University of Alberta

Geostatistics with Location-Dependent Statistics

by

David Francisco Machuca-Mory

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

Fall 2010
Edmonton, Alberta

Examining Committee

Clayton V. Deutsch, Civil and Environmental Engineering

Hooman Askari-Hasab, Civil and Environmental Engineering

Juliana Leung, Civil and Environmental Engineering

Arturo Sanchez-Azofeifa, Earth and Atmospheric Sciences

Jean-Paul Chilès, Centre de Géosciences, MINES, ParisTech

Dedico este trabajo a la felicidad y
prosperidad de la Casa Machuca Mory.

Abstract

In Geostatistical modelling of the spatial distribution of rock attributes, the multivariate distribution of a Random Function defines the range of possible values and the spatial relationships among them. Under a decision of stationarity, the Random Function distribution and its statistics are inferred from data within a spatial domain deemed statistically homogenous. Assuming stationary multiGaussianity allows spatial prediction techniques to take advantage of this simple parametric distribution model. These techniques compute the local distributions with surrounding data and global spatially invariant statistics. They often fail to reproduce local changes in the mean, variability and, particularly, the spatial continuity, that are required for geologically realistic modelling of rock attributes. The proposed alternative is to build local Random Function models that are deemed stationary only in relation to the locations where they are defined. The corresponding location-dependent distributions and statistics are inferred by weighting the samples inversely proportional to their distance to anchor locations. These distributions are locally Gaussian transformed. The transformation models carry information on the local histogram. The distance weighted experimental measures of spatial correlation are able to adapt to local changes in the spatial continuity and are semi-automatically fitted by locally defined variogram models. The fields of local variogram and transformation parameters are used in locally stationary spatial prediction algorithms. The resulting attribute models are rich in non-stationary spatial features. This process implies a higher computational demand than the traditional techniques, but, if data is abundant enough to allow a reliable inference of the local statistics, the proposed locally stationary techniques outperform their stationary counterparts in terms of accuracy and precision. These improved models have the potential of providing better decision support for engineering design.

Acknowledgements

First of all, I would like to thank my supervisor Dr. Clayton V. Deutsch for all of his guidance and support, as well as for all of the valuable things that I learned from him, which in many cases are well beyond the boundaries of Geostatistics.

I'm also very thankful to the industrial sponsors of the Centre for Computational Geostatistics for providing the funding that allowed me to pursue this research.

Among my fellow students and friends I would like to thank Dr. Oy Leuangthong, Dr. Olena Babak, Dr. Jeff Boisvert, Miguel Cuba and Mike Munroe for the fruitful discussions and exchange of ideas that I had with them during the time I was developing my research. I also appreciate the help I received from John Manchuk and Dr. Steven Lyster with solving programming problems, and from Colette Freitas for all her assistance.

Epicurus used to say that the first ingredient for a happy life is friendship. The friends I met during my time as a graduate student filled those years with happy memories. Thanks Ali, A.J., Adriana, Alex, Alison, Ander, Chonggang, Deepak, Hagop, Xiaolei (Julia), Julieta (Sol), Lisa, Magda, Marina, Mohammed (Moho), Prado, Rosa, Russell, Sabujkoli, Sawa, Serge, Steve, Talal, Taras, Tong, Zoya, and all the members of my dragon boat team, the United International Dragons. Among my friends I would like to particularly thank Alison Turner for her help in proofreading my thesis, Dr. Alexander Pswarayi for his pertinent suggestions on my presentations, and Angel Mondragon and his team for their help with data processing and mining software.

Special thanks also to Jaron and Kate Summers for making my long stay at their house most agreeable. And if someone has to be blamed for this, blame Tony Wain, for he was the one who initially facilitated my return to graduate school. Besides that, thank you very much Tony and Arda for your generosity and hospitality.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

In the geosciences there is concern for the spatial distribution of physical properties. These physical properties, also known as attributes, exhibit spatial continuity. They can be continuous within an interval or may take a categorical outcome among several possible states. Beyond the available data, an exhaustive spatial description of these attributes is often required for engineering design. This is accomplished by numerical models that represent the current knowledge, highlight the most relevant aspects and predict the spatial behaviour of the attributes at unsampled locations. These models are often used to assess the possible responses of taking different decisions, and thus for selecting the best option regarding an engineering task. For example, in mining, a model of the distribution of metal concentrations is required for assessing the viability of the exploitation of a mineral deposit and planning its extraction. Models of the petrophysical properties of petroleum reservoirs are used for deciding the placement of wells and forecasting their production. In environmental related applications, models of the spatial dispersion of a contaminant are used for identifying high-risk areas and for planning remediation. These are a few examples of the necessity of modelling the spatial distribution of attributes.

The challenge is to build such models from sparse data and incomplete knowledge of the geological setting at a scale suitable for its intended uses. Geological knowledge is used for delimiting domains where the attributes are deemed reasonably homogeneous.

Deterministic approaches regard this problem as an interpolation of scattered values between the sampled locations. This is often done by distance weighting or surface fitting approaches.

As an alternative, Geostatistics considers that the true, but unknown, attribute value at each unsampled location is one of a range of possible values. The probability of the true attribute value being within different intervals is modelled by the probability distribution of a Random Variable (RV), $Z(\mathbf{u})$, where $\mathbf{u}$ denotes the unsampled location. The values of an attribute show some spatial continuity; therefore the RVs at different

locations may be dependent. An ensemble of RVs at different locations within a homogeneous domain is known as a Random Function (RF). The multivariate probability distribution of the RF is inferred by pooling the available data within the domain and assuming a mathematical model. This pooling of data values at different locations is one aspect of the decision of stationarity.

In Geostatistics, the goal is not just to obtain the best interpolated value at every unsampled location, but to infer the local probability distributions of the attribute values. This gives Geostatistics an advantage over deterministic methods, since it provides a distribution of uncertainty. Such characterization of the uncertainty is preferred since it provides a basis to make robust decisions with respect to departures from a single deterministic estimate.

As mentioned above, one aspect of the decision of stationarity is the pooling of all data within a deemed homogeneous domain. The global statistical properties, such as the histogram and the measures of spatial continuity, are estimated from the entire pool of data. A second aspect is the invariance by translation of the global statistical properties of the RF and the mathematical model assumed for its distribution. This allows the spatial inference of the local probability distributions at unsampled locations.

In this context, the main question that this thesis addresses is: can the prediction of the distributions of uncertainty be improved by using local, rather than global, definitions of the RF? This improvement is measured in terms of increased accuracy and precision, and reduced uncertainty. To answer this question, a methodology for obtaining these prior local RF distributions and their statistics must be devised. Then, algorithms for using them in spatial prediction must be developed.

1.1. Problem Setting

Since the beginnings of Geostatistics several methodologies have been developed for dealing with variations of the local expected value of the attribute (Matheron 1969; Matheron 1970). Nowadays, this aspect of non-stationarity is well understood and the techniques for adressing it are well developed. The current approach of modelling can be summarized in two main components (McLennan 2007): (1) identification of homogenous domains and modelling of its boundaries, and, (2) if required, modelling of

a locally varying mean or trend within these domains. The subsequent estimation accounts for this structured trend component.

The inclusion of a locally varying mean in modelling may be insufficient to depict all the spatial features in a domain when the information available indicates local changes in the histogram shape, the variance, or the spatial continuity of an attribute. The effect of considering a global model of spatial continuity is illustrated in Figure 1-1. At the left side of this Figure an exhaustive image of an attribute is presented, that shows changing directions of the attribute's spatial continuity. This exhaustive image is seldom available in real life cases; instead, scattered samples may be available, as presented at the centre of Figure 1-1. In this case, the geostatistical technique called Ordinary Kriging is often used for estimating the values of the attribute between samples. This technique incorporates a global definition of the spatial continuity for such task. At the right side of the Figure, the resulting numerical model fails to reproduce the local changes in the spatial continuity and the curvilinear features observed in the exhaustive image.

Examples of geological settings where the spatial modelling of attributes may require locally changing measures of spatial continuity include those altered by processes like folding, meandering, or shearing. Attributes that show a decreasing tendency from a source may require a locally changing definition of the mean and variance. A locally changing bivariate correlation between attributes may be required when they show changing linear dependency at different locations. This is observed, for instance, in mineral deposits were the ratios between certain elements change from location to location, often in response to the temperature gradient away from the mineralization focus (Evans 1997, pp.77-79).

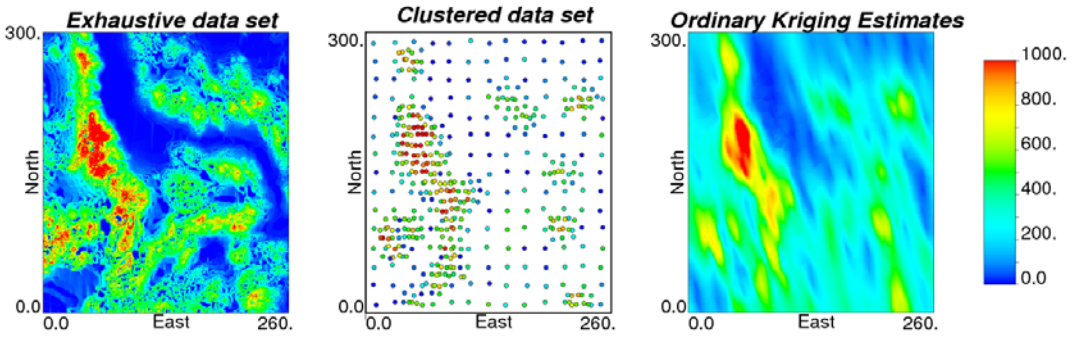

Figure 1-1: Exhaustive data set (left), clustered samples (centre) and Ordinary Kriging estimates obtained from the clustered samples (right)

There are several geostatistical techniques that assess these different aspects of non-stationarity; however, there is no comprehensive and practical methodology for dealing with them all together. The most relevant current approaches for non-stationarity are presented in Chapter 2.

1.2. The Decision of Local Stationarity

In its strictest form the decision of stationarity states that the multivariate probability distribution of the RF remains invariant if translated by any vector **h** (Matheron 1970; Deutsch & Journel 1998). This decision is associated with the adoption of a distribution model for the RF, which is often the multivariate Gaussian distribution after univariate transformation. This decision may be too rigid to accommodate local changes in the lower order distributions and their statistics, particularly when they depart from the assumed distribution model. A greater flexibility can be gained if only a few statistics of the RF are required to be invariant by translation. Relaxed forms of the desicion of stationarity used in geostatistics are the *second-order stationarity* and the *intrinsic stationarity*. The first stationarity form requires only the invariance by translation of the mean and the covariance between sample pairs separated by **h** (Chilès & Delfiner 1999, p.16). In the second, and weaker, stationarity form, only the variance of the difference between sample pairs separated by **h**, i.e. the variogram, is deemed invariant by translation (Chilès & Delfiner 1999, p.17). Both types of weak stationarity are discussed with more detail in Chapter 2.

The strict assumption of an invariant RF probability distribution is relaxed by these milder decisions of stationary; however there are still limitations and concerns: these types of weak stationarity do not allow local changes in the covariance, the variogram and other statistics relevant for spatial prediction. To overcome the limitations of the traditional forms of stationarity a *decision of local stationarity* is proposed for the definition of the RF. This amounts to strict stationarity of the RF, but only in relation to an anchor point **o**. Thus, the shape of the multivariate probability distribution of the RF and all its statistics depend of the location were this distribution is defined.

The flexibility offered by the local stationary decision comes with the price of rebuilding the statistics of the RF at many locations. Moreover, these local statistics can be reliably inferred only in presence of abundant data. Despite the increased effort that

this represents, the resultant models should incorporate more local information. This thesis demonstrates that the models built with more local information reflect better the local spatial features of the attribute under study. These models provide a more realistic assessment of the uncertainty and the potential for improved decisions.

1.3. Dissertation Outline

The central theme of this thesis is the decision of local stationarity. The chapters in this thesis develops from the reasons for proposing it, to the methodologies devised for obtaining the required location-dependent statistics and distributions and using them in spatial prediction. These chapters are outlined next.

Chapter 2 presents the concept of Random Functions and the different types of stationarity decisions in greater detail within a geostatistical context. The current methodology and techniques based on these stationarity decisions are reviewed. Their limitations are discussed and several of the currently used approaches for overcoming them are presented.

Chapter 3 develops a methodology based on distance weights for obtaining location dependent statistics and cdfs. The desirable characteristics and optimality criteria for those weights are developed and discussed. The Gaussian transformation of the local cdfs is presented. Locally weighted measures of spatial continuity are proposed and the issues concerning their inference and modelling are discussed.

Chapter 4 covers the methodology and algorithms devised for applying the location dependent statistics and distributions in estimation and simulation under the decision of local stationarity. The criteria for assessing the performance of these new algorithms in comparison to traditional techniques are presented.

Chapter 5 illustrates and tests the proposed methodology using actual examples. The practical details of the application of the developed algorithms are discussed. The resulting models are compared with those produced by traditional techniques.

Chapter 6 evaluates the advantages and disadvantages of the proposed methodologies. Issues in their practical implementation are highlighted. The place of locally stationary techniques in geostatistical modelling is discussed. The future developments related to this approach are contemplated.

Finally, an appendix containing the description of the programs developed for the implementation of the proposed methodology is included. These programs are tools for the practical application of the techniques presented in this thesis.

An integrated approach for dealing with the different aspects of non-stationarity is developed on the basis of the decision of local stationarity. This approach exploits the idea of distance weighted statistics obtained from the available samples. Smoothly changing local means and variances are able to reflect tendencies in the attribute, while locally weighted measures of spatial correlation can adapt to local changes in the anisotropy. Local Gaussian transformations facilitate Gaussian based spatial prediction techniques taking into account local changes in the mean, the variance and the histogram shape. The automatic model fitting of the local measures of correlation produce locally changing parameters of spatial continuity.

The current estimation and simulation algorithms are modified to work under the locally stationary decision by allowing them to update the required parameters at every location. The resulting models are richer in spatial features and they reflect better the real spatial distribution of the attribute.

Chapter 2
Theoretical Background

The decision of modelling the spatial distribution of an attribute deterministically or stochastically depends on the degree of uncertainty in the studied phenomena. The characterization of uncertainty would be unnecessary if the values of an attribute are known at every location. This may be true in certain cases, where the spatial distribution of an attribute can be derived with great precision from a physical law (Isaaks & Srivastava 1989, pp.195-200). However, the complexity of phenomena studied in the earth sciences makes it difficult to derive such laws and their initial and boundary conditions for different geological settings. Moreover, the processes that controlled the spatial distribution of the attributes are not completely known and the attribute values are affected by minor fluctuations in boundary conditions (Christakos et al. 2001, p.24; Isaaks & Srivastava 1989, p.197). Finally, samples are widely spaced in relation of the volume of study and they contain inevitable measurement errors (Chilès & Delfiner 1999, pp.1-2). To account for these sources of uncertainty in the spatial prediction of a geological attribute, a probabilistic approach is required (Isaaks & Srivastava 1989, pp.200-202; Christakos 2005, p.1). Thus, a probabilistic framework for modelling in the earth sciences is carried out by means of Random Variables and Random Functions.

This chapter begins with an overview of the Random Function model concept and the types of stationarity considered by classical geostatistics. The process of probabilistic spatial modelling under a Random Function framework based on the standard stationarity decisions is reviewed. This process covers the choice of statistical homogeneous domains, the inference of the Random Function statistics from data within these domains and the spatial prediction at unsampled locations. The most common geostatistical techniques for estimation and simulation are briefly presented. The limitations of these techniques in face of realistic non-stationarity are discussed. This chapter finishes with a brief overview of the recent research in non-stationary geostatistics.

2.1. The Random Function Model

The uncertainty at an unsampled location **u** is modelled by a Random Variable (RV) $Z(\mathbf{u})$. The probability of the RV taking particular outcomes within a range of possible values can be characterized by its cumulative distribution function (cdf) (Goovaerts 1997, pp.63,64):

$$F(\mathbf{u}; z) = \text{Prob}\{Z(\mathbf{u}) < z\} \tag{2.1}$$

The RVs at different locations are often spatially dependent. An ensemble of spatially correlated RVs is called a Random Function (RF) or Random Field (Christakos 2005, p.5). For K locations $\mathbf{u}_k$, k=1,...,K, the RF multivariate cdf is defined as (Deutsch & Journel 1998, p.12) :

$$F(\mathbf{u}_1,...,\mathbf{u}_k; z_1,...,z_K) = \text{Prob}\{Z(\mathbf{u}) \le z_1,...,Z(\mathbf{u}_K) \le z_K\} \in [0,1] \tag{2.2}$$

Local randomness and spatial dependence are the two characteristics identified by Matheron for a Regionalized Variable (Matheron 1970, p.5). This is a mathematical conceptualization of the spatial distribution of an attribute, whose value depends of the location **u**. Since the outcomes of the Regionalized Variable, represented by z(**u**), are unknown for most locations, it can only be studied indirectly by probabilistic methods (Journel & Huijbregts 1978, p.27) . The set of values that $Z(\mathbf{u})$ takes in a domain can be regarded as one realization within a range of possible outcomes of the RF (Journel & Huijbregts 1978, p.30).

A probability distribution model must be chosen for the RF. Due to its mathematical simplicity and flexibility, a common choice is the Gaussian distribution (Deutsch & Journel 1998, p.12). Although other probability distribution models can be used (Diggle & Ribeiro 2007; Emery & Kremer 2008), in this work the focus is on Gaussian Random Fields.

A Gaussian distributed RV, denoted by $Y(\mathbf{u})$ in order to differentiate it from the original variable $Z(\mathbf{u})$, has a univariate distribution function completely defined by its mean m_Y and variance σ_Y^2. Its univariate probability density function (pdf) is given by the well known expression:

$$g_{m_Y,\sigma_Y}(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-[(y-m_Y)/\sigma_Y]^2/2} \tag{2.3}$$

Where y, $-\infty < y < \infty$, is a realization of $Y(\mathbf{u})$. If a RF is Gaussian, the relationship between two or more of its constituent RVs is described by a multivariate Gaussian pdf.

For K RVs in the vector $\mathbf{Y}' = [Y(\mathbf{u}_1),...,Y(\mathbf{u}_n)]$, this is expressed as (Johnson & Wichern 2007, p.150):

$$g_{\mathbf{m},\boldsymbol{\Sigma}}(\mathbf{y}) = \frac{1}{(2\pi)^{k/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{(\mathbf{y}-\mathbf{m})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{m})}{2}\right] \tag{2.4}$$

where $\mathbf{y}' = [y(\mathbf{u}_1),...,y(\mathbf{u}_n)]$, with $-\infty < y(\mathbf{u}_\alpha) < \infty$, $\alpha = 1,...n$, is the vector containing the realizations of the n RVs, $\boldsymbol{\Sigma}$ is a positive definite $(n \times n)$ variance-covariance matrix, and $\mathbf{m}$ is the expected value $(n \times 1)$ vector. A RF with multivariate Gaussian distribution has some key properties that can be summarized as (Deutsch & Journel 1998, pp.139-140; Johnson & Wichern 2007, pp.156-167): (1) taking subsets of the RF or linearly combining its RV components results in new (multivariate) Gaussian distributions; (2) if the covariance is zero between two RV components they are mutually independent; and (3) when a subset of the RF is conditioned by realizations of another subset, the resulting distribution is also (multivariate) Gaussian.

In classical Geostatistics, the bivariate distribution is of special interest. For two standard Gaussian RVs, $Y(\mathbf{u})$ and $Y(\mathbf{u}+\mathbf{h})$, separated by a vector $\mathbf{h}$, the bivariate, or 2-point, cdf is defined by the covariance function $C_Y(\mathbf{h})$ and evaluated numerically (Goovaerts 1997, p.265; Deutsch & Journel 1998, p.142).

Since the values of a geological attribute very seldom follow a Gaussian distribution, a normal score transformation is usually performed to conform the variable to this model (Deutsch & Journel 1998, p.141). This univariate transformation does not assure that the bivariate distribution will be bi-Gaussian; however, this is frequently assumed (Chilès & Delfiner 1999, p.17). This assumption is allowed if there is no evidence that the transformed distributions violate the (bivariate) Gaussian distribution properties specified above (Deutsch & Journel 1998, p.144).

Defining the RF probability distribution and its summary statistics within a domain of study corresponds to the decisions of stationarity. These are presented in the next section.

2.2. Stationarity

In classical statistics, the probability distribution of a RV can be approximated by the statistics calculated from a number of repeated observations (Isaaks & Srivastava 1989,

pp.206-208). In a geological framework, where the attribute values can be considered invariant in time, repeating observations at the same location would provide information only about the sampling and measurement error distribution. Besides being impractical, such measurements would not be useful for inferring the univariate RV distribution and the spatial correlation between the RV components of the RF. Instead, the required observations are collected from the samples taken at different locations within a region or domain, *D*, assumed statistically homogeneous (Myers 1989; Journel & Huijbregts 1978, p.30). This choice of the population of samples is one aspect of stationarity. A related aspect is the invariance by translation of the RF multivariate distribution and the parameters inferred from these samples. This invariance by translation is required for statistical inference at locations where there is no direct information about its true value (Journel 1989, p.24). Depending on the parameters that are deemed invariant in space, the stationarity decision is called Strict, Second Order or Intrinsic.

2.2.1. Strict Stationarity

Under this form of stationarity the multivariate probability distribution (2.2), is assumed invariant under the translation by any vector **h** (Journel & Huijbregts 1978, p.30). Therefore, the lower order distributions and all its parameters are also invariant by translation. This can be written as follows (Goovaerts 1997, p.70):

$$F(\mathbf{u}_1,...,\mathbf{u}_k;z_1,...,z_K) = F(\mathbf{u}_1+\mathbf{h},...,\mathbf{u}_k+\mathbf{h};z_1,...,z_K) = F(z_1,...,z_K) \quad \forall\ \mathbf{u}_i,\mathbf{u}_i+\mathbf{h}\in D \quad k=1,...,K \tag{2.5}$$

In classical Geostatistics the concern is mostly on the bivariate form of the RF cdf. Thus, the hypothesis of strict stationarity can be limited to:

$$Prob\{Z(\mathbf{u})\le z_1, Z(\mathbf{u}+\mathbf{h})\le z_2\} = Prob\{Z(\mathbf{u}')\le z_1, Z(\mathbf{u}'+\mathbf{h})\le z_2\} \quad \forall\ \mathbf{u},\mathbf{u}',\ \mathbf{u}+\mathbf{h},\ \mathbf{u}'+\mathbf{h}\in D \tag{2.6}$$

Shifting the vector translation **h** from the second term of (2.1) to both terms in (2.3) is made on purpose. This allows presenting the idea that under the decision of strict stationarity the RF bivariate distribution, and its parameters, depends only on **h**, and not on the location **u**. This is the basis of Second Order Stationarity.

2.2.2. Second Order Stationarity

A milder form of stationarity is defined by not taking any assumption about the spatial invariance of the multivariate distribution, but only of the mean and the covariance. The first is considered constant within the domain and the second is a function of **h** (Chilès & Delfiner 1999, p.16):

$$\begin{cases} E\{Z(\mathbf{u})\} = E\{Z(\mathbf{u}+\mathbf{h})\} = m \\ E\{[Z(\mathbf{u})-m][Z(\mathbf{u}+\mathbf{h})-m]\} = C(\mathbf{h}) \end{cases} \quad \forall \mathbf{u}, \mathbf{u}+\mathbf{h} \in D \tag{2.7}$$

This form of stationarity implies that the variance is constant (Myers 1989), since:

$$C(0) = E\{[Z(\mathbf{u})-m]^2\} = Var\{Z(\mathbf{u})\} = \sigma^2, \qquad \forall \mathbf{u} \in D \tag{2.8}$$

Therefore, the semivariogram, $\gamma(\mathbf{h})$, and the correlogram, $\rho(\mathbf{h})$ only depend on **h**. Thus, the following relationships between these measures of spatial continuity can be established (Journel & Huijbregts 1978, pp.32-33):

$$\gamma(\mathbf{h}) = \frac{1}{2} E\{[Z(\mathbf{u})-Z(\mathbf{u}+\mathbf{h})]^2\} = C(0) - C(\mathbf{h}), \qquad \forall \mathbf{u}, \mathbf{u}+\mathbf{h} \in D \tag{2.9}$$

$$\rho(\mathbf{h}) = \frac{C(\mathbf{h})}{C(0)} = 1 - \frac{\gamma(\mathbf{h})}{C(0)} \tag{2.10}$$

If a Gaussian model is adopted for the RF multivariate cdf, second order stationarity is equivalent to strict stationarity; however there are cases where strict stationarity implies second order stationarity (Myers 1989). The reverse, second order stationarity implies strict stationarity, is possible only if the first and second order moments exist and are finite. For Earth Science attributes the mean and the variance are normally assumed existent and finite. Figure 2-1 shows a 1D Gaussian process that is both strict and second order stationary.

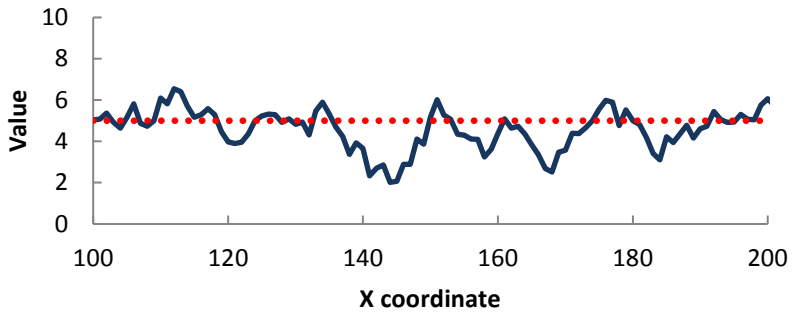

Figure 2-1: Strict and Second Order Gaussian stationary process.

2.2.3. Intrinsic Stationarity

In the intrinsic form, the stationarity of the parameters of the RF Z(**u**) is replaced by the stationarity of the increments $\{Z(\mathbf{u}+\mathbf{h})-Z(\mathbf{u})\}$ (Myers 1989). Thus, the intrinsic stationarity decision is expressed as (Matheron 1969, p.42; Chilès & Delfiner 1999, pp.17,31):

$$\begin{cases} E\{Z(\mathbf{u}+\mathbf{h})-Z(\mathbf{u})\} = m(\mathbf{h}) \\ Var\{Z(\mathbf{u}+\mathbf{h})-Z(\mathbf{u})\} = 2\gamma(\mathbf{h}) \end{cases} \qquad (2.11)$$
$$\forall \mathbf{u}, \mathbf{u}+\mathbf{h} \in D$$

The drift $m(\mathbf{h})$ is a linear function of the vector $\mathbf{h}=(h_1,...,h_n)$ with a gradient vector $\mathbf{a}=(a_1,...,a_n)$. A second order stationary RF is also an intrinsic stationary RF with $m(\mathbf{h}) = 0$. A simple form of the drift is $m(\mathbf{h}) = a_1 h_1$, which correspond to a linearly changing mean in the direction of h_1, such as: $m(\mathbf{u}) = a_0 + a_1\mathbf{u}$. This is illustrated in Figure 2-2. The stationarity of the increments can be extended to higher orders; in this case it is denoted k-order intrinsic stationarity (Matheron 1973). This allows modelling both a non-stationary mean and a non-stationary variance (Chilès & Delfiner 1999, p.247).

Another way to incorporating a non-stationary definition of the mean is to decompose the RF by a local mean $m(\mathbf{u})$ plus the residual $R(\mathbf{u})$ (Myers 1989).

$$Z(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u}) \qquad (2.12)$$

Where $R(\mathbf{u})$ is a intrinsic stationary RF with $E\{R(\mathbf{u})\}=0$. While $m(\mathbf{u})$ can be modelled parametrically or estimated from available data.

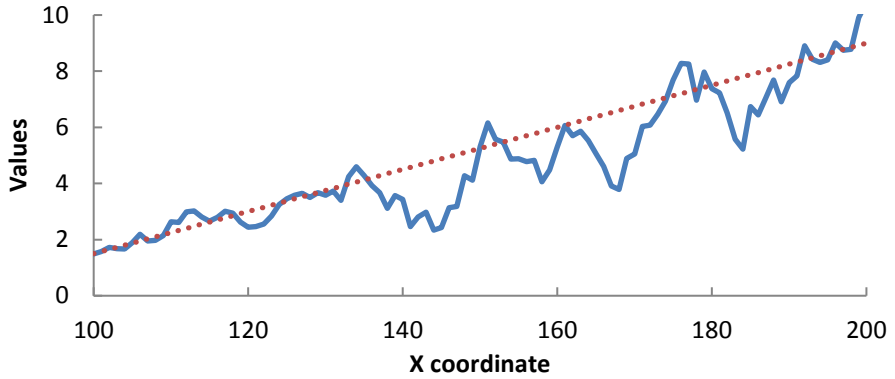

Figure 2-2: A simple 1D example of an intrinsic stationary Gaussian process with a linear drift.

2.3. Spatial Modelling

Under the decisions of stationarity described above, the standard methodology for geostatistical modelling of the spatial distribution of a continuous attribute is summarized by (McLennan 2007): (1) identification of domains that can be considered homogeneous and modelling of the geometry and nature of the boundaries between domains, (2) modelling of trends if deemed necessary, (3) inference of the RFs parameters from data within each domain, and (4) spatial prediction. Although the methodologies presented in this thesis can be applied to the stage of modelling the geometry of the domains boundaries, the focus is on the inference of the local RF's distribution parameters and the spatial prediction using them. The four stages of numerical modelling of a continuous geological attribute are briefly explained below.

2.3.1. Identification of Domains and Boundary Modelling

Since the RF distribution and its parameters are inferred from samples taken at different locations it is reasonable to select those samples from a region deemed statistically and geologically homogeneous. This subdivision is performed mainly on the basis of geological knowledge. If the geological knowledge is not enough to choose the domains, different sub-populations can be statistically compared in order to decide if they can be merged or should be kept separate. A common practice in the mining industry is to define domains based on grade cut-offs, however, this practice may exacerbate estimation errors and introduce bias in the resource estimates (Emery & Ortiz 2005).

Boundaries between domains delimit the zones in the numerical model where a stationary RF has validity. The boundary modelling stage has two aspects, first, the modelling of the geometry of the limits between domains, and second, the definition of the nature of the transitions between those domain.

Boundary Geometry Modelling

Boundary modelling can be performed by various deterministic or stochastic methods. A popular deterministic boundary modelling method consists in the wireframing between sections of interpreted boundary contours (Houlding 2000, pp.60-71). Deterministic surface interpolation techniques based on radial basis functions are becoming common (Cowan et al. 2002). These deterministic methods do not account for the uncertainty in

the geometry of the boundaries. Uncertainty is handled by drawing alternative interpretations of the domains geometry or by changing the parameters of modelling within a range of plausible values (Bárdossy & Fodor 2001).

Probabilistic boundary modelling techniques consider multiple possible geometries providing a measure of geological uncertainty. Some of the most used methods are categorical indicator sequential simulation (Rossi 2004), indicator p-field simulation (Srivastava 2005), object based simulation (Deutsch 2002, pp.223-244), and truncated Gaussian simulation (Emery 2007a; Langlais et al. 2008; Riquelme Tapia et al. 2008).

Definition of the Nature of Boundaries and their Modelling

The transition of the spatial dependence between the RVs on each side of a boundary can be a completely seamless transition, an abrupt discontinuity, or somewhere between these extremes. If the boundary allows a smooth transition of the attribute's spatial continuity between domains it is denoted as a *soft* boundary. If an abrupt discontinuity in the spatial continuity is defined along the boundary, this is referred as *hard* (Ortiz & Emery 2006). The soft or hard nature of a boundary is defined on the basis of the geological knowledge of the properties of adjacent domains and with the help of contact analysis techniques. These techniques can be divided on two groups: those based on the local expected value, and those based on measures of spatial correlation

Contact analysis consists in the analysis of the local mean of the attribute values in a one-dimensional space defined by the distance of the samples to the boundary. If the local mean changes gradually from one side to another of the boundary, this would indicate a soft transition. An abrupt change would indicate a hard boundary (McLennan 2007, pp.5-2, 5-3).

Alternatively, measures of spatial correlation such as the cross-variogram can be used to identify the presence or absence of spatial continuity across the boundary (Ortiz & Emery 2006; McLennan 2007). Thus, if spatial correlation is observed between values on each side of the boundary, it would indicate a soft boundary.

2.3.2. Trend Modeling

Once the domains and their boundaries have been defined and modelled, the attribute spatial distribution may show a large scale tendency, or trend (Chilès & Delfiner 1999,

p.165). A trend can be identified from geological knowledge or directly from data (Deutsch 2002, p.180). Unless the trend can be described by a physical law or by knowledge of the underlying process (Christakos et al. 2001, pp.33-36), the exact shape remains mostly unknown and thus, subject to uncertainty. This is usually the case in geological applications due to the complexity of the phenomena studied. Despite its uncertainty, the trend is usually modelled as a deterministic drift (Chilès & Delfiner 1999, p.233). The decision to model the trend deterministically allows separating the complex and highly uncertain local fluctuations from a simpler and more continuous large scale tendency. The RF is decomposed into a deterministic drift $m(\mathbf{u})$, usually equivalent to the local expectation of the RF, plus the stochastic residual $R(\mathbf{u})$ as presented in Equation (2.12). Therefore, the remaining uncertainty associated to the lack of knowledge of the trend is assimilated to the stochastic part of the RF. This decision comes with an unavoidable degree of subjectivity, in the sense that the amount of spatial variability attributed to each component of the RF cannot be determined uniquely (Cressie 1986). Thus, the separation of the RF $Z(\mathbf{u})$ into deterministic and stochastic components is defined by the practitioner and is dependent of the scale of modelling and the available data (Chilès & Delfiner 1999, p.165; Deutsch 2002, pp.179-180)

Methodologies for modelling the trend deterministically include hand contouring, moving window averages, distance based interpolation, and kriging (Deutsch 2002, p.182; McLennan 2007). From a practical point of view it is desirable that any technique used for trend modelling be relatively simple, since it is the large scale variability that should be reproduced. The geological knowledge and the data rarely justify a highly variable trend. Despite this smoothness, the trend model must be consistent with the available geological knowledge (McLennan 2007)

2.3.3. Inference of the RF Distribution Parameters

A decision of stationarity allows inference of the RF's global probabilistic (multivariate) distribution, and its parameters. If the RF $Z(\mathbf{u})$ has been decomposed in a deterministic drift $m(\mathbf{u})$ and a probabilistic residual $R(\mathbf{u})$ the focus is on the inference of the parameters of the residual RF. In the following notation, the parameters to be estimated refer to the whole RF $Z(\mathbf{u})$, unless otherwise specified.

Available data can be used in the inference of the RF distribution and its parameters (Goovaerts 1997, p.75). Sampling may be concentrated in areas deemed interesting

because of their high values or in areas where accessibility is not limited by logistics and other practical reasons (Borradaile 2003, p.2). This preferential sampling may lead to parameters more representative of the densely sampled areas than of the entire domain. This is translated to a bias when estimating the RF distribution. Spatial declustering techniques are used to remove this bias, they include: polygonal declustering (Isaaks & Srivastava 1989, pp.238-241; Goovaerts 1997, pp.79-80; Deutsch 2002, pp.50-51), cell declustering , (Isaaks & Srivastava 1989, pp.241-243; Deutsch 1989; Goovaerts 1997, p.81; Deutsch & Journel 1998, p.213) and global kriging weights declustering (Deutsch 1989).

By incorporating the declustering weights obtained by any of these methods, the univariate RF cdf can be inferred from n weighted observations $z(\mathbf{u}_\alpha)$ at different locations (α=1,...,n) within the domain, and for K different thresholds by (Goovaerts 1997, p.81):

$$\hat{F}(z_k) = Prob\{Z(\mathbf{u}) \le z_k\} \in \sum_{\alpha=1}^{n} w_\alpha \cdot I(\mathbf{u}_\alpha; z_k) \quad \in [0,1],\ k = 1,...,K \tag{2.13}$$

The superscript ^ indicates that this is a sample statistic. The weights w_α are the declustering weights assigned to each sample, such as $\sum_{\alpha=1}^{n} w_\alpha = 1$, and $I(\mathbf{u}_\alpha; z_k)$ is the binary indicator function that transforms each data value according if it exceeds or not a threshold z_k, k=1,...,K (Journel 1989, p.22):

$$I(\mathbf{u}_\alpha; z) = \begin{cases} 1, & \text{if } z(\mathbf{u}_\alpha) \le z_k \\ 0, & \text{otherwise} \end{cases} \qquad \forall \mathbf{u}_\alpha \in D \tag{2.14}$$

The mean and the variance are estimated from the observations that are considered to be a sample of a realization of the RF, as (Goovaerts 1997, pp.81-82):

$$\hat{m} = \sum_{\alpha=1}^{n} w_\alpha z(\mathbf{u}_\alpha) \ , \tag{2.15}$$

and

$$\hat{\sigma}^2 = \sum_{\alpha=1}^{n} w_\alpha \left[z(\mathbf{u}_\alpha) - \hat{m} \right]^2 , \tag{2.16}$$

Among the multiple 2-point statistics available (Deutsch & Journel 1998, pp.43-46), the most relevant for this thesis are the sample variogram, $\hat{\gamma}(\mathbf{h})$, covariance $\hat{C}(\mathbf{h})$ and

correlogram, $\hat{\rho}(\mathbf{h})$. The sample or experimental semivariogram is calculated by (David 1977, p.74):

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} \left[z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h}) \right]^2 \tag{2.17}$$

Where $N(\mathbf{h})$ denotes the number of sample pairs approximately separated by the vector **h**. Since sampling locations often follow irregular patterns, it is necessary to consider tolerances to the magnitude and direction angles of this vector in order to include enough number of pairs (Deutsch & Journel 1998, pp.47-50). The sample covariance is calculated by (Goovaerts 1997, p.86):

$$\hat{C}(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} z(\mathbf{u}_\alpha) \cdot z(\mathbf{u}_\alpha + \mathbf{h}) - \hat{m}_{-\mathbf{h}} \cdot \hat{m}_{+\mathbf{h}} \tag{2.18}$$

with

$$\begin{aligned} \hat{m}_{-\mathbf{h}} &= \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} z(\mathbf{u}_\alpha) \ , \\ \hat{m}_{+\mathbf{h}} &= \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} z(\mathbf{u}_\alpha + \mathbf{h}) \end{aligned} \tag{2.19}$$

In practice, these measures of spatial correlation are calculated for different directions and distances. Most geological processes have some degree of spatial continuity. That is, the values observed at small distances tend to be more similar than those observed at large distances (Tobler 1970). Thus, in presence of spatial correlation the experimental variogram increases as $|\mathbf{h}|$ increases, while the covariance decreases as $|\mathbf{h}|$ increases. The directional variograms and covariances carry information about different aspects of the spatial continuity of the attribute. These include: unstructured and short scale random variation, known as *nugget effect*, directions where the decrease of correlation with distance is less marked, known as *geometric anisotropy*, a direction with systematically lower long scale variability, known as *zonal anisotropy*, and geologically induced cyclicity (Gringarten & Deutsch 2001).

The experimental measures of continuity in different directions are jointly fitted by continuous functions defined as variogram models. The reason for doing so is mainly the necessity of defining the spatial correlation for all distances and orientations. The process of fitting the experimental variograms is helped by geological knowledge and allows filtering artifacts and other fluctuations (Goovaerts 1997, p.87; Gringarten & Deutsch 2001). Only allowable covariance and variogram models can be used. These covariances

share the mathematical property of *positive definiteness*, that is, under the assumption of second order stationarity the covariance $C(\mathbf{h})$ in the right side of the Expression 2.9 must be such that any linear combination of RVs has a positive variance (Armstrong & Jabin 1981; Goovaerts 1997, p.87):

$$\mathrm{Var}\left\{\sum_{\alpha=1}^{n}\lambda_\alpha Z(\mathbf{u}_\alpha)\right\}=\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\lambda_\alpha\lambda_\beta C(\mathbf{u}_\alpha-\mathbf{u}_\beta)\geq 0 \qquad (2.20)$$

The most used allowable variogram models linked to covariances include the *spherical*, *exponential* and *Gaussian* (Chilès & Delfiner 1999, pp.81-85). These are bounded functions in the sense that they reach a maximum value called the *sill.* These models have a covariance counterpart that can be obtained from Expression 2.9.

2.3.4. Spatial Prediction

The idea of spatial prediction in Geostatistics consists in locally conditioning the RF global cdf to neighbouring data at an unsampled location $\mathbf{u}$ (Journel 1989, p.22; Goovaerts 2000). The local conditional cumulative density function (ccdf) is expressed as:

$$F(\mathbf{u};z_k\mid n(\mathbf{u}))=\mathrm{Prob}\{Z(\mathbf{u})<z_k\mid n(\mathbf{u})\}\quad k=1,...,K \qquad (2.21)$$

Where $n(\mathbf{u})$ is the number of samples surrounding the prediction location $\mathbf{u}$. It may range from a quantity limited by a local neighbourhood centred at $\mathbf{u}$ to all the samples within the domain. Samples at locations where the variogram model indicates a higher spatial dependency in relation to the prediction location get more weight in the construction of the local ccdf. Thus, beginning from a global prior cdf corresponding to a RF model within a domain, the aim is to obtain the posterior ccdf at each unsampled location (Goovaerts 1997, p.264). This is accomplished using the punctual information provided by data, the information on spatial dependence provided by the variogram model, and the information on large scale trends provided by the trend model. Background information, such as the knowledge of the geological setting, can be incorporated in the variogram and trend models, as well as during the domain identification and boundary modelling stages. The process of spatial prediction, from inference of the global prior cdf and ending with local ccdfs, is schematically illustrated in Figure 2-3. That figure shows a 1-D set of sample values $z(\mathbf{u}_\alpha)$, α=1, 2, 4, 7, 8, and 10, with the corresponding global pdf at the extreme right. Commonly, a trend $m(\mathbf{u})$ is

fitted to the data values, and a pdf of the RV R(**u**) is inferred from the residual values. A normal score transformation of the R(**u**) cdf may be performed at this point. The variogram is also calculated from the (transformed) residuals. The fitted variogram model is used for obtaining the residual estimates $r^*(\mathbf{u}_\alpha)$ and the conditional pdfs for α=3, 5, 6, 9 and 11. A backtransformation of the ccdfs for R(**u**) is performed at this point if required. The estimates and the ccdf in original units, $z^*(\mathbf{u}_\alpha)$, can be obtained by restoring the trend. The most notable drawbacks of this common approach and some alternatives to it are discussed in Subsections 2.4.1 and 2.5.2.

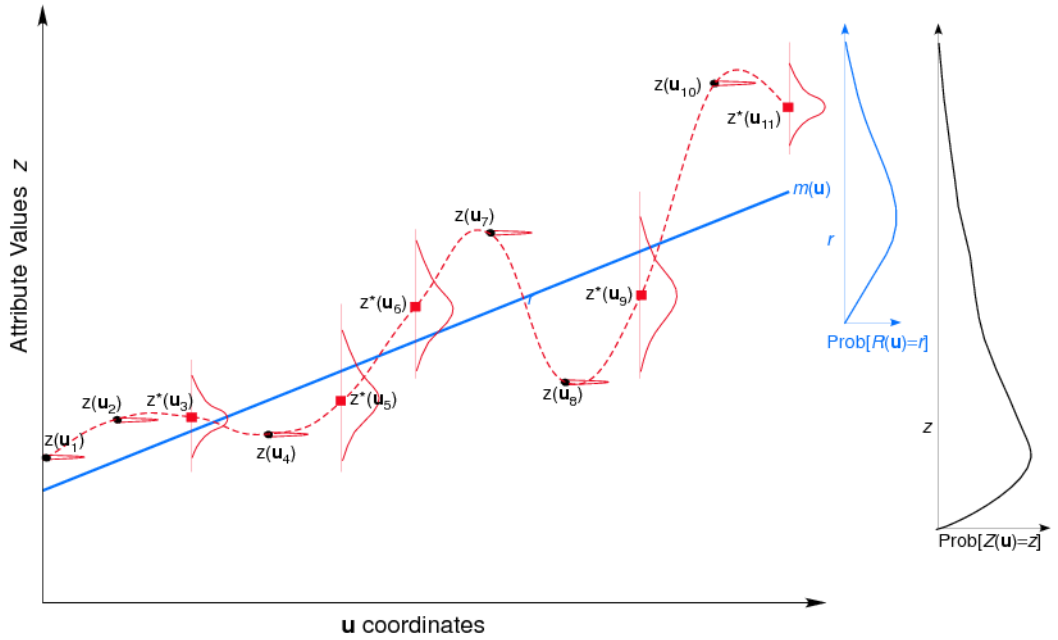

Figure 2-3: Schematic 1D representation of the Geostatistical spatial prediction process. The black curve at the right represents the global prior pdf of the RF Z(u) obtained from scattered values (black dots). The blue curve at the right represents the pdf of residuals after trend modelling. The trend model is represented as a blue solid line. The red curves represent the posterior conditional pdfs at each estimated location (red squares). The dashed red line represents the exhaustive estimation of the posterior local mean.

There are a number of techniques for spatial prediction. These can be classified as estimation or simulation techniques. Roughly speaking, the focus of estimation techniques is on the single best estimate of the ccdf center of mass, while simulation techniques are focused on providing alternate realizations within the ccdfs spread. The most relevant prediction techniques for the present work are described in the next section.

Spatial prediction is commonly performed on a regular grid discretizing the area or volume of study. Usually, for estimation techniques, the grid cell size is determined relative to the sample spacing (David 1977, p.283) and/or the goals of the attribute

numerical model (Deutsch 2002, p.84). Whereas, for simulation techniques, the cell size is constrained by the consistency requirements between the samples and predicted values supports (Journel & Kyriakidis 2004, p.24). The grid is intersected with the domain boundaries and the cell centroids coded according the domain they belong. During the spatial prediction process, the grid centroids are used to store the parameters of the posterior ccdfs (Houlding 2000, p.32).

2.4. Standard Geostatistical Techniques for Spatial Prediction.

The main goal of an estimation technique is to provide the closest approximation to the unknown true values at unsampled locations. This goal of local accuracy comes at the expense of reduced spatial variability, which is reflected in smooth estimated maps and a reduced variance in the histogram of estimates (Goovaerts 1997, pp.369-370). The histogram of estimates underestimates the frequency of large values while overestimating the frequency of small values, this is known as Conditional Bias (Goovaerts 1997, p.182). Simulation techniques try to reproduce the spatial variability and the probability distribution informed by the data. Although simulation techniques reproduce the data values, they provide less locally accurate predictions than estimation (Journel & Kyriakidis 2004, pp.14-15)

In estimation techniques, the criterion of optimality commonly consists in minimizing the mean of the square error between the unknown true values of $Z(\mathbf{u})$, and its estimator, $Z^*(\mathbf{u})$ (Journel 1989, p.27; Goovaerts 2000):

$$E\left\{\left[e(\mathbf{u})\right]^2\right\} = E\left\{\left[Z(\mathbf{u}) - Z^*(\mathbf{u})\right]^2\right\} \tag{2.22}$$

Other optimality criteria for estimation could be chosen, such as the mean of the minimum absolute error (Christakos 2005, pp.342-343). However, only the least squares criterion assures the estimator is unbiased, this is $E\left\{Z^*(\mathbf{u})\right\} = E\left\{Z(\mathbf{u})\right\}$ (Journel 1989, p.27; Christakos 2005, p.344).

The goal of simulation techniques is not the minimization of a measure of the prediction error, but the reproduction of the global statistics informed by the global histogram and variogram model.

Estimation optimality criterion is satisfied by a single estimated value at each location. The reproduction of the global statistics in simulation, however, can be satisfied by a range of possible realizations, at least approximately within ergodic fluctuations (Goovaerts 1997, pp.426-429). While the estimation optimality criteria are local, the simulation goodness criteria are global in the sense that they involve all simulated locations simultaneously.

2.4.1. Estimation Techniques

In linear estimation, the attribute value at an unsampled location **u** is obtained as a linear combination of *n* sampled values at locations $\mathbf{u}_\alpha$, $\alpha=1,...,n$. The number *n* may correspond to all the samples within the domain or may be constrained to those that fall within a neighbourhood centred at **u**. In such case, it is denoted $n(\mathbf{u})$. Thus, the Kriging estimator, $Z^*(\mathbf{u})$, is a RV built by the linear combination of $n(\mathbf{u})$ RVs $Z(\mathbf{u}_\alpha)$, $\alpha=1,...,n(\mathbf{u})$ (Journel 1989, p.10; Goovaerts 1997, p.126; Deutsch & Journel 1998, p.64). The general form of this estimator is given by:

$$\left[Z^*(\mathbf{u}) - m(\mathbf{u})\right] = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})\left[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)\right] \tag{2.23}$$

Where $m(\mathbf{u})$ and $m(\mathbf{u}_\alpha)$ are the expected values of $Z(\mathbf{u})$ and $Z(\mathbf{u}_\alpha)$, respectively, and $\lambda_\alpha(\mathbf{u})$ are the weights assigned to each data value $z(\mathbf{u}_\alpha)$ considered as realizations of the RVs $Z(\mathbf{u}_\alpha)$. The various forms of kriging differ according the form of the RV expected values (Goovaerts 1997, p.126). For all kriging forms, the weights $\lambda_\alpha(\mathbf{u})$ are such that they minimize the estimation variance (Matheron 1969, pp.44-45):

$$\sigma_E^2(\mathbf{u}) = Var\{Z^*(\mathbf{u}) - Z(\mathbf{u})\} \tag{2.24}$$

The unbiasedness condition is fulfilled for all forms of kriging:

$$E\{Z^*(\mathbf{u}) - Z(\mathbf{u})\} = 0 \tag{2.25}$$

For Simple Kriging the fulfillment of this condition is assured by the form of its estimator (Expression 2.29). For other types of kriging, such as Ordinary kriging (Expressions 2.32 and 2.33), the Condition 2.25 must be enforced during the minimization of the estimation variance.

