOTE3)
LINE 20:  10 10 1             -number of landmark points in x,y,z (evenly spaced grid nodes)
LINE 21:  -1                  -max number of dimensions to use (set -1 to use max, see NOTE4)
LINE 22:  2 30                -min, max nodes for simulation
LINE 23:  1000                -maximum search radii (a 1D isotropic distance in q dimensions)
LINE 24:  -1                  -maximum number of dimensions to use in search (-1 uses all dimensions see NOTE4)
LINE 25:  0     0             -0=SK,1=OK, simple kriging mean
LINE 26:  1 0.00              -# of nested structures, nugget effect (1D variogram)
LINE 27:  2 1 35              -it,cc,range (see NOTE5)
```

Figure B5: Parameters for SGS_LVA.exe.

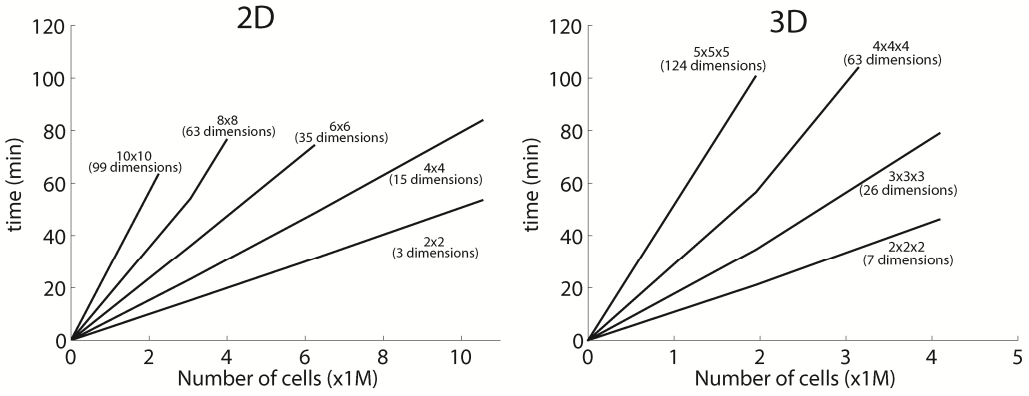


Figure B6: Total time required to generate 100 SGS realizations with 3000 data using the 20 nearest previously simulated nodes.

Basic Workflow

1) Generate LVA field (Chapter 3). It is assumed that the strike, dip, plunge and anisotropy ratios are available exhaustively in the domain of interest.
2) Calculate the experimental variogram (gamv_LVA.exe). An isotropic variogram is fit to the experimental variogram.
3) Select the MDS parameters. This requires selecting the number of landmark points and the number of offsets to use in the Dijkstra algorithm. These values should be as large as possible within CPU constraints. 50 landmark points are recommended by Izenman (2008) and it has been the authors experience that there is no need to consider more than 100.
4) Krige or simulate with the modeled variogram.

Parameter files showing the implementation of each program are provided in Figure 22.

```
------------------------------------------------------------------------------------------------------------
                  Parameters for GAMV_LVA
                  ***********************
START OF PARAMETERS:
ns_Au.out             -file with data
1 2 3 7               -columns for X, Y, Z coordinates, VAR
-998 1e21             -trimming limits
gamv.out              -file for variogram output
100                   -number of lags
1                     -lag separation distance
.75                   -lag tolerance
LVA.out               -file containing the locally varying anisotropy grid
1 2 3 4 5             -LVA columns for ang1, ang2, ang3, aniso ratio min/max, aniso ratio vert/max
100  110   5          -nx,xmn,xsiz (ESTIMATION GRID)
100   40   5          -ny,ymn,ysiz (ESTIMATION GRID)
110 -100   2          -nz,zmn,zsiz (ESTIMATION GRID)
100  110   5          -nx,xmn,xsiz (LVA GRID)
100   40   5          -ny,ymn,ysiz (LVA GRID)
110 -100   2          -nz,zmn,zsiz (LVA GRID)
1                     -noffsets for graph (number of offsets described below)
2                     -use MDS? 2=L-ISOMAP 3=read dist from  grid_cpp.out  and use L-ISOMAP
3 3 3                 -number of landmark points in x,y,z
-1                    -max number of dimensions to use (set -1 to use max)
0                     -standardize sills? (0=no, 1=yes)
------------------------------------------------------------------------------------------------------------
                  Parameters for KT3D_LVA
                  ***********************
START OF PARAMETERS:
ns_Au.out             -file with data
1  2 3 7              -columns for X,Y,Z,var
-998   1.0e21         -trimming limits
0  0                  -option (0=grid 1=cross) , col for  DHID (kriging a grid does not use the DHID)
0                     -debugging level: 0,1,2,3
kt3d_LVA.dbg          -file for debugging output
kriging_Au.ou         -file for kriged output
100  110   5          -nx,xmn,xsiz (ESTIMATION GRID)
100   40   5          -ny,ymn,ysiz (ESTIMATION GRID)
110 -100   2          -nz,zmn,zsiz (ESTIMATION GRID)
LVA.out               -file containing the locally varying anisotropy grid
1 2 3 4 5             -LVA columns for ang1, ang2, ang3, aniso ratio min/max, aniso ratio vert/max
100  110   5          -nx,xmn,xsiz (LVA GRID)
100   40   5          -ny,ymn,ysiz (LVA GRID)
110 -100   2          -nz,zmn,zsiz (LVA GRID)
1                     -noffsets for graph (number of offsets described below)
2                     -use MDS? 2=L-ISOMAP 3=read dist from  grid_cpp.out  and use L-ISOMAP
3 3 3                 -number of landmark points in x,y,z (evenly spaced grid nodes)
-1                    -max number of dimensions to use (set -1 to use max)
1 30                  -min, max data for kriging
1000                  -maximum search radii (a 1D isotropic distance in q dimensions)
-1                    -maximum number of dimensions to use in search (-1 uses all dimensions)
0     0               -0=SK,1=OK, simple kriging mean
2 0.15                -# of nested structures, nugget effect (1D variogram)
2 0.15 30             -it,cc,range
2 0.7  150            -it,cc,range
------------------------------------------------------------------------------------------------------------
```

Figure B7: Parameters used to generate the example.