The minimization of the estimation variance is equivalent to the minimization of the mean square error in Equation 2.22. If the RF distribution is modelled as Gaussian, the

best estimator in Gaussian units, $Y^*(\mathbf{u})$, under the minimum square error criterion coincides with the mean of the Gaussian ccdf at each location **u** (Goovaerts 1997, p.276; Chilès & Delfiner 1999, pp.163-164):

$$E\{Y(\mathbf{u})\,|\,(n)\} = Y^*(\mathbf{u}) = \int_{-\infty}^{+\infty} y \cdot g\left(\mathbf{u}; y\,|\,(n)\right)dy \tag{2.26}$$

and the Gaussian ccdf variance $\sigma_E^2(\mathbf{u})$ is equivalent to the estimation variance:

$$E\left\{\left(Y(\mathbf{u}) - Y^*(\mathbf{u})\right)^2\,|\,(n)\right\} = \sigma_E^2(\mathbf{u}) = \int_{-\infty}^{+\infty} \left[y - Y*(\mathbf{u})\right]^2 g\left(\mathbf{u}; y\,|\,(n)\right)dy \tag{2.27}$$

The minimization of the mean square error in order to obtain $Z^*(\mathbf{u})$, in the general case, or $Y^*(\mathbf{u})$ in the Gaussian case, under different assumptions leads to different kriging types. In the kriging types described next, these assumptions are related to the model used for accommodating non-stationarity or data transforms in the RF mean (Goovaerts 1997, p.126).

Simple Kriging

For the stationary form of the Simple Kriging (SK) algorithm, the local mean is assumed known and constant within the entire domain (Deutsch & Journel 1998, p.64)

$$E\{Z(\mathbf{u})\} = m(\mathbf{u}) = m \qquad \forall \mathbf{u} \in D \tag{2.28}$$

In this case, the stationary SK estimator takes the form:

$$Z_{SK}^*(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u})[Z(\mathbf{u}_\alpha)] + \left[1 - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u})\right] m \tag{2.29}$$

This form of kriging is adequate under the decisions of strict and second order stationarity. A locally varying mean can be incorporated in the general form of SK. In this case, the estimator is similar to the Expression 2.23. For both, the stationary and general SK, the weights are obtained from the system (Goovaerts 1997, p.128; Chilès & Delfiner 1999, p.155):

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{SK}(\mathbf{u}) C(\mathbf{u}_\alpha - \mathbf{u}_\beta) = C(\mathbf{u}_\alpha - \mathbf{u}) \qquad \alpha = 1, ..., n(\mathbf{u}) \tag{2.30}$$

The corresponding estimation variance is:

$$\sigma^2_{SK}(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u}) C(\mathbf{u}_\alpha - \mathbf{u}) \tag{2.31}$$

The difference between the stationary and general SK systems is that in the second the RF $Z(\mathbf{u})$ is decomposed in a local mean $m(\mathbf{u})$ plus a residual, $R(\mathbf{u})$, as shown in Expression 2.12. Therefore in the general SK the covariances of $Z(\mathbf{u})$ are replaced by the covariances of residuals, $C_R(\mathbf{h})$ (Goovaerts 1997, p.128).

Ordinary Kriging

Ordinary kriging (OK) bypasses the requirement of a known local or global mean by re-estimating it at each location $\mathbf{u}$ (Deutsch & Journel 1998, p.65). The mean is filtered by the unbiasedness constraint: $\sum_{\beta=1}^{n} \lambda_\beta^{OK}(\mathbf{u}) = 1$. The OK system of equations is given by (Journel & Huijbregts 1978, pp.33-34):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) C(\mathbf{u}_\beta - \mathbf{u}_\alpha) + \mu(\mathbf{u}) = C(\mathbf{u} - \mathbf{u}_\alpha), \ \alpha = 1,...,n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) = 1 \end{cases} \tag{2.32}$$

With $\lambda_\beta^{OK}(\mathbf{u})$ as the OK weights and $\mu(\mathbf{u})$ is the Lagrange parameter required to enforce the unbiasedness constraint. By re-estimating the local mean and assuming it constant within a neighbourhood, OK allows locally restricting the decision of second order stationarity. Due to this increased flexibility, OK is a commonly used geostatistical method. It can also be expressed in terms of variograms (Goovaerts 1997, pp.134-135):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) \gamma(\mathbf{u}_\beta - \mathbf{u}_\alpha) + \mu(\mathbf{u}) = \gamma(\mathbf{u} - \mathbf{u}_\alpha), \ \alpha = 1,...,n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) = 1 \end{cases} \tag{2.33}$$

The weights obtained from the equivalent OK systems expressed in 2.32 and 2.33 lead to the formulation of the OK estimator as (Deutsch & Journel 1998, p.65):

$$Z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) Z(\mathbf{u}_\alpha) \tag{2.34}$$

And the OK estimation variance is given by:

$$\sigma_{OK}^2(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) C(\mathbf{u}_\alpha - \mathbf{u}) - \mu(\mathbf{u}) \tag{2.35}$$

Kriging with a Trend Model

When the mean cannot be obtained with enough precision from data or when it is not appropriate to consider it as invariant even within small search neighbourhoods. The local mean can be provided by a trend model defined as a continuous function of the form (Goovaerts 1997, p.127; Deutsch & Journel 1998, p.69):

$$m(\mathbf{u}) = \sum_{k=0}^{K} a_k(\mathbf{u}) f_k(\mathbf{u}) \tag{2.36}$$

Where $f_k(\mathbf{u})$ are often chosen by the user as polynomial functions of the coordinates, or they can take other forms, such as trigonometric functions. The coefficients $a_k(\mathbf{u})$ are assumed constant within each search neighbourhood. They are unknown and, under the constraints imposed by Kriging with a trend model (KT), they are filtered. The KT system is expressed as (Goovaerts 1997, p.141):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{KT}(\mathbf{u}) C_R(\mathbf{u}_\beta - \mathbf{u}_\alpha) + \sum_{k=0}^{K} \mu_k^{KT}(\mathbf{u}) f_k(\mathbf{u}_\alpha) = C_R(\mathbf{u} - \mathbf{u}_\alpha) \quad \alpha = 1,...,n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{KT}(\mathbf{u}) = 1 \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{KT}(\mathbf{u}) f_k(\mathbf{u}_\beta) = f_k(\mathbf{u}) \quad k = 1,...,K \end{cases} \tag{2.37}$$

Where the Lagrange parameters, $\mu_k^{KT}(\mathbf{u})$, account for the K+1 constraints on the weights. KT can be used under the decision of intrinsic stationarity. In this case, the covariance of the residuals, $C_R(\mathbf{h})$, is required. Since the coefficients $a_k(\mathbf{u})$ remain unknown it is not straightforward to estimate the residuals and their variogram. One option for obtaining the covariance of the residuals is from semivariogram of residuals calculated in directions unaffected by the trend (Goovaerts 1997, p.142). KT also can be used for explicitly estimating the trend components $a_k(\mathbf{u})$ as linear combinations of the original data values (Goovaerts 1997, pp.145-147).

2.4.2. Simulation

A direct and widely used way to build maps that honour the data histogram and spatial correlation is by using Monte Carlo Simulation for drawing random values from correlated conditional distributions at every location**.** For such purpose, the adoption of the Gaussian distribution model in geostatistical simulation of continuous attributes is widespread since it allows a straightforward construction of the correlated ccdfs. This is particularly advantageous for the sequential simulation algorithm explained next.

Sequential Gaussian Simulation

For a multigaussian RF, the conditional expectation and variance is equivalent to the SK estimate and variance, respectively (Chilès & Delfiner 1999, pp.163-164). Due to this, SK is used at each node to obtain the local Gaussian cdf conditioned to surrounding data and previously simulated values. For each node in a randomly generated path, a random value is drawn from its corresponding ccdf and added to the data set. This sequence continues until all nodes have been simulated (Isaaks 1991, pp.15-16)

Including previously simulated values for constructing the subsequent ccdfs assures the preservation of the spatial correlation informed by the global variogram. Different realizations can be obtained by executing the sequence multiple times using different random paths. Using a single random path for all realizations has the advantage that the SK weights need to be calculated only once; however, it may result in similar realizations (Deutsch & Journel 1998, pp.145,154). This is because the single path may enforce artificial spatial continuity when only neighbouring data are considered in the inference of the local ccdfs (Goovaerts 1997, p.379). This problem may avoided by using all previously simulated nodes in the local cdf conditioning, but this would require the handling of excessively large matrices.

Sequential Gaussian simulation (SGS) relies on the assumption of strict stationarity. Although, this algorithm can accommodate a non-stationary trend by using SK with prior local means. Whatever form of kriging is used for obtaining the local ccdf means, the corresponding variance to be used must always be the SK variance (Journel 1980).

2.5. Non-Stationarity

Stationarity is a modelling assumption and not a property of the geological attribute (Myers 1989). This assumption allows grouping the data and spatial prediction; however, the spatial distribution of virtually every attribute studied in the earth sciences exhibits features that can be regarded as dependent on location. For example, contaminant concentrations fade far from the pollution sources, the geometry of geological structures is altered by various depositional and structural processes, and minerals and metals precipitate at different stages of the deposit formation resulting in zoning patterns. Therefore, it is reasonable to incorporate different kinds of non-stationary statistics in the modelling process. Some modelling aspects must be considered before proceeding in this way.

2.5.1. Some Considerations for Non-Stationary Modelling

The main aspects to be considered for incorporating non-stationary statistics in Geostatistical Modelling are the purpose of the model, the information available, and the scale of modelling.

Purpose

Depending on the goal of the model the reproduction of some non-stationary features may be important. For example, in mining, the reproduction of the curvilinear features of the mineralization, such as those that appear in vein or folded deposits, can be important for mine design, production scheduling and ore/waste classification. However, for global resources estimates required in a pre-feasibility stage, the reproduction of non-stationary features in the model may not be necessary. In the oil industry, the correct reproduction of connectivity of a particular facies can be important for evaluating the oil recovery and deciding a well location. While in terrain modelling, the correct reproduction of the continuity of high and low areas may require of non-stationary statistics that are able to adapt to changes in the orientation of valleys and mountain ranges.

Information

When sampling is scarce and no background geological knowledge is available to support the incorporation of non-stationary statistics it may be advisable to proceed under a stationary framework.

Scale

In some cases, the spatial features of an attribute may appear non-stationary within a small area, but in a larger context, they may appear as part of a repetitive pattern that can be modelled as stationary. In other instances, regional variations in the spatial continuity and tendencies in the attribute values may be isolated in sub-regions within which they can be modelled as stationary.

The scale of modelling is closely related to the model purpose. If the goal is to estimate the attribute ccdf on large scale volumes, the reproduction of small scale non-stationary patterns may not be critical.

2.5.2. Types of Non-Stationarity

In classical Geostatistics, non-stationarity normally refers to the mean. However, local variations of other statistics can also be incorporated in Geostatistical modelling. A classification of non-stationarity types based in the statistics involved, along with some examples, are presented next.

Non-Stationarity in the Mean and the Variance

Non-stationarity of the mean may be handled by the intrinsic assumption or by decomposing the RF in a local mean and a residual. A number of techniques for incorporating this kind of non-stationarity exist; some of them have been presented above. For earth sciences attributes, it is common that zones of high values also show higher variability. This translates in a positive correlation between the local mean and local variance, which is known as proportional effect (Matheron 1974; Isaaks & Srivastava 1989, pp.49-50). This positive proportionality between the local mean and variance is typical of positively skewed distributions, while increasing local variances

related to lower local means is observed in negatively skewed distributions (Journel & Huijbregts 1978, pp.186-187).

By transforming the original distribution to a normal distribution, the proportional effect is mitigated (Krige 1978, p.24), since the local fluctuations are equalized thanks to the homoscedasticity of the normal distribution. However, this direct transformation does not account for the trend in the mean. Decomposing the RF in a trend and a residual component, as in Expression 2.12, and subsequently transforming the residuals may result in violations of original attribute's constraints when back-transforming to original units. Using Stepwise Conditional Transformation (Leuangthong 2003) allows to obtain residual transforms with constant mean and variance, and when back transformation is applied, the attribute constrains are respected and the heteroscedasticity restored. This is accomplished by applying normal scores transformation, and back-transformation, of the residuals conditioned on the probability class of the local mean component (Leuangthong & Deutsch 2004)

Non-Stationarity in the Spatial Continuity

Changes in the orientation of the spatial continuity are induced by diverse structural and depositional processes such as folding, shearing, and meandering. Moreover, the magnitude of the anisotropy of the spatial distribution of an attribute may vary locally. Most of the non-stationary techniques described in the next section are aimed to incorporate local variations in the anisotropy orientation and ratio. However, other aspects of the spatial continuity models, such as the nugget effect and variogram shape, may also be considered as locally varying.

The discontinuity at the origin of the variogram, known as nugget effect, can be attributed to two sources: (1) sampling related errors, and (2) the geological variability between samples at close to zero separation (Deutsch 2002, p.116; Platten & Dominy 2001). While the first component may be independent of the location, the geological short scale variability may be location-dependent. In stockwork deposits, for example, local changes in the density and the pattern of mineralized veinlets could induce local variations in the nugget effect.

A locally changing variogram shape may be required, for example, in terrain modelling, where local changes in surface roughness are observed (Shepard et al. 2001; Lloyd & Atkinson 2002). Thus, a variogram model with high continuity at short

distances, such as the Gaussian model, can be used for modelling the smooth topography of hills, while an exponential model may reflect better the roughness of mountain areas.

Non-Stationarity in the Bivariate Correlation

Locally changing correlation patterns between variables have been incorporated in the multivariate modelling of a nickel laterite deposit (Lyall & Deutsch 2000). Similarly, when modelling a primary variable conditioned by a correlated secondary variable, such as porosity and seismic amplitude in a heavy oil reservoir (Ren 2007), or between field and satellite measurements of tree species concentrations (Pereira et al. 2000).

2.6. Current Approaches for Non-stationary Geostatistics.

Recently, particular attention has been paid to the non-stationarity of the variogram and covariance functions. The current methodologies for dealing with non-stationary spatial correlation can be grouped into local and global approaches. Local approaches restrict the stationarity of the RF locally, while global approaches deal with the spatial correlation non-stationarity at the level of the entire domain (Schabenberger & Gotway 2005, pp.421-430)

2.6.1. Local Approaches

Among the local approaches for dealing with the non-stationarity of the spatial correlation are the moving window method (Haas 1990a; Haas 1990b), segment-based variograms (Atkinson & Lloyd 2007), convolution methods (Higdon 1998; Higdon et al. 1998) and weighted stationary processes (Fuentes 2001).

Moving Windows and Segment-Based Variograms

The moving window approach calculate the experimental local variograms or covariances with only the samples that fall within a window centred on the location to be estimated. At each location, the window size is adjusted automatically in order to contain a minimum number of samples necessary for calculating experimental semivariograms that allow the convergence of its fully automatic fitting. Kriging is then applied using the

fitted local variogram model and considering only those samples within the moving window.

The required automatic fitting of local variograms at each estimated location may increase considerably the computer requirements. The main advantage of this method is the local variability information reflected by the estimation variance (Walter et al. 2001); however this approach has some important drawbacks. First, the fully automatic fitting does not allow the input of geological knowledge and the supervision of the fitted variogram models by the user. Second, when data is scarce or highly variable, the local variograms may prove difficult to fit reliably. Additionally, artifacts may be produced by including or excluding individual samples from one estimation location to the next one (Schabenberger & Gotway 2005, pp.425-426)

A related methodology based in the segmentation of the spatial variation (Lloyd & Atkinson 2000). Under this approach, the domain is divided in smaller regions, or segments, within which the variogram is assumed stationary. The segmentation may be achieved on the basis of the fractal dimension derived from the slope of the local double-log variograms (Burrough 1981). These local variogram models are obtained in a moving window process similar to the explained above, but using much smaller window sizes (Lloyd & Atkinson 2002)

Convolution Methods

The process convolution-approach proposed by Higdon (1998) decomposes the RF $Z(\mathbf{u})$ as a Gaussian process $Y(\mathbf{u})$ plus an independent error process $\mathrm{E}(\mathbf{u})$:

$$Z(\mathbf{u}) = Y(\mathbf{u}) + \mathrm{E}(\mathbf{u}) \tag{2.38}$$

$Y(\mathbf{u})$ is defined as a convolution of a zero-mean white noise $X(\mathbf{o})$ with a Gaussian convolution kernel $K_{\mathbf{u}}(\mathbf{o})$ centred at $\mathbf{u}$:

$$Y(\mathbf{u}) = \int_D K_{\mathbf{u}}(\mathbf{o}) X(\mathbf{o}) d\mathbf{o} \tag{2.39}$$

The local Gaussian kernel parameters are obtained from local Gaussian variogram models fitted to experimental variograms calculated from data within a neighbourhood. For the sake of efficiency, these local Gaussian variograms, and the kernels derived from them, are obtained only for a few locations. The kernel parameters at other location are obtained by a weighted average of the first few primary kernels. The corresponding covariance is given by:

$$\mathrm{Cov}\{Z(\mathbf{u}), Z(\mathbf{u}+\mathbf{h})\} = \int_D \mathrm{var}[X(\mathbf{o})] K_{\mathbf{u}}(\mathbf{o}, \mathrm{T}(\mathbf{u})) K_{\mathbf{u}+\mathbf{h}}(\mathbf{o}, \mathrm{T}(\mathbf{u}+\mathbf{h})) d\mathbf{o} \tag{2.40}$$

Where $\mathrm{T}(\mathbf{u}) = [\sigma_1(\mathbf{u}), \sigma_2(\mathbf{u}), \theta(\mathbf{u})]$ control the local anisotropic shape and orientation of the Gaussian kernel. Consequently, the covariance function can be made non-stationary by changing the kernel parameters at each location.

Weighted Stationary Processes

In this method it is assumed that the non-stationary RF $Z(\mathbf{u})$ can be represented as a linear combination of k stationary RFs $Z_1(\mathbf{u}), ..., Z_k(\mathbf{u})$:

$$Z(\mathbf{u}) = \frac{1}{k} \sum_{i=1}^{k} K_i(\mathbf{u} - \mathbf{o}_i) Z_i(\mathbf{u}) \tag{2.41}$$

The stationary RFs are defined at k partitions of the domain and are uncorrelated outside each partition. This is:

$$Cov\{Z_i(\mathbf{u}), Z_j(\mathbf{u}+\mathbf{h})\} = \begin{cases} C(\mathbf{h}; \boldsymbol{\theta}_i) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{2.42}$$

With $\boldsymbol{\theta}_i$ as the parameters vector of the stationary covariance model defined for each partition i=1,..,k. The kernels $K(\mathbf{u} - \mathbf{o}_i)$, are anchored at the centre $\mathbf{o}_i$ of each partition. As these stationary RF are defined locally and are uncorrelated, the non-stationary covariance can be expressed as (Fuentes 2002; Schabenberger & Gotway 2005, p.428):

$$Cov\{Z(\mathbf{u}), Z(\mathbf{u}+\mathbf{h})\} = k^{-1} \sum_{i=1}^{k} C(\mathbf{h}; \boldsymbol{\theta}_i) \cdot K_i(\mathbf{u}) \cdot K_i(\mathbf{u}+\mathbf{h}) \tag{2.43}$$

The global non-stationary covariance is constructed by weighting the local covariances by a Gaussian kernel as a function of the distance between $\mathbf{u}$ and the centre of each partition. Different from the convolution method explained above, in this method the kernel parameters are kept constant from one partition to another. Therefore, non-stationarity in the covariance is not derived from the locally changing kernels, but from the differences of the stationary covariances defined at each domain partition. This approach is affected by some of the same issues as the moving window approach, such as the determination of the partitions size and the reliability of the covariances estimated within them when data are scarce.

A different adaptive approach is to use an image analysis gradient algorithm to iteratively and progressively tune the local anisotropy ranges and directions (Stroet &

Snepvangers 2005). Beginning with traditional kriging estimation results and a global search radius, the gradient algorithm is able to identify locally the directions of maximum continuity. These directions are used to modify the anisotropy of a local linear variogram model. Then, at successive iterations shorter scale details in the anisotropy are tuned by decreasing the search radii. This approach appears to work very well in presence of dense sampling. It can be used for the modelling of 2D curvilinear structures of categorical data using indicator kriging, and for estimating continuous variables within these structures. Additionally, no local variogram fitting is required, since the single parameter needed for the linear variogram model is derived from the gradient algorithm. This approach is limited to the use of the linear variogram model and it is not suitable when the local modelling of the nugget effect, the variogram shape and sills is required. Moreover, the gradient algorithm may become unreliable if data are scarce.

2.6.2. Global Approaches

Global approaches for dealing with the non-stationarity of the spatial continuity can be classified as parametric and spatial deformation methodologies.

Parametric Derivation of the Non-Stationary Spatial Correlation Structure

A non-stationary covariance structure can be parametrically modelled using physical laws controlling the spatial dispersion of the attribute. The model parameters are estimated from available data by maximum likelihood. These models usually involve the spatial dispersion of the attribute as controlled by few sources, such as a pollutant emitting focus (Hughes-Oliver et al. 1998). In most applications related to geosciences, the major limitation of this approach is the difficulty of finding a parametric model of the spatial behaviour of a rock property.

Spatial Deformation

In this method originally proposed by Sampson and Guttorp (1992), the coordinate space is transformed by a function f in such a way that the covariance structure becomes stationary and isotropic in the transformed space. The space transformation function can be obtained using multidimensional scaling of the spatial dispersions: δ_{ij}^2. These are

defined as the variance of the differences between sample pair values measured repetitively in time:

$$\delta_{ij}^2 = \text{var}\{Z(\mathbf{u}_i, t) - Z(\mathbf{u}_j, t)\} \tag{2.44}$$

Since the dispersions can only be calculated from a repeated measurement at each sampling location, this approach can be only used for spatiotemporal problems under an assumption of temporal stationarity.

2.7. Discussion

A considerable portion of the development of geostatistical modelling techniques can be viewed as a struggle between the necessity of assuming some kind of stationarity and the requirement of mimicking the complex spatial features of the geological phenomena. Most of the methodologies reviewed in this chapter can be effective in dealing with a particular aspect of non-stationarity, while other statistics and parameters are considered as stationary. The methodology proposed in this thesis takes a local approach for defining the complete set of statistics and parameters required for spatial prediction as specific of each location. The underlying idea is to obtain the local prior distributions and their statistics by applying distance weights to available data. The next chapter develops this methodology.

Chapter 3
Location Dependent Distributions and Statistics

The standard geostatistical approach for uncertainty modelling relies on a domain-wide stationary RF model that is conditioned locally by data during spatial prediction (Journel 1986; 1989, pp.8-9). In order to account for the types of non-stationarity discussed in Chapter 2, a locally defined prior RF model is proposed. This requires an assumption of local stationarity. This chapter presents the definition of a locally stationary RF. A distance weighting approach is proposed for the inference of the RF prior local cdf and statistics. The criteria for choosing the distance weighting function is discussed. The inference of weighted location-dependent cdfs and their corresponding 1-point and 2-point statistics are presented and discussed. Locally stationary spatial prediction under the multiGaussian assumption requires the modelling of local normal scores transformation. The fitting of location-dependent variograms required by locally stationary simulation and estimation techniques is also described.

Two main datasets are used for illustrating the proposed methodologies and algorithms and for comparing them with the traditional techniques. The first dataset consists of a 1-D silver grades profile corresponding to a single hole drilled during the exploration campaign of a hydrothermal deposit in the Peruvian Andes. This dataset is used when the location-dependent statistics and the impact of applying them in spatial prediction are better appreciated in one dimension. The second dataset is the well known Walker Lake 2-D dataset presented by Isaaks and Srivastava (1989). This data is derived from surface elevation measurements. This second data set is used mainly for illustrating the 2-point statistics and their capability to track the non-stationarity features of the spatial continuity. The units for the Walker Lake 2-D are set as generic. An image of this 2-D dataset is shown in Figure 3-1. Additional datasets obtained by simulation or by image processing are also used for particular examples.

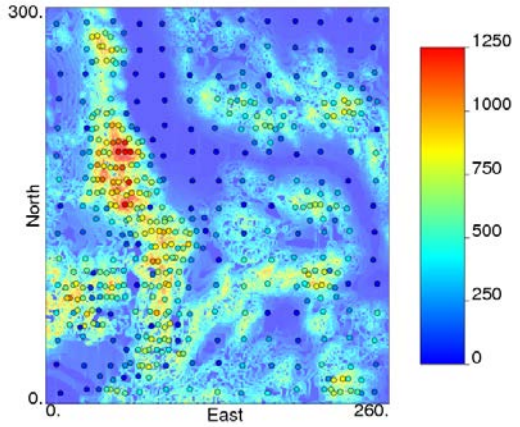

Figure 3-1: Walker Lake clustered data set (dots) superimposed on the exhaustive data set (background). The color scale indicates the elevation in meters.

3.1. The Assumption of Local Stationarity

The proposed assumption of local stationarity amounts to strict stationarity defined in relation to a reference point. Thus, the RF multivariate distribution is invariant by translation within the domain D when anchored to a reference point $\mathbf{o}$:

$$Prob\{Z(\mathbf{u}_\alpha) < z_1, ..., Z(\mathbf{u}_n) < z_K; \mathbf{o}_i\} = Prob\{Z(\mathbf{u}_\alpha + \mathbf{h}) < z_1, ..., Z(\mathbf{u}_n + \mathbf{h}) < z_K; \mathbf{o}_j\}$$
$$\forall\ \mathbf{u}_\alpha, \mathbf{u}_\beta + \mathbf{h} \in D, \text{ and only if } i=j \tag{3.1}$$

Or, expressed only for the bivariate cdf:

$$Prob\{Z(\mathbf{u}) \le z_1, Z(\mathbf{u}+\mathbf{h}) \le z_2; \mathbf{o}_i\} = Prob\{Z(\mathbf{u}') \le z_1, Z(\mathbf{u}'+\mathbf{h}) \le z_2; \mathbf{o}_j\}$$
$$\forall\ \mathbf{u}, \mathbf{u}',\ \mathbf{u}+\mathbf{h},\ \mathbf{u}'+\mathbf{h} \in D, \text{and only if } i=j \tag{3.2}$$

Anchoring the definition of the bivariate cdf allows the local inference of the univariate and bivariate parameters directly from the statistics calculated with all available data. In the assumption of local stationarity, as in the standard stationary framework, every sample is taken into account in the inference of the local cdf and its statistics; however, the individual contribution of each datum depends on its closeness to the anchor point $\mathbf{o}$. Thus, the assumption of local stationarity can be regarded as a local adaptation of the stationary RF. A way to achieving this local adaptation is using distance weighting functions. The alternative of restricting the stationarity decision to a neighbourhood makes the inference of the RF cdf and its parameters in areas with scarce data quite difficult and unstable.

Spatial prediction under the decision of local stationarity is an extension of the trend modelling idea to the RF cdf and all its statistics and parameters. Figure 3-3 shows a schematic 1-D representation of the idea that all parameters can be varied locally.

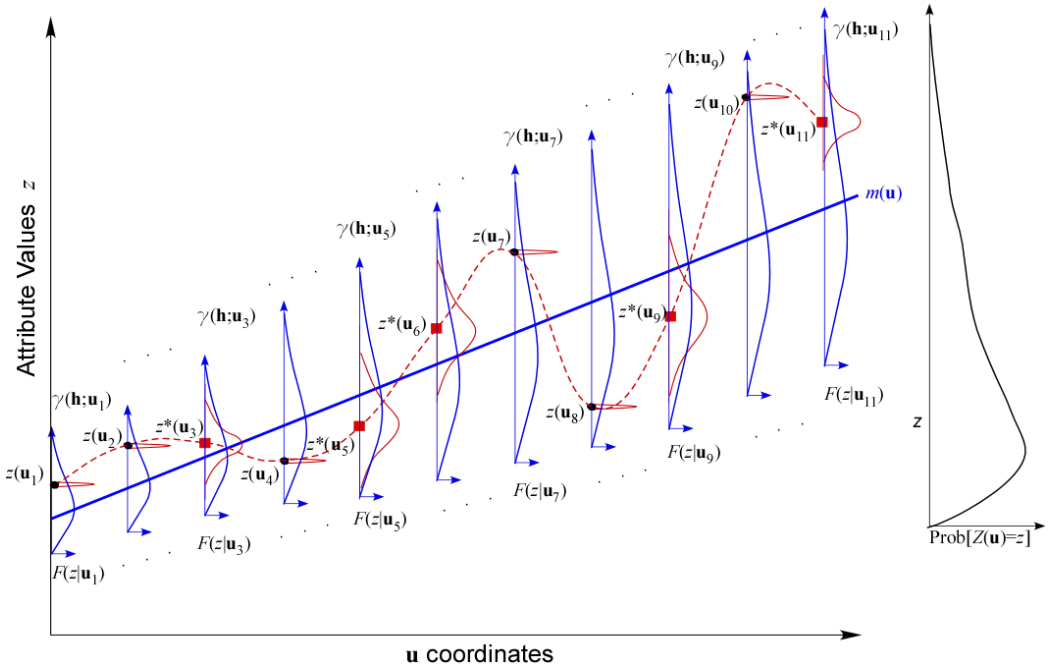

Figure 3-2: Schematic illustration of spatial prediction with location-dependent distributions and parameters. The black curve at the right represents the global prior pdf of the RF Z(u) obtained from scattered values (black dots). The blue curves at each sample location represent the local prior pdf's with local means represented by the straight blue line. At each location a variogram model is specified. The red curves represent the posterior conditional pdfs at each estimated location (red squares). The dashed red line represents the exhaustive estimation of the posterior local mean.

3.2. A Distance Weighting Approach

The idea of using distance weights for the calculation of local statistics has appeared in spatial statistics as the methodology of Geographically Weighted Regression (Fotheringham 1997; Brunsdon et al. 1998; 2002; Fotheringham et al. 2002). A similar approach within a geostatistical framework is developed below.

Several aspects must be considered in the design of a distance weighting approach for location-dependent cdfs and their statistics. These include: (1) the desirable mathematical properties of distance-based weights, (2) weighting functions that fulfill these properties, (3) the inclusion of other desirable features, such as anisotropic distances and declustering, (4) the combination of 1-point weights in order to obtain higher order weights, and (5) the criteria for selecting the anchor point locations.

3.2.1. Properties of Weights

The weights assigned to samples at locations $\mathbf{u}_\alpha$ (α=1,…,n) in relation to an anchor point located at **o** are intended to allow the unbiased inference of local statistics. They also should be inversely proportional to distance, smoothly changing, strictly positive, globally consistent for all statistics, and independent of units. These desirable properties are discussed next.

Smooth Distance-Decay

Within a domain, and in absence of discontinuities, weights should decay continuously as the distance to the anchor point increases. This represents the idea that closer samples should have a greater contribution to the corresponding cdf and its statistics.

This decay should be smooth, that is, if the separation between two samples is very small, $d\varepsilon$, the weights assigned to each of them should be similar:

$$\omega(\mathbf{u}_\alpha + d\varepsilon;\mathbf{o}) \cong \omega(\mathbf{u}_\alpha,\mathbf{o}) \quad \forall \mathbf{u}_\alpha \in D,\ \alpha = 1,...,n \tag{3.3}$$

A weighting function with a steep decay or that ends in a discontinuity may produce abrupt changes in the weights assigned to adjacent locations. This would result in instability of the local statistics. Smooth decreasing functions that are differentiable in $[0,\infty]$ are preferred.

The weights assigned to a sample in relation to anchor points separated by a very small distance, $d\varepsilon'$, should be similar:

$$\omega(\mathbf{u};\mathbf{o}_j) \cong \omega(\mathbf{u},\mathbf{o}_j + d\varepsilon') \quad \forall \mathbf{o}_j \in D,\ i = 1,...,P \tag{3.4}$$

Properties 3.3 and 3.4 are fulfilled simultaneously if a constant smooth distance-decaying function is used at all anchor points. The difference between the two properties becomes relevant when varying distance functions are used at each anchor point. Locally varying distance functions are presented in Subsection 3.2.4.

Strict Positivity and Unbiasedness

The distance based weights assigned to each sample are proportional to their probability contribution to the local cdf. Therefore, the sum of contributions from all samples should add to 1. Additionally, all probability contributions must be positive within a domain. These two properties, unbiasedness and strict positivity, are expressed as:

$$\begin{cases} \omega(\mathbf{u}_\alpha;\mathbf{o}) > 0 \\ \sum_{\alpha=1}^{n} \omega(\mathbf{u}_\alpha;\mathbf{o}) = 1 \end{cases} \qquad \forall \mathbf{o},\ \forall \mathbf{u}_\alpha \in D,\ \ \alpha = 1,...,n \tag{3.5}$$

The strict positivity property ensures licit local probabilities and allows the inclusion of all data in the calculation of location dependent statistics. Considering all samples avoids the artifacts caused by using only the closer samples to an anchor point and decreases the instability of local statistics obtained at areas with low sampling density. Moreover, giving positive weight to the samples within a window and zero weight to those outside of it would preclude the inference of the local 2-point statistics for separation distances larger than the window size.

Global Consistency and Independence of Units

The weights should be the same for 1-point and 2-point location-dependent statistics. This global consistency of weights allows using the local 1-point statistics in the calculation of 2-point statistics. Moreover, when the separation distance in the 2-point statistics becomes zero, the resulting statistics are equivalent to the 1-point statistics. Additionally, the weights should be independent of the distance units used, but dependent only on the relative distances.

3.2.2. Distance Weighting Functions

Two functions that produce weights fulfilling all or most of the specified properties are inverse distance weighting and the Gaussian Kernel. Other kernel functions commonly used in nonparametric statistics, such as Uniform, Triangle, Epanechnikov, and Cosine (Li & Racine 2007) are not considered here because they are discontinuous when they reach a bandwidth distance. This may cause artifacts and unwarranted fluctuations in the calculation of location-dependent statistics; particularly in areas were samples are very sparse.

Inverse distance and Gaussian kernel are two continuously decreasing functions of the Euclidean norm of the vector $\overline{\mathbf{uo}}$ formed by an anchor point **o** and a sample location **u**:

$$d(\mathbf{u};\mathbf{o}) = \left\|\overline{\mathbf{uo}}\right\| = \sqrt{(u_x - o_x)^2 + (u_y - o_y)^2 + (u_y - o_y)^2} = \sqrt{d_x^2 + d_y^2 + d_z^2} \tag{3.6}$$

Inverse Distance Weighting

The inverse distance weighting (IDW) method (Shepard 1968) is very popular in spatial interpolation. The weights decay according the inverse of a power of distance, *b*. The higher the *b* value, the less smooth are the estimates. The inverse distance weighting function is expressed as:

$$\omega_{IDW}(\mathbf{u}_\alpha;\mathbf{o}) = \frac{\dfrac{1}{\left(d(\mathbf{u}_\alpha;\mathbf{o}) + c\right)^b}}{\displaystyle\sum_{\alpha=1}^{n} \frac{1}{\left(d(\mathbf{u}_\alpha;\mathbf{o}) + c\right)^b}} \qquad \alpha = 1,...,n \tag{3.7}$$

Where the offset, *c*, is a constant that avoids computational problems when $d(\mathbf{u}_\alpha;\mathbf{o})$ is very small. If the *c* value is comparable or larger than the sample spacing, it also influences the smoothness of the estimates; the smaller the *c* offset, the closer the interpolated surface approaches the data values at $d(\mathbf{u}_\alpha;\mathbf{o}) = 0$. For small *b* power values, particularly for *b*<1, the estimated values vary smoothly; whereas, if *b* is higher, the interpolation shows a steep gradient near data values.

Gaussian Kernel

Gaussian kernels are used in non-parametric regression and spatial statistics (Härdle 1992, p.36; Schabenberger & Gotway 2005, p.111; Lloyd 2007, p.80). The well known shape of the Gaussian function allows a gentle decay for the weights assigned to samples located near the anchor point (see Figure 3-3, right). The distance-decay is controlled by the bandwidth, or standard deviation, value *s*. The Gaussian kernel (GK) weighting function is given by:

$$\omega_{GK}(\mathbf{u}_\alpha;\mathbf{o}) = \frac{\varepsilon + exp\left\{-\dfrac{\left(d(\mathbf{u}_\alpha;\mathbf{o})\right)^2}{2s^2}\right\}}{n\varepsilon + \displaystyle\sum_{\alpha=1}^{n} exp\left\{-\frac{\left(d(\mathbf{u}_\alpha;\mathbf{o})\right)^2}{2s^2}\right\}} \qquad \alpha = 1,...,n \tag{3.8}$$

This is similar to the Nadaraya-Watson estimator (Wasserman 2006, p.71). The background constant ε is included here in order to avoid computational problems when

$d(\mathbf{u}_\alpha;\mathbf{o})$ is large and also for controlling the smoothness of the estimated statistics (see Figure 3-3). The denominator in Expression 3.8 normalizes the weights to sum to one and allows adapting the weights to the local data density (Härdle 1992, p.32). Thus, if **o** is located in a low density area the weights given to individual samples are higher than those in densely sampled areas.

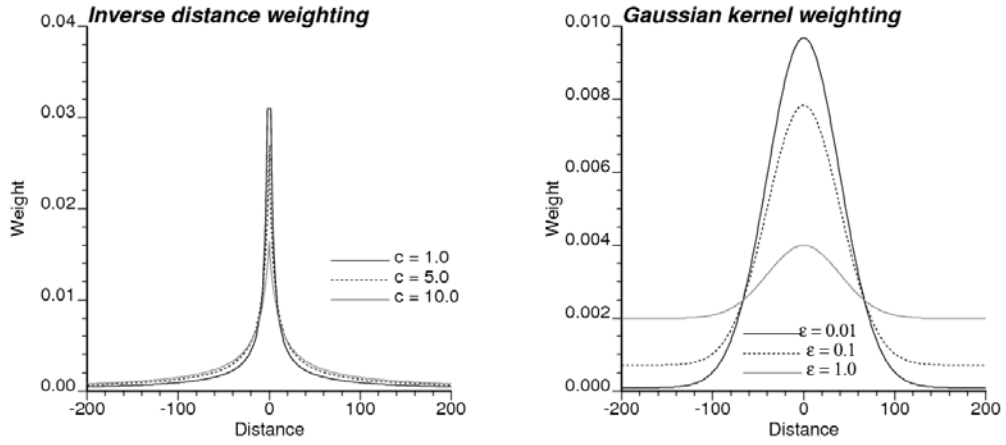

Figure 3-3: Inverse Distance (left) and Gaussian Kernel (right) weighting functions. The IDW profiles were calculated using a power of 1 and different values for offset *c*. The GK profiles were calculated using a bandwidth of 40 and for different background *ε* values.

3.2.3. Selection of Distance Weighting Parameters

The distance functions used for location-dependent statistics should not be exact interpolators; there is no necessity to reproduce the data exactly. Location-dependent weights should reflect the local trend variations informed by the data without overfitting. Therefore, the minimization of an average squared error is not a good criterion for selecting the optimal distance weighting parameters, since the resulting local estimates will match the data. Instead, subjective criteria can be used for choosing the distance function parameters that yield to smoothly varying statistics. Trend modelling criteria incorporates such subjectivity (McLennan 2007).

A basic requirement to assure the unbiasedness of local statistics is that the expected value of the local means $m(\mathbf{o})$ over the domain D must approach the RF global mean m:

$$\frac{1}{D}\int_D m(\mathbf{o})d\mathbf{o} \approx E\{Z(\mathbf{u})\} = m \tag{3.9}$$

Once this is verified, the total variance, σ^2, of the decomposed RF in the local mean, $m(\mathbf{u})$, and the residual, $R(\mathbf{u})$ (see Equation 2.12), is expressed as the sum of the

variance of the local means, $Var\{m(\mathbf{u})\}$, the variance of residuals, $Var\{R(\mathbf{u})\}$, and the covariance between the local means and the residuals, $Cov\{m(\mathbf{u}), R(\mathbf{u})\}$:

$$Var\{Z(\mathbf{u})\} = Var\{m(\mathbf{u})\} + Var\{R(\mathbf{u})\} + 2Cov\{m(\mathbf{u}), R(\mathbf{u})\} \quad (3.10)$$

If the trend is smooth, the variance of the local means represents a limited fraction of the total variance. It has been suggested that, as a rule of thumb, this fraction should not exceed 50% (McLennan 2007, pp.6-4,6-5). Additionally, the absolute value of the covariance between the local means and the residuals should be minimized. While the coefficient of correlation between $Z(\mathbf{u})$ and the local means should be maximized. These three criteria can be summarized as:

$$Var\{m(\mathbf{u})\}/Var\{Z(\mathbf{u})\} < 0.5 \quad (3.11)$$

$$\min\{Cov\{m(\mathbf{u}), R(\mathbf{u})\}\} \quad (3.12)$$

$$\max\left\{\frac{Cov\{m(\mathbf{u}), Z(\mathbf{u})\}}{\sqrt{Var\{m(\mathbf{u})\} \cdot Var\{Z(\mathbf{u})\}}}\right\} \quad (3.13)$$

The first criterion (3.11) avoids overfitting of data by the local mean. As this ratio approaches 1, the local mean more closely matches the data. Criterion 3.12 is important when treating the residuals as independent of the mean. If the covariance between the local means and the residuals is significant, artifacts, such as values exceeding the natural limits of the attribute, may appear when the trend is restored after spatial prediction with the residuals (Deutsch 2002, p.186). For the purpose of locally stationary modelling, this correlation is irrelevant, since the RF is not decomposed, but locally modified by the distance function weights. Criterion 3.13 is aimed to adjust the local mean to the local data variations. The practitioner, guided by the appearance of resulting location-dependent statistics maps, must find a reasonable balance between these measures.

These criteria involve only the local mean; however, the weights must fulfil the property of global consistency. Those weights deemed reasonable for the local mean are also considered adequate for other location-dependent statistics.

Figure 3-4 shows an example for modelling the trend of silver grades mean in a drill hole using IDW (left) with a background value of 3 and different power parameters, and GK (right) with a background value of 0.01 and for different bandwidths. The trends modelled by IDW show a higher sensitivity to extreme values and appears less smooth than the ones obtained from GK, even when the power value is as small as 0.5. The spikes in the IDW trend model are not a matter of appearance only, but they may translate

in unwarranted fluctuations of the location dependent moments. GK weighting, instead, produces smooth trends models. In this example, a bandwidth of 40m can be reasonably chosen between the excessively smooth trends produced using an 80m bandwidth and the fluctuating trend obtained with a 5m bandwidth.

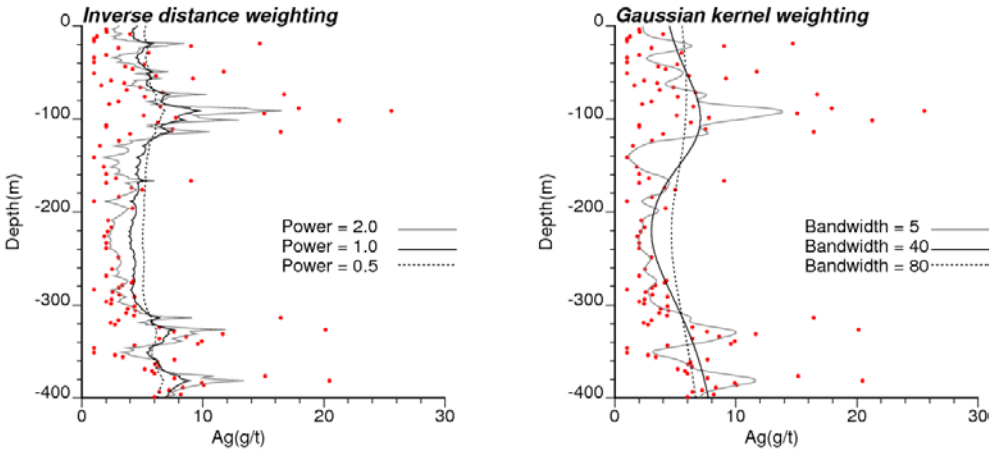

Figure 3-4: Trend modelling of silver grades (dots) in a drillhole using Inverse Distance (left) and Gaussian Kernel (right) weighting functions.

Using GK bandwidth of 40m, the trend-data variance ratio is just 0.085, the trend - residuals covariance is 0.967, and the trend- data correlation is 0.40. Figure 3-5 shows the progression of the variance ratios and the coefficient of correlation between the data and the trend for different parameters in IDW and GK weighting. For IDW, power values below 1 yield very flat trends when using a background value comparable to the samples separation, 3m in this case. This is reflected in the very low trend variance /data variance ratio for the low power parameters in Figure 3-5, left. Increasing the power parameter increases the correlation between the trend model and data but at the price of the spikes observed at Figure 3-4. For the GK (Figure 3-5, right) a reasonable choice lies between closely fitting the data using a narrow bandwidth and the quasi-flat trends produced by excessively wide bandwidths.

The choice of an offset equivalent or larger than the sample separation in IDW has a considerable smoothing effect (See Figure 3-6, left). In GK weighting, the background constant represents the major part of the weight assigned to distant samples (see Figure 3-3). The smoothing effect of this constant becomes noticeable when it is higher than 0.01 (see Figure 3-6, left).

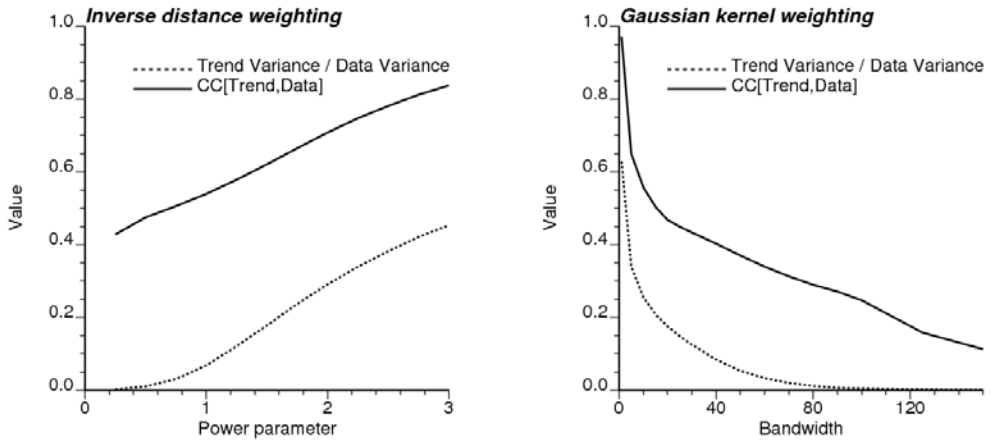

Figure 3-5: Progression of trend and data variance ratio (dashed lines) and the coefficient of correlation between data and the trend (continuous lines) according the power parameter of IDW (left) and the GK bandwidth (right).

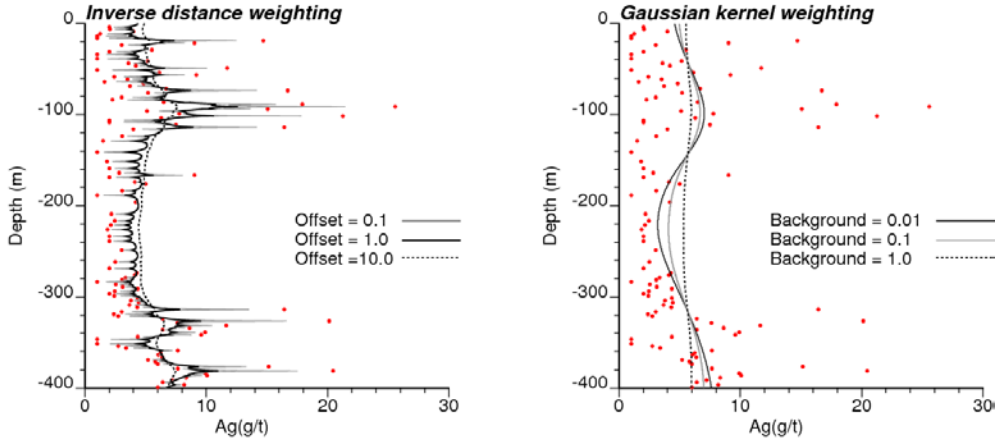

Figure 3-6: Background value effect on the local means for inverse distance (left) and the Gaussian kernel (right) weighting.

Additional considerations for choosing the distance weighting function parameters are the data density and the scale of modelling. If sampling is dense it is possible to use smaller bandwidths in order to resolve the local statistics. In this sense, another advantage of the kernel distance methods is that the bandwidth can be selected in relation to the sampling separation and the scale of modelling. Moreover, if the scale of the non-stationary features can be obtained from secondary information or abundant data, it can also be used for tuning the distance function parameters. An example of this is when the spatial distribution of an attribute may be controlled by topography and exhaustive topographical information is available. In this case, the parameters of the weighting function can be selected in relation to the spatial features observed in the terrain model.

3.2.4. Anisotropy, Declustering and Local Adaptation

Distance weights can be modified to account for the presence of a dominant anisotropy orientation, preferential sampling and changes in local data density. These modifications are achieved by changing the distance function shape or by correcting the distance weights making them inversely proportional to the local data density.

Use of Anisotropic Distances in the Weighting Functions

Strong anisotropy can be present within many geological domains. This anisotropy is related to the geologic processes and may be considered. The anisotropy definition used for distance weighting function should likely be weaker than the true anisotropy of the attribute. A very strong anisotropy for the distance function may mask the local changes in the anisotropy of the true spatial distribution. Using anisotropic distances is equivalent to modifying the amplitude of the weighting function in a particular direction. This translates in an increased variability of the trend in the minor anisotropic direction.

The small example presented in Figure 3-7 illustrates this idea using samples on a quasi-regular 7 by 7 pixels grid. The geological image has a strong anisotropy parallel to the X-axis. The global variogram of this image has an anisotropy ratio of 2.3 for the short scale structure, and about 4.4 for the long range structure. The three trend models presented in Figure 3-7 were built applying different anisotropy ratios in the distance weighting. In this figure (top right), it can be observed that, isotropic distances in the weighting function may induce isotropic features in the trend model. Contrarily, imposing an anisotropy ratio similar or higher to that shown by the variogram model inhibits the depiction of local anisotropic features that depart from the global anisotropy (see bottom right map in Figure 3-7). A mildly anisotropic distance weighting function can be used to avoid these two extremes, particularly when dealing with sparse data.

Figure 3-8 shows the coefficient of correlation between the trend and data, and the trend variance/data variance ratio for different anisotropy ratios applied on the Gaussian kernel bandwidth. Both statistics increases as the anisotropy of distance weighting function gets stronger. This indicates a reduced smoothing in the direction of the minor kernel bandwidth.

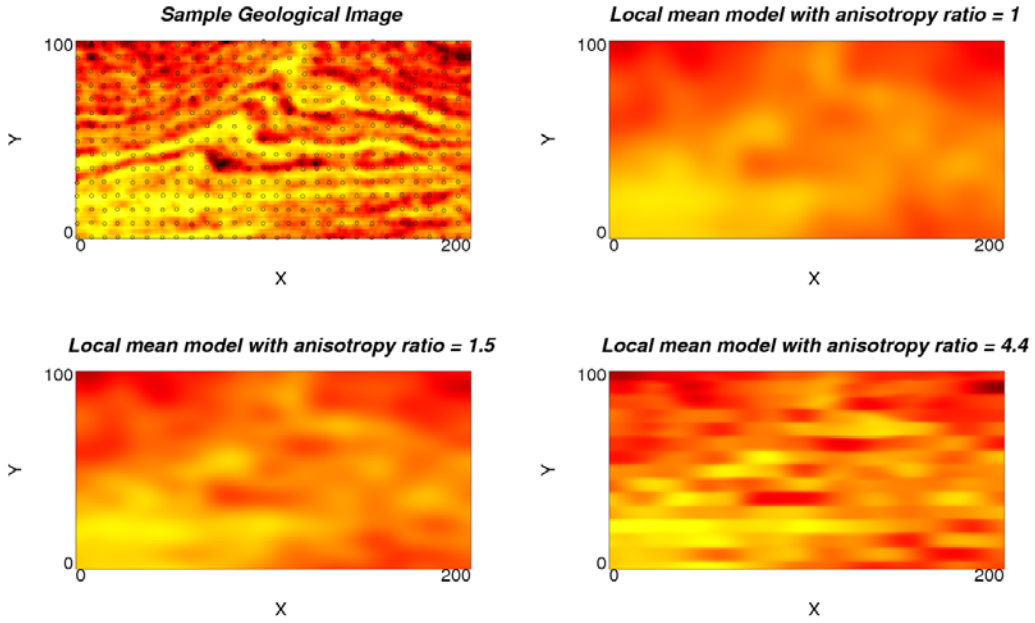

Figure 3-7: Geological image sampled in a 7 x 7 pixels grid (top left), local means model produced using Gaussian kernels with anisotropy ratio of 1 (top right), 1.5 (bottom left) and 4.4 (bottom right)

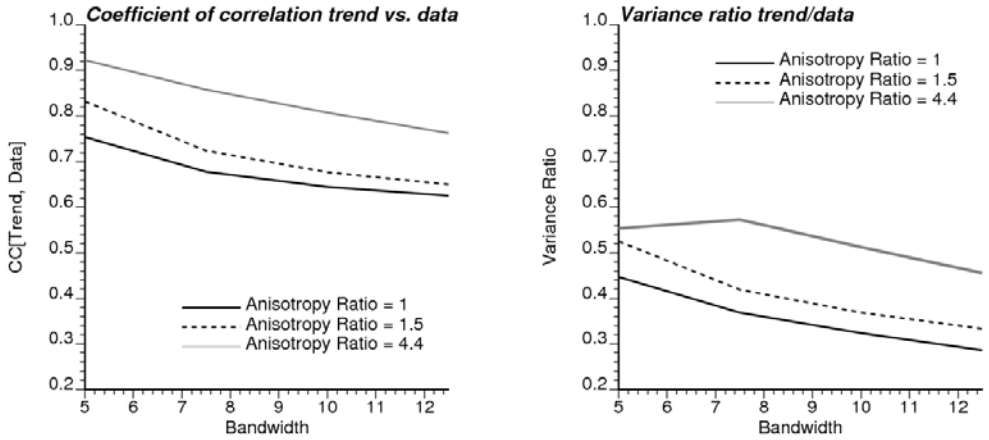

Figure 3-8: Effect of the distance weighting anisotropy ratio on the coefficient of correlation between trend and data (left) and on the trend/data variance ratio (right)

Geometric anisotropy is modelled as a tri-axial ellipsoid defined by 3 radii and three rotation angles (Gendzwill & Stauffer 1981). Incorporating an anisotropic definition in distance weighting is equivalent to enlarge the components of the vector $\overline{\mathbf{uo}}$ parallel to the two shorter anisotropy radii, while fixing the component parallel to the major anisotropy radius. A coordinate rotation can also be considered if the axes of anisotropy ellipsoid are not parallel to the coordinate axes. The following transformation matrix is used for this purpose (Leuangthong et al. 2008, pp.62-69)

$$[\mathbf{T}] = [\mathbf{S}][\mathbf{R}_3][\mathbf{R}_2][\mathbf{R}_1]$$
$$= \begin{bmatrix} a_1/a_2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a_1/a_3 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta_3 & 0 & -\sin\theta_3 \\ 0 & 1 & 0 \\ \sin\theta_3 & 0 & \cos\theta_3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_2 & \sin\theta_2 \\ 0 & -\sin\theta_2 & \cos\theta_2 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(3.14)

The angles and radii that define the anisotropy ellipsoid follow the GSLIB convention (Deutsch & Journel 1998, pp.27-28). This, $[\mathbf{R}_1]$ is the rotation matrix corresponding to the clockwise angle of rotation θ_1 around the z axis, $[\mathbf{R}_2]$ corresponds to the rotation angle θ_2 around rotated X-axis and positive upwards, $[\mathbf{R}_3]$ corresponds to a rotation of $_3$ degrees around the rotated Y-axis, which is positive clockwise. $[\mathbf{S}]$ is the scaling matrix where a_1, a_2 and a_3 are the radii of the ellipsoid parallel to the rotated Y, X and Z axes, respectively.

The vector $\overline{\mathbf{uo}} = (d_x, d_y, d_z)$ presented in Expression 3.6 is modified by:

$$\begin{bmatrix} d'_x \\ d'_y \\ d'_z \end{bmatrix} = [\mathbf{T}] \begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} \tag{3.15}$$

Thus, the anisotropic distance between the anchor points and the data becomes:

$$d'(\mathbf{u};\mathbf{o}) = \sqrt{d'^2_x + d'^2_y + d'^2_z} \tag{3.16}$$

Correction by Declustering Weights

Due to sampling design or logistical reasons, clusters of data are common in diverse spatial datasets related to the earth sciences (Davis 2002, p.299; Borradaile 2003, p.14; Sinclair & Blackwell 2002, pp.81-83; Webster & Oliver 2007, p.32). For the purpose of increasing the information in interesting areas, this clustering is practical; however, these practices may introduce bias in the global statistics (Deutsch 2002, p.50), as well as in the distance weighted location-dependent statistics. In order to remove this bias in the location-dependent statistics, the set of distance-based weights should fulfil the following property:

$$\bar{\omega}(\mathbf{u}_\alpha) = \frac{1}{P}\sum_{j=1}^{P} \omega(\mathbf{u}_\alpha;\mathbf{o}_j) = w_\alpha \qquad \forall\ \alpha = 1,...,n \tag{3.17}$$

That is, the average of the distance weights assigned to a sample $\mathbf{u}_\alpha$ in relation to all P anchor points must be equal to its corresponding declustering weight, w_α. A correction is required because distance weighting functions do not necessarily yield weights fulfilling this property.

Common declustering methods include cell declustering (Deutsch & Journel 1998, pp.213-214; Isaaks & Srivastava 1989, pp.241-242), polygonal declustering (Isaaks & Srivastava 1989, pp.238-239; Deutsch 2002, p.51), and global kriging weights (Deutsch 1989). The declustering weights obtained from these methods can be imposed on the distance weights by a correction factor:

$$\hat{\omega}(\mathbf{u}_\alpha;\mathbf{o}) = \omega(\mathbf{u}_\alpha;\mathbf{o})\frac{w_\alpha}{\overline{\omega}(\mathbf{u}_\alpha)} \qquad \forall\ \alpha = 1,...,n \tag{3.18}$$

Therefore, the Property 3.17 is fulfilled:

$$\frac{1}{P}\sum_{j=1}^{P}\hat{\omega}(\mathbf{u}_\alpha;\mathbf{o}_j) = w_\alpha \qquad \forall\ \alpha = 1,...,n \tag{3.19}$$

A new regularization of the weights is required after this operation in order to make the sum of weights to be equal to one. The 1-D dataset presented in Figure 3-4 is used to illustrate the effect of the correction by declustering weights on the distance weights and the modeled trend. For each datum, the corresponding declustering weight is proportional to the distance between the midpoints of the adjacent sampled segments. This is akin to polygonal declustering in 1-D. The left side of Figure 3-9 shows that, for narrow bandwidths, the average GK weight assigned to each sample approximates the corresponding declustering weight. Wider bandwidths are less sensitive to changes in local data density. In mining exploration it is common not to assay the intervals that do not show visual evidence of mineralization. This is also the case for the 1-D dataset corresponding to a single drill hole considered here. Thus, the section outside the mineralized structure at the centre of the drill hole has a lower sample density, and the declustering weights are much higher for the samples there.

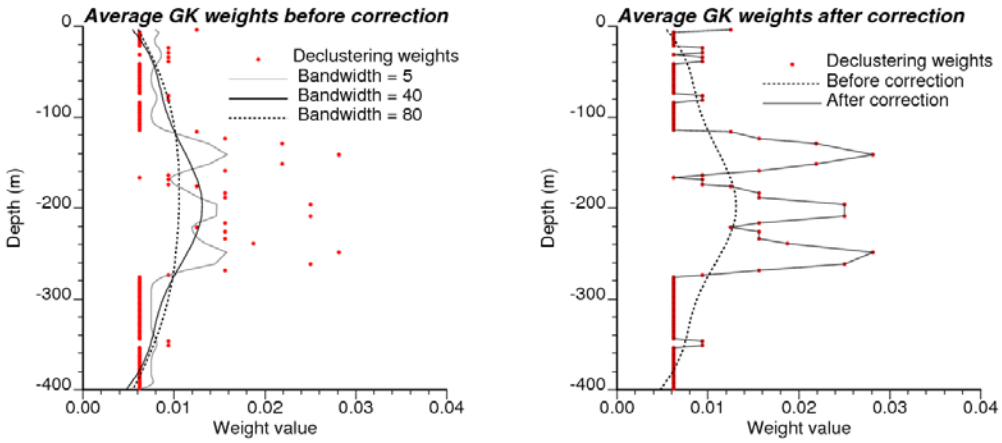

Figure 3-9: Left, declustering weights and average Gaussian kernel weights assigned to each sample for different bandwidths. Right, average sample weights after declustering correction.

After the correction, the average distance weight assigned to each sample matches corresponding declustering weight, this is: $\bar{\omega}(\mathbf{u}_\alpha)=w_\alpha,\ \alpha=1,...,n$ (see Figure 3-9, right). The resulting trend model after correction by declustering weights is slightly shifted to the low grade values (see Figure 3-10, left). This shift is consistent with the correction of the bias caused by dense sampling in the high grade intervals of the drillhole. The correction for declustering comes with the price of discontinuous weight profiles around the anchor points (see Figure 3-10, right). Consequently, condition 3-3 is no longer satisfied. These discontinuities are caused by the scaling of original distance weights by the declustering weights. Thus, the contribution of samples located in low density areas is increased for the inference of local statistics, and the continuity of the weights assigned to contiguous samples in relation to the same anchor point may be broken. However, the weights assigned to each individual sample in relation to contiguous anchor points still varies smoothly (see Figure 3-11), fulfilling the Property 3.4. The preservation of this property after the correction by declustering weights allows generating smoothly varying local statistics. The sum of corrected weights related to an anchor point may not be equal to one, thus a new standardization is required in order to fulfill Property 3.5. These standardized 1-point weights will be denoted as $\omega'(\mathbf{u}_\alpha;\mathbf{o}),\ \alpha=1,...,n$, $\sum_{i=1}^{n}\omega'(\mathbf{u}_\alpha;\mathbf{o})=1$.

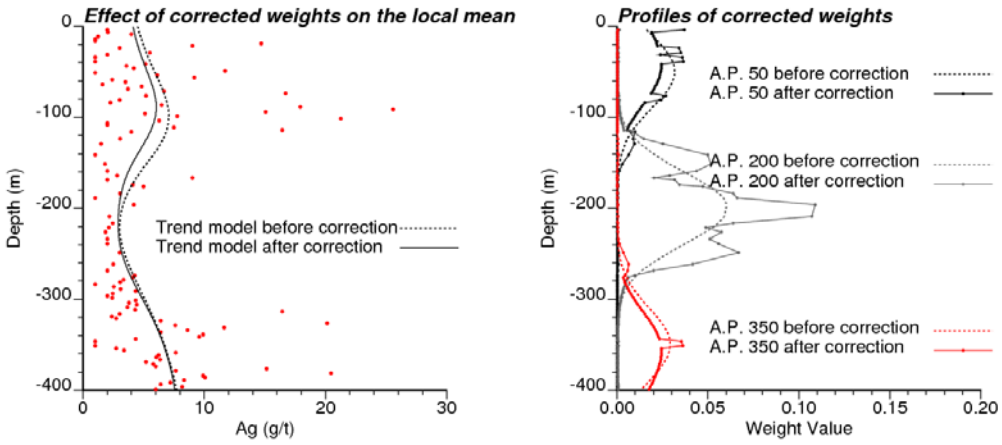

Figure 3-10: Left, effect of declustering correction on the local means model. Right, weight profiles at three anchor points before and after declustering weights correction.

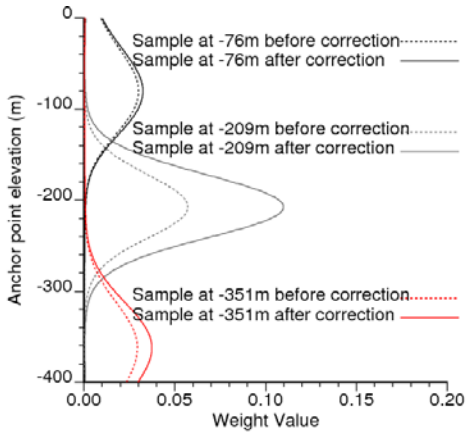

Figure 3-11: Distance weights assigned to three individual samples before and after correction by declustering weights. The selected samples correspond to the highest spikes in the weights profiles of Figure 3-10.

Dynamic Kernel Bandwidth

In areas with low sample density, the inference of local statistics may become difficult. This may cause unwarranted fluctuations in the local statistics in such areas. While densely sampled areas may contain enough information to provide a more detailed description of the local statistics. The idea is to increase the bandwidth in low sample density areas and to narrow it in highly sampled regions.

The denominator of Expression 3.8 is directly proportional to the data density, note Figure 3-12, left. The local data density calculated with a moving window is presented along with the Gaussian kernel denominator value. A locally changing bandwidth, s'_0 can be obtained by approximating the denominator of Expression 3.8 to the average

denominator obtained from all anchor points using a constant bandwidth *s*. Thus, the locally adapting Gaussian kernel becomes:

$$\tilde{\omega}_{GK}(\mathbf{u}_\alpha;\mathbf{o}) = \frac{\varepsilon + exp\left\{-\frac{\left(d(\mathbf{u}_\alpha;\mathbf{o})\right)^2}{2s_{\mathbf{o}}'^2}\right\}}{n\varepsilon + \sum_{\alpha=1}^{n} exp\left\{-\frac{\left(d(\mathbf{u}_\alpha;\mathbf{o})\right)^2}{2s_{\mathbf{o}}'^2}\right\}} \cong \frac{\varepsilon + exp\left\{-\frac{\left(d(\mathbf{u}_\alpha;\mathbf{o})\right)^2}{2s_{\mathbf{o}}'^2}\right\}}{n\varepsilon + \frac{1}{P}\sum_{\beta=1}^{P}\sum_{\alpha=1}^{n} exp\left\{-\frac{\left(d(\mathbf{u}_\alpha;\mathbf{o}_\beta)\right)^2}{2s^2}\right\}} \quad (3.20)$$

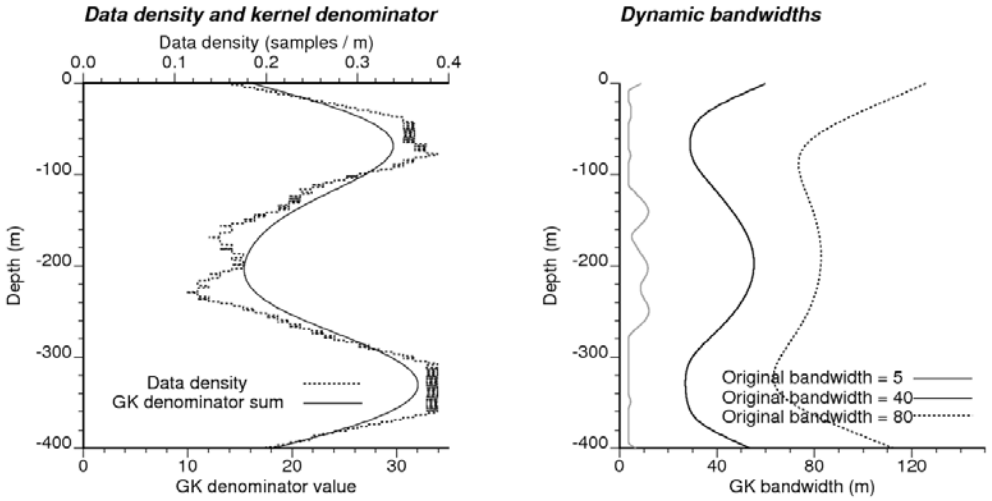

Figure 3-12: Left, local data density calculated with a moving window and local kernel denominator value. Right, local Gaussian kernel bandwidths after regularization of the kernel denominator value for different original bandwidths.

Figure 3-13, right shows the resulting local bandwidths widened in low density areas and narrowed in highly sampled areas. This translates to smoother local statistics in low density areas, while the statistics in highly sampled areas show more local detail (see Figure 3-13, left). In Figure 3-13, right, the weight profiles at three separated anchor points for the silver drillhole assays are shown before and after the use of a locally changing bandwidth.

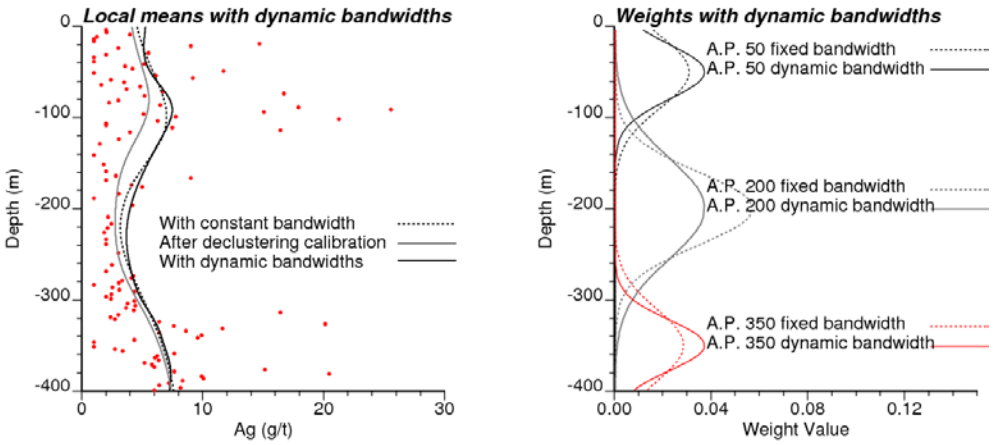

Figure 3-13: Effect of the dynamic bandwidth in the local means model (left). Weights profiles for three anchor points before and after regularization of the kernel denominator value (right).

3.2.5. From 1-Point to 2-Point Weights

Calculation of 2-point statistics requires weights corresponding to a pair of samples simultaneously. The resulting 2-point weights are obtained by applying a mixture rule to the weights corresponding to the individual samples (See Figure 3-14). Mixture rules have been defined by Korvin (1982) for obtaining the value of a composite material attribute formed by two different components or phases. Given two phases with corresponding property values g_1 and g_2, and volume fractions ϕ, and 1-ϕ, respectively, the general mixture rule is expressed as:

$$M(g_1, g_2, \phi, t) = \left[\phi g_1^t + (1-\phi) g_2^t\right]^{1/t}, \quad \text{with } t \neq 0, t \in \square \tag{3.21}$$

For obtaining 2-point weights, it can be assumed that the volume fractions of each sample are equivalent, that is $\phi = 0.5$, and the individual sample weights are the property values: $g_1 = \omega(\mathbf{u}_\alpha; \mathbf{o})$, $g_1 = \omega(\mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})$. Thus, the mixture rule for combining the weights corresponding to a sample pair separated by the vector **h** becomes:

$$\omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) = \left[\frac{\omega^t(\mathbf{u}_\alpha; \mathbf{o}) + \omega^t(\mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})}{2}\right]^{1/t} \tag{3.22}$$

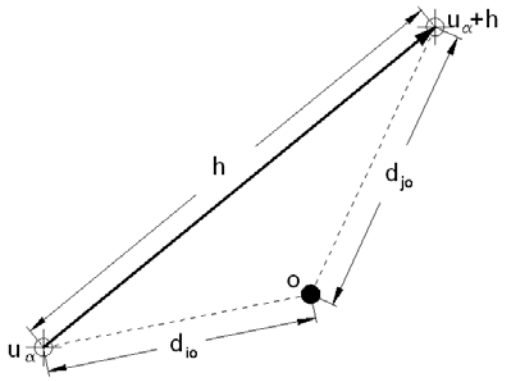

Figure 3-14: Weighting schema for sample pairs

For t=1, the mixture rule is the arithmetic average of weights:

$$\omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) = \frac{\omega(\mathbf{u}_\alpha; \mathbf{o}) + \omega(\mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})}{2} \tag{3.23}$$

Whereas, if t approaches 0, the expression converges to the geometric average:

$$\omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) = \sqrt{\omega(\mathbf{u}_\alpha; \mathbf{o}) \cdot \omega(\mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})} \tag{3.24}$$

These two particular cases of the mixture rule are simple. Nevertheless, t can take any real value. For $t \in (-\infty, +\infty)$, Expression 3.22 is bounded by the tail and head sample weights:

$$Min\left(\omega(\mathbf{u}_\alpha; \mathbf{o}), \omega(\mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})\right) \leq \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) \leq Max\left(\omega(\mathbf{u}_\alpha; \mathbf{o}), \omega(\mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})\right)$$

As t increases the 2-point weights approach the closer sample weight, while the 2-point weights is lowered closer to the farther sample weight as t decreases. This is illustrated in Figure 3-15, where the tail sample location has been fixed at the origin and its weight standardized to 1, while the head sample location is changing according to the red curve.

The value of the exponential parameter in the mixture rule has little impact in the 2-point statistics for small separations **h**. However, for longer separations, the increasing of the value of inferred statistics when a positive t value is used, and its reduction when t is negative, becomes noticeable. As discussed in Subsection 3.3.5, a choice of the t value that avoids these extremes in the inference of local measures of spatial continuity is 0.

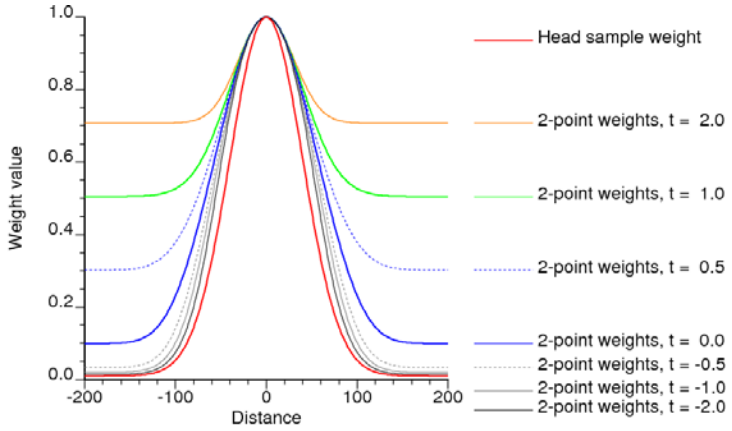

Figure 3-15: 2-Point weight profiles for different values of the power *t* in the mixture rule. The tail sample location has been fixed in the origin and has a weight of 1, while the head sample is allowed to move.

It is important to remark that whatever the *t* value is, if the sample separation **h** is 0, the 2-point weight reverts to a 1-point weight. As explained in Section 3.3.4, this property allows consistency between 1-point and 2-point statistics. These 2-point weights fulfil the properties specified for 1-point weights, except the unbiasedness property. This can be achieved by:

$$\omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) = \frac{\omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})}{\sum_{i=1}^{N(\mathbf{h})} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o})} \tag{3.25}$$

Where $N(\mathbf{h})$ is the number of sample pairs separated by the vector **h**.

If the corrected weights, $\hat{\omega}(\mathbf{u}_\alpha; \mathbf{o})$, $\alpha = 1, ..., n$, are used for building the 2-point weights, these also will be able to minimize the bias caused by preferential sampling. Thus, for sample pairs located at similar distances to the anchor point, those located in scarcely sampled areas will be associated with higher 2-point weights.

3.2.6. Choosing the Anchor Point Locations

Ideally, the location-dependent statistics should be inferred at every location in the model. For 1-point statistics this can be accomplished straightforwardly. However, for 2-point statistics, this would be very demanding in computer resources. Additionally, checking the local statistics and the models fitted to them would be tedious if it is done at every location. Alternatively, the 1-point and 2-point local statistics obtained for a limited number of anchor points can be interpolated for all other locations. Any interpolation

method that honours the data values at their locations and is able to produce smoothly varying maps is adequate for this task. The spacing of the anchor points must be such that the interpolated statistics and parameters closely follow those that would be inferred directly at every location. Since the local mean and variance are relatively straightforward to infer at every location, these statistics are used for finding an adequate anchor point separation. Normally, the error between the interpolated statistics and those inferred at every location increases with the anchor point separation, particularly when this exceeds the kernel bandwidth. Thus, the practitioner must find a balance between computational efficiency and the minimization of the error introduced by the interpolation of local parameters.

For example, in Figure 3-16, left, it can be observed that the exhaustively inferred local mean model can be closely reconstructed by interpolating the values obtained at anchor points separated up to a distance equivalent to the kernel bandwidth. The difference is negligible at such separation (see Figure 3-16, right). For larger separations between anchor points, this difference grows quickly. Ordinary kriging with a spherical variogram with isotropic range equal to the anchor point separation and zero nugget effect was used for interpolating the local statistics in this case.

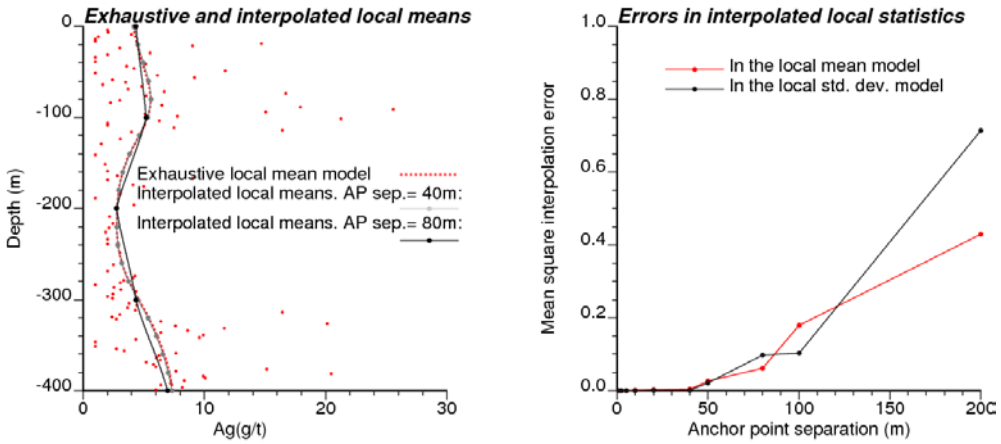

Figure 3-16: Interpolated local mean models between anchor points of different separations (left). Mean square errors between the exhaustively inferred local statistics and the interpolated statistics for different separations of anchor points (right). A Gaussian Kernel with 40m bandwidth of 40m was used for the inference of the local means and standard deviations.

In the case of the 2-D dataset, the location-dependent statistics were obtained using a Gaussian kernel with 20m bandwidth and 0.1 background, with calibration by declustering weights. A 20m x 20m anchor points mesh was used for the local inference of the local statistics. Figure 3-17 shows the interpolated values of these local statistics

between the anchor points. The interpolation method for used this case was also ordinary kriging with a spherical variogram model with 20m isotropic range and zero nugget effect.

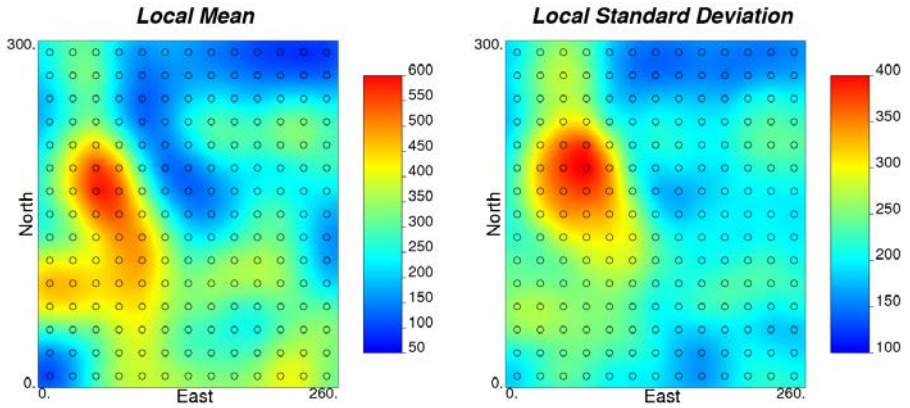

Figure 3-17: Local mean and local standard deviation fields inferred for the elevation attribute in the 2-D dataset. The circles mark the location of the anchor points.

3.3. Location-Dependent Distributions and Statistics

Once a set of 1-point and 2-point distance weights are obtained, these can be used for the inference of locally varying 1-point and 2-point statistics. The expressions of location-dependent statistics are such that they revert to their stationary form when the distance weights are constant. The normal scores transformation of the local cdf by incorporating the distance weights is presented in this Section. There is a trade-off between the capability of experimental location-dependent measures of spatial continuity to capture the local anisotropy and the stability of these local 2-point statistics. This trade-off is controlled by the choice of the distance weighting function parameters. The availability of closely spaced samples is critical for the correct inference of the location-dependent 2-point statistics. These issues are also discussed in the present Section.

The fact that the location dependent statistics are sample statistics is emphasized by the use of the notation ^ on them.

3.3.1. Location-Dependent Distributions

The Expression 2.13 is used for building the stationary cdf from the values obtained at different locations $\mathbf{u}_\alpha$ (α=1,...,n). The local univariate cdf anchored at $\mathbf{o}$ can be inferred from a similar expression by incorporating the standardized weights $\omega'(\mathbf{u}_\alpha;\mathbf{o})$:

$$Prob\{Z(\mathbf{u}) \le z_k \mid \mathbf{o}\} \cong \hat{F}(\mathbf{u}; z_k; \mathbf{o}) = \sum_{\alpha=1}^{n} \omega'(\mathbf{u}_\alpha; \mathbf{o}) \cdot I(\mathbf{u}_\alpha; z_k) \ \in [0,1]$$
$$\forall \mathbf{u}_\alpha \in D, \ k = 1, ..., K \tag{3.26}$$

Where the indicator function $I(\mathbf{u}_\alpha; z_k)$ is defined by the expression (2.14), and $\sum_{i=1}^{n} \omega'(\mathbf{u}_\alpha; \mathbf{o}) = 1$. Similarly, the location dependent bivariate cdf can be inferred from:

$$\text{Prob}\left\{Z(\mathbf{u}) \le z_k, Z(\mathbf{u}+\mathbf{h}) \le z_{k'} \mid \mathbf{o}\right\} \cong \hat{F}(\mathbf{u}, \mathbf{u}+\mathbf{h}; z_k, z_{k'}; \mathbf{o})$$
$$= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) \cdot I(\mathbf{u}_\alpha; z_k) \cdot I(\mathbf{u}_\alpha + \mathbf{h}; z_{k'}) \tag{3.27}$$

This is equivalent to the location-dependent noncentered indicator (cross) covariance (Deutsch & Journel 1998, p.11), with $\sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}; \mathbf{o}) = 1$.

Figure 3-18 shows the global univariate cdf for the drillhole silver assays presented before, along with the local cdfs at three different anchor points.

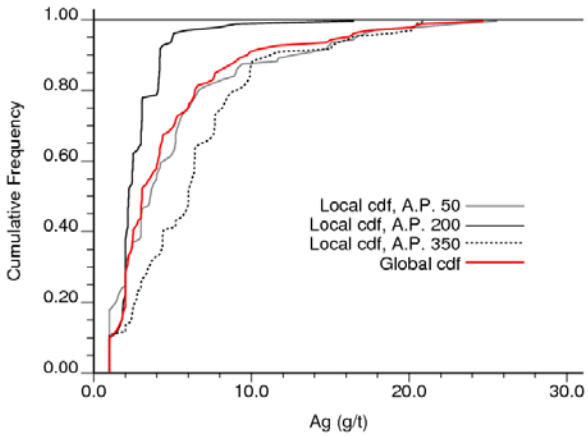

Figure 3-18: Location-dependent cdfs obtained using Gaussian Kernel distance weights at anchor points located at elevations of -50, -200 and -350 along the drillhole presented previously.

3.3.2. Local Normal Scores Transformations

In order to conform to the multiGaussian model, the location-dependent univariate cdfs must be transformed to univariate normal cdfs. This requires a normal scores transformation (Goovaerts 1997, pp.266-271; Deutsch & Journel 1998, p.141) at each location. This methodology allows accounting for trends in the mean and local variations of the variance and histogram shape (McLennan & Deutsch 2008).

Given $G(y)$ as the standard Gaussian normal cdf, with $m_Y = 0$ and $\sigma_Y = 1$, and a location-dependent cdf $\hat{F}(\mathbf{u}; z; \mathbf{o})$ of arbitrary shape, the local normal scores transform,

$$y_p = G^{-1}\left(\hat{F}(\mathbf{u}, z_p; \mathbf{o})\right) = \varphi_Y(z_p; \mathbf{o}), \tag{3.28}$$

allows matching the values z_p with y_p, such as:

$$\hat{F}(\mathbf{u}, z_p; \mathbf{o}) = G(y_p) = p \quad \forall p \in [0,1] \tag{3.29}$$

At each anchor point, the n ordered pairs (z_j, y_j) are stored in a transformation lookup table (Deutsch & Journel 1998, pp.223-226). These tables can be used for the non-parametric modelling of the local back transformation function:

$$z_j = \hat{F}^{-1}\left(G(y_j); \mathbf{o}\right) = \varphi_Z(y_j; \mathbf{o}) \qquad j = 1, \ldots, n \tag{3.30}$$

Figure 3-19 shows examples of local normal scores transformation functions built by local transformation lookup tables at the same three anchor points considered in Figure 3-18. As it can be observed, the transformed values in these tables do not lead to a standard normal distribution. If the transformation is performed in a low value region, the distribution of transformed values is shifted to the positive side, and vice versa. Only with the application of distance weights, the local transformed distributions become standard Gaussian.

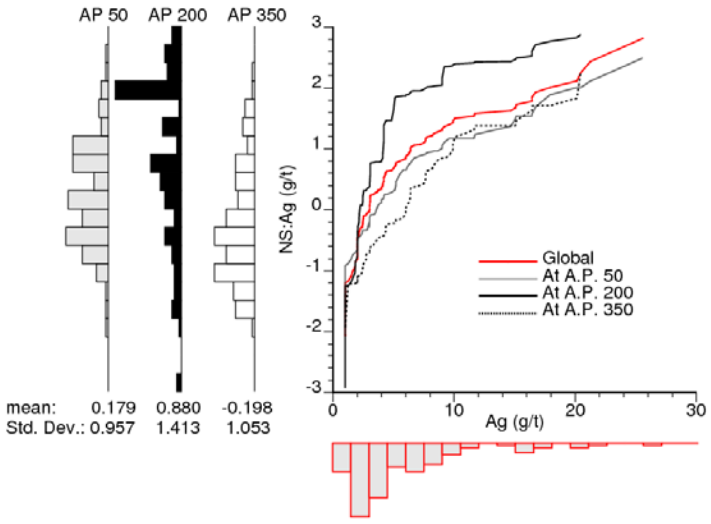

Figure 3-19: Local normal scores transformation functions at three different anchor points and their corresponding transformed distributions (1-D dataset).

In Figure 3-20, left, the cdfs of locally transformed values appear different from the Gaussian cdf and in inverse order when compared with Figure 3-18. Local transforms closely conform to the normal Gaussian distribution when weighted by the distance

weights (Figure 3-20, right). However, deviations from the standard Gaussian cdf are common in the tails of locally transformed cdfs. These are caused by the presence of isolated very low or very high values that cause discontinuities at the extremes of the original local cdfs.

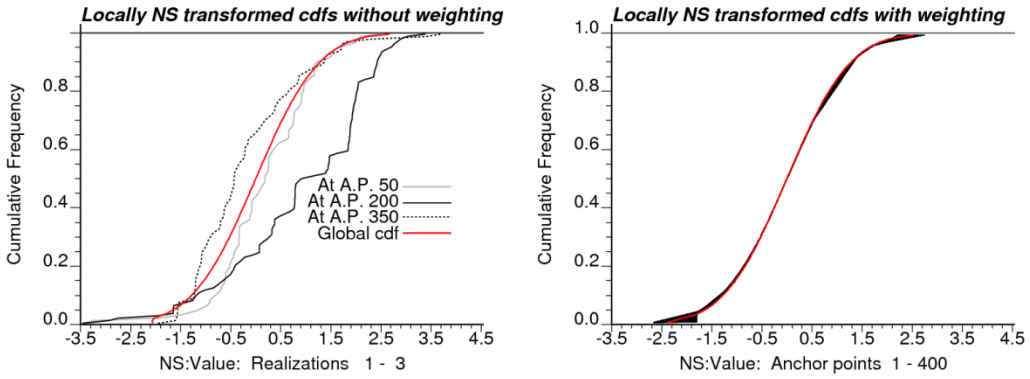

Figure 3-20: Left, three locally transformed cdfs plotted without incorporating the distance weights used in the construction of their respective transformation functions. Right, 400 locally weighted transformed cdfs (1-D dataset).

3.3.3. Location-Dependent 1-Point Statistics

The use of local means, variances and other 1-point statistics calculated by distance based weights has been proposed for the exploratory analysis of spatial data (Brunsdon et al. 2002). Here, the location-dependent 1-point statistics are presented within a geostatistical framework as quantities summarizing the prior definition of the locally stationary RF cdf.

Location-Dependent Mean and Quantiles

The location-dependent mean is estimated as the distance weighted average of sample values (C. Brunsdon et al. 2002; Borradaile 2003, p.39):

$$\hat{m}(\mathbf{o}) = \sum_{\alpha=1}^{n} \omega'(\mathbf{u}_\alpha;\mathbf{o}) \cdot z(\mathbf{u}_\alpha) \tag{3.31}$$

Any location-dependent p-quantile, $\hat{z}_p(\mathbf{o})$, can be obtained by finding the corresponding threshold in the inferred local distribution, such as:

$$\hat{F}(\mathbf{u}, z_p;\mathbf{o}) = \sum_{\alpha=1}^{n} \omega'(\mathbf{u}_\alpha;\mathbf{o}) \cdot I(\mathbf{u}_\alpha; z_p) = p \ \in [0,1] \tag{3.32}$$

Location-Dependent Measures of Spread

The location-dependent variance is estimated as a weighted average of the square of differences between the values and the location-dependent mean (Brunsdon et al. 2002). This is obtained from:

$$\hat{\sigma}^2(\mathbf{o}) = \sum_{\alpha=1}^{n} \omega'(\mathbf{u}_\alpha;\mathbf{o})\left[z(\mathbf{u}_\alpha) - \hat{m}(\mathbf{o})\right]^2 \tag{3.33}$$

Another measure of spread is the location-dependent interquartile range (LDIQR). As in the stationary case (Isaaks & Srivastava 1989, p.20), it is given by the difference between the upper and lower location-dependent quartiles:

$$LDIQR(\mathbf{o}) = \hat{z}_{0.75}(\mathbf{o}) - \hat{z}_{0.25}(\mathbf{o}) \tag{3.34}$$

At the left side of Figure 3-21 an example of local standard deviation model is presented along with the local mean model. In positively skewed distributions, the location-dependent standard deviation tends to be proportional to the location-dependent mean, as observed in this figure. An example of location-dependent quartiles is presented for the same dataset in Figure 3-21, right.

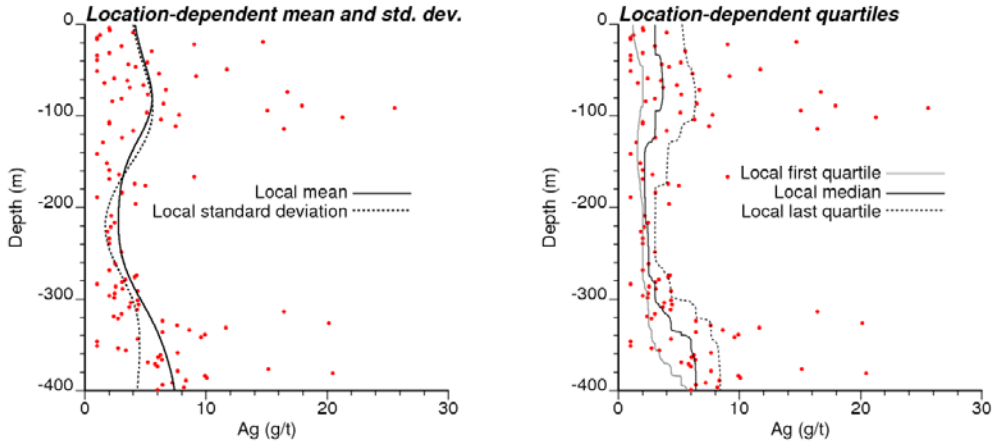

Figure 3-21: Location-dependent mean and variance (left) and location-dependent p25, p50 and p75 quartiles (right) for the 1-D dataset.

Location-Dependent Coefficient of Bivariate Correlation

Local changes in the linear correlation between two different RVs *X* and *Z* may be important. The location-dependent coefficient of correlation can be estimated by (Ren 2007, p.43)

$$\hat{\rho}_{XZ}(\mathbf{o}) = \frac{\hat{C}_{XZ}(\mathbf{o})}{\hat{\sigma}_X(\mathbf{o}) \cdot \hat{\sigma}_Z(\mathbf{o})} = \frac{\sum_{\alpha=1}^{n} \omega'(\mathbf{u}_\alpha; \mathbf{o}) \cdot x(\mathbf{u}_\alpha) \cdot z(\mathbf{u}_\alpha) - \hat{m}_X(\mathbf{o}) \cdot \hat{m}_Z(\mathbf{o})}{\hat{\sigma}_X(\mathbf{o}) \cdot \hat{\sigma}_Z(\mathbf{o})} \quad \in [-1,1] \tag{3.35}$$

For this local statistic it is important to note that the weighting function parameters must be chosen in regard to both variables. In practice, however, these parameters can be chosen considering the most important or the most populated of the two variables.

3.3.4. Location-Dependent 2-Point Statistics

The idea of using smoothly changing distance weights to obtaining the local statistics can be extended to the 2-point measures of spatial continuity. In this case, the 2-point distance weights defined in Subsection 3.2.5 are used in weighted measures of spatial continuity. The forms used under the standard assumptions of stationarity can be regarded as a particular case of the locally weighted measures of spatial continuity where the distance weights are similar for all samples within the domain. The locally weighted or location-dependent semivariogram, covariance and correlogram are presented below.

The same one-dimensional single drillhole dataset is used for illustrating the location-dependent measures of spatial continuity. Although the notation for the attribute values in original units, $z(\mathbf{u})$, is used for the sake of generality in the expressions of these 2-point measures, most of the examples shown next use the local normal scores transforms of the drillhole silver grades. Globally or locally normal scores transformed values are preferred in the inference of location-dependent 2-point statistics since these minimize the effect of extreme values and their corresponding models are needed as the input of locally stationary multiGaussian estimation and simulation techniques.

Samples are not always located in perfectly regular grids, thus when calculating the experimental 2-point statistics for a given lag distance and orientation, distance and angular tolerances must be specified in order to collect enough sample pairs (Deutsch & Journel 1998, p.49). The lag distances, directions and tolerances parameters can be chosen in relation to the global experimental variogram. These chosen parameters are assumed to be suitable for the inference of the location-dependent 2-point statistics.

Location-Dependent Variogram

The location-dependent experimental variogram is defined as the weighted semi average of the squared difference between values $z(\mathbf{u}_\alpha)$ and $z(\mathbf{u}_\alpha+\mathbf{h})$ separated by a vector $\mathbf{h}$, such as:

$$\hat{\gamma}(\mathbf{h};\mathbf{o})=\frac{1}{2}\sum_{\alpha=1}^{N(\mathbf{h})}\omega'(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})\cdot\left[z(\mathbf{u}_\alpha)-z(\mathbf{u}_\alpha+\mathbf{h})\right]^2 \tag{3.36}$$

Where $\omega'(\mathbf{u}_\alpha,\mathbf{u}_\alpha+\mathbf{h};\mathbf{o})$ is the standardized 2-point distance weight, as given in Expression 3.25, for the samples at locations $\mathbf{u}_\alpha$ and $\mathbf{u}_\alpha+\mathbf{h}$, and in relation to the anchor point $\mathbf{o}$. The experimental location-dependent variograms at three anchor points in a single drillhole are shown in Figure 3-22. Although they were calculated on locally normal scores transformed values, these local variograms exhibit different sills (Figure 3-22, left). Dividing the calculated experimental variogram values by the local sill standardizes the location-dependent variograms (Figure 3-22, right).

The sill of the location-dependent variogram tends to the 2-point weighted semi average of all combinations of pair squares differences:

$$\hat{S}(\mathbf{o})\approx\frac{1}{2\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\omega(\mathbf{u}_\alpha,\mathbf{u}_\beta;\mathbf{o})}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N}\omega(\mathbf{u}_\alpha,\mathbf{u}_\beta;\mathbf{o})\cdot\left[z(\mathbf{u}_\alpha)-z(\mathbf{u}_\beta)\right]^2 \tag{3.37}$$

This is not equal to the location-dependent variance presented in Expression 3.33. Figure 3-27 shows that while the local variance of locally normal scores transformed values is very close to 1 everywhere, the variogram sill may vary considerably from one place to another.

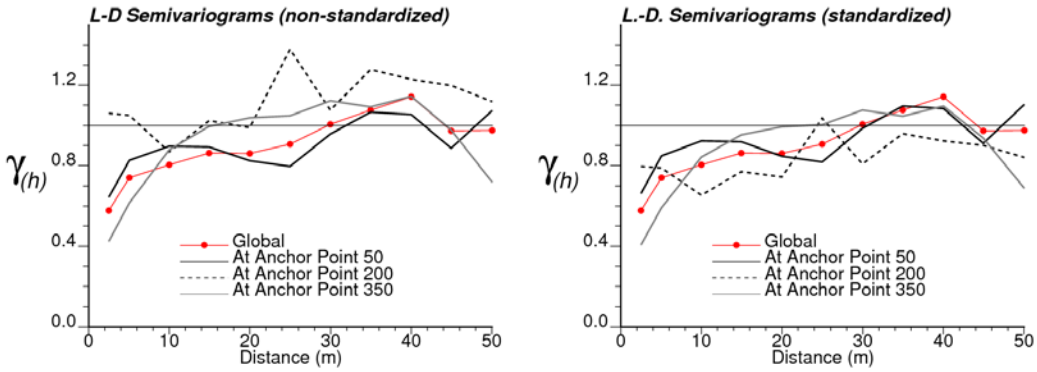

Figure 3-22: Location-dependent experimental variograms before (left) and after (right) standardization by the local sill. These variograms were calculated on locally normal scores transformed values of the silver grades in the 1-D dataset.

Location-Dependent Covariance

Similarly, the experimental location-dependent covariance is calculated as:

$$\begin{aligned}\hat{C}(\mathbf{h};\mathbf{o}) &= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot z(\mathbf{u}_\alpha) \cdot z(\mathbf{u}_\alpha + \mathbf{h}) - \hat{m}_{-\mathbf{h}}(\mathbf{o}) \cdot \hat{m}_{+\mathbf{h}}(\mathbf{o}) \\ &= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \left[z(\mathbf{u}_\alpha) - \hat{m}_{-\mathbf{h}}(\mathbf{o}) \right] \left[z(\mathbf{u}_\alpha + \mathbf{h}) - \hat{m}_{+\mathbf{h}}(\mathbf{o}) \right]\end{aligned} \tag{3.38}$$

With $m_{-\mathbf{h}}(\mathbf{o})$ and $m_{+\mathbf{h}}(\mathbf{o})$ being the location-dependent tail and head means, which are obtained from:

$$\begin{aligned}\hat{m}_{-\mathbf{h}}(\mathbf{o}) &= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot z(\mathbf{u}_\alpha) \ , \\ \hat{m}_{+\mathbf{h}}(\mathbf{o}) &= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot z(\mathbf{u}_\alpha + \mathbf{h})\end{aligned} \tag{3.39}$$

The use of 2-point weights for calculating the location-dependent 2-point tail and head means allows the equivalence expressed by Equation 3.38. Moreover, if **h** = 0, $\omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) = \omega'(\mathbf{u}_\alpha;\mathbf{o})$, thus, $\hat{m}_{-\mathbf{h}}(\mathbf{o}) = \hat{m}_{+\mathbf{h}}(\mathbf{o}) = \hat{m}(\mathbf{o})$, and the location-dependent covariance becomes the 1-point location-dependent variance:

$$\hat{C}(0;\mathbf{o}) = \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha;\mathbf{o}) \cdot \left[z(\mathbf{u}_\alpha) \right]^2 - \hat{m}^2(\mathbf{o}) = \hat{\sigma}^2(\mathbf{o}) \tag{3.40}$$

The local tail and head mean values filter the variations in the local mean at different distances. This allows the location-dependent covariance to be a more robust measure of spatial continuity than the location-dependent variogram. Figure 3-23, left, shows the experimental location-dependent covariances of locally normal scores transformed values at the same anchor points as the location-dependent variograms in Figure 3-22.

Location-Dependent Correlogram

The location-dependent covariance accounts for departures of location-dependent lag means from the local mean inferred from the Expression 3.31. The location-dependent correlogram also takes into account the departures of the lag variances. This is calculated as:

$$\hat{\rho}(\mathbf{h};\mathbf{o}) = \frac{\hat{C}(\mathbf{h};\mathbf{o})}{\sqrt{\hat{\sigma}^2_{-\mathbf{h}}(\mathbf{o}) \cdot \hat{\sigma}^2_{+\mathbf{h}}(\mathbf{o})}} \in [-1,+1] \tag{3.41}$$

Where $\hat{\sigma}^2_{-\mathbf{h}}(\mathbf{o})$ and $\hat{\sigma}^2_{+\mathbf{h}}(\mathbf{o})$ are the tail and head variances, respectively. As for the location-dependent tail and head means, the location-dependent tail and head variances are calculated using 2-point weights, such as:

$$\begin{aligned} \hat{\sigma}^2_{-\mathbf{h}}(\mathbf{o}) &= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot \left[z(\mathbf{u}_\alpha) - \hat{m}_{-\mathbf{h}}(\mathbf{o}) \right]^2 , \\ \hat{\sigma}^2_{+\mathbf{h}}(\mathbf{o}) &= \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot \left[z(\mathbf{u}_\alpha + \mathbf{h}) - \hat{m}_{+\mathbf{h}}(\mathbf{o}) \right]^2 \end{aligned} \tag{3.42}$$

The use of the 2-point weights in the location-dependent head and tail variances is in concordance with the weighting of the location-dependent head and tail means.

Among the three location-dependent measures of spatial continuity presented so far, the location-dependent correlograms are the most robust. This is due to the incorporation of the local means and variances in its calculation. Moreover, its interpretation is straightforward. Figure 3-23 shows the experimental location-dependent covariances (left) and correlograms (right) at the same anchor points as considered before. As it can be observed in this figure, when calculated on locally normal scores transformed values, the location-dependent correlograms are very similar to the location-dependent covariances.

The product of tail and head local means tend to approach the square of the local mean as the lag distance increases (see Figure 3-24, left). In the same way, the geometric average of the tail and head local variances approaches the location-dependent variance for long lag separations (see Figure 3-24, right). These fluctuations of the experimental head and tail local means and variances at short lag distances are contrary to the expected from Expressions 3.39, 3.40 and 3.42. In theory, the local head and tail means and variances should approach the local 1-point mean and variances as the modulus of vector **h** tends to zero. However, the lack of sample pairs separated by short distances causes the fluctuations observed for the short lags in Figure 3-24, particularly in the less sampled areas.

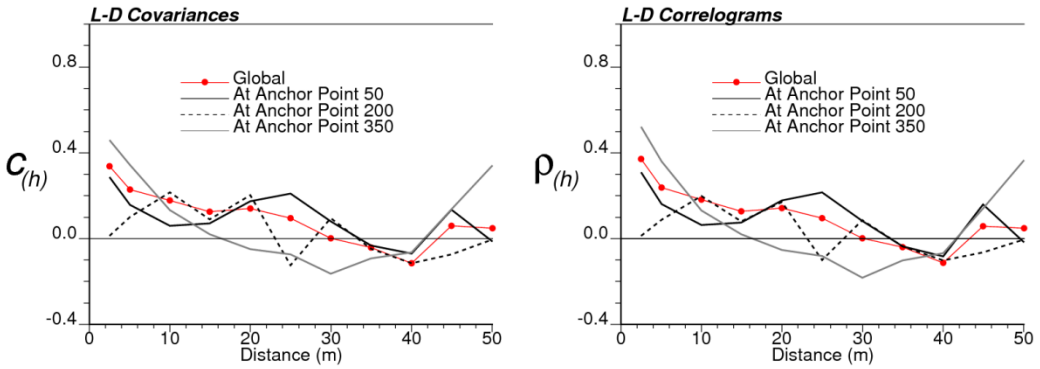

Figure 3-23: Location-dependent correlograms at anchor points located at elevations of -50, -200 and -350 along the drillhole presented previously.

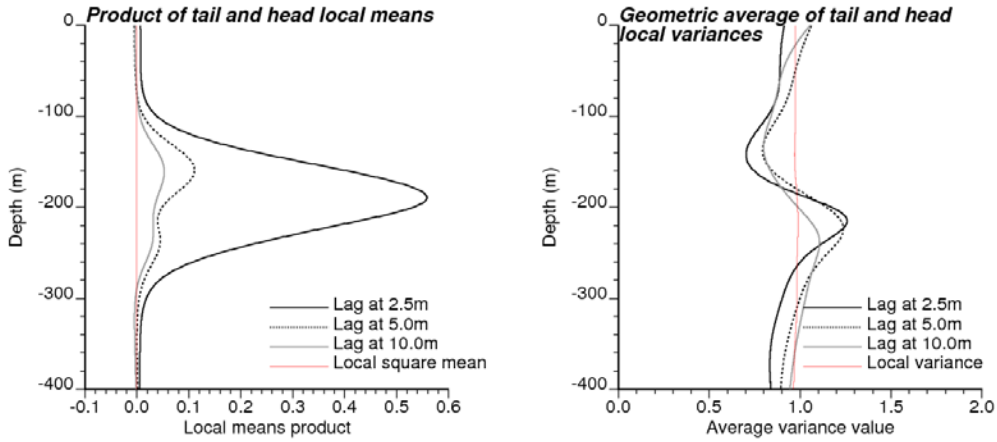

Figure 3-24: Product of tail and head local means (left) and geometric average of tail and head local variances (right) for the first three lags (1-D dataset).

Location-Dependent Indicator 2-Point Statistics

The formulation of location-dependent indicator variograms, covariances and correlograms is focused on categorical variables, since categorical attributes can be defined in a very limited number of states. These can be coded by a categorical indicator function of the form:

$$I(\mathbf{u}_\alpha; s_k) = \begin{cases} 1, & \text{if } z(\mathbf{u}_\alpha) = s_k \\ 0, & \text{otherwise} \end{cases} \tag{3.43}$$

This indicator function replaces $z(\mathbf{u}_\alpha)$ and $z(\mathbf{u}_\alpha + \mathbf{h})$ in the calculation of the location-dependent indicator semivariogram and covariance. In the last, the location-dependent lag means in Expression 3.39 are replaced by the location-dependent lag proportions:

$$\hat{F}_{-\mathbf{h}}(s_k;\mathbf{o}) = \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot I(\mathbf{u}_\alpha; s_k)$$
$$\hat{F}_{+\mathbf{h}}(s_k;\mathbf{o}) = \sum_{\alpha=1}^{N(\mathbf{h})} \omega'(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h};\mathbf{o}) \cdot I(\mathbf{u}_\alpha + \mathbf{h}; s_k) \quad (3.44)$$

While, in the calculation of the location-dependent correlogram, the indicator variances are calculated as:

$$\hat{\sigma}^2_{-\mathbf{h}}(s_k;\mathbf{o}) = \hat{F}_{-\mathbf{h}}(s_k;\mathbf{o})\left[1 - \hat{F}_{-\mathbf{h}}(s_k;\mathbf{o})\right]$$
$$\hat{\sigma}^2_{+\mathbf{h}}(s_k;\mathbf{o}) = \hat{F}_{+\mathbf{h}}(s_k;\mathbf{o})\left[1 - \hat{F}_{+\mathbf{h}}(s_k;\mathbf{o})\right] \quad (3.45)$$

3.3.5. Behaviour of the Location-Dependent 2-Point Statistics

The location-dependent 2-point statistics are able to capture the local changes in the spatial continuity; however, they are sensitive to the parameters of the distance weighting function, the mixture rule used to obtain the 2-point weights and data density. As for the 1-point statistics, narrow kernel bandwidths can result in highly fluctuating 2-point statistics, but very wide bandwidths can result in the oversmoothing of the local variations of the spatial continuity. The choice mixture rule parameter impacts the height at which the variogram stabilizes. A value of 0 for this parameter is suggested to minimize the underestimation or overestimation of the local sill height. Location-dependent correlograms can be more resilient than location-dependent variograms to the choice of the distance function parameters and the mixture rule, as well as the use of locally vs. globally transformed values. Data scarcity hinders the inference of the local 2-point statistics, particularly at short lag distances. This subsection discussed these issues.

Effect of Incorporating Declustering Weights and Dynamic Bandwidths

The correction of distance weights by declustering weights and the use of locally adapting bandwidths can modify the shape of local measures of spatial continuity. The experimental location-dependent variograms in sparsely sampled areas are particularly sensitive to these modifications. As shown on Figure 3-25, when distance weights are corrected by declustering weights, the experimental local variograms are mainly altered by the effect of the distance weights modifications on the local sill. Generally, location-dependent correlograms prove to be less sensitive to these modifications of the distance weights.

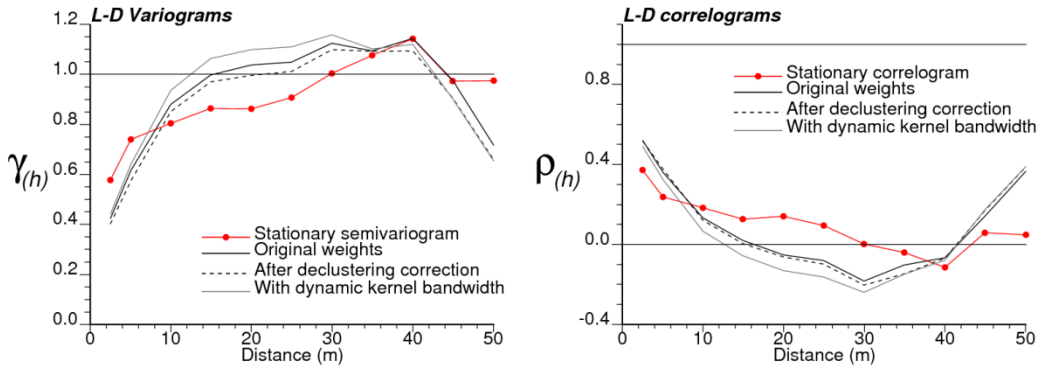

Figure 3-25: Location-dependent variograms using unmodified Gaussian kernel weights and weights modified by declustering correction and dynamic kernel bandwidths (1-D dataset).

Effect of the Mixture Rule Parameter

Location-dependent semivariograms calculated using different values of the mixture rule parameter *t* (see Expression 3.23) are very similar at short lag distances. As the separation given by the norm of vector **h** increases, the experimental location-dependent semivariograms calculated with different values of *t* increasingly diverge (see Figure 3-26, left). This divergence can be explained by considering that, as explained in Subsection 3.2.5, for positive *t* values the 2-point weights approach the closest sample weight. Thus, when the norm of **h** increases the contribution of pairs formed by samples located at very dissimilar distances to the anchor point also increases. This normally translates in higher values of the location-dependent variogram, and also of the local sill calculated by Expression 3.37, since samples located far away tend to have very different values. Contrarily, when a negative *t* parameter is used, the contribution of pairs formed by samples at very dissimilar distances to the anchor point is diminished. This may translate in a lower local sill, which approaches to the local 1-point variance. Figure 3-27 shows the effect of the *t* parameter on the local sill. When these local sills are used to standardize the location-dependent variogram, they may overestimate the standardized local sill if *t* is lower than zero, and underestimate it if *t* is positive (see Figure 3-26, right). A *t* value of zero, which corresponds to the geometric mean of the tail and head sample weights, is a balanced choice that minimizes the issues related to the overestimation or underestimation of the 2-point local variability.

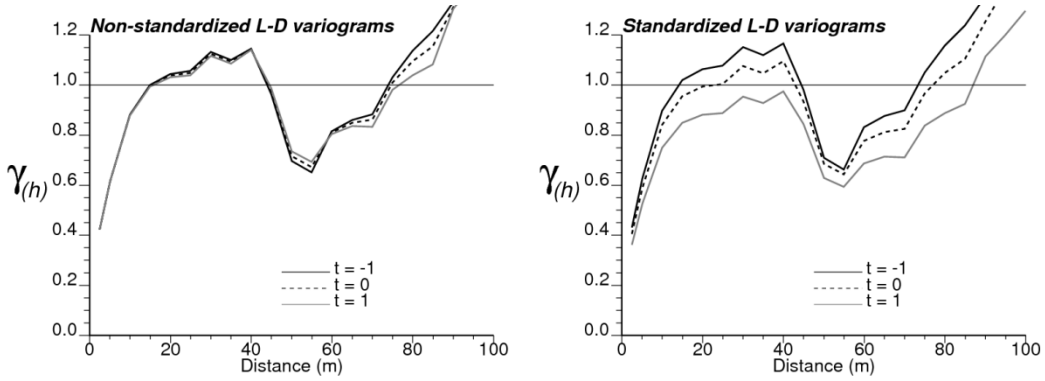

Figure 3-26: Non-standardized (left) and standardized (right) location-dependent variograms for the anchor point located at z =-350 using different exponential parameters for the mixture rule (1-D dataset).

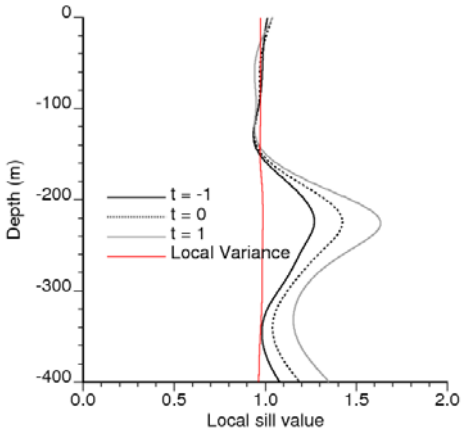

Figure 3-27: Location-dependent variance and semivariogram sill calculated for the locally normal scores transformed Ag values in the 1-D dataset and using different exponential parameters for the mixture rule.

The location dependent correlograms are not highly affected by the choice of the mixture rule parameter choice in the local sill (see Figure 3-28). This is because the correlogram value is standardized at every lag by its corresponding tail and head variances, while the standardization of location-dependent variograms is performed with the average of all lag values.

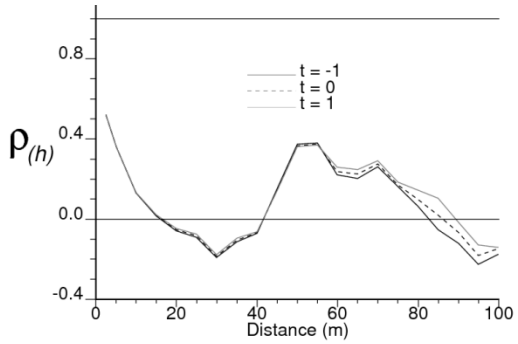

Figure 3-28: Location-dependent correlograms using different exponential parameters for the mixture rule (1-D dataset).

Behaviour at Short Distances

Local changes in the short scale spatial variability are reflected in the shorter lag distances. Aspects like the discontinuity at the origin and the slope are crucial for choosing variogram model parameters (Isaaks & Srivastava 1989, p.376; Sinclair & Blackwell 2002, p.199). Due to the importance of the short distance experimental points in the fitting of the variogram model (Deutsch 2002, p.134) a closer look to these in the location-dependent 2-point statistics is considered below.

The nugget effect is usually inferred in relation to the discontinuity of the experimental variogram close to the origin. This discontinuity is calculated from very closely spaced sample pairs. Usually dense sampling is performed preferentially in high grade zones. In that case, the short lag values of location-dependent 2-point statistics are not representative locally. If additionally, the low value areas exhibit a continuous behaviour while the high values areas present more variability, the bias is increased and the local 2-point statistics may appear excessively variable at short distances. This issue is illustrated on the same single drillhole dataset as presented before, which shows densely sampled high grade intervals at the bottom and top, and a sparsely sampled low grade interval at the middle. Figure 3-29, left, shows that the shorter scale lag values at this section are excessively high in relation to the low local sill at the middle of the drillhole. Thus, when the location-dependent variograms are standardized, anomalously high short scale experimental variogram values appear in the mid region (see Figure 3-29, right).

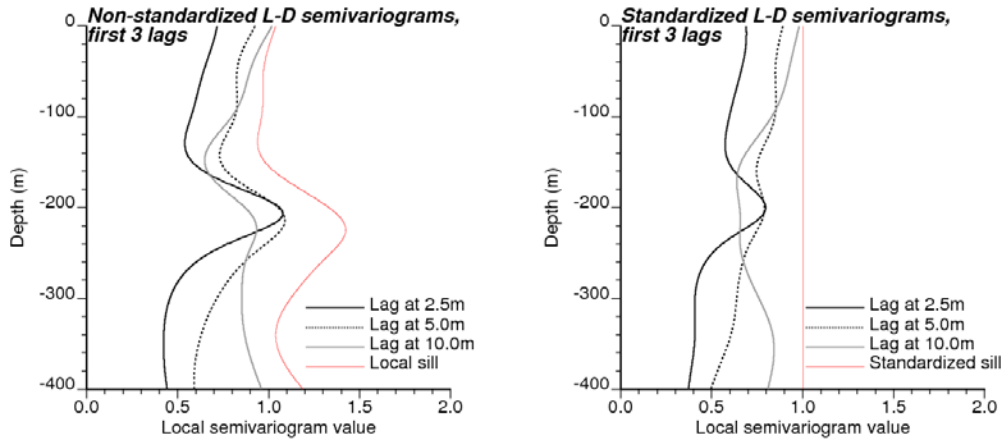

Figure 3-29: Values of non-standardized (left) and standardized (right) location-dependent semivariograms for the first three lags (1-D dataset).

In the location-dependent covariances and correlograms, the shortest lag bias caused by the local absence of closely spaced data is manifested as anomalously low short-scale values. The slope of the local covariances and correlograms becomes positive for this example (see Figure 3-30). When fitting the location-dependent variograms, a way to take into account this issue is to neglect or diminish the relevance of the locally biased experimental values. This idea is developed in Subsection 3.4.2.

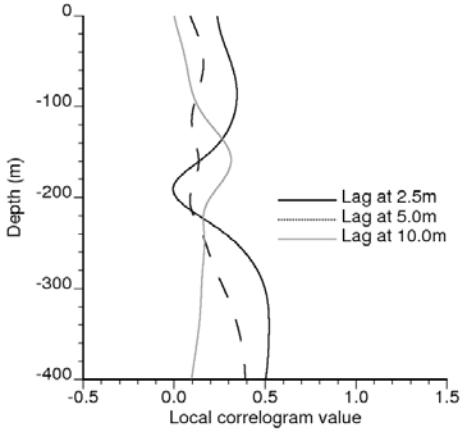

Figure 3-30: Values of location-dependent correlograms for the first three lags (1-D dataset).

Measures of Spatial Continuity on Globally and Locally Transformed Values

Location-dependent semivariograms calculated on globally transformed values differ from those calculate on locally transformed values mostly in the sill. The local variances and sills calculated on globally transformed values are not necessarily close to 1 (see Figure 3-31, left). For locally transformed values, the local variance approaches 1

everywhere, but this is not necessarily the case for the local sill (see Figure 3-31, right). The local sill values can be standardized, doing so minimizes the difference between location-dependent variograms obtained from locally and globally transformed values (see Figure 3-32, left).

In the case of location-dependent correlograms the use of local or global normal scores normally yields to very similar results (see Figure 3-32, right).

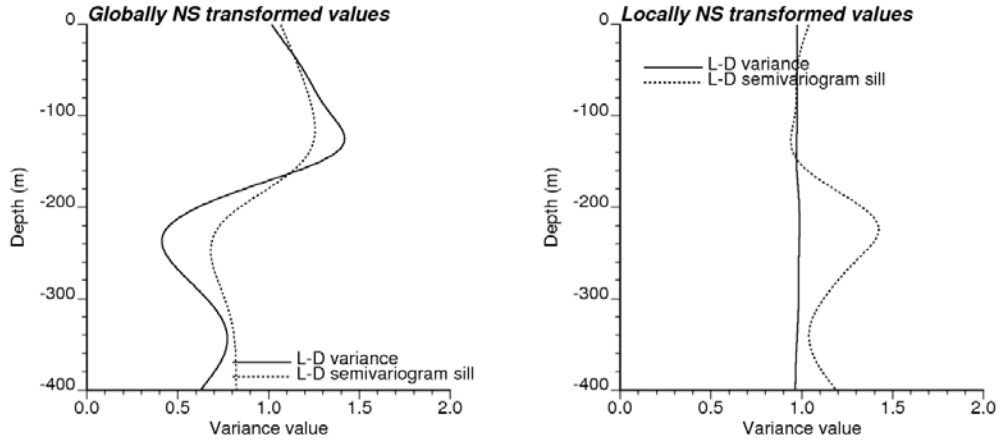

Figure 3-31: Local variances and sill for globally (left) and locally (right) normal scores transformed values of Ag grades in the 1-D dataset.

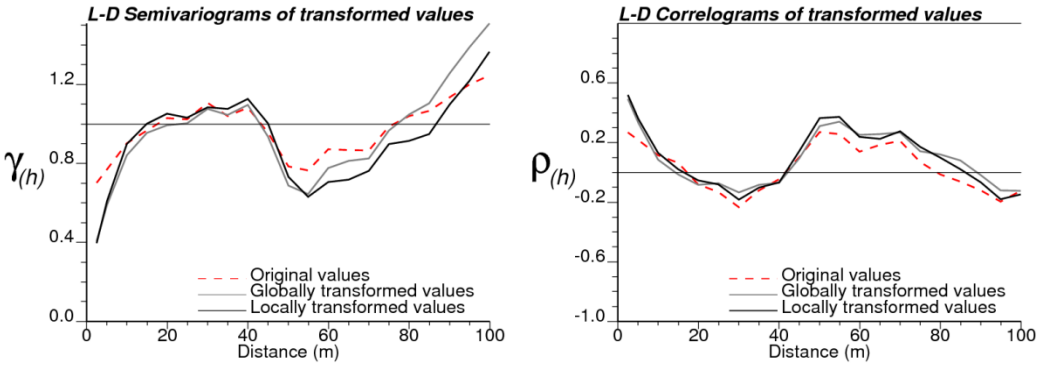

Figure 3-32: Location-dependent semivariograms (left) and correlograms (right) calculated on globally and locally normal scores transformed values of Ag grades in the 1-D dataset.

Capability for Identifying the Local Anisotropy

As shown above, in presence of closely spaced data, local 2-point statistics are sensitive to local changes in the spatial continuity of an attribute. Here, the capability of the location-dependent measures of spatial continuity to detect changes in the local anisotropy of the attribute's spatial distribution is assessed with the help of a small synthetic example of 100 x 250 pixels size. This example was generated using Sequential

Gaussian Simulation and contains two anisotropic areas at the east and west, and an isotropic transition area at the middle (see Figure 3-33). The anisotropy ratio of the east and west areas is 2.5/1. The orientation of the major anisotropic axis is NS in the east area and EW in the west area. The image was sampled in a semi regular grid of 5x5 pixels. The resulting dataset was used for calculating location-dependent variograms and correlograms weighted by Gaussian kernel functions of different bandwidth.

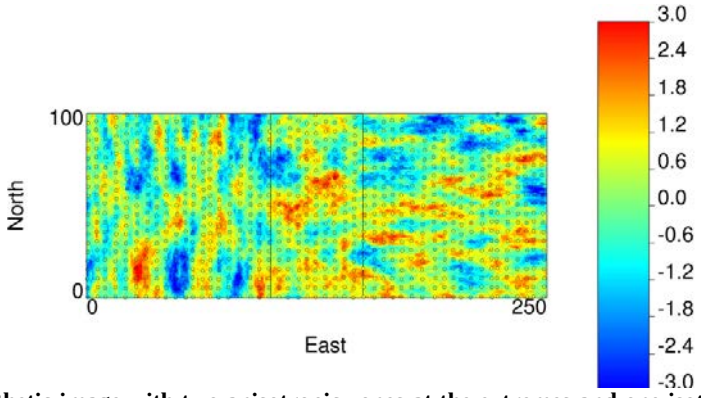

Figure 3-33: Synthetic image with two anisotropic zones at the extremes and one isotropic zone in the middle. The dots in the image correspond to data locations, while the vertical lines delimit the three zones.

Figure 3-34 shows the local correlograms calculated using three different Gaussian kernel bandwidths. These experimental location-dependent correlograms appear as dashed lines coloured by the X-location of their corresponding anchor point. When a narrow bandwidth is used (Figure 3-34, top), the local correlograms exhibit great versatility to adapt to local changes, fluctuating around the "true" experimental correlograms in each region (shown as continuous lines). But this versatility comes with the price of increased noise. A wide bandwidth (Figure 3-34, bottom), leads to local correlograms that approach to the global experimental correlogram (black dots in Figure 3-34). An intermediate bandwidth allows flexible and smoothly changing correlograms (Figure 3-34, middle).

Figure 3-35 shows the average absolute errors between the location-dependent correlograms and semivariograms obtained at different bandwidths from the gridded dataset and the "true" correlograms and semivariograms calculated on each region of the exhaustive image. The u-shape of the curves presented indicate that the precision of the location-dependent 2-point statistics is diminished when narrow bandwidths are used due the increased noise, while for very wide bandwidths the precision is lost due to the increased homogenization. A bandwidth wide enough to allow accommodating roughly

50% to 75% of the Gaussian kernel function within each anisotropic region yields the lowest errors.

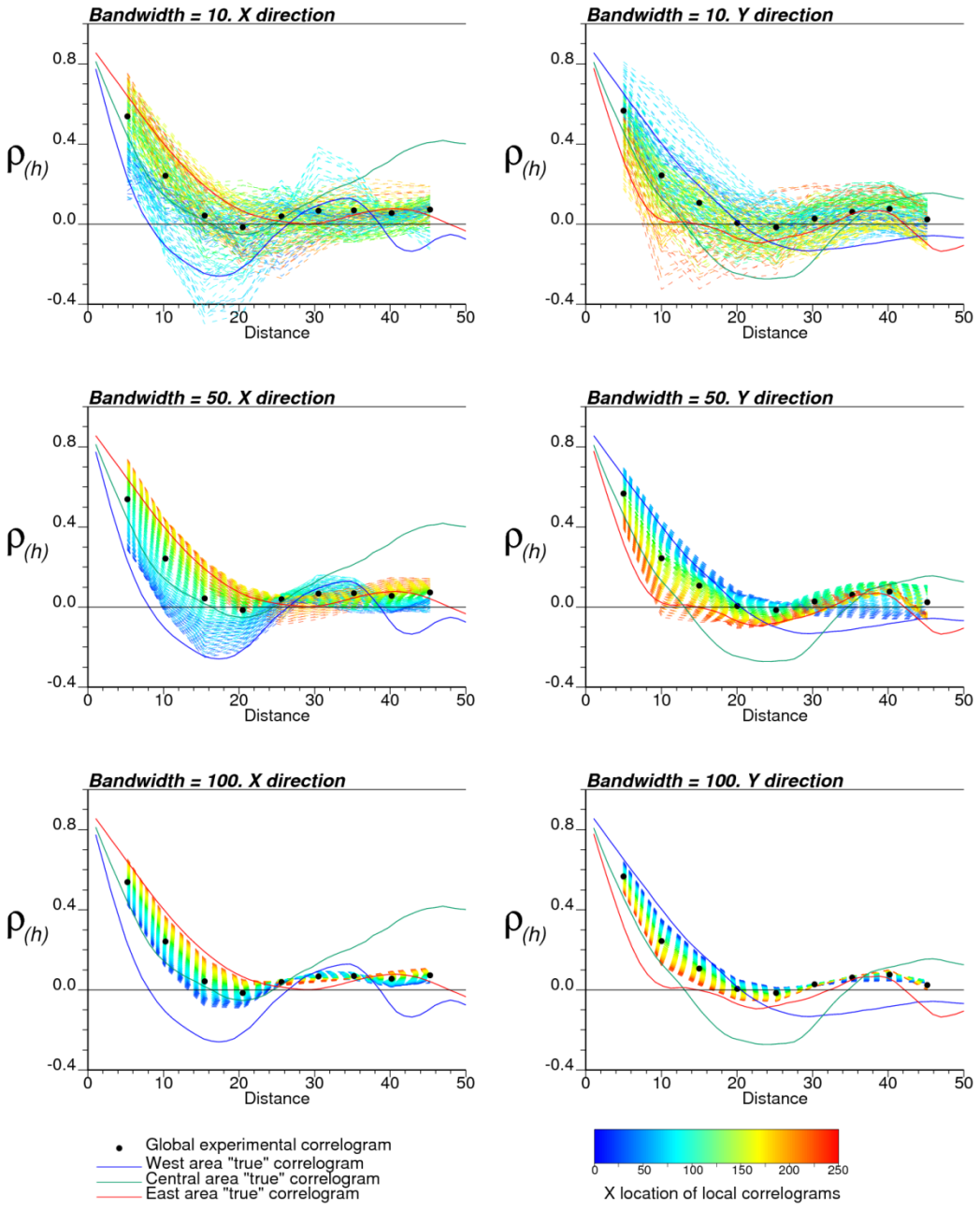

Figure 3-34: Experimental location-dependent correlograms (dashed lines) along with the "true" experimental correlograms calculated directly from the exhaustive image at each region (continuous lines), and the global experimental variograms calculated from the gridded dataset (black dots).

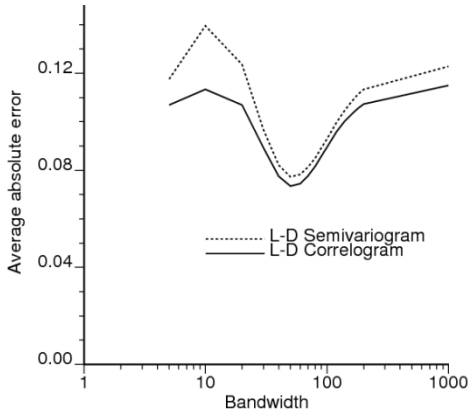

Figure 3-35: Average absolute errors of the location-dependent semivariograms and correlograms in function of the kernel bandwidth.

Impact of Data Density on the Capability for Identifying the Local Anisotropy

Data density is critical in the reliable inference of location-dependent statistics. If data are too sparse, the local statistics may reflect the individual values of nearby samples rather than local tendencies. The ideal case is when samples are close enough to inform the local features with certain redundancy, but their number is limited by a reasonable sampling cost. For 2-point statistics, the impact of low sample density is twofold: first, it precludes the inference of these statistics at short separations; second, it increases the variability of these statistics due to the increased relative contribution of individual samples in the total value of the inferred statistics. This is illustrated with the help of the synthetic image presented in Figure 3-33, which was sampled in semi regular grids of size 1 x 1, 2.5 x 2.5, 5 x 5, 7.5 x 7.5, and 10 x 10 pixels. The corresponding location-dependent experimental correlograms and variograms were calculated using a Gaussian kernel of 50 pixels bandwidth. Figure 3-36 shows the location-dependent correlograms in the East-West and North-South directions obtained from the datasets sampled in grids of size 1 x 1, 5 x 5 and 10 x 10 pixels. The location-dependent correlograms for the sparsest dataset (Figure 3-36, bottom) show the difficulties in identifying the local spatial correlation mentioned above.

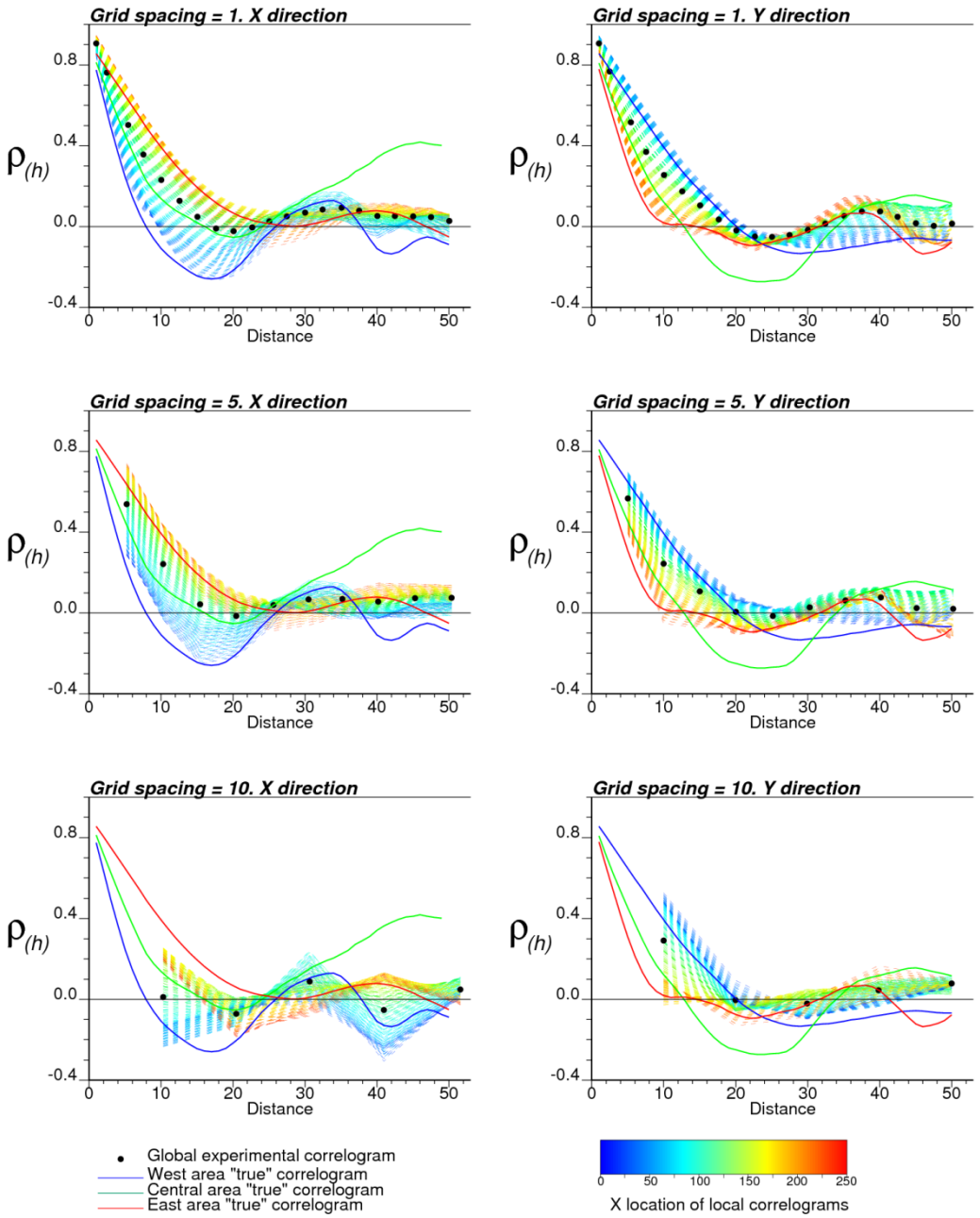

Figure 3-36: Experimental location-dependent correlograms (dashed lines) along with the "true" experimental correlograms calculated directly from the exhaustive image at each region (continuous lines), and the global experimental variograms calculated from the gridded dataset (black dots) at different sampling spacing.

Figure 3-37 show the average absolute errors between the experimental location-dependent variograms and correlograms obtained from the gridded datasets and the experimental stationary variograms and correlograms calculated within each of the three zones in Figure 3-33. The increase in the error as the grid size increases may be relatively

small, particularly if sampling spacing is still short enough to allow redundancy in the inference of the local measures of spatial correlation. Note also that for the 1x1 grid size dataset, which is equivalent to the exhaustive image itself, there is still an error in the reproduction of the local variograms and correlograms. This is because the samples in the three anisotropy zones in the image were treated as belonging to a single domain. The resulting transitional local correlograms observed in Figure 3-36, and their variogram counterparts, lead to this error even with dense data. In real world cases, the best practice, if the information is abundant enough to identify different domains, is to delimit them and treat them separately. Locally weighted statistics are best used when non-stationary changes occur gradually within a rather homogenous domain or when transitions between domains are not clear and can be assumed to be smooth.

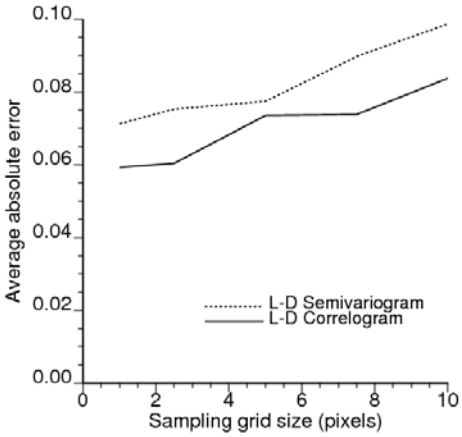

Figure 3-37: Average absolute errors of the location-dependent semivariograms and correlograms in function of the sampling grid spacing.

3.4. Modelling the Location-Dependent Parameters

Models of location-dependent statistics allow describing different non-stationary features. Instead of storing the local Gaussian transforms of each data value, only a limited number of local Hermite coefficients are required for modelling the location-dependent Gaussian transformations. For location-dependent variograms, the local anisotropy, the short scale variability and other local features of the spatial continuity are described by a few variogram model parameters. The local normal scores transformation models must minimize the introduction of any bias in the local cdfs, while the local variogram models must produce non-negative variances for any 2-point combination.

The modelling of local normal scores transformations is performed automatically using Hermite polynomial series with location-dependent coefficients. Fully automatic modelling of variograms has been normally discouraged in favour of manual techniques that allow the incorporation of geological sense by the practitioner (Gringarten & Deutsch 2001; Webster & Oliver 2007, pp.101-102). However, when dealing with location-dependent variograms calculated at multiple anchor point locations the efficient option is to resort to a semiautomatic algorithm for fitting the variograms under parameter constraints and guides imposed by the user.

As mentioned in Section 3.2.6, the local cdfs and measures of spatial continuity are inferred and modelled at a limited number of anchor points for the sake of increased efficiency. The local parameters required for locally-stationary spatial prediction techniques are subsequently interpolated between anchor points.

3.4.1. Modelling the Local Normal Scores Transformation

The local nonparametric modelling of the normal scores transformation and back-transformation tables requires the Gaussian equivalent values for all n samples in the dataset at each anchor point. This requires a significant amount of computer memory, particularly for large datasets. A more efficient way is to store the equivalent Gaussian values for a limited number of locally inferred $\hat{z}_p(\mathbf{o})$ quantiles. Usually, between 100 and 200 quantiles are adequate for nonparametric modelling of the local transformation functions.

Another efficient way is to approximate these functions by a series of Hermite polynomials. At each anchor point a number P of local quantiles is normal scores transformed. This transformation is approximated as (Journel & Huijbregts 1978, pp.472-478; Wackernagel 2003, pp.238-249):

$$\hat{z}_p(\mathbf{o}) = \varphi_Z(y_p;\mathbf{o}) \cong \sum_{q=0}^{Q} \phi_q(\mathbf{o}) H_q[y_p] \tag{3.46}$$

The reasons for using the local quantiles instead of the actual data values in the modelling of the local transformation are twofold. First, the local quantiles already embed information about the local cdf shape without need of applying the distance weights again, and second, since the quantiles in Gaussian units do not change, there is no need of recalculating the Hermite polynomial at each anchor point. Only the Hermite coefficients,

$\phi_q(\mathbf{o})$, change . Thus, the location-dependent coefficients $\phi_q(\mathbf{o})$ are obtained from the sum (Oz et al. 2002):

$$\phi_q(\mathbf{o}) = \sum_{p=2}^{P} (\hat{z}_{p-1}(\mathbf{o}) - \hat{z}_p(\mathbf{o})) \cdot \frac{1}{\sqrt{q}} H_{q-1}(y_p) \cdot g(y_p) \tag{3.47}$$

Where $\phi_0(\mathbf{o}) = \hat{m}(\mathbf{o})$, and $g(y_p)$ is the Gaussian probability density function. The Hermite polynomials are developed as (Rivoirard 1990, p.27):

$$\begin{aligned} &H_0(y) = 1 \\ &H_1(y) = -y \\ &H_{q+1}(y) = -\frac{1}{\sqrt{q+1}} y H_q(y) - \sqrt{\frac{q}{q+1}} H_{q-1}(y) \end{aligned} \tag{3.48}$$

The Hermite polynomials are independent of the location. For any integer $q > 0$, and if $Y(\mathbf{o})$ is a standard Gaussian distributed variable, they have zero mean and variance equal to one (Rivoirard 1990, p.28):

$$\begin{aligned} &E\{H_q(Y(\mathbf{o}))\} = \int H_q(y) g(y) dy = 0 \\ &Var\{H_q(Y(\mathbf{o}))\} = E\{H_q(Y(\mathbf{o}))^2\} = 1, \qquad \forall q \geq 1 \end{aligned} \tag{3.49}$$

The number of polynomials, *Q*, must be chosen in order to model the variability of *Z*:

$$\hat{\sigma}_Z^2(\mathbf{o}) \cong \sum_{q=1}^{Q} \phi_q^2(\mathbf{o}) \tag{3.50}$$

Figure 3-38, left, shows the Hermite polynomials fitting to the location-dependent transformation function at one anchor point in the drill hole profile previously presented. A higher number of polynomials yields to a closer fit of the nonparametric transformation function, but, unwarranted fluctuations may appear. These are common when the transformation function shows spikes and breaks. This issue is akin to the problems encountered when fitting the indicator function by Hermite polynomials series (Chilès & Delfiner 1999, pp.416-417). Thus, the practical implementation of the modeling of location-dependent normal scores transformation by Hermite polynomials require the incorporation of the correction of order relation deviations. This procedure is similar to the one applied for the correction of order relation deviations in cdfs built by indicator kriging (Deutsch & Journel 1998, pp.81-86). In regard to these issues, and depending of the histogram smoothness, a good practice is to use between 20 and 40 Hermite polynomials (Vann & Sans 1995; Wackernagel 2003, p.247). Thus, the transformation

function can be closely fitted and the order relation problems are minimized, with the additional benefit of limiting the storage requirements for the complete set of location-dependent Hermite coefficients. Figure 3-38, right, show the values of the local Hermite coefficients one to five. The values of the local Hermite coefficients at two adjacent anchor points are very similar despite small fluctuations. The continuous change of the Hermite coefficients between very closely separated anchor points suggest this variation can be reconstructed by interpolation if the anchor points are located within reasonable separations from each other.

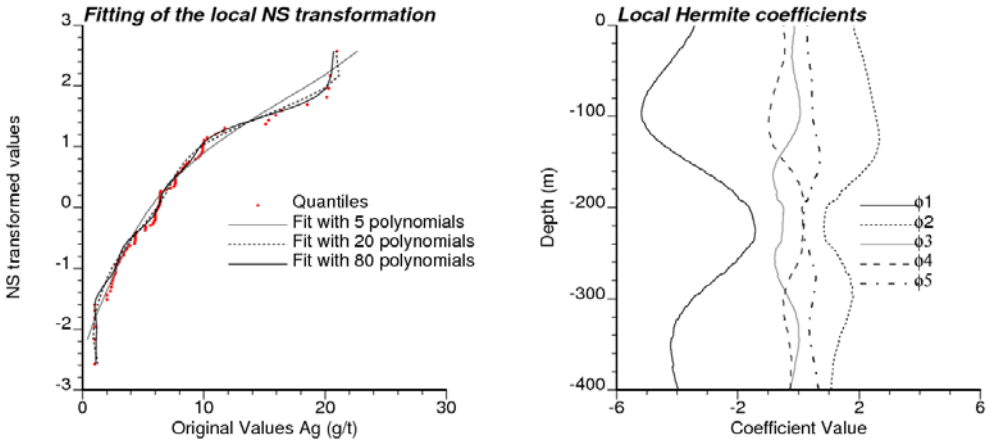

Figure 3-38: Left, Hermite polynomials fitting to a local normal score transformation function. Note the difficulty to fit the spikes in the transformation function. Right, values of the local coefficients one to five resulting from the Hermite polynomials fitting of the local NS transformation functions defined for the 1-D dataset.

3.4.2. Location-Dependent Variogram Models

Location-dependent experimental measures of spatial correlation must be modelled. As with their stationary counterparts, the reasons for this include (1) the need of a complete definition of the spatial correlation for all distances and orientations, and (2) the requirement of positive definiteness for the covariance used in estimation and simulation (Gringarten & Deutsch 2001). Variogram models can be useful for incorporating geological information related to the spatial distribution of the attribute. While fitting the local variograms models, this geological knowledge can be incorporated as locally changing orientations of the spatial continuity derived from geological interpretations or field measurements. If available, this information can be used to guide the fitting of the local anisotropy parameters.

The list of permissible models for fitting the location-dependent measures of spatial correlation is the same as for stationary experimental variograms. The only difference with the stationary variograms is that local variogram models incorporate locally changing parameters.

A common choice is the spherical model (Matheron 1969, p.41); its local adaptation is given by:

$$\gamma(\mathbf{h};\mathbf{o}) = \gamma^t(\mathbf{h}';\mathbf{o}) = \begin{cases} c(\mathbf{o}) \cdot \left[1.5|\mathbf{h}'| - 0.5|\mathbf{h}'|^3 \right], & \text{if } |\mathbf{h}'| \le 1 \\ c(\mathbf{o}), & \text{if } |\mathbf{h}'| \ge 1 \end{cases} \tag{3.51}$$

Where $\gamma^t(\mathbf{h}';\mathbf{o})$ is the local isotropic and dimensionless variogram model with range equal to 1 and $c(\mathbf{o})$ is its sill contribution. The vector $\mathbf{h}' = (h_{x'}, h_{y'}, h_{z'})^T$ is obtained from a transformation of $\mathbf{h} = (h_x, h_y, h_z)^T$ with components in the original coordinate system. This transformation is similar to the given by Expressions 3.14 and 3.15, but in this case the scaling matrix $[\mathbf{S}]$ is replaced by:

$$[\mathbf{S}] = \begin{bmatrix} 1/a_{x'}(\mathbf{o}) & 0 & 0 \\ 0 & 1/a_{y'}(\mathbf{o}) & 0 \\ 0 & 0 & 1/a_{z'}(\mathbf{o}) \end{bmatrix} \tag{3.52}$$

Where $a_{y'}(\mathbf{o})$ and $a_{x'}(\mathbf{o})$ are the local range parameters parallel to the rotated y and x directions, and $a_{z'}(\mathbf{o})$ is the local range in the rotated vertical direction. Moreover, the rotation matrices in Expression 3.14 are modified by the local anisotropy angles $\theta_1(\mathbf{o})$, $\theta_2(\mathbf{o})$ and $\theta_3(\mathbf{o})$.

Greater flexibility in the fitting of location-dependent measures of correlation can be achieved by using a model that is able to change its shape. The stable model (Chilès & Delfiner 1999, pp.88-90) offers this capability:

$$\gamma(\mathbf{h};\mathbf{o}) = c(\mathbf{o}).\left[1 - \exp\left(-\left(3|\mathbf{h}'| \right)^{b(\mathbf{o})} \right) \right] \qquad 0 < b(\mathbf{o}) \le 2 \tag{3.53}$$

This variogram model contains the exponential and Gaussian models, when $b(\mathbf{o}) = 1$ and $b(\mathbf{o}) = 2$, respectively (see Figure 3-39). In areas where spatial continuity is smooth a higher value for the parameter $b(\mathbf{o})$, would be reasonable. Highly discontinuous regions would require lower $b(\mathbf{o})$ values. Since the major shape change of the stable variogram

model function occurs at short lags, a good local fitting of this model requires many closely spaced pairs of samples.

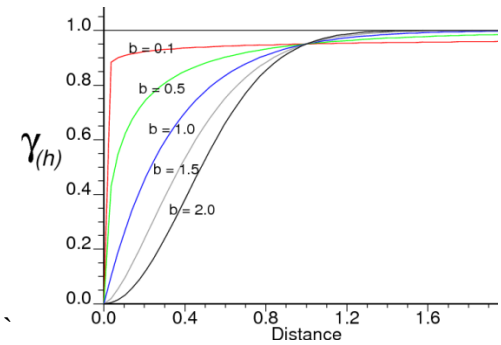

Figure 3-39: Stable model shape according to different power values.

Other models that can be considered for locally changing variogram shapes are the Matérn model (Stein 1999, pp.31-32; Minasny & McBratney 2005) and the generalized Cauchy model (Chilès & Delfiner 1999, pp.85-86; Chilès 2004, p.5).

A locally changing nugget effect $C_0(\mathbf{o})$ may also be considered. The nugget effect is related to spatial variation at a scale shorter than the smallest sampling separations (Journel & Huijbregts 1978, p.39). Some of the sources of this variation, such as measurement errors, may not be spatially correlated. Therefore, the fitting of location-dependent nugget effect must be controlled or kept at a constant value.

3.4.3. Semi-Automating Fitting of Location-Dependent Variograms

The local spatial correlation is required for all 2-point separation distances, orientations, and all locations where locally-stationary spatial prediction is performed. Thus, the fitting of local variogram models must produce fields of locally changing variogram parameters at the resolution required for spatial prediction. Since the number of prediction locations may range from the thousands to millions, the local variogram models are fitted only for a limited number of anchor points. The resulting local variogram parameters are interpolated between those anchor points. Nevertheless, even the number of anchor points can be too large to consider a manual fitting. Therefore, a semiautomatic variogram fitting algorithm is required to fit the local variogram models at the anchor points. Currently available semi-automatic fitting algorithms based on the minimization of the square errors can be used for this task.

In absence of geologically supported discontinuities the local variogram model parameters should vary smoothly with the location-dependent statistics and within reasonable limits. Smoothly changing local variogram parameters allow a consistent reconstruction of the local parameters by interpolation between the anchor point locations. However, the semiautomatic and independent fitting of local variogram models from one anchor point to the next may result in excessively low or high parameters values or in unwarranted abrupt changes in the local parameters. Abruptly changing local variogram parameters may occur when the local experimental variogram points are highly discontinuous, particularly at short lags. Moreover, if the values of the local experimental variogram points at the shorter lags are high, this may result in geologically unrealistic local nugget effect values. The presence of the hole effect (Isaaks & Srivastava 1989, p.156; Pyrcz & Deutsch 2001) in the local experimental variograms can also result in excessively high and abruptly fluctuating local variogram parameters when fitting monotonic models, such as the spherical or the stable. In these cases, a model with a high nugget effect and very long range can yield a similar least square fitting error as a model with a low nugget effect and short range. Figure 3-40, left, shows the local nugget effect and variogram range values fitted to the local experimental correlograms using the 1-D dataset. The very high local nugget effect values and the abrupt changes in both local parameters are due to the issues discussed above. Figure 3-40, right, shows the minimized average square error of the local variogram model fitting.

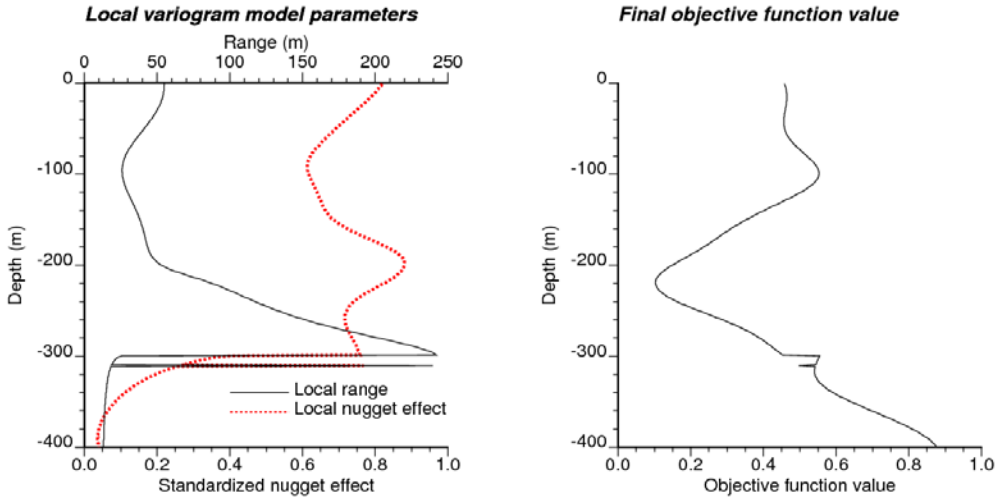

Figure 3-40: left, resulting local variogram model parameters fitted separately to the location-dependent experimental correlograms calculated at each anchor point using the 1-D dataset. Right, least square error of the local variogram fitting.

When data is abundant and dense enough in relation to the scale of the true spatial continuity within the domain, the issues discussed above may occur only at a few anchor point locations. In such case, a manual fitting of the local variogram models at problematic locations may suffice.

Another option is to use a fitting algorithm that enforces the consistency between local variogram models and allows an increased control of the parameter values. Thus, the main criteria for a proposed fitting algorithm for series of local variogram models is twofold: (1) the minimization of the mean square differences between the local experimental variograms and their corresponding proposed models, and (2) the consistency between model parameters fitted at contiguous anchor points. Both criteria must be achieved while respecting the limits of tolerable parameter values imposed by the user and honouring the local anisotropies derived from the background geological knowledge, if they are available.

Starting from the global variogram model, the fitting algorithm adds small random deviations to the parameters at each anchor point. This is repeated many iterations. If the average square error between the local variogram model and the local experimental variogram points diminishes, the corresponding local parameters are kept for the next iteration. Penalty functions are used to control the departure of the local parameters from the intervals imposed by the user and to enforce smoothly changing local parameters. The main aspects of the proposed algorithm are described below.

Least-Squares Criteria for Fitting the Location-Dependent Variogram Models

The least-squares criterion for the fitting of variogram models has been proposed since the initial years of computational Geostatistics (David 1977, p.119). Since each of the experimental variogram points carry different amounts of information and those calculated at the shorter lags distances are more important, weighted least-squares criteria has been commonly adopted (Goovaerts 1997, p.105). The weights assigned to the experimental variogram points can be directly proportional to the number of sample pairs involved in the calculation of these points, and inversely proportional to the squared model value at the corresponding lags (Cressie 1985) or to the corresponding lag distances (Zhang et al. 1995).

Direct supervision of the variogram models obtained using this criterion is straightforward when fitting a few global variogram models, but not with many local

variogram models. In this case, the fitted local parameter values can be constrained within intervals defined by the user. These intervals can be absolute or relative to values predefined from geological background knowledge.

In a series of iterations, the optimization algorithm slightly and randomly alters the variogram model parameters in order to provide the best fit to the experimental variogram. After each iteration, the goodness of this fit is assessed by the weighted average of the square differences between the proposed model and the experimental points calculated at multiple directions and lag distances. The weighting is proportional to the importance given to the experimental points and the quantity of information used for their inference. Since the experimental values at shorter lags are the most important for depicting the local changes in spatial continuity, thus the weights assigned to these points can be set inversely proportional to their lag distance $|\mathbf{h}|=h$ (Larrondo et al. 2003):

$$\lambda_{dist}(\mathbf{h}_k) = \frac{\frac{1}{h_k}}{\sum_{k=1}^{n_{\text{exp.points}}} \frac{1}{h_k}} \tag{3.54}$$

The number of experimental points, $n_{\text{exp.points}}$ is equal to the total number of experimental points calculated at different directions and for different lag distances. If the same directions and lags separations are used for the calculation of the experimental measures of spatial continuity at all anchor points, these weights are independent of the anchor point locations.

At different anchor points, the experimental location-dependent variogram points at the same lag distance and orientation can have the same number of sample pairs involved in their calculation; however, the availability of close sample pairs change from anchor point to anchor point. Those experimental points inferred with a greater number of close samples should have higher relevance during the fitting of local variograms. Therefore, it is reasonable to weight the experimental points by the sum of the 2-points weights assigned to the sample pairs involved in their calculation. This weighting criterion based on the information available, λ_{inf}, is expressed as:

$$\lambda_{\inf}(\mathbf{h}_j,\mathbf{o}) = \frac{\sum_{\alpha=1}^{N(\mathbf{h}_j)} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}_j; \mathbf{o})}{\sum_{j=1}^{n_{\text{exp,points}}} \sum_{\alpha=1}^{N(\mathbf{h}_j)} \omega(\mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h}_j; \mathbf{o})} \tag{3.55}$$

With $N(\mathbf{h}_j)$ as the number of pairs used to calculate the experimental point at the distance and orientation corresponding to $\mathbf{h}_j$. Notice that these are 2-point weights before the normalization performed by Expression 3.25.

In order to prevent the parameters from taking values that largely exceed the limits judged as reasonable by the user a penalty function is considered. The penalty function can take different forms; this one is a simple quadratic function. Thus if a number n_{par} of local variogram parameters, $b_\beta(\mathbf{o})$, $\beta = 1,...n_{par}$, is to be controlled rather than fixed, a penalty is applied to those values, that exceed a range $(b_{\beta,\min}, b_{\beta,\max})$ imposed by the user:

$$W_\beta\left(b_\beta(\mathbf{o})\right) = \begin{cases} k \cdot (b_\beta(\mathbf{o}) - b_{\beta,\min})^2 & \text{if } b_\beta(\mathbf{o}) \le b_{\beta,\min} \\ 0 & \text{if } b_{\beta,\min} < b_\beta(\mathbf{o}) < b_{\beta,\max} \\ k \cdot (b_\beta(\mathbf{o}) - b_{\beta,\max})^2 & \text{if } b_\beta(\mathbf{o}) \ge b_{\beta,\max} \end{cases} \tag{3.56}$$

The factor *k* controls the strength of the penalty. The higher this value is, the harder is for the algorithm to produce local variogram parameters that exceed the predefined range. Usually, a penalty factor higher than one avoids the occurrence of parameter values beyond the range $(b_{\beta,\min}, b_{\beta,\max})$. A penalty factor smaller than one can be used if the practitioner decides to allow some flexibility in the parameter limits, and thus, to avoid the hard capping of the parameter values.

The minimum and maximum allowable parameter values can be set as absolute or as relative tolerances to previously defined local values. The second form can be useful for allowing a certain degree of flexibility in the fitting of local anisotropy angles guided by values taken from field measurements or the geological interpretation of the deposit.

Thus, given an experimental spatial correlation measure, $\hat{\gamma}(\mathbf{h};\mathbf{o})$, and the proposed model value at the same lag $\mathbf{h}$, $\gamma(\mathbf{h};\mathbf{o})$, the optimization criterion for semi-automatic fitting of the local variogram at an anchor point $\mathbf{o}$ is to minimize the next objective function:

$$O(\mathbf{o}) = \frac{1}{n_{\text{exp.points}}} \sum_{j=1}^{n_{\text{exp.points}}} \lambda(\mathbf{h}_j, \mathbf{o}) \left(\gamma(\mathbf{h}_j; \mathbf{o}) - \hat{\gamma}(\mathbf{h}_j; \mathbf{o}) \right)^2 + \frac{1}{n_{\text{par}}} \sum_{\beta=1}^{n_{\text{par}}} W_\beta \left(b_\beta(\mathbf{o}) \right) \quad (3.57)$$

The weights $\lambda(\mathbf{h}_j, \mathbf{o})$ can take the form of either Expression 3.54 or 3.55, or they can be built as the product of both. The penalties $W_\beta \left(b_\beta(\mathbf{o}) \right)$ are as in Expression 3.56. This minimization criterion does not assure that the parameters of the variogram models fitted at contiguous anchor point will be consistent with each other. Thus, it needs to be complemented with criterions that enforce smoothly changing variogram parameter values.

Consistency Between Variogram Models Fitted at Contiguous Anchor Points

If a wide enough kernel bandwidth is used to obtain the distance weights, the local experimental variograms change smoothly from one anchor point to another (see Figure 3-34); therefore, the models fitted on them should also vary smoothly. In order to enforce the consistency between the models fitted at contiguous anchor points two measures are implemented in the iterative fitting algorithm: (1) penalization of locally anomalous parameter values, and (2) control of the local objective function convergence.

A way to impose continuity in the variation of the local parameters is by penalizing their departure from a local mean. Thus, a penalty function for locally anomalous parameter values is proposed as:

$$W'_\beta \left(b_\beta(\mathbf{o}) \right) = k' \cdot (b_\beta(\mathbf{o}) - \bar{b}_\beta(\mathbf{o}))^2 \quad (3.58)$$

The constant *k'* controls the strength of the penalty, and $\bar{b}_\beta(\mathbf{o})$ is the average local parameter value within a small moving neighbourhood, V(**o**), of anchor points centered at **o**. This neighbourhood, defined by the user, should pick only a limited number of the closest anchor points, included the one located at **o**, in order to avoid oversmoothing of the local variogram parameters. This penalty is incorporated in the minimization of the local objective function, such as:

$$\min\left\{O(\mathbf{o})\right\} =$$
$$\min\left\{\frac{1}{n_{\text{exp.points}}}\sum_{j=1}^{n_{\text{exp.points}}}\lambda(\mathbf{h}_j,\mathbf{o})\left(\gamma(\mathbf{h}_j;\mathbf{o})-\hat{\gamma}(\mathbf{h}_j;\mathbf{o})\right)^2+\sum_{\beta=1}^{n_{\text{par}}}\left[W_\beta\left(b_\beta(\mathbf{o})\right)+W'_\beta\left(b_\beta(\mathbf{o})\right)\right]\right\} \tag{3.59}$$

Additionally, as it can be observed in Figure 3-40, abrupt changes in the local parameters are reflected by abrupt variations in the values towards the local objective functions converge. Since the aim is to minimize the objective function, a procedure for identifying and correcting local extremely high convergence values is implemented. A simple outlier detection criterion is used to identify those values. This criterion is based in the Q test for small datasets (Wellmer 1998, pp.60-61). After a number of iterations, the updated values of the local objective functions within a moving neighbourhood are assembled and ordered in decreasing order. Thus, a Q statistic is obtained from:

$$\mathrm{Q}=\frac{O(\mathbf{o}_a)-O(\mathbf{o}_b)}{O(\mathbf{o}_a)-O(\mathbf{o}_z)} \tag{3.60}$$

Where $O(\mathbf{o}_a)$ is the maximum current value of the objective function within the neighbourhood V($\mathbf{o}$), $O(\mathbf{o}_b)$ is the second highest value, and $O(\mathbf{o}_z)$ is the minimum value within the same neighbourhood. If $O(\mathbf{o}_a)=O(\mathbf{o}_z)$ it means that the convergence values of the objective function are the same for all anchor points in the neighbourhood; therefore, no outlier detection procedure is needed. Otherwise, if the Q statistic exceeds a predefined threshold, the maximum value is considered as an outlier. Dean and Dixon (1951) tabulated the values of this threshold in relation to the number of observations. For the sake of simplicity and versatility the next curve provides a close fit to these thresholds:

$$\mathrm{Q}'\cong 1.9622[n_{\mathrm{V}(\mathbf{o})}]^{-0.687} \tag{3.61}$$

With $n_{\mathrm{V}(\mathbf{o})}$ as the number the anchor points in the neighbourhood V($\mathbf{o}$). Thus, if $\mathrm{Q}>\mathrm{Q}'$, the parameter values fitted at the anchor point $\mathbf{o}$ are replaced by the local averages of the parameters fitted at surrounding anchor points. Figure 3-41 shows the tabulated and fitted values.

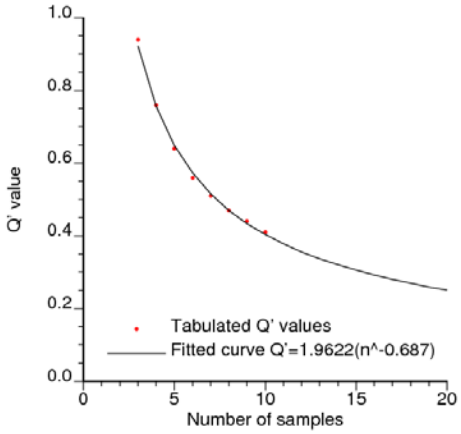

Figure 3-41: Tabulated Q' values and approximation by a power function.

Figure 3-42, left, shows the local nugget effect and local range parameters of the exponential models fitted on the location-dependent local variograms obtained from the single drill hole dataset. The nugget effect values above 0.45 where penalized during the iterative fitting using a strength factor of 0.1, no penalization for abrupt changes in the local parameter values was applied. The final values of the final objective function presented at the right side of Figure 3-42, are considerably lower than those presented at the left side of Figure 3-40. This is explained by the weighting of the local experimental points by the sum of the 2-point distance weights used in their calculation, while no weighting was considered in the results presented in Figure 3-40. The global exponential variogram model parameters fitted to the normal score transformed values of 1-D dataset are 0.41 and 21.1m for the nugget effect and the range, respectively.

Figure 3-43 shows the interpolated local exponential variogram model parameters fitted to the experimental correlograms obtained from the locally transformed values of the 2-D dataset. In this case, the local variogram models were fitted individually at each anchor point. Inverse lag distance weighted was applied to the experimental points. The few abrupt variations in the parameters were corrected by manual fitting. It is interesting to note the correspondence between the features of the location-dependent variogram model parameters and the spatial features of the exhaustive image in Figure 3-1.

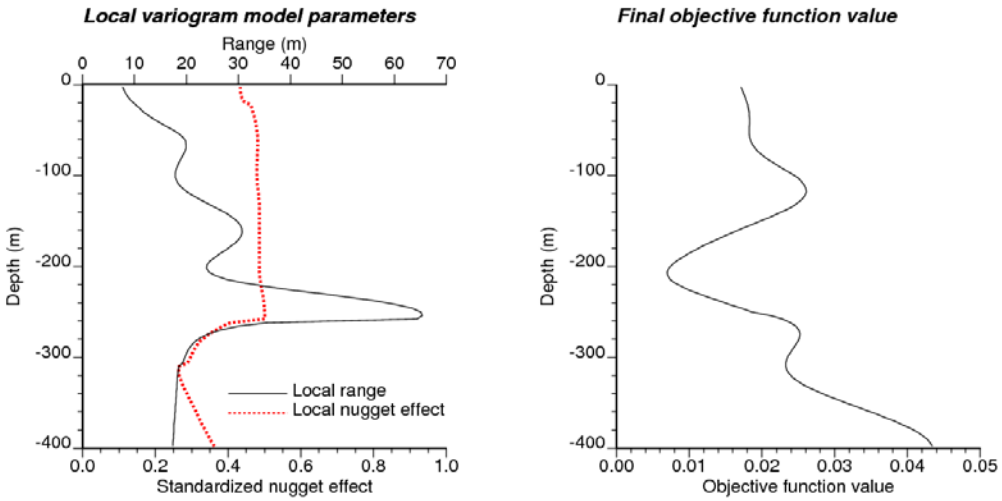

Figure 3-42: Left, local nugget effect and range of the local exponential variogram models fitted on the location-dependent experimental correlograms of the 1-D dataset. The nugget effect was penalized above the threshold of 0.45 during the iterative fitting. Right, final values of the local variogram fitting objective function.

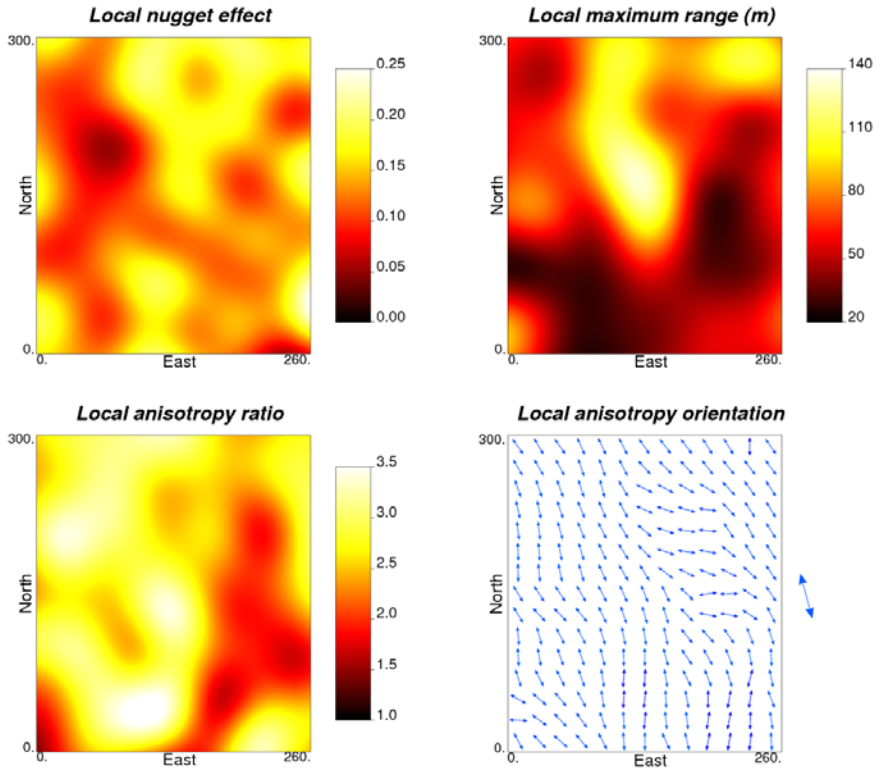

Figure 3-43: Location-dependent exponential variogram model parameters fitted on the local experimental correlograms from the locally transformed values of the 2-D dataset. For this example the local variograms were fitted independently at each anchor point, local variogram models with anomalous parameters were fitted manually. The large arrow outside the domain limits represent the global anisotropy direction.

3.5. Discussion

A methodology for the inference of location-dependent cdfs and their statistics based on distance weighting functions has been developed in this chapter. These statistics provide prior definitions of the locally stationary RFs at every location. The local accuracy and precision of the location-dependent cdfs are expected to be higher than the global stationary cdf, but lower than what should be obtained after spatial prediction. The parameters of the local RF cdfs are expected to change smoothly. The variation of local statistics should respond to tendencies in the data at a scale larger than the samples separation. Choosing the degree of smoothness in the spatial variation of local statistics remains mostly a subjective exercise. Numerical criteria, such as the maximization of the correlation between the trend and the data, are considered as secondary to the visual assessment of the local statistics models. However, these criteria are central in the subsequent stage of spatial prediction with location-dependent statistics.

Location-dependent variograms and correlograms are able to adapt to local changes in the spatial continuity. Nevertheless, any abrupt change unrelated to the presence of domain boundaries or previously defined discontinuities is suspicious. It may be caused by the presence of local outliers, by the overfitting of the local statistics, or by artifacts in the inference and modelling of the local statistics. Smoothly varying local 1-point statistics are relatively straightforward to obtain. In the case of local 2-point statistics, that require the fitting of a model, the inference of their corresponding model parameters may be problematic. Location-dependent correlograms are more robust than other local measures of spatial continuity, but still, model fitting on them by the method of least squares may result in locally anomalous variogram parameter models. In such cases, close checking is required. Additionally, applying penalties in the semiautomatic fitting algorithm for anomalous variogram parameters can reduce the occurrence of anomalous fits. However, beyond these improvements to the least squares method, a more robust method for variogram modelling is required.

The use of the location-dependent statistics and models in spatial prediction under the assumption of local stationarity is developed in the next chapter.

Chapter 4
Spatial Prediction under the Decision of Local Stationarity

Locally varying statistics and distributions allow introducing a local, rather than global basis for the spatial prediction of an attribute. The inferred statistics and their parameters are required at each prediction location. Estimation and simulation techniques are applied in the same fashion as standard techniques, but with the locally varying parameters. The variogram model and prior cdf change from location to location.

The first two sections of this chapter cover the adaptation of the traditional estimation and simulation techniques to local stationarity. The impact of incorporating locally changing statistics in the estimates and posterior uncertainty distributions is analysed. Locally stationary techniques can be applied without the assumption of any type of distribution. Nevertheless, particular attention is given to the stationary multiGaussian model. This is due to the congenial properties of this model, see Chapter 2, and the capability of the local normal scores transformation to embed changes in not only the local cdf mean, but also in its variance and shape. The locally stationary approach can be extended to indicator based techniques. Its application to the simulation of categorical variables is presented. The last section discusses cross validation, accuracy plots and other methodologies for validating the locally stationary models in relation to the input data and parameters.

4.1. Locally Stationary Estimation

The assumption of local stationarity amounts to a strict stationarity assumption restricted to each prediction point. Therefore, locally stationary Simple Kriging becomes the main estimation algorithm. Ordinary kriging could also be adapted to work under the assumption of local stationarity. Simple Kriging is required for Gaussian based estimation and simulation, since it yields the correct moments for the conditional distributions. Locally multiGaussian estimation is performed on locally transformed

values. As shown in Chapter 3, an efficient way to model and store the local normal score transformation table is by means of Hermite polynomials. Another advantage of the Hermitian modeling is that it allows a straightforward implementation of a local change of support model for providing block estimates within a multiGaussian framework.

4.1.1. Simple and Ordinary Kriging

Under the assumption of local stationarity the kriging estimator becomes:

$$Z^*_{LSSK}(\mathbf{o}) = \sum_{\alpha=1}^{n(\mathbf{o})} \lambda_\alpha^{(LSSK)}(\mathbf{o}) \cdot Z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(\mathbf{o})} \lambda_\alpha^{(LSSK)}(\mathbf{o})\right] m(\mathbf{o}) \tag{4.1}$$

The mean, $m(\mathbf{o})$, is specific to the estimated point **o**. The notation of the estimated point location is changed from **u** to **o**, this is done to stress that **o** is the estimated location and also the point where all the location-dependent statistics and parameters are defined. The locally stationary simple kriging (LSSK) system of equations allows the use of location-dependent means and covariances, $C(\mathbf{h};\mathbf{o})$:

$$\sum_{\beta=1}^{n(\mathbf{o})} \lambda_\beta^{(LSSK)}(\mathbf{o}) C(\mathbf{u}_\beta - \mathbf{u}_\alpha;\mathbf{o}) = C(\mathbf{o} - \mathbf{u}_\alpha;\mathbf{o}) \quad \alpha = 1,...,n(\mathbf{o}) \tag{4.2}$$

This can also be expressed in terms of correlograms (Goovaerts 1997, p.129):

$$\sum_{\beta=1}^{n(\mathbf{o})} \lambda_\beta^{(LSSK)}(\mathbf{o}) \rho(\mathbf{u}_\beta - \mathbf{u}_\alpha;\mathbf{o}) = \rho(\mathbf{o} - \mathbf{u}_\alpha;\mathbf{o}) \quad \alpha = 1,...,n(\mathbf{o}) \tag{4.3}$$

The impact of using location-dependent variograms is better appreciated in a 2-D example. Figure 4-1 shows the local surface roughness estimates using SK and LSSK on the clustered samples of Walker Lake dataset (Isaaks & Srivastava 1989). The location-dependent exponential model parameters shown in Figure 3-43 were used for generating the LSSK estimates plotted at right side of Figure 4-1. The traditional SK estimates, shown at the left side of Figure 4-1, show clearly the uniform spatial continuity tendency imprinted by the use of a global variogram model. LSSK estimates, on the contrary, show changing orientations of spatial continuity more akin to the terrain morphology (see Figure 4-1, right).

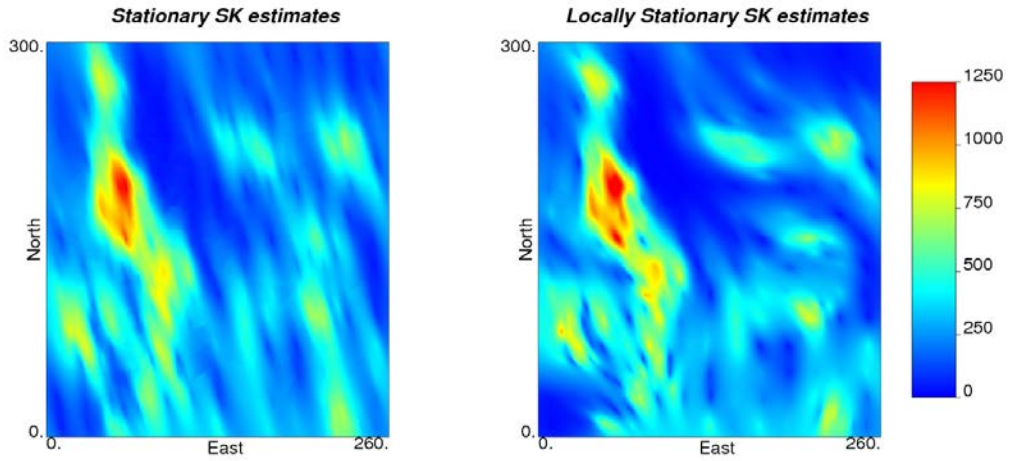

Figure 4-1: 2-D comparison between SK (left) and LSSK (right) elevation estimates.

While the stationary SK estimates tend to the global mean away from the samples, the LSSK estimates tend to the local-dependent means. This use of the local mean can also be achieved by traditional SK with local means. The same set of distance based weights modifies all the required statistics at each location, including the local variance, which is incorporated in the estimation variance. The LSSK estimation variance is given by:

$$\sigma^2_{LSSK}(\mathbf{o}) = C(0;\mathbf{o})\left[1 - \sum_{\alpha=1}^{n(\mathbf{o})} \lambda_\alpha^{(LSSK)}(\mathbf{o})\rho(\mathbf{o}-\mathbf{u}_\alpha;\mathbf{o})\right] \tag{4.4}$$

Where $C(0;\mathbf{o})$ is the location-dependent variance (see Equation 3.40). Thus, the LSSK variance is locally conditioned not only by the data availability and arrangement around the estimated point, but also by the local variability.

The capability of the LSSK variance to take into account the local variance is illustrated in Figure 4-2 by the estimation variances for the 1-D dataset presented in Chapter 3. The impact of incorporating the location-dependent variogram models on the estimation variance is illustrated in Figure 4-3. The differences between the SK and LSSK estimation variances are pronounced. Areas where the SK variance (Figure 4-3, left) is low due to high sampling density may show an increased LSSK variance due to high local data variability (Figure 4-3, right), while areas where the values are very continuous show much lower LSSK variances (compare with Figure 3-17, left). Moreover the pattern of the LSSK variance spatial distribution changes locally not only in relation to data configuration but also in relation to the local variogram. Thus, the

LSSK variance becomes richer in information than the SK variance that is defined only by the global variogram model and by the data configuration.

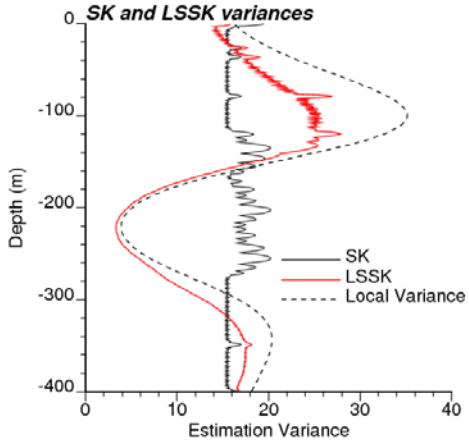

Figure 4-2: SK and LSSK variances obtained using the Ag grades in the 1-D dataset. For the sake of clarity, the estimation variances at data locations are not plotted.

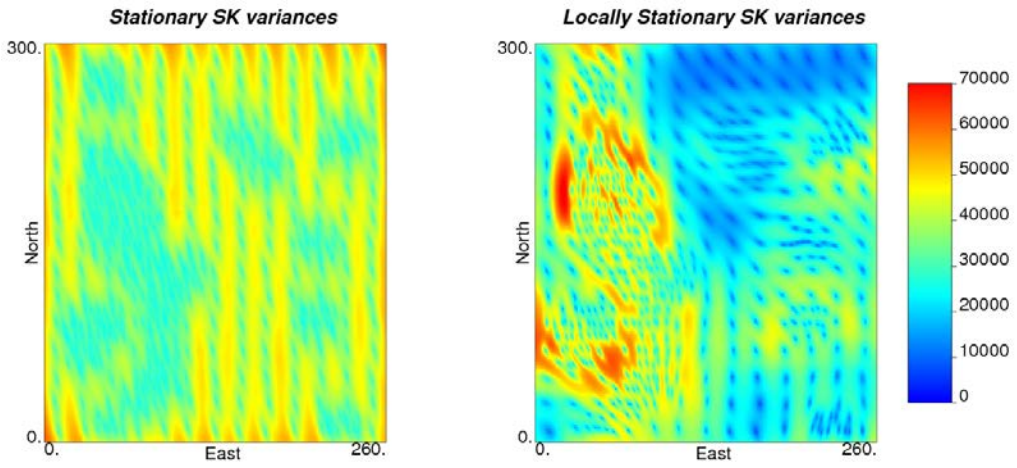

Figure 4-3: SK (left) and LSSK (right) variances obtained using the elevation attribute in the 2-D dataset.

The locally stationary ordinary kriging (LSOK) estimator is similar to the traditional ordinary kriging estimator:

$$Z^*_{LSOK}(\mathbf{o}) = \sum_{\alpha=1}^{n(\mathbf{o})} \lambda_\alpha^{(LSOK)}(\mathbf{o}) Z(\mathbf{u}_\alpha) \tag{4.5}$$

And the correlograms in the LSOK system are location-dependent:

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{o})} \lambda_\beta^{(LSOK)}(\mathbf{o})\rho(\mathbf{u}_\beta - \mathbf{u}_\alpha;\mathbf{o}) + \mu(\mathbf{o}) = \rho(\mathbf{o} - \mathbf{u}_\alpha;\mathbf{o}) \quad \alpha = 1,...,n(\mathbf{o}) \\ \sum_{\beta=1}^{n(\mathbf{o})} \lambda_\beta^{(LSOK)}(\mathbf{o}) = 1 \end{cases} \tag{4.6}$$

As in the case of LSSK, the LSOK variance is enhanced by the incorporation of the location-dependent variance and correlograms:

$$\sigma^2_{LSOK}(\mathbf{o}) = C(0;\mathbf{o})\left[1 - \sum_{\alpha=1}^{n(\mathbf{o})} \lambda_\alpha^{(LSOK)}(\mathbf{o}) \cdot \rho(\mathbf{o} - \mathbf{u}_\alpha;\mathbf{o})\right] - \mu(\mathbf{o}) \tag{4.7}$$

The extension to block estimation of these two point estimators is straightforward; this is explained below.

Locally Stationary Simple and Ordinary Block Kriging

Usually the block volume $v(\mathbf{o})$ is small compared to the domain volume D, so the changes of the location-dependent moments within blocks may be considered negligible. So, the assumption is that the location-dependent cdf and statistics centered at $\mathbf{o}$ correspond to all points within $v(\mathbf{o})$. This amounts to extending the locally stationary decision from the point support to the entire block volume:

$$Prob\{Z(\mathbf{u}_i) < z_1,...,Z(\mathbf{u}_n) < z_K;\mathbf{o}_i\} = Prob\{Z(\mathbf{u}_i + \mathbf{h}) < z_1,...,Z(\mathbf{u}_n + \mathbf{h}) < z_K;\mathbf{o}_j\}$$
$$\forall\ \mathbf{u}_\alpha, \mathbf{u}_\alpha + \mathbf{h} \in D, \text{ and only if } \mathbf{o}_i \text{ and } \mathbf{o}_j \in v(\mathbf{o}) \tag{4.8}$$

A straightforward way to obtain the block estimates is to discretize the block volume, perform point estimation using the same location-dependent parameters for all the discretization points, and average the estimated values (Goovaerts 1997, p.153). Alternatively, the block values may be estimated directly. In this case, the only change required in the LSSK and LSOK systems is to replace the location-dependent covariances, or correlograms, between the data points and the estimated point, $C(\mathbf{o} - \mathbf{u}_\alpha;\mathbf{o})$, by an approximation of the location-dependent point-to-block covariance (Journel & Huijbregts 1978, p.54) of the form:

$$\bar{C}(\mathbf{u}_\alpha, v(\mathbf{o}); \mathbf{o}) \approx \sum_{j=1}^{M} C(\mathbf{o}_j - \mathbf{u}_\alpha; \mathbf{o}) \quad \forall \mathbf{o}_j \in v(\mathbf{o}) \tag{4.9}$$

Where M is the number of discretization points of the volume $v(\mathbf{o})$. In the same way, for calculating the LSSK and LSOK estimation variances, the location-dependent covariance at lag 0, $C(0;\mathbf{o})$, may be replaced by the average:

$$\bar{C}(v(\mathbf{o}), v(\mathbf{o}); \mathbf{o}) \cong \sum_{i=1}^{M} \sum_{j=1}^{M} C(\mathbf{o}'_j - \mathbf{o}'_i; \mathbf{o}) \quad \forall \mathbf{o}'_i, \mathbf{o}'_j \in v(\mathbf{o}) \tag{4.10}$$

Selection of the Estimation Neighbourhood

In practice, kriging is performed within a search neighbourhood centred on the estimate location and using only a limited number of surrounding samples. Two of the reasons for doing so respond to computational efficiency and the lack of reliable knowledge on the large scale spatial continuity (Deutsch & Journel 1998, p.32). There is no rigorous methodology for defining the optimum dimensions of the search boundaries and the number of data to consider (Chilès & Delfiner 1999, pp.201-202). Proposed criteria include the minimization of parameters such as the conditional bias, the weight assigned to the mean in simple kriging (Rivoirard 1987), the proportion and magnitude of negative weights, and the kriging variances (Vann et al. 2003). For locally stationary estimation, these criteria would require the design and implementation of locally changing search neighbourhoods. This may be demanding in professional time and computer effort. An ad hoc alternative consists in using a uniform search window with radii equivalent to at most three times the bandwidth parameter of the Gaussian Kernel used for the inference of the local statistics and to consider all the samples that fall within such window. The logic behind such decision is to include in the estimation all the samples that had an important contribution in the inference of the corresponding local statistics. Another option is to select the maximum number of conditioning data to 40 and 60, for 2-D and 3-D estimation, respectively. Those numbers have been proposed to assure robust kriging results at a reasonable computational effort (Guo & Deutsch 2008).

4.1.2. MultiGaussian Kriging

Locally Stationary MutiGaussian Kriging (LSMGK) considers not only location-dependent measures of spatial correlation, but also local changes in the distribution shape.

This is accomplished with local normal scores transforms. These transformation functions are modelled by a series of Hermite polynomials with local coefficients, as explained in Section 3.4.1. By smoothly interpolating these coefficients, the transformation function can be defined over the entire domain.

LSMGK is equivalent to LSSK using the local normal scores transforms and the location-dependent correlograms obtained from these values. As in the stationary multiGaussian kriging, the LSSK estimate and variance define the posterior local ccdf in Gaussian space. The posterior ccdf in original units, $F(\mathbf{o}; z_p(\mathbf{o}) \mid n(\mathbf{o}))$ is built, by back-transforming P local quantiles $y_p(\mathbf{o})$ of the posterior ccdf in Gaussian units, $G'(\mathbf{o}; y_p(\mathbf{o}) \mid n(\mathbf{o}))$ with mean equal to the LSSK estimator Y_{LSSK} and standard deviation equal to the square root of LSSK variance, $\sigma_{LSSK}(\mathbf{o})$:

$$z_p(\mathbf{o}) = \varphi_Z(y_p(\mathbf{o}); \mathbf{o}) = \sum_{q=0}^{Q} \phi_q(\mathbf{o}) H_q[y_p(\mathbf{o})] \cong \sum_{q=0}^{Q} \phi_q(\mathbf{o}) H_q[Y_{LSSK}(\mathbf{o}) + \sigma_{LSSK}(\mathbf{o}) \cdot t_p] \tag{4.11}$$

And: $F(\mathbf{o}; z_p(\mathbf{o}) \mid n(\mathbf{o})) = G'(\mathbf{o}; y_p(\mathbf{o}) \mid n(\mathbf{o})) = G(t_p) = p$, where $G(\cdot)$ and t_p are the standard Gaussian cdf and quantile, respectively. The location-dependent coefficients are obtained from the approximation of local Gaussian transformation by series of Hermite polynomials, as shown in section 3.4.1. The estimator in original units can be approximated by the average of the local quantiles, $z_p(\mathbf{o})$, given that their number P is big enough, in practice between 100 and 200:

$$Z^*_{LSMGK}(\mathbf{o}) \cong E[z_p(\mathbf{o})] = E[\varphi_Z(y_p(\mathbf{o}); \mathbf{o})] = \frac{1}{P} \sum_{p=1}^{P} \sum_{q=0}^{Q} \phi_q(\mathbf{o}) H_q[Y_{LSSK}(\mathbf{o}) + \sigma_{LSSK}(\mathbf{o}) \cdot t_p] \tag{4.12}$$

Besides the reproduction of the piecewise linear features informed by location-dependent variograms, LSMGK estimates show the impact of the incorporation of local prior cdfs (see Figure 4-4, left, for the 1-D dataset and Figure 4-5, right, for the 2-D dataset). The differentiation between low and high grade values is more marked. This higher differentiation results in an improved accuracy of the model only if the prior location-dependent cdfs and its statistics follow the local changes in attribute values. If the local statistics are inferred with few neighbouring samples with an excessive

contribution of local outliers, their incorporation in locally stationary kriging may result in diminished accuracy.

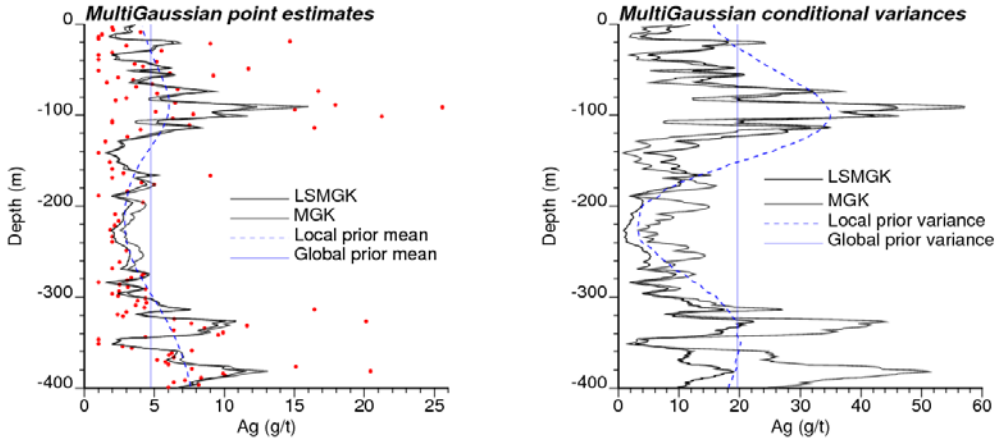

Figure 4-4: MultiGaussian point estimates (left) and estimation variances (right) for the 1-D dataset.

The backtransformation of the ccdfs obtained by multiGaussian kriging allows imposing the data variability on the conditional variances by restoring the proportional effect. Thus, high values areas will show higher conditional variances if the distribution in original units is positively skewed, and vice versa if it is negatively skewed. In LSMGK instead, the incorporation of the local variability in the conditional variances is due mainly to the use of local backtransformations and local variogram models. The restitution of the proportional effect is still an important component of the final conditional variance, but this is controlled by the spread and shape of the local prior cdf. This can be observed in Figure 4-4, right, where the LSMGK conditional variance in original units seldom exceeds the local prior variance. Figure 4-6 show the conditional variances for the 2-D dataset after the backtransformation of the ccdfs obtained with stationary MGK (left) and LSMGK (right). The impact of the local variogram models on the structure of the conditional variances is clear in the right of Figure 4-6.

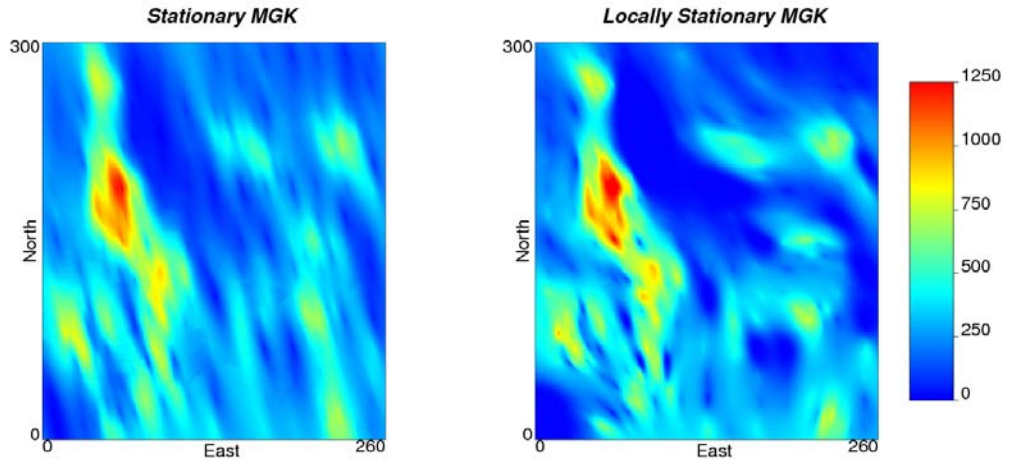

Figure 4-5: Traditional (left) and locally stationary multiGaussian (right) kriging elevation estimates for the 2-D dataset.

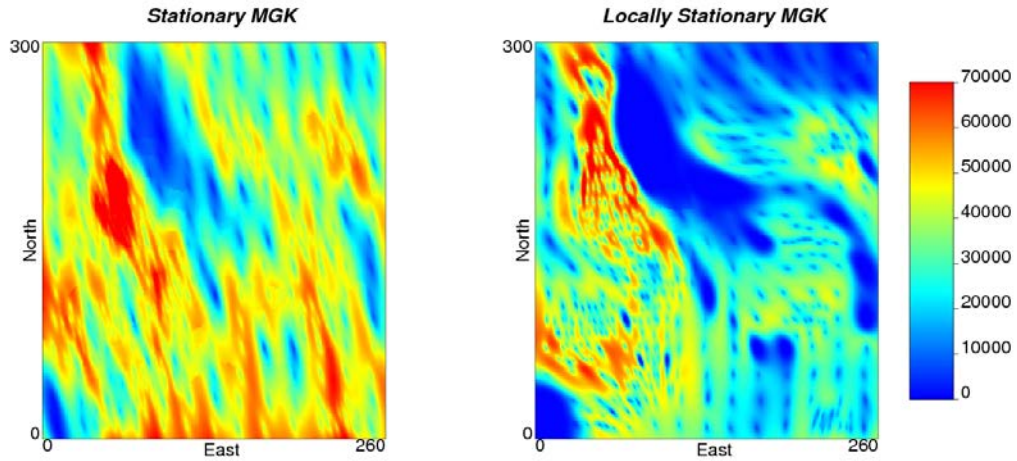

Figure 4-6: Traditional (left) and locally stationary multiGaussian (right) conditional variances for the 2-D dataset.

MultiGaussian Block Kriging

When working with a multiGaussian model, block estimates in original units require a change of support model. The local normal scores transformation function can be assumed constant within the block volume $v(\mathbf{o})$, if it is relatively small compared with the entire domain. Thus, any randomly located sample within the volume centred at $\mathbf{o}$ is transformed by the same function:

$$Z(\mathbf{o}_i) = \varphi_Z(Y(\mathbf{o}_i);\mathbf{o}) \qquad \forall \mathbf{o}_i \in v(\mathbf{o}) \tag{4.13}$$

The block grade can be estimated as the average of M point estimates within the block. Therefore, the posterior block support ccdf can be built for P cut-offs z_p, p=1,…,P, from (Chilès & Delfiner 1999, p.435):

$$
\begin{aligned}
F_v(\mathbf{o}; z_p \mid n(\mathbf{o})) &= \text{Prob}\left\{ \frac{1}{M} \sum_{i=1}^{M} Z(\mathbf{o}_i) \le z_p \mid Z(\mathbf{u}_\alpha) : \alpha = 1, \ldots, n(\mathbf{o}) \right\} \\
&= \text{Prob}\left\{ \sum_{i=1}^{M} \varphi\big(Y(\mathbf{o}_i); \mathbf{o}\big) \le M \cdot z_p \mid Y(\mathbf{u}_\alpha) : \alpha = 1, \ldots, n(\mathbf{o}) \right\} \\
&= E\left[I\left(\sum_{i=1}^{M} \varphi\big(Y(\mathbf{o}_i); \mathbf{o}\big); M \cdot z_p \right) \mid Y(\mathbf{u}_\alpha) : \alpha = 1, \ldots, n(\mathbf{o}) \right]
\end{aligned}
\tag{4.14}
$$

Where $I[\cdot]$ is the indicator function presented in Section 2.3.4. In Gaussian space, the ccdfs at the locations $\mathbf{o}_i$ are fully defined by the SK estimate and variance. Therefore, the posterior block support ccdf can be expressed as (Emery 2007b):

$$
F_v(\mathbf{o}; z_p \mid n(\mathbf{o})) = \int I\left[\sum_{i=1}^{M} \varphi\Big(Y_{LSSK}(\mathbf{o}_i) + \sigma_{LSSK}(\mathbf{o}_i) \cdot t; \mathbf{o} \Big); M \cdot z_p \right] g(t) dt \tag{4.15}
$$

This expression is approximated numerically by drawing a large number N of standard Gaussian distributed random numbers, and averaging the results (Verly 1984):

$$
F_v(\mathbf{o}; z_p \mid n(\mathbf{o})) \cong \frac{1}{N} \sum_{j=1}^{N} I\left[\sum_{i=1}^{M} \varphi\Big(Y_{LSSK}(\mathbf{o}_i) + \sigma_{LSSK}(\mathbf{o}_i) \cdot t_j; \mathbf{o} \Big); M \cdot z_p \right] \tag{4.16}
$$

Building the complete block support ccdf requires this numerical calculation for different cut-offs. Thus, this approach may be computationally demanding if the block support ccdf is required in detail.

A more efficient approach is given by the Discrete Gaussian Model. For this change of support model, the point support RVs are considered randomly located at points $\underline{\mathbf{o}}$ within the blocks. The point and block support Gaussian transformed RVs $Y(\underline{\mathbf{o}})$ and $Y_v(\mathbf{o})$ are assumed bigaussian with location-dependent correlation $r(\mathbf{o})$. The Gaussian transformation functions for both variables are given by (Emery 2005):

$$
\begin{cases}
Z(\underline{\mathbf{o}}) = \varphi(Y(\underline{\mathbf{o}}); \mathbf{o}) \quad = \sum_{q=0}^{Q} \phi_q(\mathbf{o}) H_q\big(Y(\underline{\mathbf{o}})\big) \\
Z_v(\mathbf{o}) = \varphi_v(Y_v(\mathbf{o}); \mathbf{o}) = \sum_{q=0}^{Q} \phi_q(\mathbf{o}) r^q(\mathbf{o}) H_q\big(Y_v(\mathbf{o})\big)
\end{cases}
\tag{4.17}
$$

Where the location-dependent change of support coefficient $r(\mathbf{o})$ is obtained from (Rivoirard 1990, p.64):

$$\text{var}[Z_v(\mathbf{o})] = \gamma_Z(\infty;\mathbf{o}) - \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}\gamma_Z(\mathbf{o}_i - \mathbf{o}_j;\mathbf{o}) = \sum_{q=1}^{Q}\phi_q^2(\mathbf{o})r^{2q}(\mathbf{o}) \tag{4.18}$$

With $\gamma_Z(\mathbf{h};\mathbf{o})$ as the non-standardized variogram in original units for the location **o**. Working with correlograms instead of variograms is desirable for greater stability and more direct interpretation. In this case the Expression 4.18 is equivalent to:

$$\text{var}[Z_v(\mathbf{o})] = C_Z(0;\mathbf{o})\left(1 - \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}\left[1 - \rho_Z(\mathbf{o}_i - \mathbf{o}_j;\mathbf{o})\right]\right) = \sum_{q=1}^{Q}\phi_q^2(\mathbf{o})r^{2q}(\mathbf{o}) \tag{4.19}$$

The experimental location-dependent variograms, or correlograms, calculated from normal scores transformed values are more stable than those in original units, and thus easier to model. Thus it is preferred to calculate and model these measures of spatial correlation on the transformed values and then back-transform them to original units. This assures the consistency between the variograms used in LSMGK and those in original units. The back-transformation of the normal scores correlograms is achieved by considering the relationship between the covariances in normal scores, $C_Y(\mathbf{h};\mathbf{o})$, and original units, $C_Z(\mathbf{h};\mathbf{o})$. This is given by (Guibal 1987):

$$C_Z(\mathbf{h};\mathbf{o}) = \sum_{q=1}^{Q}\phi_q^2(\mathbf{o})\left[C_Y(\mathbf{h};\mathbf{o})\right]^q \tag{4.20}$$

Or if location-dependent variograms are used instead:

$$\gamma_Z(\mathbf{h};\mathbf{o}) = \sum_{q=1}^{Q}\phi_q^2(\mathbf{o})\left[1 - \left(1 - \gamma_Y(\mathbf{h};\mathbf{o})\right)^q\right] \tag{4.21}$$

In practice, the location-dependent change of support coefficient is calculated only at the anchor point locations and subsequently interpolated between them. The block support posterior ccdf is built for different local quantiles, $z_p(v(\mathbf{o}))$, by:

$$z_p(v(\mathbf{o})) = \varphi_v(y_p(v(\mathbf{o}));\mathbf{o}) \cong \sum_{q=0}^{Q} \phi_q(\mathbf{o}) \cdot r^p(\mathbf{o}) \cdot H_q[y_p(v(\mathbf{o}))]$$
$$= \sum_{q=0}^{Q} \phi_q(\mathbf{o}) \cdot r^p(\mathbf{o}) \cdot H_q[Y_{LSSK}(v(\mathbf{o})) + \sigma_{LSSK}(v(\mathbf{o})) \cdot t_p]$$

(4.22)

As in Expression 4.12, the estimate at block support is approximated numerically by the average of the P local quantiles $z_p(v(\mathbf{o}))$, provided that P is big enough; this is between 100 and 200 quantiles.

Figure 4-7, left, shows the prior local point-support variances approximated from the sum of square local Hermite coefficients and the prior block-support variances obtained from Expression 4.19. The global and local change of support coefficients are presented in the right side of Figure 4-7. Given a block size, the greater the spatial continuity specified by the location-dependent variogram models, the closer the change of support coefficient will be to one. Figure 4-8 presents the MGK and LSMGK block support estimates for the 1-D dataset at the left, and the corresponding block support conditional variances at the right.

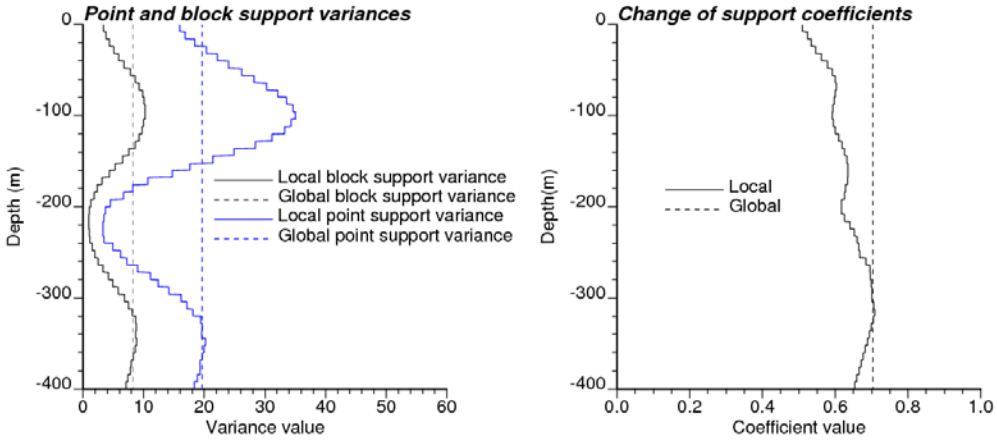

Figure 4-7: Point and block support prior local variances (left) and local change of support coefficients (right) inferred on the drillhole silver grades.

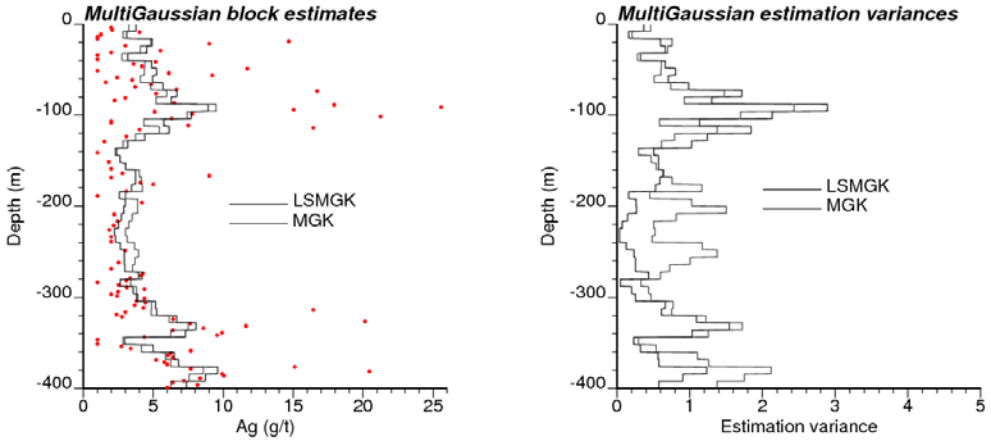

Figure 4-8: MGK and LSMGK block estimates (left). MGK and LSMGK block conditional variances (right) for the 1-D dataset.

Figure 4-9 shows the locally stationary multiGaussian estimates and variances at a block support of 10x10 units for the elevation attribute in the 2-D dataset. The block estimates are very similar to the point estimates presented at the left side of Figure 4-5. The conditional block variances, however, are very different from the obtained for point estimates (Figure 4-6, left). The reference grade-tonnage curves in Figure 4-10 were calculated by averaging the elevation attribute of the exhaustive 2-D dataset within 10x10 units blocks. In the same figure and for the same support the resulting-grade tonnage curves for block MGK and LSMGK are shown. The average elevation above cut-off appears to be closer to the reference curve for LSMGK estimates. The locally stationary method also yield to an improved reproduction of the fraction of total blocks above the lower cut-offs. For higher cut-off this fraction is comparable for MGK and LSMGK.

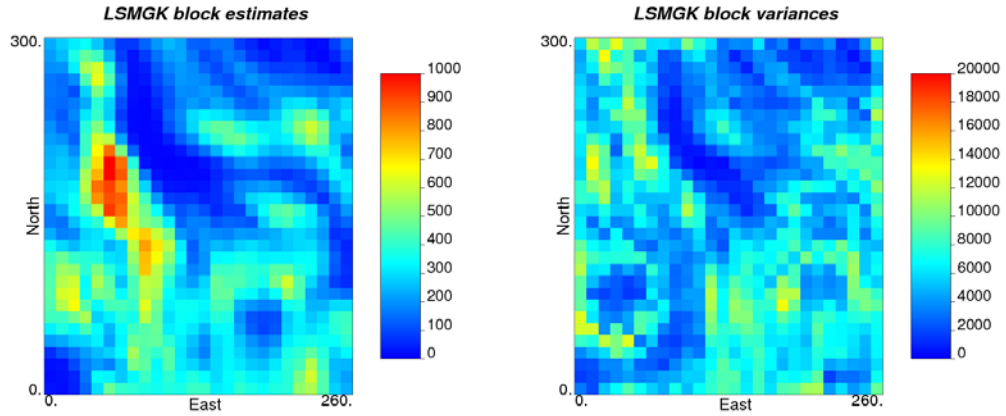

Figure 4-9: Locally stationary multiGaussian block estimates and variances for the elevation attribute in the 2-D dataset.

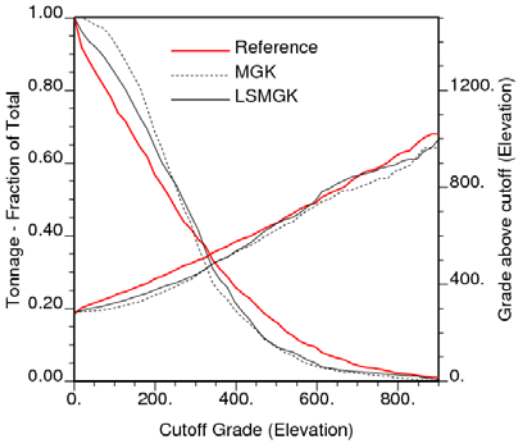

Figure 4-10: Grade tonnage curves for the elevation attribute in the 2-D dataset at a 10x10 units block support. The reference curves (red) were obtained from the averages of the exhaustive values (see Figure 3-1) in blocks of 10x10 units.

4.2. Simulation

Locally stationary simulation algorithms based on the multiGaussian assumption use the local normal scores transformation modelled by Hermite polynomials. Similarly to multiGaussian kriging, the local transformations account for non-stationarity of the RF parameters. In the case of locally stationary indicator simulation the local categorical pdf is required. Non-stationarity in the spatial correlation is informed by the location-dependent variogram models.

4.2.1. Sequential Gaussian Simulation

Local normal scores transformation of inverse distance weighted cdfs have been already proposed to account for trends in Sequential Gaussian Simulation (McLennan & Deutsch 2008). The idea is to modify the global cdf by the inverse distance weights at each simulated location, perform the normal score transformation of the weighted cdf keeping the transformation table, draw a simulated value on the conditional distribution, and backtransform it using the local transformation table. The algorithm proposed here requires that the Gaussian transformation function be defined prior to the simulation. This is done to decrease the processing demand of rebuilding the complete transformation tables at each location and because the weights used for locally weighting the cdf are also used for inferring the local measures of spatial continuity.

Locally Stationary Sequential Gaussian Simulation Algorithm

The algorithm of Locally Stationary Sequential Gaussian simulation (LSSGS) proceeds in the following steps:

a) Read and store the local Hermite coefficients and the location-dependent variogram parameters for every simulation node.
b) Visit each simulation node in a random path. Search for the surrounding conditioning data and previously simulated grid nodes.
c) Construct the location-dependent transformation function with the local Hermite coefficients. Perform the local Gaussian transformation of surrounding data and previously simulated nodes.
d) Obtain the mean and variance of the local ccdf by locally stationary Simple Kriging with the location-dependent variogram model informed by the corresponding local variogram parameters.
e) Perform Monte Carlo simulation to obtain a simulated value from that ccdf.
f) Back transform the simulated value according to the local Gaussian transformation function. Add the simulated value in original units to the data set.
g) Go to the next node in the random path and loop from step c until all nodes are simulated.

The impact of using locally changing variogram models is clear when comparing a map produced by LSSGS (Figure 4-11, right) with an equivalent map obtained using the traditional SGS (Figure 4-11, left). Figure 4-12 shows the E-type estimates obtained from 100 SGS and LSSGS realizations performed on the 2-D dataset. The locally stationary SGS e-type map (Figure 4-12, right) clearly shows the changing anisotropic orientations informed by the location-dependent variogram models and a sharper differentiation of the zones of low and high values. As for the locally stationary kriging variances, the conditional variances obtained from the LSSGS realizations (Figure 4-13, right) show the strong influence of the local prior variances and the location-dependent variogram models.

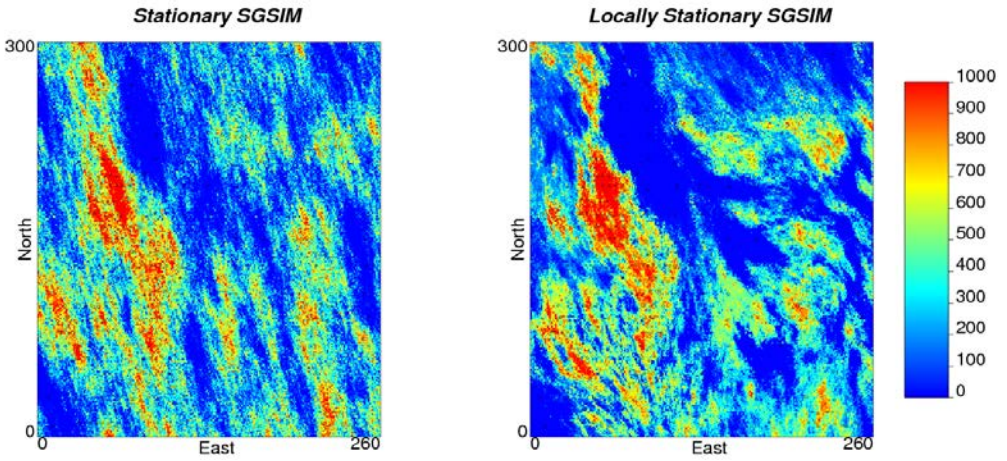

Figure 4-11: Example realizations of SGS (left) and LSSGS (right) using the terrain roughness attribute in the 2-D dataset.

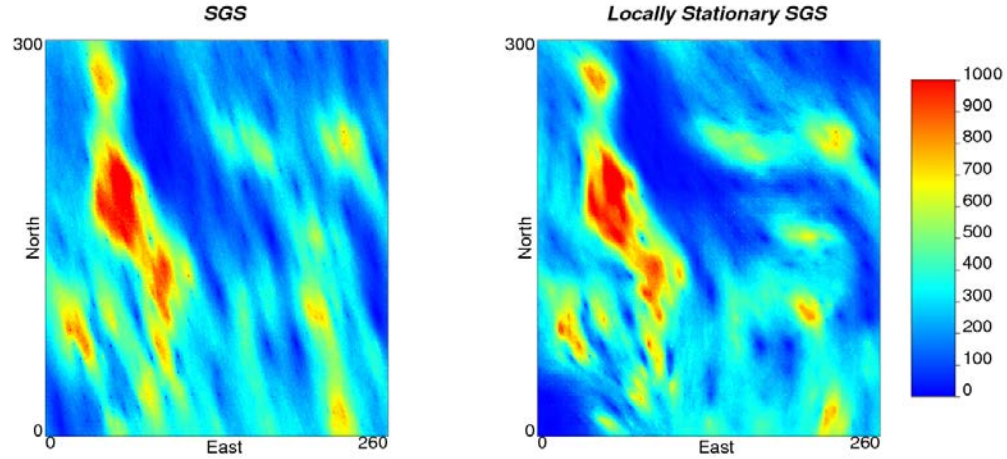

Figure 4-12: Posterior local means obtained by averaging 100 realizations of SGS (left) and LSSGS (right).

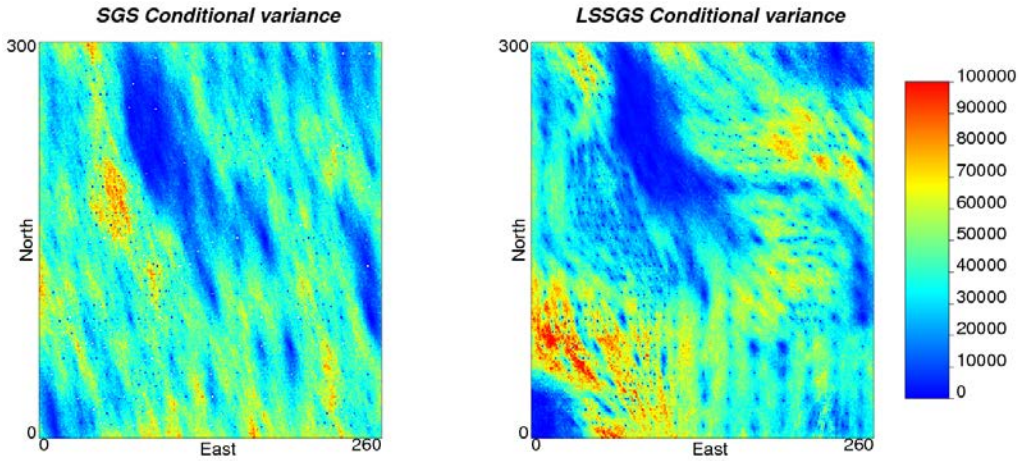

Figure 4-13: Conditional variances obtained from 100 of SGS (left) and LSSGS (right) realizations.

LSSGS demands more computer resources in terms of memory and processing time than SGS. The increase in memory demand is because the LSSGS algorithm requires all the location-dependent variogram parameters and local Hermite coefficients to be stored in the RAM for each node. The increase in computation time is due to the reconstruction of the covariance matrix at each node in the random path. For this example, the generation of 100 realizations with LSSGS required took 2.5 times longer than the generation of the same number of realizations with SGS.

4.2.2. Sequential Indicator Simulation

The Locally Stationary Sequential Indicator Simulation (LSSIS) algorithm uses local proportions of categorical values, which can be obtained by Equation 3.24, and location-dependent indicator variograms. As in LSSGS, a covariance lookup table is no longer used and the covariances are updated locally at each simulation node used in the construction of the local ccdf. Figure 4-14 shows a comparison between simulated maps obtained stationary and locally stationary SIS. The second map shows a better reproduction of curvilinear features. This often entails increased connectivity within categories.

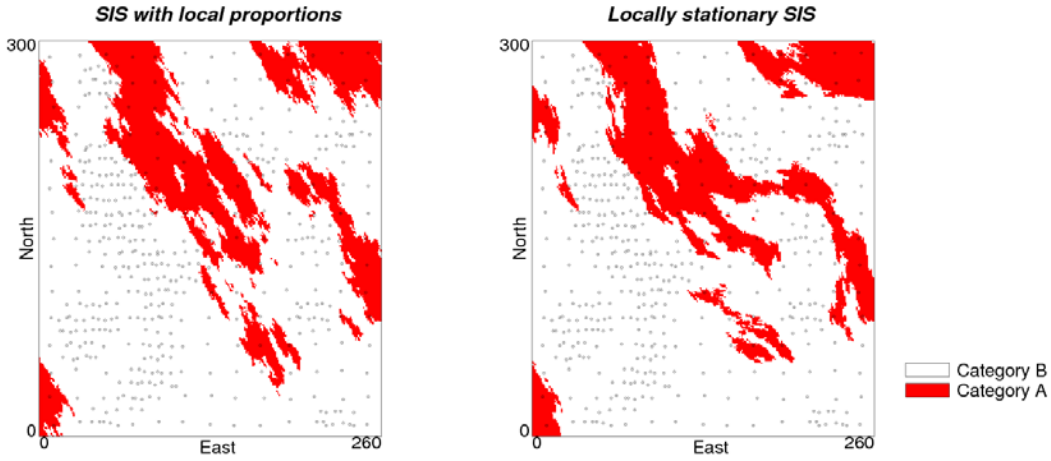

Figure 4-14: Example realizations of SIS (left) and LSSIS (right) using the categorical variable in the 2-D dataset.

A single set of location dependent indicator variograms is required for the simulation of two categories, as presented in Figure 4-14. In theory, more than two categories or thresholds can be simulated using the locally stationary algorithm. In practice, this

demands the inference and modelling of location dependent indicator variograms for all categories or thresholds at all anchor point locations. This can be a tedious process in traditional Geostatistics; it can be even more tedious in locally stationary modelling.

As for LSSGS, the LSSIS algorithm requires more computational resources than its stationary counterpart. For this example, the generation of 100 LSSIS realizations took 4.3 times longer than the generation of the same number of realizations with SIS.

4.2.3. Checking the Realizations

Traditional simulation techniques must fulfill three minimal criteria of validity. These criteria consist in the reproduction of (1) data values, (2) the input variogram model and (3) the declustered global cdf (Delfiner 1976; Leuangthong et al. 2004). Similar criteria are valid for locally stationary simulation. If sample values are assigned to the closest node, data reproduction is normally satisfied unless particular circumstances, such as the presence of multiple samples within a cell, or if the samples are flagged as outliers, or they fall outside the grid model.

Although the input of locally stationary simulation is the set of locally transformed cdfs rather than the global cdf, the reproduction of the latter is still the aim. A considerable divergence between the global cdfs of realizations and the global cdf of original data may indicate a bias introduced by the location-dependent distributions. These local distributions, instead, do not need to be reproduced, since they are used only as prior models of local uncertainty, which are updated by the surrounding data and the location-dependent variogram models. Figure 4-15 shows the histogram reproduction for the 1-D dataset in Gaussian (left) and original units (right) after backtransformation. The histogram shape and the global mean are closely reproduced by the realizations. However the modelling of the transformation function can introduce slight bias, particularly if the original cdf contains spikes and outliers (see Section 3.3.1).

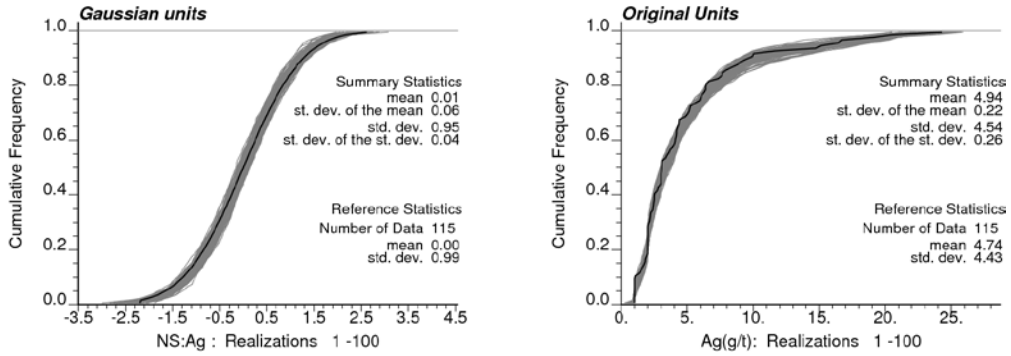

Figure 4-15: Global cdf reproduction of 100 LSSGS realizations performed on the 1-D dataset. The realizations cdfs in Gaussian units are presented in the left figure, and the backtransformed cdfs in the right figure. The black curves correspond to the data cdf, while the grey curves correspond to the realizations cdfs.

The local variogram models, rather than the global, must be reproduced. However, due to their locally stationary nature, the reproduction of these models can only be verified for short lag distances. As shown in Figure 4-16, the average of non-standardized variograms calculated on the realizations closely follows the variogram model and the experimental local variogram at short lags. At longer lags, the mismatch between these measures increases.

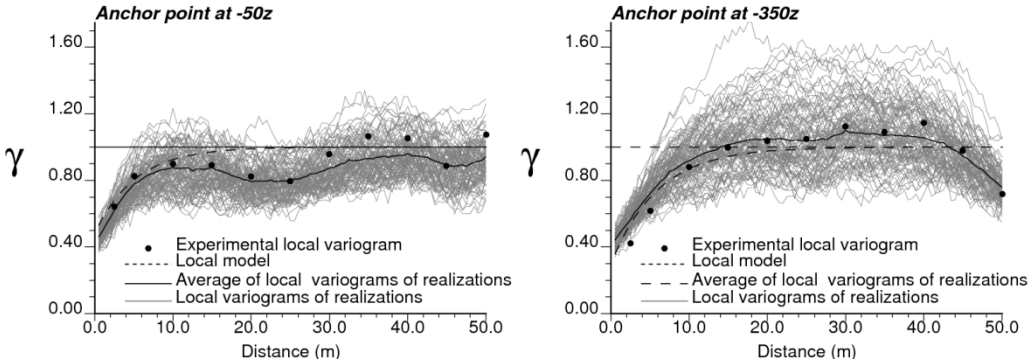

Figure 4-16: Reproduction of the non-standardized location-dependent variograms at two Anchor points for 100 LSSGS realizations generated using the 1-D dataset.

In a very large domain, the local variations of simulated values compensate variations in other areas, allowing the convergence of the realization statistics to the stationary RF statistics. This property of the RF model is called ergodicity (Luster 1985, p.205). When inferring the locally weighted statistics of realizations any fluctuation of the simulated values in the vicinity of an anchor point will have a high impact on the local statistics. Due to higher weight assigned to the closer locations, the effect of such fluctuations may not be compensated by fluctuations in other parts of the domain. Thus, in the same way

as ergodic fluctuations in the statistics of traditional SGS realizations increase as the domain size decreases (Deutsch 1995, pp.243-252), the ergodic fluctuations of the location-dependent statistics of realizations are expected to be higher due to the areal restrictions imposed by the weighting function bandwidth.

Checking the local variogram reproduction may be tedious, since the experimental variograms must be calculated at different anchor points for multiple realizations and directions. Alternatively, the visual verification of the reproduction of the anisotropic features of local variogram models in realizations maps can be performed as a quick check. Figure 4-17 shows the E-type estimates map of 100 realizations performed using the 2-D dataset superimposed by the directions of anisotropy of the location-dependent variograms.

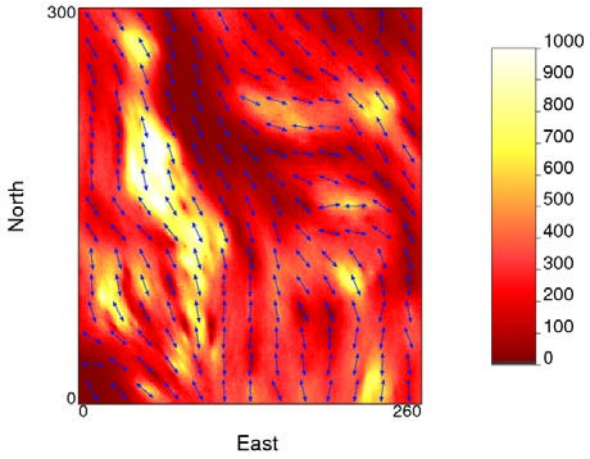

Figure 4-17: Anisotropy directions of the 2-D dataset local variogram models superimposed on the resulting e-type estimates of 100 LSSGS realizations.

The criteria for checking LSSIS realizations are similar to the locally stationary Gaussian simulation. Categories must be honoured at their sampled locations. The local spatial continuity informed by the local variograms need to be reproduced, as well as the global categorical cdf. For the 2-D dataset, the 100 LSSGS realizations show a reasonable reproduction of the global cdf (Figure 4-18, left). However, for the lower cut-offs, the cumulative frequencies are systematically higher in the realizations cdfs than in the global cdf. This bias may have different possible causes: Inaccuracies introduced by the Hermite transformation function modelling or the excessive weight given to low value samples, among others. This small bias does not appear in the realizations of the 1-D dataset (Figure 4-15). The reproduction of the categorical global proportions by LSSIS realization is presented in Figure 4-18, right.

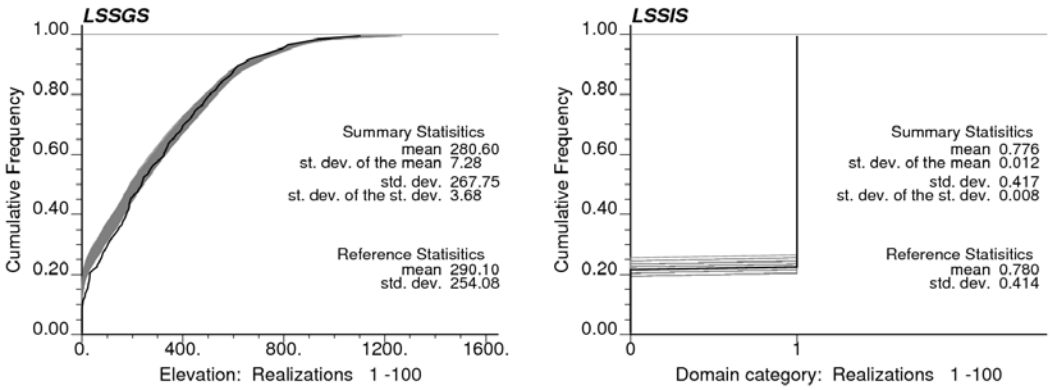

Figure 4-18: Reproduction of the global cdf for elevation values (left) and category proportions (right) after LSSGS and LSSIS using the 2-D datset. The black curves correspond to the data cdf, while the grey curves correspond to the realizations cdfs.

4.3. Criteria for the Validation of Locally Stationary Models.

A numerical model is valid if it is consistent with the data, the parameters, the assumptions and the background knowledge that were used for building it (Oreskes et al. 1994). Simulated locally stationary models are required to honour data values, the global data cdf and reflect the local spatial continuity informed by local variogram models. Estimated locally stationary models are expected to be unbiased in relation to the input data mean, to reproduce the data values at their locations, and also to reflect the general features of the local spatial continuity informed by the local variogram models. Beyond these checks for internal consistency, the locally stationary techniques are required to provide accurate and precise probabilistic models with a fair uncertainty.

Accuracy is defined as the closeness of predictions to the true values. For estimates of the ccdf mean, accuracy can be checked by the mean square error and other statistics between estimates and true values. The accuracy of local probabilistic models provided by the ccdfs can be assessed by their capability to accommodate the true values within symmetric probability intervals (Deutsch 2002, p.300). Precision of ccdfs is only relevant if they are accurate (Deutsch 1996). This can be assessed by the capability of the ccdf to reproduce the proportions of true values for different symmetric probability intervals. Local uncertainty is assessed by the width of the ccdf. Generally, a narrow uncertainty model is preferred, but only if it can accommodate all the possible outcomes in relation to the data, the model assumptions and parameters and the background information. If it is too narrow, it may not be able to cover the range of all possible realizations. At the other

extreme, if the ccdf spread is too large, it can lead to useless, although accurate, uncertainty models (Taylor 1997, pp.5-6). Thus, a fair uncertainty model is one that can provide useful information about the range of possible outcomes in regard to the available information and model assumptions.

As discussed in Section 2.4, estimation techniques aim for local accuracy by minimizing the mean square estimation error, and this may come with the price of smoothing. Although the main aim of simulated models is to reproduce the input statistics, it is also desirable to improve the accuracy and precision of simulated models while the local uncertainty is kept within reasonable limits. In this section some numerical criteria for checking the accuracy of local predictions, and the accuracy, precision and fairness of uncertainty distributions are discussed. These are presented as means to validate the input parameters with available data; however, they can also be used to confirm the models with additional data. Additional quantitative and qualitative criteria for checking the consistency of the locally stationary models with the input parameters and the geological knowledge base are also discussed.

4.3.1. Accuracy of Estimates

Cross Validation is traditionally used to assess the impact of the use of different stationary models and kriging parameters in the estimates (Davis 1987). In this methodology, each datum is removed one at a time and the value at its location is estimated using the remaining samples within a neighbourhood (Goovaerts 1997, p.105). In the Jackknife, a dataset is divided in different non-overlapping subsets, the statistics and parameters are inferred for one of these subsets and subsequently used to re-estimate the values of the other subset (Deutsch 2002, pp.298-299). For both Cross Validation and the Jackknife the error is calculated as the difference between the re-estimated values and the true values:

$$e(\mathbf{u}_\alpha) = z^*(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha) \qquad \alpha = 1,...n \tag{4.23}$$

In order to fulfill the condition of global unbiasedness (see Expression 2.25), the average of this error is expected to be close to zero. Local accuracy is assessed by the mean square error:

$$MSE = \sum_{\alpha=1}^{n} e^2(\mathbf{u}_\alpha) \tag{4.24}$$

The MSE should be minimized. The covariance between re-estimated and true values, $Cov\left\{Z^*, Z\right\}$, should be maximized. Commonly, the correlation coefficient between the true and the re-estimated values is used to assess cross-validation results:

$$\rho_e = \frac{Cov\left\{Z^*, Z\right\}}{\sigma_{Z^*}\sigma_Z} \tag{4.25}$$

Where and σ_{Z^*} and σ_Z are the standard deviations of the re-estimated and true values, respectively. However, this metric can be misleadingly high if the estimates are oversmoothed and, consequently, their variance very low. A simple way to assess the magnitude of the conditional bias is by considering the slope of the linear regression model adjusted to the cloud (Olea 1999, pp.141,145).

In Gaussian space, local confidence intervals for kriging estimates can be derived from the square root of SK variances σ_{SK} (Delhomme 1978; Journel 1986; Chilès & Delfiner 1999, p.164). Narrow confidence intervals may indicate higher precision of estimates, but only if the ccdfs tend to contain the true values (Deutsch 1996; Goovaerts 2001). In traditional Geostatistics confidence intervals cannot be derived from SK variances in original units, since they do not depend on conditioning values. However, they are consistent with the homoscedastic property of the multiGaussian model. Contrarily, LSSK variances are affected by the variability of neighbouring data; hence, they have the potential to provide more meaningful confidence intervals than those provided by traditional kriging variances in original units.

Additionally, in original units, an increase in the average of variance of estimates indicates a reduction of the smoothing of the estimation. The variance of the estimated values can be checked considering all estimated nodes in the model, rather than only for the re-estimated values in cross validation.

When cross validating locally stationary models, the local cdf and statistics should carry more information than in the case of stationary models. This should result in a reduced estimation error in areas where the local statistics are robustly inferred. Higher errors appear in highly variable areas that, for positively skewed distributions, are commonly associated to high grade zones. If these zones are preferentially sampled, the local prior cdf in scarcely sampled low grade zones will be biased, resulting in an

increased cross validation error. Moreover, if the spatial continuity changes, increased cross validation errors can also arise because of local variograms that are not representative. Location-dependent statistics reflect the non-stationary variations informed by data. If the available data is not enough to provide reliable information about underlying trends of the attribute, cross-validation may yield optimistic results that could not be necessarily confirmed by additional sampling.

Due to these issues, cross validation statistics should be taken with caution. These statistics can be used to detect areas in the locally stationary models where the poor definition of the local statistics is likely to worsen the estimates. Cross validation results can also be used to compare the locally stationary models with their stationary equivalents and decide if the incorporation of local statistics improves the estimates. Locally stationary models built on scarce data and without the backing of geological knowledge are likely to perform worse than models built with traditional stationary techniques. Figure 4-19 shows the scatterplots of true values versus the conditional cdf means obtained from MGK and LSMGK on the terrain roughness attribute of the 2-D dataset. LSMGK cross validation estimates show a higher correlation with the true values compared with the stationary multiGaussian Kriging. This is due to the higher covariance between the true values and the estimates, and despite the reduced smoothing evidenced by the higher variance of LSMGK conditional means. Moreover a substantial reduction in the mean squared error is observed for the cross validation results obtained with the locally stationary technique.

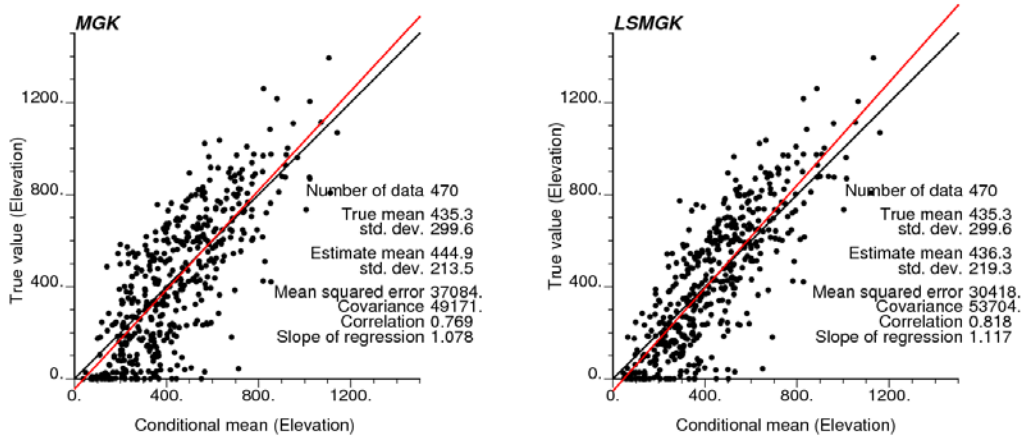

Figure 4-19: True vs. estimated scatterplots for MGK (left) and LSSMGK (right) of the 2-D dataset. Declustering weights not considered. Values on the black 45° slope line are predicted with zero error. The red line is the linear regression model adjusted to the estimates vs. true values cloud.

The Jackknife of locally stationary models is feasible only for large datasets, since the diminished number of samples in each subset makes the inference of local statistics difficult. If data are abundant, this validation technique can be used for assessing the robustness of the location-dependent statistics and the parameters of models fitted on them prior to the estimation. This approach is seldom used due to its high demand of computational and professional effort. In the case of locally stationary models, this demand is even higher.

4.3.2. Accuracy and Precision of Uncertainty Distributions

Local distributions of uncertainty are accurate if they are centered close to the true values, while they are precise if their spread is narrow. Given different symmetric intervals of the local distributions of uncertainty, if the true values fall within them in a proportion equal or higher than the width of these intervals, the local distributions are considered to be accurate. The closer these proportions are to the probability intervals width, the higher is the precision of the local uncertainty distributions (Deutsch 1996). Accuracy and precision are checked on the local uncertainty distributions at cross validated locations. These distributions can be obtained from multiGaussian models or from indicator based models. Accuracy plots (Deutsch 1996; 2002, pp.299-309) are useful for checking the data proportions versus the widths defined on the uncertainty distributions.

Figure 4-20 shows the accuracy plots for the 1D and 2D datasets obtained from MGK and LSMGK cross validation. An increased accuracy of the local uncertainty distributions is expected if the local prior cdfs and the local spatial continuity are reliably inferred. The precision of the posterior local distributions is, in great measure, inversely proportional to the width of the prior local cdfs and directly proportional to the continuity of the local variogram models. Areas with high local spatial continuity will result in narrower posterior distributions. Thus, although desirable, locally stationary models do not necessarily provide higher precision than their stationary equivalents. However, as observed in Figure 4-20, left, the spread of the uncertainty distributions may reflect better the local variability of grades.

An overall measure of the uncertainty of a stochastic model can be obtained from the average of the variances of all local ccdfs in the domain. Between two accurate and precise models, the one with lower average ccdf variance is preferred (Deutsch 2002, p.302).

Table 4-1 presents the comparative statistics of the MGK and LSMGK estimation using the 2-D dataset along the declustered data statistics. For this particular dataset, LSMGK shows practically no bias in the average of estimates. The variance of the LSMGK estimates is higher than the obtained from MGK, this indicates less smoothing in the locally stationary maps. In Gaussian units, the average ccdf variance is slightly higher after LSMGK, but after backtransformation to original units the average locally stationary ccdf variances is decreased considerably in relation to those obtained with MGK.

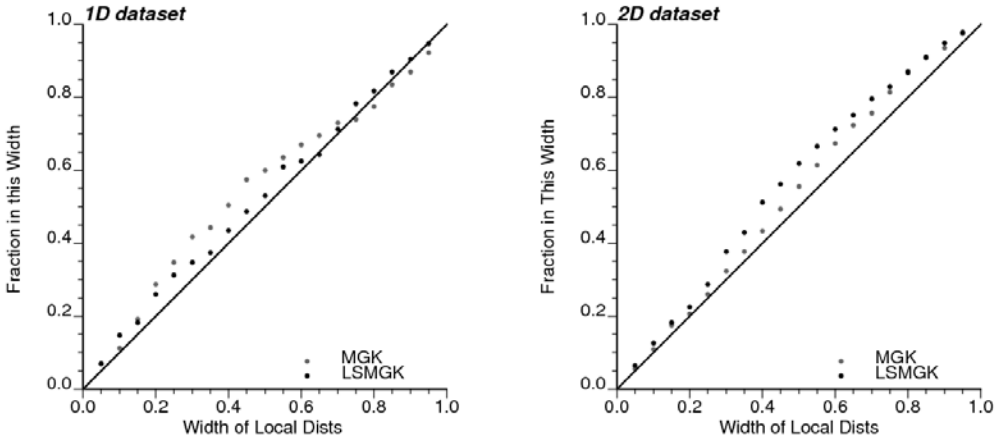

Figure 4-20: Accuracy plots comparing the uncertainty distributions obtained with MGK and LSMGK of the 1-D dataset (left), and the 2-D dataset (right).

Table 4-1: Comparative statistics for declustered 2-D data values, MGK and LSMGK estimation results

	Mean (Elevation)	Variance	Average ccdf variances in Gaussian units	Average of ccdf variances in original units
Data	290.09	64562		
MGK	306.01	33263	0.6	39444
LSMGK	290.47	35794	0.63	29592

4.3.3. Other Relevant Criteria

Beyond the important requirements of accuracy, precision and fair uncertainty, which can be checked in relation to the input data, it is important to check the conformity of the resulting models with geological knowledge. Local variograms models may be reasonably reproduced within the limitations discussed in Section 4.2.3, but the non-

stationary features of the spatial continuity in the resulting models should follow the changes indicated by geological knowledge.

In some cases, such as vein and facies modelling, the 2-point spatial connectivity (Journel & Alabert 1989) within a given class or category can be reliably recognized within certain distances during the interpretation of the geological setting. A desired property of locally stationary indicator based models is the improvement in the reproduction of such connectivity, particularly if it is curvilinear and it can be described at a scale larger than the sample separation.

Figure 4-21, left, shows an E-type map indicating the probability of being within the category A obtained from 100 LSSIS realizations using the categorical variable in the 2-D dataset (see Figure 4-14). The points A and B plotted at the north and east limits of this map are known to be connected; a path can be defined between them within the same domain. Figure 4-21, right, shows the histogram of connected cells between A and B for the SIS and LSSIS realizations. Only 2 out of 100 SIS realizations present connectivity between A and B. Contrarily, 50% of the LSSIS realizations result in connected regions of different size between those points. The number of cells in each of the connected regions within category A was calculated using the algorithm **`geo_obj`** (Deutsch 1998).

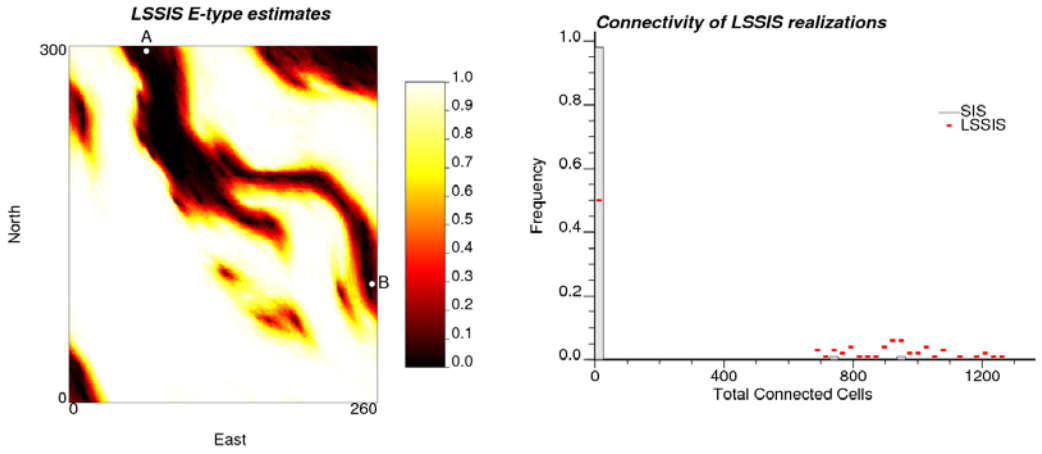

Figure 4-21: Left, E-type estimates of 100 LSSIS realizations indicating the probability of being within the category A of the 2-D dataset. Right, histogram of the number of connected cells between points A and B (see left side of the Figure) for 100 SIS and LSSIS realizations of the categorical variable in the 2-D dataset.

4.4. Discussion

The implementation of location-dependent statistics in estimation and simulation requires relatively minor modifications to the traditional algorithms. The necessary changes are mostly related to changing the histogram and covariance model from location to location. This may require an increase in computational effort.

The local normal scores transformations account for local changes in the mean, variance and histogram shape. Their modelling by Hermite polynomials allows an efficient storage of the required transformation functions at the scale of estimation and simulation. This also allows a straightforward implementation of local discrete Gaussian change of support model. A change of support model with locally changing parameters allows adapting the block support variance reduction to the changes in spatial continuity informed by the local variogram models. However, the Hermitian modelling of the local transformation functions may introduce small deviations in the reproduction of the transformation tables, particularly when spikes and outliers are present in the local distributions.

Cross validation can be used to compare the performance of locally stationary estimation versus traditional algorithms. If the location-dependent statistics are robustly inferred and effectively reflect the different aspects of non-stationarity within a domain, estimates are expected to be more accurate. However, if the local statistics are inferred with scarce data, or if the local outliers have a high contribution, locally stationary estimation may yield worse results. Cross validation error maps can be used to identify problem areas and check possible localized biases.

The reproduction of the global input cdf by realizations of locally stationary simulation can be achieved within ergodic fluctuations. The averages of local variograms of realizations should match closely the input local variogram models and the experimental local variograms at short lag distance. However, ergodic fluctuations are increased due to the much higher weights assigned to locations in the vicinity of the anchor points. If the local statistics are reliably inferred, locally stationary sequential simulation can provide accurate posterior distributions of uncertainty. Local precision will depend on the width of the prior local cdfs and the spatial continuity informed by the local variograms. The latter can show a shorter spatial continuity than the global variogram model at certain locations, causing a lower precision.

Other criterion for assessing the goodness of locally stationary numerical models is the reproduction of curvilinear features in agreement with geological knowledge. Additionally, the capability of location-dependent variograms to follow changes in the orientation and degree of anisotropy can be translated to an improved reproduction of the curvilinear connectivity informed by the geological interpretation of drillhole intercepts.

In the next chapter, the complete process, from the inference to the use of location-dependent statistics in estimation and simulation, is developed using datasets from the mining industry. The performance of local stationary estimation and simulation techniques is compared with their traditional equivalents.

Chapter 5
Case Study

This case study shows the predictive performance of locally stationary numerical models in a realistic scenario where the attributes of interest are highly variable and data are scarce. A 2-D dataset of more than 160,000 underground channel samples of the Ventersdorp contact reef is considered. In order to mimic the limited information available in a typical drillhole exploration campaign, only a few hundred samples are used for the inference of the location-dependent statistics and for locally stationary prediction.

After a brief description of the Ventersdorp dataset, this chapter continues with the selection of the Gaussian kernel parameters and the anchor point locations. The location-dependent 1-point and 2-point statistics are inferred for the accumulated gold and the reef width. These statistics are used into the locally stationary multiGaussian kriging algorithm. The performance of the locally stationary technique is compared with the traditional multiGaussian kriging. The complete dataset is used for comparing the prediction capability between both locally stationary and traditional methodologies.

5.1. 2-D Case Study: The Ventersdorp Contact Reef

The Ventersdorp Contact Reef (VCR) is one of the goldfields located at the northern fringes of the West Rand Group in the Witwatersrand Basin. This is mainly formed by mineralized conglomerates arranged in well-defined channels (Moon et al. 2006, pp.320-322). Sedimentological factors control the spatial distribution of gold grades in the conglomerates; however, a detailed geological interpretation of this reef is not available. This case study relies on the information provided by samples collected from the underground operations in the reef.

5.1.1. The Dataset

The complete VCR dataset consist of the registers corresponding to 161176 underground channel samples. These contain the 2-D coordinates of the sample locations, the assayed gold grades, the reef width and the Facies indicator. Four different facies are present in the dataset; however, the geological description of these was not provided. Only Facies 1 and 2, located at the centre of the lease, are considered in this case study. Figure 5-1 shows the accumulated gold values and the reef width obtained from interpolating the samples in Facies 1 and 2. As for the width map (Figure 5-1, right), the reef appears to be wider at the West and North of the domain. These spatial trends are also observable in the accumulated gold map. Figure 5-2, left, shows the distribution of Facies 1 and 2 within the domain obtained from the complete dataset.

The complete underground sampling dataset is sampled in a 200m x 200m quasi regular grid simulating drillhole intercepts in the reef. This grid spacing is within the typical range for exploration drilling meshes of the deep gold reefs in the Witwatersrand (du Pisani & Vogt 2004). Figure 5-2, right, shows only the intercepts located in Facies 1 and 2, these correspond to 181 simulated drillholes. This number is comparable to the number of holes drilled at other properties in neighbouring goldfields (Rance et al. 2006). The closest four samples around each drillhole were collected. This was done to simulate the wedge deflections commonly drilled in order to increase the number of reef intercepts with a minimum number of additional drilling (Magri 1987). The number of samples corresponding to the simulated ddh is 782, including those obtained in the deflections.

The complete dataset contains 143445 samples located in Facies 1 and 2. Table 5-1 summarizes the declustered indicator statistics for the Facies attribute in the drillhole and the complete datasets.

Table 5-1: Indicator statistics for the Facies attribute

Dataset:	Drillhole	Complete
Number:	782	143445
Proportion of facies 1	0.236	0.240
Indicator variance:	0.180	0.182

Figure 5-3 shows the cumulative probability graphs for the accumulated gold and reef width obtained from the simulated ddh samples in Facies 1 and 2. The reef tends to be wider and the accumulated gold content higher in Facies 2. Due to this reason and for the sake of brevity the analysis of the continuous attributes, accumulated gold and reef width,

are performed only for Facies 2. A scatterplot between the values of both attributes in Facies 2 is presented in Figure 5-4.The correlation coefficient is 0.44.

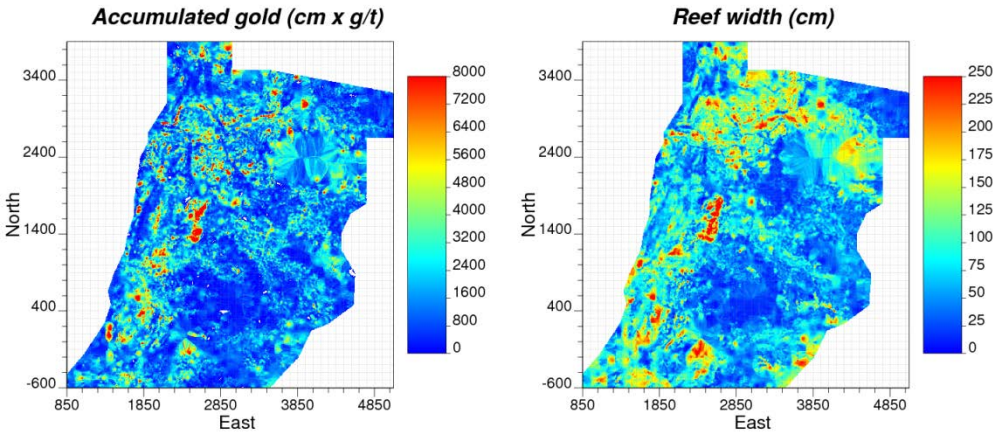

Figure 5-1: Gold grades (left) and reef width (right) maps obtained by the interpolation of the complete dataset.

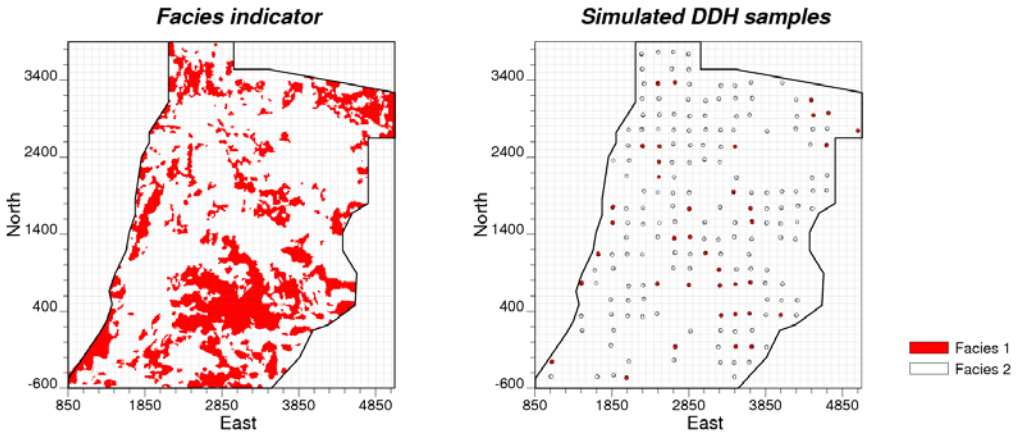

Figure 5-2: Facies locations obtained from the complete dataset (left) and locations of the simulated ddh samples (right)

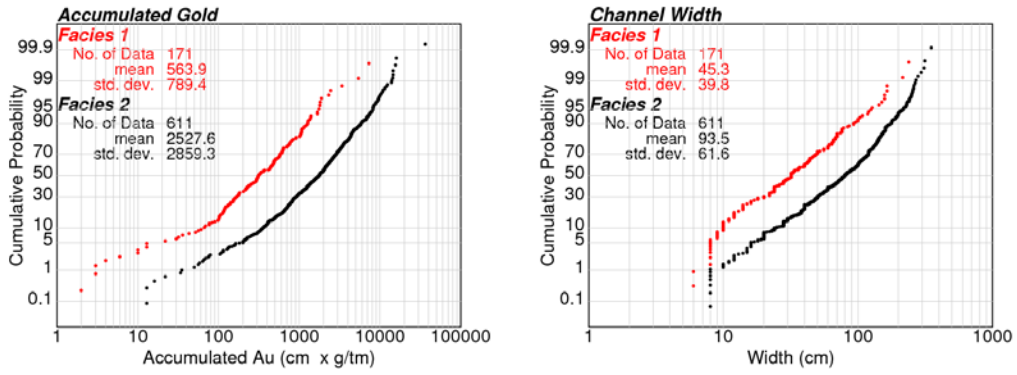

Figure 5-3: Accumulated gold (left) and reef width (right) probability plots in Facies 1 and 2

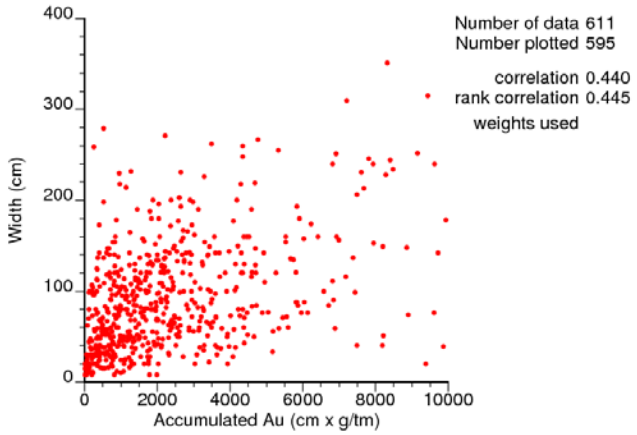

Figure 5-4: Scatterplot between the accumulated gold and the reef width attributes in Facies 2

5.1.2. Calculation and Modelling of the Location-Dependent Statistics

Selection of the Distance Weighting Function Parameters

As discussed in Chapter 3, the distance weighting function parameters are chosen by judging the smoothness of the resulting local statistics maps. If these maps show features that are controlled by very few samples, this may indicate the overfitting of the local statistics. Figure 5-5 shows the local mean models built using Gaussian kernels of different bandwidth. For a 200m bandwidth, the resulting map shows several details constrained to the proximity of few drillhole locations. 500m and higher bandwidth produce excessively smooth maps. The maps obtained using bandwidths of 300m and 400m look very similar and they present continuous features spreading to several drillhole locations. While the zoning of values is still distinguishable in the statistics, maps with a larger bandwidth will be preferred since they yield more stable local statistics.

As for the bandwidth value, the choice of the background value is done mainly visually, although it can be backed by checking the model /data variance ratio and coefficient of correlation statistics. The variance ratio of the local means vs. data values drops considerably from background values higher than 0.001 (see Figure 5-6, left). The coefficient of correlation of the local mean model vs. data values is more resilient to the background value, except for smaller bandwidths (see Figure 5-6, right). A background of

0.001 is chosen to avoid over smoothing while allowing the contribution of farther samples in the inference of the location-dependent statistics.

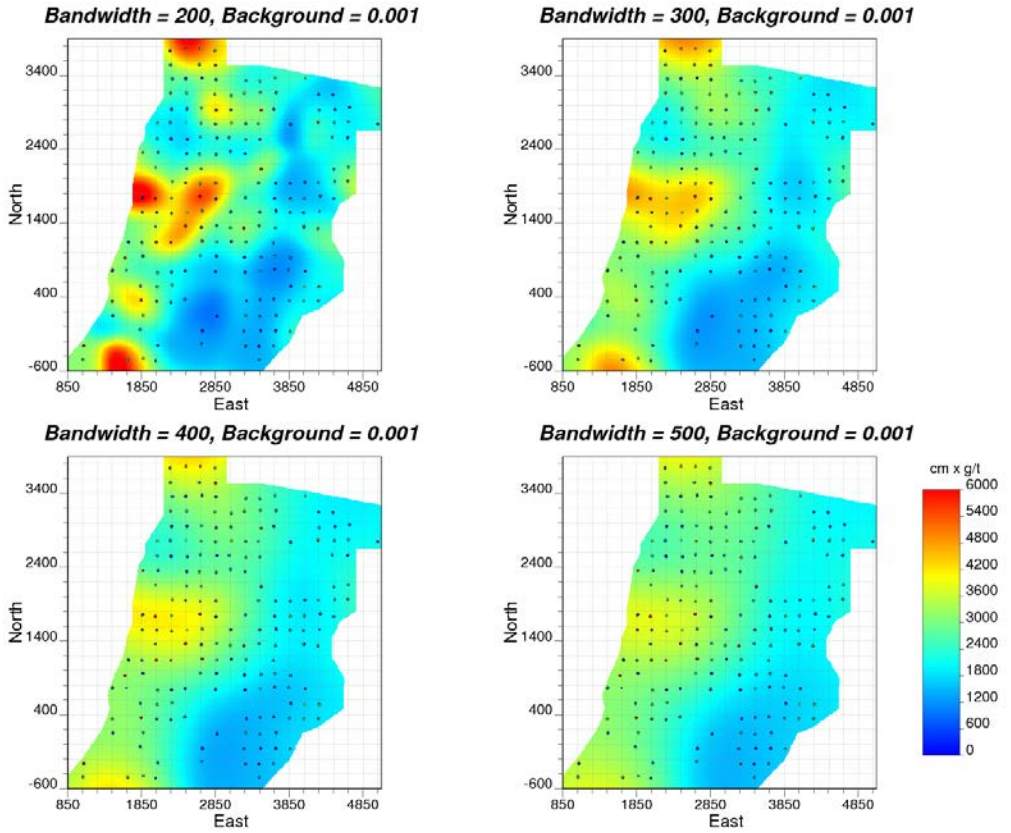

Figure 5-5: Local means obtained with different bandwidths of the Gaussian kernel.

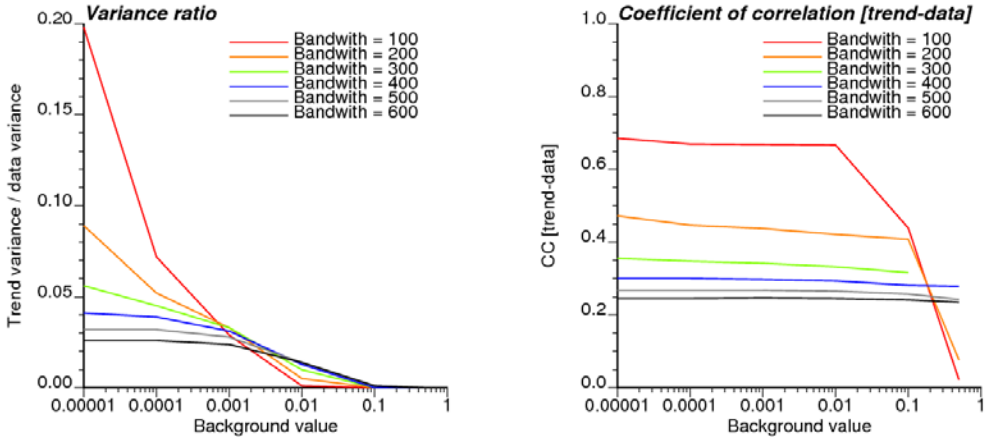

Figure 5-6: Data versus local mean model variance ratio (left) and coefficient of correlation (right)

Selection of the Anchor Point Number and Separation

As discussed in Chapter 2, the anchor points arrangement is designed for the sake of computational efficiency. A very narrow anchor points grid yields a more complete definition of the inferred location-dependent statistics, but at the price of high computational effort, specially for the calculation and fitting of location-dependent measures of spatial continuity. Thus, the number and locations of anchor points must be such that they allow the interpolation of the local parameters to approach the parameters directly inferred at every location. The local 1-point statistics for accumulated gold attribute in Facies 2 were used as referent for choosing the anchor point separations. Since this attribute is obtained from the product of the gold grade and the reef width, the anchor points separation chosen for the accumulated gold attribute is also used for the inference of the reef width local statistics.

The left side of Figure 5-7 shows the absolute error of the interpolated local means and standard deviations in relation of the anchor point separation. These errors increase quickly when the anchor points mesh has a separation greater than 200m. At this separation distance 317 anchor points are required within the domain. Increasing the anchor point separation to 300m allows reducing the number of required anchor points to 189, but at the price of increasing the interpolation error by more than four times. The errors are still relatively very low for a 300m anchor point separation; they amount for around 2% of the global mean and standard deviation. However, since the number of anchor points required for a 200m x 200m mesh is still manageable, this separation is preferred. For this anchor point mesh, the interpolation errors for the local means and standard deviations are around 0.5% of the global mean and standard deviation. The right side of Figure 5-7 shows the selected anchor points configurations in a 200m x 200m mesh. As it can be observed in this figure, the configuration of the anchor points mesh is such that collocation with sampling points is avoided.

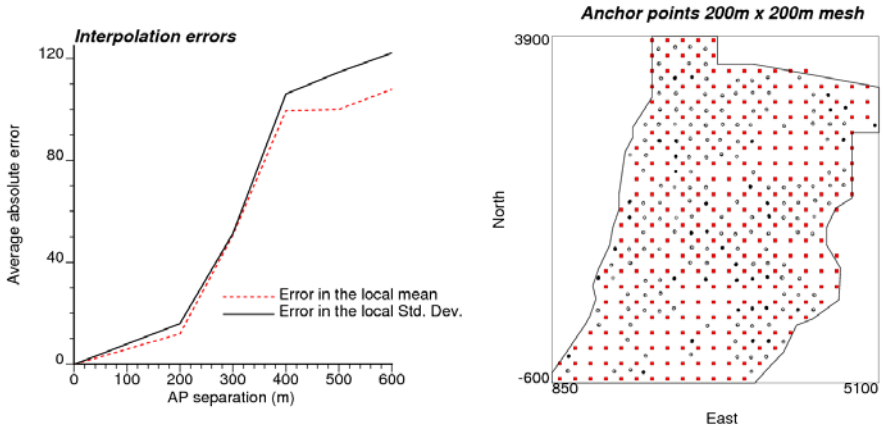

Figure 5-7: Interpolation errors of the local mean and standard deviation of the accumulated gold for different separations of the anchor point grid (left). Locations of the anchor points in the 200m x 200m grid (right). The anchor point locations are represented as red squares, while the data points are shown as circles.

Location-Dependent Distributions and Normal Scores Transformation

A set of distance weights for each anchor point was generated using a Gaussian kernel of 400m bandwidth and 0.001 background value. The same distance weights obtained for the accumulated gold attribute were used to infer the location dependent statistics for the width attribute at each anchor point. The main reason for doing this is that since the accumulated gold and reef width are related attributes, it is desirable to keep mutual consistency between their location-dependent statistics.

Figure 5-8 shows the interpolated local means (left) and local standard deviations (right) for the accumulated gold value. A region of higher local means and standard deviations of the accumulated gold attribute appears in the west side of the domain. The interpolated local statistics for this last attribute are shown in Figure 5-9. The interpolated local statistics of the width attribute reflects the wider reef regions on the west and north areas of the domain. The grid for the interpolation of the location-dependent statistics has the same resolution as the grid that is used for the estimation of the attributes. This is a 5m x 5m grid.

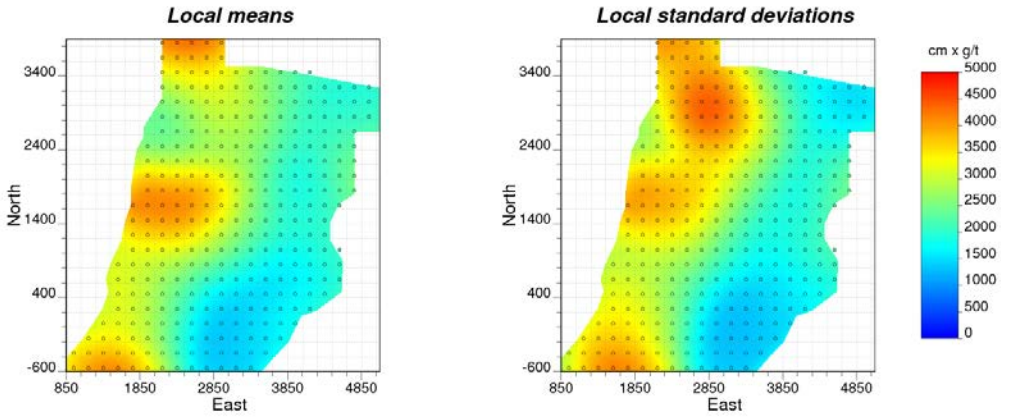

Figure 5-8: Interpolated local means (left) and local standard deviation (right) between anchor point locations for the accumulated gold.

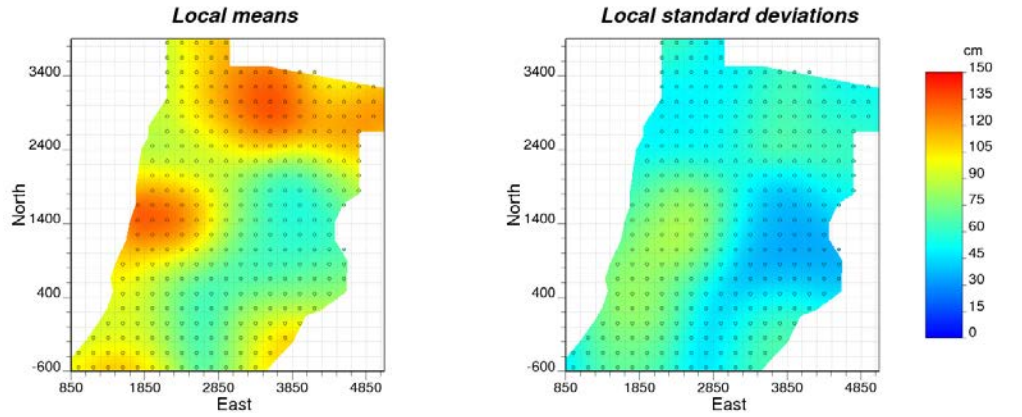

Figure 5-9: Interpolated local means (left) and local standard deviation (right) between anchor point locations for the reef width.

Figure 5-10 shows the local cdfs obtained at each anchor point using the same distance weights as for the location dependent 1-point statistics. For each of these cdfs, 200 percentiles were used to build the local normal scores transformation tables. These tables were used for fitting the Gaussian transformation functions using a series of Hermite polynomials with 36 terms. Although a higher number of Hermite coefficients could be used to improve the fitting of the Gaussian transformation functions, that number was judged as a reasonable trade-off between the fitting precision and the storage and processing requirements of the Hermite coefficients at every location. As for the other location-dependent parameters, the 36 Hermite coefficients were interpolated at the resolution of the final numerical model.

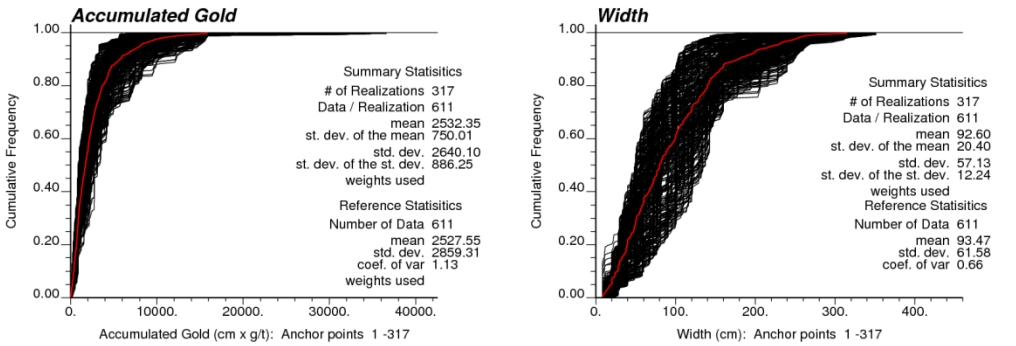

Figure 5-10: Location dependent cdfs for the accumulated gold and reef width attributes obtained using a Gaussian kernel with 400m bandwidth at 317 anchor points.

Location-Dependent Correlograms Fitting

The experimental location-dependent correlograms use the locally transformed Gaussian values of both accumulated gold and width attributes. At each anchor point 6 directions and 12 lag distances were considered. The angular separations and angular tolerances are 30 and 20 degrees, respectively, and the lag distances and distance tolerances are 200m and 150m, respectively.

The fitting of the location dependent correlograms was performed semi-automatically at 317 anchor point locations giving higher weight to the experimental correlogram points at shorter lags. A single exponential structure was used as the variogram model at every anchor point. Local variogram models with anomalous high values for the fitted nugget effect and the anisotropy ratio were identified and fitted manually or automatically using different criteria. For the accumulated gold, the automatic local variogram fitting at 14% of anchor points yielded locally anomalous or abruptly changing parameters, particularly for the anisotropy orientation and the ranges. At those locations the original fit was replaced by alternative fits obtained by weighting the experimental correlograms points proportionally to the sum of the involved 2-point weights, or by those obtained without applying any weights to the experimental points. The criteria for choosing between these alternate fits were the minimization of the square error and the coherence with the parameters fitted at neighbouring locations. For a few anchor point locations, manual variogram modelling was required. Only 5% of the local variogram models for the reef with attribute yielded locally anomalous and abruptly changing parameters and required fitting by alternate criteria or manual fitting.

Figure 5-11shows the location-dependent variogram model parameters fitted for the accumulated gold attribute, while Figure 5-12 shows the same local parameters for the width attribute. These graphs represent the local variogram parameter fields that are used in locally stationary estimation and simulation.

The global experimental correlograms for both attributes were calculated for the same orientations and lag distances and using the same angular and distance tolerance parameters as for the locally stationary ones. Table 5-2 presents the fitted model parameters for both attributes

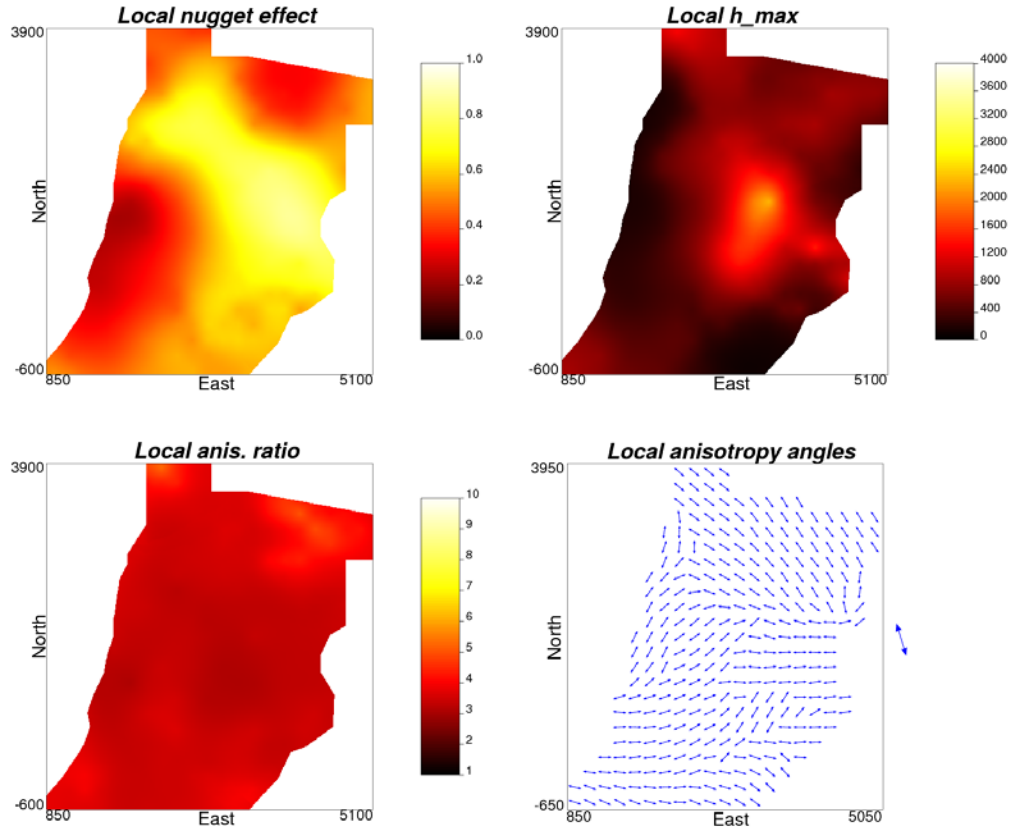

Figure 5-11: Local exponential variogram model parameters for the accumulated gold.

Table 5-2: Model parameters for the global correlograms

Attribute	Nugget Effect	Sill Contribution	Anisotropy Orientation	Maximum Range (m)	Minimum Range (m)
Accumulated gold	0.5	0.5	130	730	230
Width	0.25	0.75	125	880	290

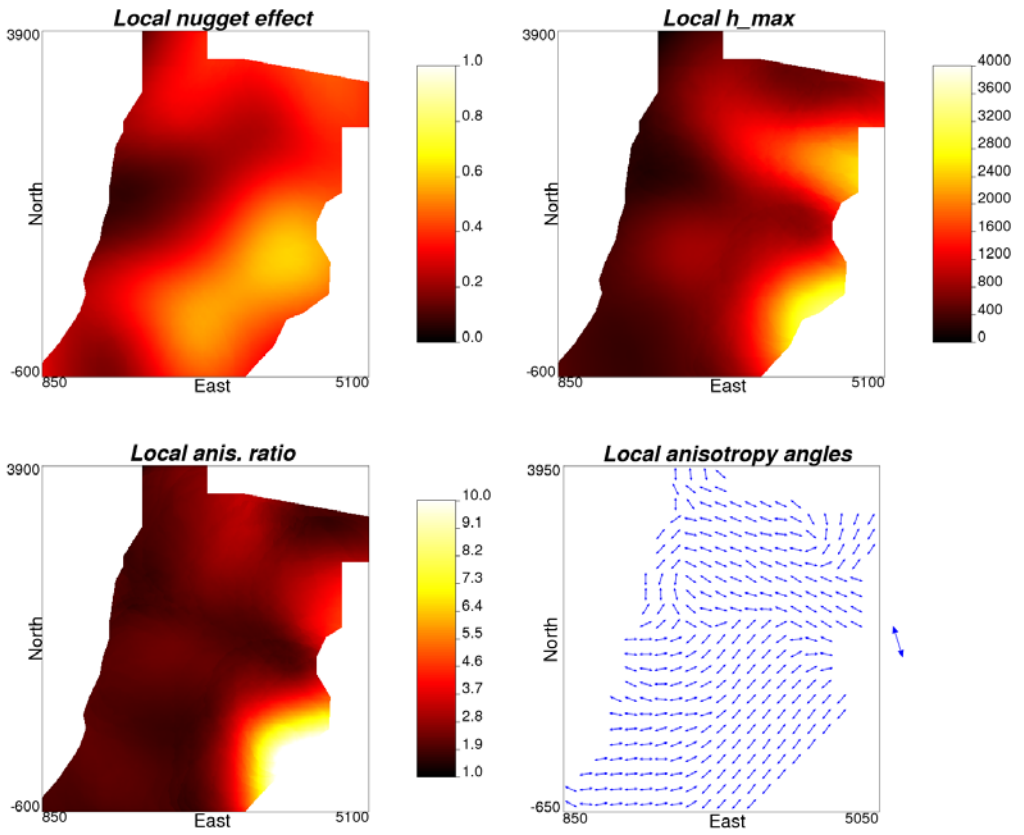

Figure 5-12: Local exponential variogram model parameters for the reef width

5.1.3. Locally Stationary Estimation

As described in Chapter 4, locally stationary multiGaussian kriging (LSMGK) uses the gridded local variogram model parameters and Hermite coefficients to modify the covariance matrix and the normal scores transformation table at each estimation location. The performance of the LSMGK algorithm in point estimation with its stationary counterpart is assessed by comparing the cross-validation results and the classification errors on the confirmatory dataset. Each sample group corresponding to a simulated ddh and its deflections was taken out and their locations re-estimated during cross-validation. The high variability of the attribute values and the low sample density of the simulated ddh intercepts results in poor Cross-validation statistics for both, the accumulated gold and the width attributes.

Estimation of Accumulated Gold Values

As it is shown in Figure 5-13, LSMGK yields a noticeably lower mean squared error, a much higher covariance, and a higher coefficient of correlation between estimated and true values than the equivalent stationary algorithm. Moreover, the incorporation of local variogram models and locally transformed results in reduced smoothing of the estimates; this is evidenced by the larger standard deviation obtained for LSMGK. Underestimation of extreme values are slightly less pronounced in the locally stationary cross-validation results. The absolute average of the four worst overestimations is 3779 cm x g/t for MGK while it is 3865 cm x g/t for LSMGK. The absolute averages of the four worst underestimations are 12711 cm x g/t and 11281 cm x g/t for MGK and LSMGK, respectively.

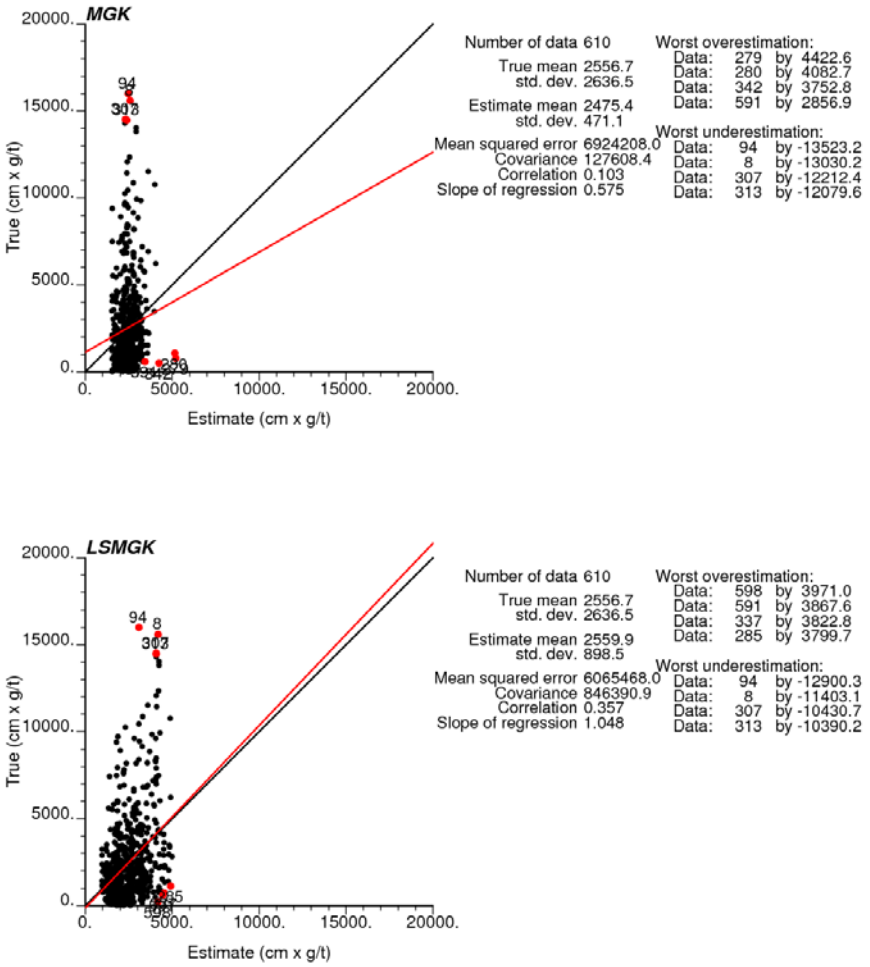

Figure 5-13: Cross-validation results for accumulated gold in Facies 2 of stationary multiGaussian kriging (top) and locally stationary multiGaussian kriging (bottom). One true value above 20000 cm x g/t has been trimmed from this figure.

The slope of the regression in the MGK cross-validation results (also in Figure 5-13, top) is lower than one, 0.575. This relatively low slope value is due to the sensitivity to outliers in the regression. For LSMGK the slope of the regression is slightly higher than 1. Figure 5-14 shows the resulting maps for MGK and LSMGK estimates of the accumulated gold values. The map produced by the stationary algorithm shows a uniform pattern of spatial continuity, while the impact of locally changing variogram models is clearly visible in the LSMGK map.

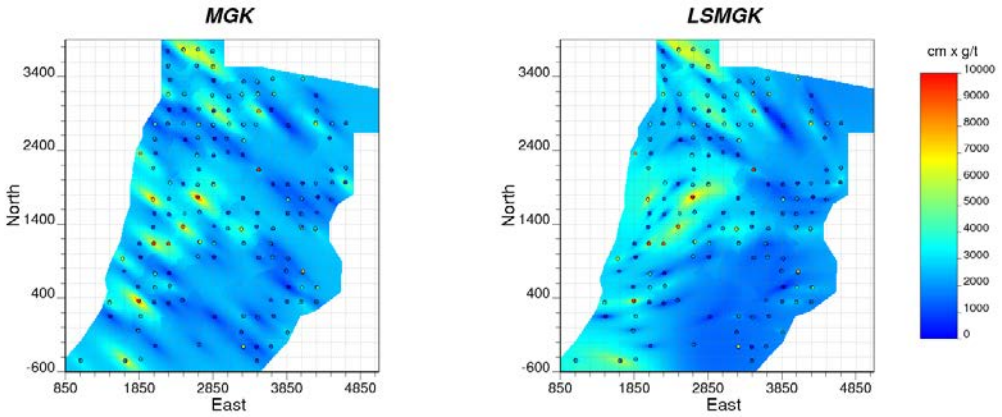

Figure 5-14: Estimates maps of the accumulated gold obtained for stationary multiGaussian kriging (left) and locally stationary multiGaussian kriging (right)

Figure 5-15, left, presents a histogram of the differences between LSMKG and MGK estimates of the accumulated gold. These differences are due not only to the incorporation of local variogram model parameters but also to the use of locally normal score transformed cdfs. The quartiles in the histogram of differences between estimates define four classes whose extend is shown in Figure 5-15, right. The upper class, where the LSMGK are much larger than the MGK estimates, is prevalent in the west side of the domain, where the local means are higher than the global mean. By contrasts, the lower class of estimates differences is prevalent at the east side, where the local means are smaller than the global mean. Thus, while the stationary MGK estimates fluctuate around the global mean, spatial prediction with LSMGK allows incorporating the trend in the mean.

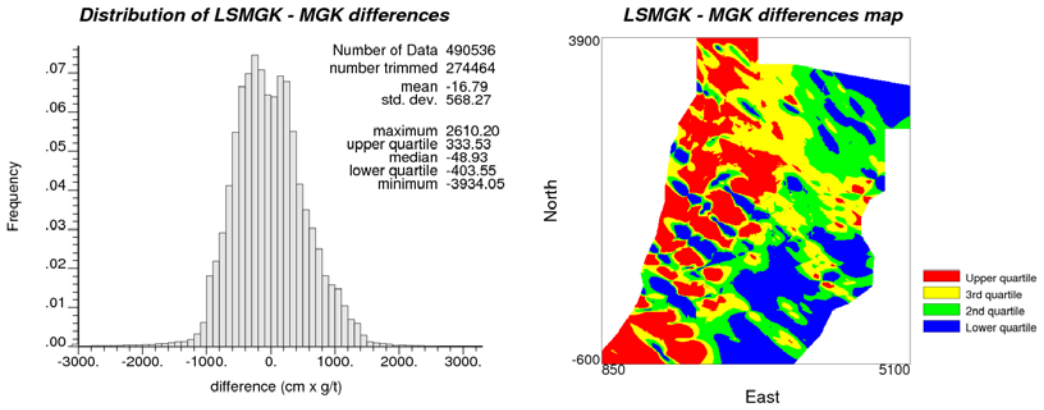

Figure 5-15: histogram of differences between LSMGK and MGK estimates (left) and location of the four classes defined by the quartiles of the histogram of differences.

The performance of the stationary and locally stationary estimation techniques in the correct classification of estimates as being above or below a given threshold is assessed by comparing the estimates with the complete dataset. Table 5-3 shows the classification errors taking the median of the sample values, this is 1722 cm x g/t, for ore and waste selection. There is a noticeable increase (33%) in the correct identification of waste locations for the LSMGK estimates in comparison with the MGK estimates, but the locations correctly identified as ore decrease slightly (-2%). The number of locations misclassified as ore is reduced in 7%, but the ore locations misclassified as waste are increased in 16%. Overall, a reduction of 3.8% in the misclassified locations is observed.

Table 5-3: Classification errors above the median for the accumulated gold estimates

MGK

TRUE (cm x g/t)	ESTIMATES (cm x g/t) Waste	Ore
Waste	9665	46487
Ore	6986	45854

LSMGK

TRUE (cm x g/t)	ESTIMATES (cm x g/t) Waste	Ore
Waste	12833	43319
Ore	8115	44725

Estimation of the Reef Width

Cross-validation results for the reef width attribute show an increased accuracy for the LSMGK technique (see Figure 5-16). This is shown by the lower mean square error and increased covariance between the true values and the locally stationary estimates. The higher covariance and the slightly lower standard deviation translate to a coefficient of correlation higher than the observed for the MGK cross-validation. The slope of the estimates versus true values regression line for the stationary technique is around 0.81,

while for LSMGK it is slightly greater than one. Additionally, the averages of the worst overestimation and underestimation are slightly smaller for the locally stationary technique. Figure 5-17 presents the maps of MGK and LSMGK estimates. The locally changing anisotropic directions informed by the local variogram model parameters can be observed in the right side of this figure.

As for the accumulated gold estimation, the complete dataset is used to compare the classification performance of the stationary and locally stationary methods. As presented in Table 5-4, the second results in a small reduction of the misclassification (3.3%) above the median value (80cm).

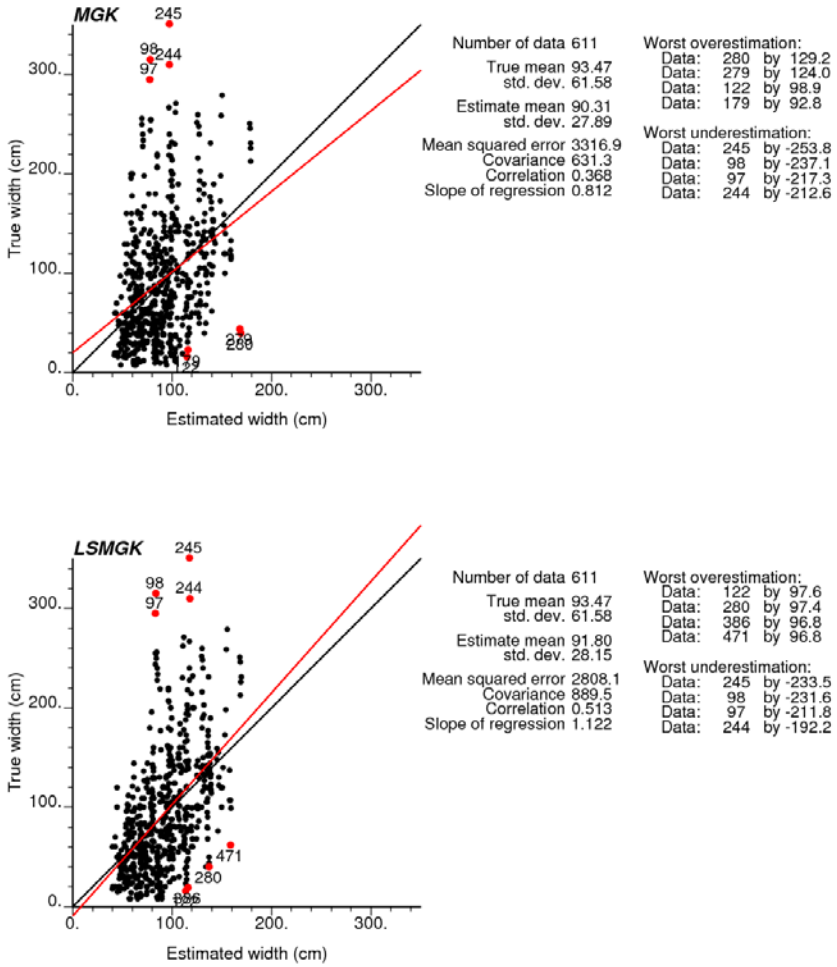

Figure 5-16: Cross-validation results for the reef width in Facies 2 of stationary multiGaussian kriging (top) and locally stationary multiGaussian kriging (bottom)

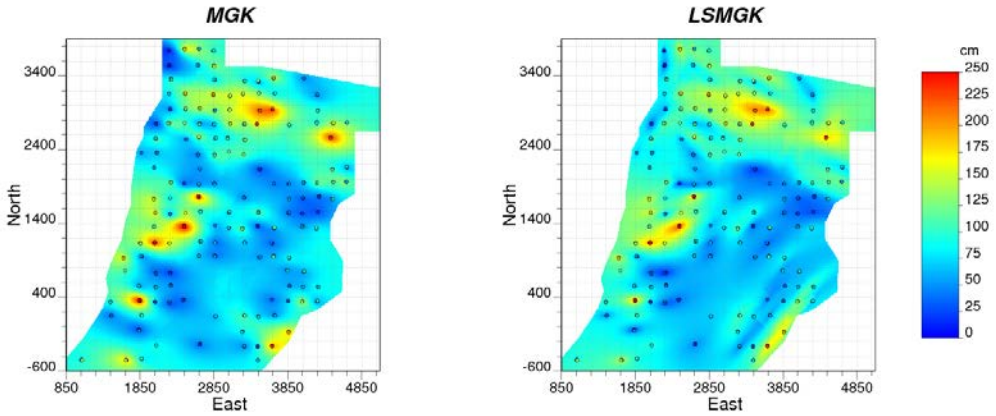

Figure 5-17: Estimates maps of the reef width obtained for stationary multiGaussian kriging (left) and locally stationary multiGaussian kriging (right)

Table 5-4: Classification errors above the median for the reef width estimates

MGK

TRUE (cm)	ESTIMATES (cm) Waste	Ore
Waste	30252	20412
Ore	17915	40413

LSMGK

TRUE (cm)	ESTIMATES (cm) Waste	Ore
Waste	29128	21536
Ore	15524	42804

5.2. Discussion

The impact of incorporating location dependent statistics is clearly visible in the resulting estimates maps. The piecewise linear features obtained contrast with the uniform patterns produced by stationary techniques. These non-stationary features can be validated in relation to the geological background knowledge, if it is available. Beyond the locally changing patterns of spatial continuity that can be assessed visually, the locally stationary numerical models can offer improved performance in the accuracy of estimates and in the correct classification above a given cut-off. For the dataset used in the present case study, these improvements are rather modest, particularly in the ore/waste classification. This is due to data scarcity and the high variability of the attributes values. This makes the inference of location-dependent statistics difficult. It is worthwhile to note that for this particular dataset the improvements of the locally stationary estimation are much clearer in the cross-validation results than in the confirmation of the estimates with true values. This indicates that the location-dependent statistics adapt well to the changes informed by available samples, but they do not necessarily reflect the true non-stationary features of the attribute.

Although the improvements obtained by the application of locally stationary techniques may seem rather modest, even a small reduction of misclassification may result in substantial profits when translated into the framework of a mining operation.

Chapter 6
Conclusions and Future Work

The idea of conditioning the statistics by distance weighting functions in order to capture different aspects of non-stationarity has appeared in different contexts and for separated applications (Fotheringham et al. 2002; Ren 2007; McLennan 2007). This thesis proposes to use it under the assumption of local stationarity for a comprehensive treatment of the non-stationarity of the RF cdf and its parameters. Other techniques focus on particular aspects of non-stationarity, such as the trend or the local anisotropy. The proposed methodology is demonstrated to be a viable comprehensive alternative for non-stationary geostatistical modelling.

The methodology for locally stationary modelling presented in this thesis spans from the inference and modelling of the local statistics to their use in spatial prediction. Several algorithms were developed to implement the different stages of this approach (see Appendix). These algorithms are ready to be used by practitioners; however, their application must be undertaken keeping in mind important remarks on the proposed methodology. Several areas related to this methodology require further research. This last chapter discusses these remarks and presents areas of future research.

6.1. Concluding Remarks

The local stationarity assumption is the basis of the proposed methodology. Under this assumption, the Random Function cdf and all the required statistics are assumed invariant by translation only in relation to an anchor point. If the anchor point location changes the Random Function cdf changes.

This methodology can be subdivided into different sub-processes: selection of the distance weighting functions, inference of the location-dependent cdfs and its statistics, modelling of the local cdfs and variograms, local normal scores transformation, spatial prediction and model validation. The following remarks are grouped accordingly.

On the Selection of the Distance Weighting Function and the Inference of the Location-Dependent Statistics

A key aspect of locally stationary modelling under the locally stationary assumption is the inference of the anchored RF cdf and its statistics. These local statistics are intended to model the different aspects of non-stationarity. The approach explored in this thesis for the inference of the local statistics is based on the use of anchored distance weighting kernel functions applied to the available data. A critical aspect is the choice of the distance weighting function parameters. Although this choice can be supported by numerical measures, it is mostly left to the practitioner's judgement. When choosing these parameters the practitioner must be aware of the related variance/bias trade-off. A narrow bandwidth and very low background constant may capture smaller non-stationary features but they may render the local statistics unstable and cause overfitting. On the other hand, a very wide bandwidth could result in excessive smoothing of the local statistics that may mask non-stationary features. The selection of the bandwidth also depends on data density. The distance weighting function must capture local trends informed by groups of samples rather than reflect the local influence of a few individual values. Therefore, if data is sparse, using a wider bandwidth would be preferred.

The Gaussian kernel was shown to be a reasonable distance weighting function for the inference of the local statistics. This allows the inference of smoothly changing local statistics using all available data, while fulfilling desirable properties, such as strict positivity, unbiasedness, continuous decrease with distance, and independence of units. Other forms of smoothly decreasing kernel functions with the same properties could be considered. A correction by declustering weights allows correcting for preferential sampling in the inference of the location-dependent statistics.

At each anchor point location the same distance weights are used for inferring the local cdf and the 2-point statistics. Proceeding in this way assures the consistency between all the required statistics. 1-point distance weights are used for the inference of the local marginal cdf and the 1-point statistics. These 1-point weights are combined in pairs to form 2-point weights that are used for the inference of location-dependent measures of spatial continuity. The geometric average of pairs of 1-point weights is preferred for obtaining the 2-point weights because this avoids possible biases caused by other mixture rules in the inference of the local measures of spatial continuity, particularly for large 2-point separations. Additionally, the 2-point weights revert to 1-

point weights when the distance between them goes to zero, allowing the consistency between the1-point and 2-point statistics.

The locally weighted measures of spatial continuity are able to adapt to changes in the anisotropy range and orientation informed by the data. The capability of location-dependent 2-point statistics to adapt to the changes in the local anisotropy is diminished when the distance kernel bandwidth is too narrow or too wide. When a very narrow kernel bandwidth is used, only a few samples in the vicinity of the anchor point contribute significantly to the inference of the experimental local measures of spatial continuity. This translates to high fluctuations of the experimental values, particularly when local outliers are present in data. When a very wide bandwidth is used instead, the local changes in the anisotropy are not captured efficiently. The practitioner must judge a suitable kernel bandwidth based on visual inspection of the experimental location-dependent measures of spatial continuity. These should vary smoothly from one anchor point to another. Among these experimental statistics, location-dependent correlograms appear to be more robust than location-dependent variograms when data presents high variability and local outliers.

On the Modelling of the Location-Dependent Variograms and the Local Normal Scores Transformation

At the different anchor points the experimental measures of spatial continuity are fitted using one of the allowable variogram models. The local variogram model parameters identify local changes in the anisotropy range, orientation, and short scale continuity. If sampling is dense enough in relation to the local ranges of spatial continuity, a model that allows a locally varying shape, such as the Stable Model can be considered.

Since the experimental 2-point statistics are normally inferred for multiple anchor point locations, it is necessary to employ a semiautomatic variogram fitting algorithm. However, several factors, such as the high variability of data values, the presence of locally anomalous values and local data scarcity may cause abrupt fluctuations and unwarranted anomalies in the fitted variogram model parameters. When data is dense and continuous the occasional abrupt fluctuations and unrealistic parameter values obtained from currently available variogram fitting algorithms can be fixed by manual fitting. In other cases, a robust methodology for semiautomatic locally stationary variogram modelling that allows more control by the user is required. The algorithm proposed is

based in the iterative joint minimization of the weighted square errors between the variogram model and the experimental points at multiple anchor points. The weights for variogram fitting can be set inversely proportional to the lag distance, directly proportional to the amount of information used to infer the experimental points, or both. A penalty function allows the user to minimize the occurrence of local parameters that exceed previously defined thresholds. Another penalty function penalizes the local parameters that strongly depart from the parameters fitted at surrounding anchor points within a previously defined neighbourhood. Additionally, the fitting of local variogram model parameters such as the anisotropy angles and ratios can be improved by geological knowledge, if available.

Gaussian-based estimation and simulation techniques require the transformation of the original distribution into a standard Gaussian distribution. The local normal scores transformation incorporates locally changing means, variances and distribution shapes. The Hermite modelling of the local normal scores transformation is more efficient than storing the local transformation tables for all anchor points, particularly for large datasets. Nevertheless, the Hermitian models can introduce small inaccuracies in the reproduction of the backtransformed cdfs. These are caused by fluctuations of the Hermitian transformation function model, this particularly occurs when the experimental transformation function has gaps or spikes.

The processes of location-dependent variogram inference and modelling, and local normal scores transformation and modelling can be very demanding in time and computer resources if performed at all the locations to be estimated or simulated. Therefore, this is performed only at a limited number of anchor points. The locations of these anchor points are chosen in order to allow the reconstruction of the smoothly changing local statistics and parameters by interpolating their inferred values between those points. The local variogram parameters and local Hermite coefficients do not necessarily average linearly, but it is reasonable to reconstruct their variation between anchor points by interpolation if they change smoothly from one anchor point to another. An adequate anchor point separation minimizes the number of required anchor points while keeping the error introduced by the interpolation within tolerable limits. Since local 1-points statistics are relatively straightforward to infer, these are used to assess the trade-off between the number of anchor points and the error introduced by interpolation.

On Locally Stationary Spatial Prediction

The modified estimation and simulation algorithms use the interpolated local variogram parameter values and Hermite polynomials under the assumption of local stationarity. The normal scores transformation and the covariance matrix are updated at every location. The local normal scores transformation allows the incorporation of the trend in the mean and other local changes in the distribution into locally stationary Gaussian estimation and simulation. Additionally, the use of Hermite polynomials for modelling the local normal scores transformation allows a straightforward implementation of a local change of support model for block support estimation

The incorporation of locally changing variogram models results in numerical models that are richer in local features of spatial continuity. The resulting changing patterns of spatial continuity contrast with the uniform pattern observed in models built by traditional methods. Moreover, the variances obtained from locally stationary estimation respond not only to the data availability and configuration, but they are also enriched with information on the local variability. These locally changing features should be validated by geological interpretation if it is available.

On the Validation of Resulting Models

Beyond the richer spatial features resulting from locally stationary modeling, it has the potential of improved accuracy, precision and selectivity, and also in increased connectivity. This occurs between the extreme scenarios of having very low or very high sampling density. If samples are very scarce, the location-dependent statistics are unreliable because they are highly affected by individual sample values rather than by local trends. If samples are very dense, they have more influence than the statistics and parameters used in the spatial prediction algorithm. So, in the first case, the locally stationary methods may actually perform worse than their stationary counterparts due to the poor definition of the local moments, particularly if supporting geological or secondary information is missing. In the second case, although the location-dependent statistics can be robustly inferred, stationary and locally stationary methods will tend to perform similarly. It is between these two extreme cases where locally stationary techniques based on distance weighted statistics can improve the numerical modelling of geological attributes.

When comparing cross validation results between traditional and locally stationary estimation techniques, the latter can show improved results in terms of the minimization of the mean square error and the increase of the correlation between true and estimated values, among other metrics. However, when the resulting estimated maps are confirmed against exhaustively sampled values, the locally stationary estimation does not necessarily outperform the traditional kriging. This may occur when data is not abundant enough to allow the robust inference of the local statistics.

A practical disadvantage of the locally stationary approach is the increased computational effort that its application requires. These algorithms demand higher memory storage and processing time than their stationary counterparts.

6.2. Future Work

Several possible avenues of future research related to the different components of the proposed methodology are delineated in this section.

Assessment of the Location-Dependent Statistics

The ability of location-dependent statistics to improve the modelling of rock attributes is assessed after locally stationary estimation or simulation and with the help of cross-validation and accuracy plots. An a priori evaluation of the location-dependent statistics without the need to complete the spatial prediction and the required previous steps would be useful for highlighting areas with different degrees of robustness. Metrics like local data density, cross-validation and bootstrapping errors during the inference of the location-dependent statistics could be used for such purpose. These metrics may lead to uncertainty measures for the local statistics that can be translated to a more complete uncertainty assessment of the attribute.

Use of Non-Euclidean Distances

The weighting kernels presented for the inference of the location-dependent statistics are based on straight-line Euclidean distances. A reliable inference of the location-dependent statistics becomes difficult if samples are separated by distances that exceed the non-linear features of the geological setting. In some geological settings where the non linear patterns can be recognized from the knowledge base, it may be reasonable to use non-

Euclidean distance measures (Christakos et al. 2001, pp.19-24; Curriero 2006) for the kernel weighting function. The non-Euclidean metrics could also be used to transform the original space to an Euclidean space. Since this transformation may filter out local variations in the anisotropy orientation, the weighting kernel would be mainly used for inferring the cdf and the remaining local parameters in the transformed space. Locally stationary spatial prediction would be performed in the transformed space, and subsequently backtransformed to original space. This alternative contrasts with approaches that use a dissimilarity distance to transform a non-Euclidean space into a high-dimensional Euclidean isotropic space (Sampson & Guttorp 1992; Boisvert et al. 2009).

Alternatives to the Direct Inference of the Location-Dependent Statistics

Alternative ways to obtain the location-dependent statistics that require further development include their inference with the support of exhaustively sampled secondary information and their inference from representative training images. In this second idea, a training image containing patterns that are deemed as representative of a particular geological setting is used to extract the location-dependent statistics and parameters, particularly those required for local variogram modelling. The resulting histograms of these local parameters are used for co-simulating different possible scenarios of the regionalized spatial distribution of the parameters. These parameter maps are subsequently used in locally stationary estimation or simulation with the available data values. Proceeding in this way could allow for the reproduction of the non-stationary patterns of the attribute, but the resulting models would not necessarily be more precise or accurate.

Relation with Multiple Point Statistics

The mixture rules used for obtaining the weights to be used in the inference of the local 2-point statistics can also be used to obtain combinations of weights assigned to multiple samples. These multiple point weights could be used for the inference of local higher order measures of spatial continuity (Christakos 2000, pp.107-108; Dimitrakopoulos et al. 2010; Mustapha & Dimitrakopoulos 2010). A related area of future research is to compare the results of locally stationary simulation with those obtained using algorithms based on multiple point statistics (Strebelle 2002; Lyster 2009).

Robust Inference of the Local 2-point Statistics

A robust estimation of the location-dependent experimental measures of continuity is a key aspect of the proposed methodology. So far, this task has been accomplished by the standard practice of grouping the data pairs according to the lag and angular separations and their respective distance and angular tolerances. The parameters that define these groups are chosen in relation to the global experimental variogram and used at all anchor points. In the locally stationary framework this practice may yield unreliable estimates of the local variograms in scarcely sampled areas. An alternative may be to obtain the local experimental variograms by smoothing the local variogram cloud directly rather than within bins defined by lag distances and orientations. One way to achieve this is by using Bernstein polynomials (Manchuk & Leuangthong 2008).

Improved Modelling of the Local Variograms

The proposed joint local variogram fitting algorithm still requires further testing and debugging with complex 2-D and 3-D datasets. Consistent variogram modelling in 3-D is still a difficult task, particularly when no geological knowledge is available for guiding the fitting of anisotropy orientations and ranges. Alternatives to the weighted least squares criterion for parametric variogram modelling, such as the use of Fourier-Bessel matrices (Genton & Gorsich 2002), should be considered. These could provide a more flexible option than the modelling of local variograms by imposing a unique variogram model for all anchor point locations.

Several other aspects of local variogram modelling require further research. Among them are the fitting of local variogram models in the presence of locally varying short scale anisotropies but global large scale zonal and geometric anisotropies, and the effect of cyclicity on the attribute values.

Multivariate Location-Dependent Statistics

Multivariate variogram modelling can be a complicated task in stationary Geostatistics, particularly when multiple variables are involved; this could be much more challenging in the proposed locally stationary framework. Therefore, the integration of local multivariate statistics with the locally stationary framework requires further development and research. The difficulties related to the inference and modelling of the locally stationary

cross variograms or covariances can be alleviated by recurring to the Markov-type approximation (Journel 1999). With this model the location-dependent cross correlograms between the primary and collocated secondary variables can be approximated by the product between the local correlation coefficients and the local correlograms. A known problem with the use of the Markov-type assumption in collocated cokriging is the resulting variance inflation, which hinders the histogram reproduction in simulated realizations. The development of locally stationary Gaussian cosimulation should consider alternatives capable to solve this problem, such as the intrinsic model of coregionalization (Wackernagel 2003, pp.154-157; Babak & Deutsch 2009).

Locally changing correlation coefficients have been implemented by Ren (2007, pp.39-47) for the enhanced integration of different scale data by Bayesian updating. This technique could be further improved by the incorporation of location-dependent variograms and cdfs.

Validation of the Locally Stationary Models

Checking the global histogram reproduction by locally stationary simulation results is straightforward. Contrarily, verifying the reasonable reproduction of the local spatial continuity informed by the local variogram models may be difficult. This is not only because it can be tedious to plot and check the locally stationary variograms for multiple realizations, anchor points and directions, but also because of the difficulty to reproduce the local variogram beyond the closer lags and the increased ergodic fluctuations. These aspects require further research; practical criteria for checking the local variogram reproduction are required. A related issue to the validation of locally stationary simulations is the mathematical consistency of the updated RF. The locally stationary RF may be well defined in relation to the location where it is anchored, but its global consistency requires further investigation.

The locally stationary estimation variances are richer in spatial information than their traditional counterparts (see Figure 4-3). Further studies are required to determine if the locally stationary variances can be used as reliable measures of the uncertainty of estimates.

The incorporation of local cdfs and variograms do not always result in a reduced updated local uncertainty when compared with those obtained from traditional methods

(see Figure 4-20). More testing with different datasets is required to determine under which conditions the locally stationary ccdfs are narrower or wider and establish if these differences yield a more realistic characterization of the uncertainty. Some related subjects to be investigated involve the impact of the location-dependent statistics on the optimum sample location and in the resource classification based in the quantification of uncertainty.

Although the proposed locally stationary approach requires more research and testing, the set of methodologies and tools presented in this thesis are ready to be applied on real case studies.

Bibliography

Armstrong, M. & Jabin, R., 1981. Variogram models must be positive-definite. *Mathematical Geology*, 13(5), 455-459.

Atkinson, P.M. & Lloyd, C.D., 2007. Non-stationary variogram models for geostatistical sampling optimisation: An empirical investigation using elevation data. *Computers & Geosciences*, 33(10), 1285-1300.

Babak, O. & Deutsch, C.V., 2009. An intrinsic model of coregionalization that solves variance inflation in collocated cokriging. *Computers & Geosciences*, 35(3), 603-614.

Bárdossy, G. & Fodor, J., 2001. Traditional and new ways to handle uncertainty in geology. *Natural Resources Research*, 10(3), 179–187.

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

Borradaile, G., 2003. *Statistics of Earth Science Data: Their Distribution in Time, Space, and Orientation*, Berlin; New York: Springer.

Brunsdon, C., Fotheringham, A.S. & Charlton, M., 1998. Geographically weighted regression—modelling spatial non-stationarity. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 47(3), 431–443.

Brunsdon, C., Fotheringham, A.S. & Charlton, M., 2002. Geographically weighted summary statistics—a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, 26(6), 501–524.

Burrough, P.A., 1981. Fractal dimensions of landscapes and other environmental data. *Nature*, 294(5838), 240-242.

Chilès, J., 2004. *La Modélisation Géostatistique de la Variabilité Spatiale et ses Applications*. Habilitation à diriger des recherches. Paris: Université Pierre et Marie Curie.

Chilès, J. & Delfiner, P., 1999. *Geostatistics : Modeling Spatial Uncertainty*, New York: John Wiley & Sons Inc.

Christakos, G., 2000. *Modern Spatiotemporal Geostatistics*, New York: Oxford University Press.

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

Christakos, G., Bogaert, P. & Serre, M.L., 2001. *Temporal GIS: Advanced Functions for Field-Based Applications*, Berlin; New York: Springer.

Cowan, E. et al., 2002. Rapid geological modelling. In S. Vearncombe, ed. *Applied Structural Geology for Mineral Exploration and Mining International Symposium*. Kalgoorlie, Western Australia: West Perth : The Institute.

Cressie, N., 1985. Fitting variogram models by weighted least squares. *Mathematical Geology*, 17(5), 563-586.

Cressie, N., 1986. Kriging nonstationary Data. *Journal of the American Statistical Association*, 81(395), 625-634.

Curriero, F., 2006. On the use of non-Euclidean distance measures in geostatistics. *Mathematical Geology*, 38(8), 907-926.

David, M., 1977. *Geostatistical Ore Reserve Estimation*, Amsterdam: Elsevier Scientific Pub. Co.

Davis, B.M., 1987. Uses and abuses of cross-validation in geostatistics. *Mathematical Geology*, 19(3), 241–248.

Davis, J., 2002. *Statistics and Data Analysis in Geology* 3rd ed., New York: John Wiley & Sons Inc.

Dean, R.B. & Dixon, W.J., 1951. Simplified statistics for small numbers of observations. *Analytical Chemistry*, 23(4), 636-638.

Delfiner, P., 1976. Linear estimation of nonstationary spatial phenomena. In M. Guarascio, C. J. Huijbregts, & M. David, eds. *Advanced Geostatistics in the Mining Industry*. NATO advanced study institutes series : Series C, Mathematical and physical sciences. 49-68: Springer, pp. 49-68.

Delhomme, J.P., 1978. Kriging in the hydrosciences. *Advances in Water Resources*, 1(5), 251-266.

Deutsch, C.V., 1995. *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data*. PhD Thesis. Stanford, CA: Stanford University.

Deutsch, C.V., 2005. *Check histogram reproduction (histpltsim)*, Centre for Computational Geostatistics (CCG).

Deutsch, C.V., 1989. DECLUS: A fortran 77 program for determining optimum spatial declustering weights. *Computers & Geosciences*, 15(3), 325-332.

Deutsch, C.V., 1996. Direct assessment of local accuracy and precision. In E. Baafi & N. Schofield, eds. *Geostatistics Wollongong '96*. Fisrt International Geostatistics Congress. Dordrecht: Kluwer Academic, pp. 115-125.

Deutsch, C.V., 1998. FORTRAN programs for calculating connectivity of three-dimensional numerical models and for ranking multiple realizations. *Computers & Geosciences*, 24(1), 69-76.

Deutsch, C.V., 2002. *Geostatistical Reservoir Modeling*, Oxford University Press.

Deutsch, C.V., 2007. *Variogram of irregularly spaced data (gamv2004)*, Centre for Computational Geostatistics (CCG).

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

Deutsch, C.V. & Zanon, S., 2002. UltimateSGSIM: Non-stationary sequential gaussian cosimulation by rock type. In *Center for Computational Geostatistics, Report 4*. Edmonton: University of Alberta, p. Paper 51.

Diggle, P. & Ribeiro, P.J., 2007. *Model-Based Geostatistics*, New York: Springer.

Dimitrakopoulos, R., Mustapha, H. & Gloaguen, E., 2010. High-order statistics of spatial random fields: exploring spatial cumulants for modeling complex non-Gaussian and non-linearphenomena. *Mathematical Geosciences*, 42(1), 65-99.

Emery, X., 2005. Simple and ordinary multiGaussian kriging for estimating recoverable reserves. *Mathematical Geology*, 37(3), 295-319.

Emery, X., 2007a. Simulation of geological domains using the plurigaussian model: New developments and computer programs. *Computers and Geosciences*, 33(9), 1189–1201.

Emery, X., 2007b. Two ordinary kriging approaches to predicting block grade distributions. *Mathematical Geology*, 38(7), 801-819.

Emery, X. & Kremer, F., 2008. A survey of random field models for simulating mineral grades. In *Proceedings of the Eight International Geostatistics Congress*. GEOSTATS 2008. Santiago, Chile: University of Chile, pp. 157-166.

Emery, X. & Ortiz, J.M., 2005. Estimation of mineral resources using grade domains: critical analysis and a suggested methodology. *Journal of The South African Institue of Mining and Metallurgy*, 105(04), 247-256.

Evans, A., 1997. *An Introduction to Economic Geology and its Environmental Impact*, Oxford; Malden MA: Blackwell Science.

Fotheringham, A.S., 1997. Trends in quantitative methods 1: stressing the local. *Progress in Human Geography*, 21, 88–96.

Fotheringham, A.S., Brunsdon, C. & Charlton, M., 2002. *Geographically Weighted Regression*, John Wiley & Sons Inc.

Fuentes, M., 2001. A new high frequency kriging approach for nonstationary

environmental processes. *Environmetrics*, 12(5), 469–483.

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

Gendzwill, D. & Stauffer, M., 1981. Analysis of triaxial ellipsoids: Their shapes, plane sections, and plane projections. *Mathematical Geology*, 13(2), 135-152.

Genton, M.G. & Gorsich, D.J., 2002. Nonparametric variogram and covariogram estimation with Fourier-Bessel matrices. *Computational Statistics & Data Analysis*, 41(1), 47-57.

Goovaerts, P., 2000. Estimation or simulation of soil properties? An optimization problem with conflicting criteria. *Geoderma*, 97(3-4), 165–186.

Goovaerts, P., 2001. Geostatistical modelling of uncertainty in soil science. *Geoderma*, 103(1-2), 3–26.

Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*, Oxford University Press US.

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

Guibal, D., 1987. Recoverable reserves estimation at an Australian gold project. In G. F. Matheron & M. Armstrong, eds. *Geostatistical Case Studies*. Quantitative geology and geostatistics. Dordrecht: Springer, pp. 149-168.

Guo, H. & Deutsch, C.V., 2008. Choosing and adequate number of conditioning data for kriging. In *Centre for Computational Geostatistics, Report 10*. Edmonton: University of Alberta, p. Paper 122.

Haas, T.C., 1990a. Kriging and automated variogram modeling within a moving window. *Atmospheric Environment. Part A. General Topics*, 24(7), 1759-1769.

Haas, T.C., 1990b. Lognormal and moving window methods of estimating acid deposition. *Journal of the American Statistical Association*, 85(412), 950-963.

Härdle, W., 1992. *Applied Nonparametric Regression*, Cambridge, UK: Cambridge University Press.

Higdon, D., 1998. A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, 5(2), 173–190.

Higdon, D., Swall, J. & Kern, J., 1998. Non-stationary spatial modeling. *Bayesian Statistics*, 6, 761-768.

Houlding, S., 2000. *Practical Geostatistics Modeling and Spatial Analysis*, Berlin; New York: Springer.

Hughes-Oliver, J.M. et al., 1998. Parametric nonstationary correlation models. *Statistics & Probability Letters*, 40(3), 267-278.

Isaaks, E.H., 1991. *The Application of Monte Carlo Methods to the Analysis of Spatially Correlated Data*. PhD Thesis. Stanford, CA: Stanford University.

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

Johnson, R. & Wichern, D., 2007. *Applied Multivariate Statistical Analysis* 6th ed., Upper Saddle River N.J.: Pearson Prentice Hall.

Journel, A.G., 1989. *Fundamentals of Geostatistics in Five Lessons*, Washington D.C.: American Geophysical Union.

Journel, A.G., 1986. Geostatistics: Models and tools for the earth sciences. *Mathematical Geology*, 18(1), 119-140.

Journel, A.G., 1999. Markov models for cross-covariances. *Mathematical Geology*, 31(8), 955–964.

Journel, A.G., 1980. The lognormal approach to predicting local distributions of selective mining unit grades. *Mathematical Geology*, 12(4), 285–303.

Journel, A.G. & Alabert, F., 1989. Non-Gaussian data expansion in the Earth Sciences. *Terra Nova*, 1(2), 123-134.

Journel, A.G. & Huijbregts, C.J., 1978. *Mining Geostatistics*, London: Academic Press Inc.

Journel, A.G. & Kyriakidis, P.C., 2004. *Evaluation of Mineral Reserves: A Simulation Approach*, Oxford University Press.

Korvin, G., 1982. Axiomatic characterization of the general mixture rule. *Geoexploration*, 19(4), 267-276.

Krige, D.G., 1978. *Lognormal-de Wijsian Geostatistics for Ore Evaluation*, South African Institute of Mining and Metallurgy.

Langlais, V., Beucher, H. & Renard, D., 2008. In the shade of the truncated gaussian simulation. In *Proceedings of the Eight International Geostatistics Congress*. GEOSTATS 2008. Santiago, Chile: University of Chile, pp. 799-808.

Larrondo, P.F., Neufeld, C. & Deutsch, C.V., 2003. VARFIT: A program for semi-automatic variogram modelling. In *Centre for Computational Geostatistics, Report 5*. Edmonton: University of Alberta, p. Paper 404.

Leuangthong, O., 2003. *Stepwise Conditional Transformation for Multivariate Geostatistical Simulation*. PhD Thesis. Edmonton: University of Alberta.

Leuangthong, O. & Deutsch, C.V., 2004. Transformation of residuals to avoid artifacts in geostatistical modelling with a trend. *Mathematical Geology*, 36(3), 287-305.

Leuangthong, O., Khan, K.D. & Deutsch, C.V., 2008. *Solved Problems in Geostatistics*, Hoboken N.J.: Wiley.

Leuangthong, O., McLennan, J.A. & Deutsch, C.V., 2004. Minimum acceptance criteria for geostatistical realizations. *Natural Resources Research*, 13(3), 131–141.

Li, Q. & Racine, J.S., 2007. *Nonparametric Econometrics: Theory and Practice*, Princeton N.J.: Princeton University Press.

Lloyd, C.D., 2007. *Local Models for Spatial Analysis*, Boca Raton: CRC/Taylor & Francis.

Lloyd, C.D. & Atkinson, P.M., 2000. Interpolating elevation with locally-adaptive kriging. In P. M. Atkinson & D. Martin, eds. *GIS and geocomputation*. Innovations in GIS. CRC Press.

Lloyd, C.D. & Atkinson, P.M., 2002. Non-stationary approaches for mapping terrain and assessing prediction uncertainty. *Transactions in GIS*, 6(1), 17-30.

Luster, G.R., 1985. *Raw Materials for Portland Cement: Applications of Conditional Simulation of Coregionalization*. PhD Thesis. Stanford, CA: Stanford University.

Lyall, G. & Deutsch, C.V., 2000. Geostatistical modeling of multiple variables in presence of complex trends and mineralogical constraints. In *Proceedings of the Sixth International Geostatistics Congress*. GEOSTATS 2000. Cape Town,South Africa: Document Transformation Technologies, pp. M-17.

Lyster, S.J., 2009. *Simulation of Geologic Phenomena Using Multiple-Point Statistics in a Gibbs Sampler Algorithm*. PhD Thesis. Edmonton: University of Alberta.

Magri, E.J., 1987. Economic optimization of the number of boreholes and deflections in deep gold explomtion. *Journal of the South African Institute of Mining and Metallurgy*, 87(10), 307-321.

Manchuk, J. & Leuangthong, O., 2008. Experimental variogram calculation using bernsteing polynomials. In *Centre for Computational Geostatistics, Report 8*. Edmonton: University of Alberta, p. Paper 303.

Matheron, G.F., 1969. *Cours de Géostatistique*, Centre de Morphologie Mathématique de Fontainebleau: Ecole des Mines de Paris.

Matheron, G.F., 1974. *Effet proportionnel et lognormalité ou: le retour du serpent de mer*, Fontainebleau: Centre de Géostatistique.

Matheron, G.F., 1969. *Le Krigeage Universel*, Centre de Morphologie Mathématique de Fontainebleau: Ecole des Mines de Paris.

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

Matheron, G.F., 1970. *Théorie des Variables Régionalisées et ses Applications*, Centre de Morphologie Mathématique de Fontainebleau: Ecole des Mines de Paris.

McLennan, J.A., 2007. *The Decision of Stationarity*. PhD Thesis. Edmonton: University of Alberta.

McLennan, J.A. & Deutsch, C.V., 2008. A new approach to SGS with a trend: A non-stationary Gaussian transformation. In *Proceedings of the Eight International Geostatistics Congress*. GEOSTATS 2008. Santiago, Chile: University of Chile, pp. 419-428.

Minasny, B. & McBratney, A.B., 2005. The Matérn function as a general model for soil variograms. *Geoderma*, 128(3-4), 192-207.

Mohammadhassanpour, R., 2007. *Tools for Multivariate Modeling of Permeability Tensors and Geometric Parameters for Unstructured Grids*. MSc Thesis. Edmonton: University of Alberta.

Moon, C., Whateley, M. & Evans, A.M., 2006. *Introduction to Mineral Exploration* 2nd ed., Malden, MA: Wiley-Blackwell.

Mustapha, H. & Dimitrakopoulos, R., 2010. A new approach for geological pattern recognition using high-order spatial cumulants. *Computers & Geosciences*, 36(3), 313-334.

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

Neufeld, C. & Deutsch, C.V., 2004. Developments in semiautomatic variogram fitting. In *Center for Computational Geostatistics, Report 6*. Edmonton: University of Alberta, p. Paper 404.

Olea, R.A., 1999. *Geostatistics for Engineers and Earth Scientists*, Boston: Kluwer Academic Publications.

Oreskes, N., Shrader-Frechette, K. & Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147), 641-646.

Ortiz, J.M., 2000. *Scale histogram with discrete Gaussian model (dgm)*, Centre for Computational Geostatistics (CCG).

Ortiz, J.M. & Emery, X., 2006. Geostatistical estimation of mineral resources with soft geological boundaries: a comparative study. *Journal of The South African Institue of Mining and Metallurgy*, 106(8), 577-584.

Oz, B., Ortiz, J.M. & Deutsch, C.V., 2002. Determining the shape of conditional

distributions with Hermite polynomials and Disjunctive Kriging. In *Center for Computational Geostatistics, Report 4*. Edmonton: University of Alberta, p. Paper 16.

Pereira, M., Soares, A. & Rosario, L., 2000. Characterization of forest resources with satellite spot images by using local models of co-regionalization. In *Proceedings of the Sixth International Geostatistics Congress*. GEOSTATS 2000. Cape Town,South Africa: Document Transformation Technologies, pp. E-17.

du Pisani, P. & Vogt, D., 2004. Borehole radar delineation of the Ventersdorp Contact Reef in three dimensions. *Exploration Geophysics*, 35(4), 319-323.

Platten, I.M. & Dominy, S.C., 2001. The occurrence of high-grade gold pockets in quartz reefs at the Gwynfynydd mine, Wales, United Kingdom: A geological explanation for the nugget effect. *Exploration and Mining Geology*, 10(4), 249-272.

Pyrcz, M. & Deutsch, C.V., 2001. The whole story on the hole effect. In *Centre for Computational Geostatistics, Report 3*. Edmotnon: University of Alberta, p. Paper 11.

Rance, D. et al., 2006. *Technical Report on the Feasibility Study for the Burnstone Gold Project*, Toronto: Great Basin Gold LTD.

Ren, W., 2007. *Scale Consistent Geostatistical Modelling for Reservoir Characterization*. PhD Thesis. Edmonton: University of Alberta.

Riquelme Tapia, R., Le Loc'h, G. & Carrasco, P., 2008. Truncated gaussian and plurigaussian simulations of lithological units in Mansa Mina deposit. In *Proceedings of the Eight International Geostatistics Congress*. GEOSTATS 2008. Santiago, Chile: University of Chile, pp. 819-828.

Rivoirard, J., 1990. *Introduction to Disjunctive Kriging and Nonlinear Geostatistics*, Fontainebleau: Centre de Géostatistique.

Rivoirard, J., 1987. Two key parameters when choosing the kriging neighborhood. *Mathematical Geology*, 19(8), 851-856.

Rossi, M., 2004. Comparing simulated and interpreted geologic models. In *2004 SME Annual Meeting and Exhibit*. SME Anual meeting. Denver, CO: Society for Mining, Metallurgy & Exploration.

Sampson, P.D. & Guttorp, P., 1992. Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 87(417), 108–119.

Schabenberger, O. & Gotway, C.A., 2005. *Statistical Methods for Spatial Data Analysis* 1st ed., Chapman & Hall/CRC Press.

Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data.

In *Proceedings of the 1968 23rd ACM national conference*. New York: Association for Computing Machinery ACM, pp. 517–524.

Shepard, M.K. et al., 2001. The roughness of natural terrain: A planetary and remote sensing perspective. *Journal of Geophysical Research*, 106(E12), 32,777–32,795.

Sinclair, A.J. & Blackwell, G.H., 2002. *Applied Mineral Inventory Estimation*, Cambridge, UK: Cambridge University Press.

Srivastava, R.M., 2005. Probabilistic modeling of ore lens geometry: An alternative to deterministic wireframes. *Mathematical Geology*, 37(5), 513-544.

Stein, M.L., 1999. *Interpolation of Spatial Data: Some Theory for Kriging* 1st ed., New York: Springer-Verlag.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34(1), 1-21.

Stroet, C.B.M.T. & Snepvangers, J.J.J.C., 2005. Mapping curvilinear structures with local anisotropy kriging. *Mathematical Geology*, 37(6), 635-649.

Taylor, J.R., 1997. *An introduction to error analysis: the study of uncertainties in physical measurements*, University Science Books.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit Region. *Economic Geography*, 46, 234-240.

Vann, J., Jackson, S. & Bertoli, O., 2003. Quantitative kriging neighbourhood analysis for the mining geologist—a description of the method with worked case examples. In *5th International Mining Geology Conference*. Bendigo, Victoria: Australasian Institute of Mining and Metallurgy, pp. 17–19.

Vann, J. & Sans, H., 1995. Global resource estimation and change of support at the Enterprise Gold Mine, Pine Creek, Northern Territory - Application of the geostatistical discrete gaussian model. In *Applications of Computers and Operations Research in the Mineral Industry*. APCOM XXV. Brisbane, pp. 171-179.

Verly, G., 1984. The block distribution given a point multivariate normal distribution. In M. David, A. G. Journel, & A. Marechal, eds. *Geostatistics for Natural Resources Characterization*. NATO Science Series Committee. Dordrecht: Springer, pp. 495-515.

Wackernagel, H., 2003. *Multivariate Geostatistics* 3rd ed., Berlin Heildelberg: Springer.

Walter, C. et al., 2001. Spatial prediction of topsoil salinity in the Chelif Valley, Algeria, using local ordinary kriging with local variograms versus whole-area variogram. *Australian Journal of Soil Research*, 39(2), 259–272.

Wasserman, L., 2006. *All of Nonparametric Statistics*, New York: Springer Science

Business Media Inc.

Webster, R. & Oliver, M.A., 2007. *Geostatistics for Environmental Scientists*, West Sussex, UK: John Wiley & Sons Inc.

Wellmer, F., 1998. *Statistical Evaluations in Exploration for Mineral Deposits* 1st ed., Berlin; New York: Springer.

Zhang, X.F., Van Eijkeren, J.C.H. & Heemink, A.W., 1995. On the weighted least-squares method for fitting a semivariogram model. *Computers and Geosciences*, 21(4), 605–608.

Appendix A
Software Implementation

This appendix covers the algorithms needed for the practical application of spatial prediction under the assumption of local stationarity. These algorithms have been implemented as FORTRAN programs and most of them were developed from their equivalents already used in stationary modelling. The output of the distance weights generator, **LDWgen**, provides the basis for the algorithms used in the inference of the location-dependent statistics. **LDWgen** output consists of the matrix of distance weights assigned to the samples in relation to each one of the anchor points. The program **nscore_loc** uses these weights to build the local normal scores transformation tables. The program **herco_loc** is used to model the local transformation functions described by these tables by a series of Hermite polynomials. The program **gamvlocal** takes the matrix of weights and the local normal scores transforms to produce location-dependent experimental variograms, covariances and correlograms at each anchor point location, either in original units or in Gaussian units. The program **globfit** is used for the joint fitting of all the experimental local measures of correlation by allowable variogram models. The resulting coefficients of the local Hermite models and parameters of the local variogram models can be interpolated at the resolution of the final numerical model using any interpolation algorithm capable of producing smoothly changing maps. **kt3d_LMG** program reads these maps for locally stationary point or block estimation, while **ultimateSGSim v.2** and **SISim_loc** use them for locally stationary sequential Gaussian simulation and sequential indicator simulation, respectively.

A.1. Generation of Distance Weighted Datasets: LDWgen

The **LDWgen** program is used for generating a matrix of weights for all samples in a dataset in relation to multiple anchor points locations. It can operate in parameter calibration mode or in distance weights generation mode. In calibration mode the

program can be used to assess the smoothness of the local mean and variance models in relation to the original sample values. In distance weighting mode, the program will produce a $n \times P$ matrix of distance weights, with n as the number of samples and P as the number of anchor points. Although the Gaussian kernel is preferred as a weighting function, the program also allows inverse distance weighting and uniform kernel weighting. Figure A-1 shows an example of the parameter file for this program, the details of which are below.

```
                  Parameters for LDWgen
                  *********************

START OF PARAMETERS:
DDH73.dat                 - file with data
2  3  4   5  6            - columns for X, Y, Z coordinates variable and weight
0     1.0e21               - trimming limits
1                          - including declustering weights by 0=scaling 1=self calibration
0                          - parameter calibration mode? 0=no, 1=yes
0                          - anchor points in file =0, in grid =1
425  855  10.0             - nx,xmn,xsiz
475  -845  10.0            - ny,ymn,ysiz
1    0.0     1.0           - nz,zmn,zsiz
APddh73.dat               - file with anchor points locations
1  2  3                    - columns for X, Y, Z anchor point coordinates
2  20  0.01                - weigting function (1=ID, 2=GK, 3=UK)  and parameters
0                          - dinamic kernel bandwidth: 0=no, 1=yes
 0.0  0.0  0.0             - anisotropic search angles
 400.0  400.0  400.0       - anisotropic search radius
LDW-DDH73.out             - file for output
LDW-DDH73.dbg             - file for weights output
LDW-DDH73.sum             - file for debugging
```

Figure A-1: An example parameter file for LDWgen

- **datfl:** a data file in GSLIB format.
- **ixl**, **iyl**, **izl**, **ivr**, and **iwt**: the column numbers for the x, y and z coordinates, values and declustering weights of the samples in the data file.
- **tmin** and **tmax**: values smaller than **tmin** and greater than **tmax** are ignored.
- **idecinc**: form of including the declustering weights. If **idecinc** is set to zero, the distance weights are scaled by the declustering weights. If **idecinc** is set to one, the distance weights are corrected by the methodology presented in Section 3.2.4.
- **ical:** the output, debugging and summary files below will change depending on wheter **ical** = 0, i.e. weights generation mode, or **ical** = 1, i.e. weights calibration mode.
- **iapgrid:** anchor points can be located arbitrarily or in a regular grid. In the first case, **iapgrid** = 0, a file in GSLIB format containing their coordinates and identification numbers must be provided. In the second case, **iapgrid** = 1, the specifications of the anchor point grid must be entered.

- **apxn**, **apxmn**, and **apxsiz**: the number of anchor points in the x direction, the x coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the X axis.
- **apyn**, **apymn**, and **apysiz**: the number of anchor points in the y direction, the y coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the Y axis.
- **apzn**, **apzmn**, and **apzsiz**: the number of anchor points in the z direction, the z coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the Z axis.
- **apfl**: a file in GSLIB format containing the coordinates of anchor point locations.
- **uxl**, **uyl**, and **uzl**: the column numbers for the anchor points x, y and z coordinates.
- **wfunc**, **para**, and **parb**: type of distance weighting function and its parameters. If **wfunc** = 1, an inverse distance function is used and **para** and **parb** are the power and offset parameters, respectively. If **wfunc** = 2, a Gaussian kernel is used. If **wfunc** = 3, a uniform kernel is used. In the last two cases **para** and **parb** are the bandwidth and background parameters, respectively.
- i**dyn**: this parameter is set as 1 for allowing a dynamic kernel bandwidth.
- **ang1**, **ang2** and **ang3**: the rotation angles for an anisotropic kernel bandwidth.
- **aa**, **aa1** and **aa2**: the anisotropic kernel ellipsoid radii. Their actual lengths are irrelevant, but what is taken into account is the anisotropy ratios **aa**/**aa1** and **aa**/**aa2**.
- **outfl**: the output file. If parameter calibration mode is selected, i.e. **icv** =1, this file will contain the local mean values at the data values locations. Otherwise, the output file contains the matrix of distance weights assigned to the individual samples in relation to all anchor points.
- **dbgfl**: this debugging output file contains the local means, standard deviations, quartiles and other 1-point local statistics. Additionally, it provides the sum of distance weights at each anchor point for control purposes. It also contains the size of the dynamic kernel bandwidth.
- **sumfl**: this file stores the summary statistics of the relation between data values and the local mean model.

A.2. Local Normal Scores Transformation and Modelling

nscore_loc is constructs the local normal scores transformation tables using the distance weights provided in the output of the **LDWgen** program. **herco_loc** is used for modelling the local normal scores transformation function by Hermite polynomial series. This is done for the sake of efficiency; storing normal scores transformation tables for large datasets and feeding them to Gaussian estimation and simulation algorithms can be very demanding in computer memory and storage.

A.2.1. Local Normal Scores Transformation: nscore_loc

This program was developed from the FORTRAN program **nscore** (Deutsch & Journel 1998, pp.223-226). The new **nscore_loc** uses the data values and the distance weights from the **LDWgen** output file to build the local cdfs. At each anchor point, the contribution of each sample in the local cdf is proportional to its corresponding distance weight. The normal scores transformation procedure on the local cdfs results in local transformation lookup tables. Figure A-2 shows the parameter file required for the **nscore_loc** program and the parameters are detailed below.

```
                  Parameters for NSCORE_Loc
                  *************************

START OF PARAMETERS:
../LDWgen/LDW-DDH73.out          - file with data and distance weights
2  3  4   5  6                   - columns for X, Y, Z coordinates variable and weight
400                              - number anchor points
0       1.0e21                   - trimming limits
0                                - 1=transform according to specified ref. dist.
../histsmth/histsmth.out         - file with reference dist.
1   2                            - columns for variable and weight
nsLDW-DDH73.out                  - file for output
nsLDW-DDH73.trn                  - file for output transformation table
```

Figure A-2: An example parameter file for nscore_loc.

- **datfl**: this is a **LDWgen** output file. It must contain $n \times P$ entries.
- **ixl**, **iyl**, **izl**, **ivr**, and **iwt**: the column numbers for the x, y and z coordinates, and the values and declustering weights of the samples in the **LDWgen** output file.
- **nap**: *P*, this is the number of anchor points.
- **tmin** and **tmax**: sample values smaller than **tmin** and greater than **tmax** are ignored.
- **ismooth**: if this parameter is set to 1 a smoothed distribution will be considered, otherwise the distribution is built directly from data. In its current version

nscore_loc allows a smoothed distribution only if operating for single distributions, i.e. $P = 1$.

- **ismoothfl**: the file containing the smoothed distribution.
- **icolvr** and **icolwt**: the column numbers corresponding to the values and weights in the smoothed distribution file.
- **outfl**: the file containing sample coordinates, original values and locally normal scores transformed values for all anchor points. This output file can be used as an input for **gamvlocal** in order to generate the experimental local measures of spatial correlation in local Gaussian space.
- **transfl**: this file contains the transformation lookup tables for all anchor points. The first column in the file corresponds to the original values and the second to the locally normal scores transformed values. The program **herco_loc** uses this file to model the local Gaussian transformation function.

A.2.2. Modelling the Local Normal Scores Transformation Function with Hermite Polynomials: herco_loc

The program **herco_loc** was developed from the discrete Gaussian change of support program **DGM** (Ortiz 2000). It uses the transformation lookup table file produced by **nscore_loc** to model the local Gaussian transformation function. The output file contains the Hermite coefficients of the polynomial fits of the Gaussian transformation functions at all the required anchor point locations. The fitting of the local normal score transformation function is improved when the input is the local normal scores transformation table obtained from a few hundreds of quantiles obtained from the local distributions. These local quantiles can be calculated using the **hispltsim** program (Deutsch 2005). The parameters required for this small program are described below and the file containing them is shown in Figure A-3.

- **transfl**: this is the file containing the Gaussian local transformation lookup tables produced by **nscore_loc.**
- **aplfl**: a file containing the coordinates of the anchor points locations.
- **uxl**, **uyl**, and **uzl**: the column numbers for the anchor points x, y and z coordinates.
- **nherco**: number of Hermite polynomials and coefficients to be considered in the fitting of the local normal scores transformation function.

- **outfl**: This file is generated for checking purposes. It contains the original values and their approximated values after the Hermite polynomials fitting. It also contains the locally normal scores transformed values.
- **dbgfl**: This file contains the coordinates of the anchor point locations and all the local Hermite coefficients used at such locations. It also contains the local variances approached by the Hermite polynomial series.

```
                Parameters for HERCO_loc
                ************************

START OF PARAMETERS:
nsQLDW-DDH73.trn                - file with input transformation table
aploc.dat                       - file with anchor point coordinates
1  2  3                         - X,Y,Z coordinates for anchor points
40                              - number of Hermite polynomials to use (e.g. np=20)
herco_loc40.out                 - file for point scale output to check the anamorphosis
herco_loc40.dbg                 - file with Hermite Coefficients, phi(p)
```

Figure A-3: An example parameter file for `herco_loc`.

A.3. Location-Dependent Variograms

The set of programs for location dependent variograms includes a calculator of location-dependent measures of correlation, **`gamvlocal`**, a generator of local variogram maps, **`varmap_loc`** and an algorithm for semiautomatic fitting of the location-dependent measures of correlation, **`globfit`**.

A.3.1. Calculation of Location-Dependent Experimental Variograms: `gamvlocal`

The **`gamvlocal`** program was developed from **`gamv2004`** program (Deutsch 2007), which is a modified version of the GSLIB program for the calculation of experimental measures of correlation with irregularly spaced data, **`gamv`** (Deutsch & Journel 1998, pp.53-54). The program reads and stores the sample coordinates, the sample values, and the distance weights assigned to each sample in relation to all anchor points from a file produced either by **`LDWgen`** or **`nscores_loc`** programs. The weights assigned to individual samples are combined for the pairs that fall within the tolerances specified for different lag distances and directions. The program is able to generate location-dependent variograms, covariances or correlograms for continuous and categorical variables. The output file contains the experimental values of the chosen local measure of spatial

continuity for the selected directions and lag separations. An example of the required parameter file is presented in Figure A-4. The parameters in this file are described below.

```
                  Parameters for GAMVLOCAL
                  ************************

START OF PARAMETERS:
LDWG-wlc-g20des.out             - output file from LDWgen or from Nscores_loc
2   3   4   6                   - columns for X, Y, Z coordinates and weight
1   5   0                       - number of variables,column numbers
-1.0e21      1.0e21             - trimming limits
0                               - local NS transformation 0 = no, 1 = yes
1                               - anchor points in file = 0, in grid = 1
50    5.0     10.0              - nx,xmn,xsiz
50    5.0     10.0              - ny,ymn,ysiz
1     1.0     1.0               - nz,zmn,zsiz
apoints.dat                     - file with anchor points location
1   2   3                       - columns for X, Y, Z coordinates
0                               - Mixture rule exponent: 0= geometric, 1=arithmetic
lv-wlc-g20des.out               - file for variogram output
lv-wlc-g20des.dbg               - file for debugging
4                               - number of directions
0  20 500 0 15 500              - Dir 01: azm,atol,bandh,dip,dtol,bandv
12   10  5                      -   nlag,xlag,xtol
45 20 500 0 15 500              - Dir 02: azm,atol,bandh,dip,dtol,bandv
12   10  5                      -   nlag,xlag,xtol
90 20 500 0 15 500              - Dir 03: azm,atol,bandh,dip,dtol,bandv
12   10  5                      -   nlag,xlag,xtol
135 20 500 0 15 500             - Dir 03: azm,atol,bandh,dip,dtol,bandv
12   10  5                      -   nlag,xlag,xtol
1                               - standardize sills? (0=no, 1=yes)
1   1   1                       - tail var., head var., variogram type
```

Figure A-4: An example parameter file for `gamvlocal`

- **datfl**: This is a **`LDWgen`** or **`nscores_loc`** output file. It must contain $n \times P$ entries.
- **ixl**, **iyl**, **izl**, and **iwt**: the column numbers for the x, y and z coordinates and distance weights assigned to the values in **datfl**.
- **nvar** and **ivar(1)** ... **ivar(nvar)**: The number of variables and their column order in **datfl**. In the current version of **`gamvlocal`** the maximum value for **nvar** is 2.
- **tmin** and **tmax**: values smaller than **tmin** and greater than **tmax** are ignored.
- **itrans**: if this parameter is set to 1, extra memory is allocated for the locally normal scores transformed values.
- **iapgrid**: if this parameter is set to 0, the program will try to read the anchor point coordinates from the file **apfl**. Otherwise it will take the anchor point coordinates from the grid specification.
- **apxn**, **apxmn**, and **apxsiz**: the number of anchor points in the x direction, the x coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the X axis.
- **apyn**, **apymn**, and **apysiz**: the number of anchor points in the *y* direction, the y coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the Y axis.

- **apzn**, **apzmn**, and **apzsiz**: the number of anchor points in the z direction, the z coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the Z axis.
- **apfl**: a file in GSLIB format containing the coordinates of anchor point locations.
- **uxl**, **uyl**, and **uzl**: the column numbers for the anchor points x, y and z coordinates.
- **mrexp**: the value for the exponential in the mixture rule for building the 2-point weights. If this parameter is set to 0, a geometric average of the weights assigned to the individual samples in each pair is performed.
- **outfl**: a file containing the output for the experimental measure of correlation ordered by anchor point number and then by direction. For each anchor point there is a line containing the anchor point number and coordinates. For each direction four heading lines are included containing the azimuth and dip angles and its angular tolerances, the number of lags, the lag distances and the lag tolerance, and the type of variogram. The following **nlag** lines contain the fields:
    1. Lag number.
    2. Average separation distance for the current lag.
    3. Value of the measure of spatial correlation.
    4. Number of pairs involved in the calculation of such value.
    5. Sum of the 2-point weights assigned to the pairs involved.
    6. Variance of the variogram cloud.
    7. Locally weighted mean of the data contributing to the tail.
    8. Locally weighted mean of the data contributing to the head

  When the location-dependent correlograms is chosen, the following columns are added:

    9. Locally weighted variance of the data contributing to the tail.
    10. Locally weighted variance of the data contributing to the head
- **dbgfl**: a file containing local statistics of the values and weights for checking purposes.
- **ndir**: the number of directions for the calculation of the chosen experimental local measure of spatial correlation. The following two lines of parameters are repeated for the **ndir** directions.

- **azm**, **atol**, **bandwh**, **dip**, **dtol**, and **bandwd**: the azimuth angle, the azimuth angular tolerance, the azimuth bandwidth, the dip angle, the dip angular tolerance, and the dip bandwidth.
- **nlag**, **xlag**, **xtol**: the number of lags to compute, the unit lag separation distance, and the lag distance tolerance.
- **isill**: if set to 1, the local semivariogram values will be divided by the local sill.
- **ivtail**, **ivhead**, and **ivtype**: the tail and head variable numbers, and the variogram type to compute. For direct variograms the **ivtail** is the same as in **ivhead**, for cross variograms they are different. The parameter numbers for the location-dependent variogram types allowed are:
    1. Variogram
    2. Cross Variogram
    3. Covariance
    4. Correlogram
    9. Indicator variogram for a continuous variable
    10. Indicator variogram for a categorical variable.
- **cut**: This parameter must be added at the end of the previous line when the **ivtype** value corresponds to an indicator variogram. Its value is a cut-off for continuous variables or a category code for discrete variables.

A.3.2. Location-Dependent Variogram Maps: `varmap_loc`

This program generates locally weighted variogram maps for multiple anchor points. Additionally, it calculates an anisotropy ellipsoid for such local variogram maps using the moments of inertia tensor method presented by Mohammadhassanpour (2007, pp.45-53). This program was developed from **`varmap`** (Deutsch & Journel 1998, pp.55-57). The parameters required for **`varmap_loc`** are separated into three blocks: the data file parameters, the anchor point parameters, and the variogram parameters. The following are the parameters in the first block:

- **datfl**: This is a **`LDWgen`** or **`nscores_loc`** output file.
- **nvar** and **ivar(1)** ... **ivar(nvar)**: The number of variables and their column order in **datfl**. In the current version of **`gamvlocal`** the maximum value for **nvar** is 2.
- **tmin** and **tmax**: values smaller than **tmin** and greater than **tmax** are ignored.

- **itrans**: if this parameter is set to 1, extra memory is allocated for the locally normal scores transformed values.
- **igrid**: this parameter is set to 0 if the **datfl** contains the coordinates of scattered samples. If it is set to 1, then the data must be arranged in a grid.
- **nx**, **ny**, **nz**: if **igrid** is 1 then these are the number of cells in the x, y and z directions.
- **xsiz**, **ysiz**, **zsiz**: if **igrid** is 1 then these are the size of the cells parallel to the x, y and z directions.
- **ixl**, **iyl**, **izl**, and **iwt**: if **igrid** is 0, these are the column numbers for the x, y and z coordinates and distance weights assigned to the values in **datfl**.

  The parameters for the block specifying the anchor point locations are listed next:
- **iapgrid**: if this parameter is set to 0, the program will try to read the anchor point coordinates from the file **apfl**. Otherwise it will take the anchor point coordinates from the grid specification.
- **apxn**, **apxmn**, and **apxsiz**: the number of anchor points in the x direction, the x coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the X axis.
- **apyn**, **apymn**, and **apysiz**: the number of anchor points in the *y* direction, the y coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the Y axis.
- **apzn**, **apzmn**, **apzsiz**: the number of anchor points in the *z* direction, the *z* coordinate of the anchor point located in the southwest corner at the bottom of the grid, and the anchor points separation parallel to the Z axis.
- **apfl**: a file in GSLIB format containing the coordinates of anchor point locations.
- **uxl**, **uyl**, and **uzl**: the column numbers for the anchor points x, y and z coordinates.

  The parameters in the variogram specification block are the following:
- **outfl**: This output file contains the 2-D o 3-D mapped local variogram values, local head and tail means and local head and tail variances for all anchor points. These values are written sequentially and sorted by the anchor point number first, followed by the z direction, then the y direction and, finally, the x direction.
- **angfl**: the output file for the angles and radii of the anisotropy ellipsoid fitted to the local variogram maps or volumes at each anchor point.
- **wmass**: if set to 1, the variogram values close to the origin will have a higher weight for the calculation of moments of inertia.

- **nxlag**, **nylag**, and **nzlag**: the number of lags to compute in the x, y and z directions.
- **dxlag**, **dylag**, and **dzlag**: the lag tolerances or variogram map/volume "cell sizes" in the x, y and z directions.
- **minpairs**: the minimum number of values to calculate a value of the variogram map.
- **mrexp**: the value for the exponential in the mixture rule for building the 2-point weights. If this parameter is set to 0, a geometric average of the weights assigned to the individual samples in each pair is performed.
- **isill**: if set to 1, the local semivariogram values will be divided by the local sill.
- **ivtail**, **ivhead**, and **ivtype**: the tail and head variable numbers, and variogram type to compute. For a correct calculation of the local anisotropy parameters using the moments of inertia method, the **ivtype** value must be 3 or 4. These values indicate the calculation of covariances or correlograms.

  An example of the parameter file for **`varmap_loc`** is presented in Figure A-5.

```
                  Parameters for VARMAP_loc
                  *********************

START OF DATA PARAMETERS:
LDW-w1-G20-0-40.out          - output file from LDWgenerator or nscores-loc
1   5                        - number of variables: column numbers
-1.0e21     1.0e21           - trimming limits
0                            - local NS transformation 0 = no, 1 = yes
0                            - 1=regular grid, 0=scattered values
 50   50    1                - if =1:  nx,     ny,   nz
1.0  1.0  1.0                -         xsiz, ysiz, zsiz
2   3   4   6                - columns for X, Y, Z coordinates and weight

START OF ANCHOR POINTS PARAMETERS:
1                            - anchor points in file = 0, in grid = 1
7  10  40.0                  - if  =1: nx,xmn,xsiz
8  10  40.0                  -         ny,ymn,ysiz
1  0.0  1.0                  -         nz,zmn,zsiz
apLQzalta.dat                - if  =0 file with anchor points locations
2   3   4                    -        columns for X, Y, Z coordinates and weight

START OF VARIOGRAM PARAMETERS:
varmap_loc.out               - file for variogram output
varmap_loc-1h.ang            - file for angles output
1                            - Weighting the VARMAP mass by lag distance? 0= no, 1= yes
 10    10     0              - nxlag, nylag, nzlag
5.0   5.0   1.0              - dxlag, dylag, dzlag
15                           - minimum number of pairs
0                            - Weights averaging: 0= geometric, 1=arithmetic
1                            - standardize sill? (0=no, 1=yes)
1   1   4                    - tail, head, variogram type
```

Figure A-5: an example parameter file for `varmap_loc`

A.3.3. Joint Fitting of Location-Dependent Variogram Models: `globfit`

The program **`globfit`** was developed from the FORTRAN program for semiautomatic fitting of experimental variograms **`varfit`** (Larrondo et al. 2003; Neufeld & Deutsch 2004). **`globfit`** is intended for the conjoint semiautomatic fitting under user defined

constraints of local variogram models for the experimental local measures of spatial continuity. The required parameter file is organized in five main blocks: the main parameters, the experimental local variogram files specification, the global model parameters, the anchor point location parameters, and the advanced options parameters. The parameters in the first block are the following:

- **nvario**: the number of experimental variograms to consider.
- **nst**: the number of nested structures to include in the variogram model. A maximum of two structures is allowed for location-dependent variogram models.
- **conang**: if set to 1 the same rotation angles are used for the anisotropic ellipsoid of all structures. If it is 0, the anisotropy definition of each structure is independent from the others.
- **idiswt**: if set to 1, the experimental variogram points are weighted inversely proportional to their lag distance.
- **inpwt**: if set to 1, each experimental variogram point is weighted by the sum of all the 2-point weights of the pairs involved in its calculation.
- **ivvwt**: if set to 1, each experimental variogram point is weighted inversely proportional to the variance of the variogram cloud at its corresponding lag.
- **npmin**: the experimental variogram points calculated with less than npmin pairs are not considered during the variogram fitting.
- **penfac** and **penpow**: the value of the coefficient and the power of the penalty function for extreme variogram model parameters.
- **penfit**: the coefficient of the quadratic penalty function for locally anomalous variogram parameters.
- **psfl**: the prefix for the names of the postscript files containing the graphic output of the fitted local variogram models. The program will add the corresponding anchor point number and the extension .ps at the right end of the prefix.
- **varfl**: the output file for the local variogram model parameters. These are ordered sequentially according the anchor point number.
- **sumfl**: a summary output file containing the anchor point coordinates, the local variogram model parameters and the fitting mean square error.
- **pagetitle**: a character string containing the title in the output postscript files.

  An example of the main block of parameters is presented in Figure A-6.

```
                  Parameters for GLOBFIT
                  **********************

START OF MAIN PARAMETERS:
3                     - number of variograms
1                     - number of nested structures
1                     - constant angle between structures
1                     - inverse distance weighting (0=no, 1=yes)
0                     - pairs  weighting (0=no, 1=yes)
0                     - variogram variance weighting (0=no, 1=yes)
10                    - minimum number of pairs to use
0.0  0.0              - penalty constant and exponent for extreme parameter values
0.0                   - penalty factor for local odd fits
./graphs/corfit       - file for PostScript output
corrfit.var           - file for variogram model
corrfit.sum           - file for summary file
Local variograms      - project title
```

Figure A-6: Example of the main block of parameters for `globfit`.

The parameters specifying the experimental local variogram files and the related options are listed next:

- **iytoz**: if set to 1, the experimental location dependent measure of spatial continuity must be calculated using locally normal scores transformed values. In this case the program will back-transform the experimental points to original units and use them for variogram fitting.
- **lhercofl**: a file containing the Hermite coefficients obtained from the modelling of the normal scores transformation function.
- **npol**, **ipol**: the number of Hermite coefficients and initial column number in the file specified by **lhercofl**.
- **datfl**: an output file generated by **`gamvlocal`** containing the experimental values of the location-dependent measure of spatial correlation at different lags, directions and for all anchor points.
- **ivario**: the order of the variogram to be picked in **datfl**.

The last two parameters must be repeated for the number of variograms specified by nvario. Figure A-7 shows an example of this block of parameters.

```
START OF EXPERIMENTAL VARIOGRAMS SPECIFICATION:
0                          - Fitting Z variograms from Y values? (no=0, yes=1)
herpol.dat                 - file with Hermite coefficients
30 1                       - number of hermite polynomials and column for phi(0)
1-rDDH75_lns_40.out        - variogram #1 file
   1                       - variogram number in file
1-rDDH75_lns_40.out        - variogram #1 file
   2                       - variogram number in file
1-rDDH75_lns_40.out        - variogram #1 file
   3                       - variogram number in file
```

Figure A-7: An example of the experimental variograms block of parameters for `globfit`.

The parameters for the specification of the global variogram model are listed next:

c0g and **alfag**: the nugget effect value, and the exponent of the stable model.

- **itg**, **ccg**, **ang1g**, **ang2g**, and **ang3g**: the type of the structure, the sill contribution of the structures, and the angles defining the geometric anisotropy.
- **aag**, **a2g**, and **a3g**: the maximum horizontal range, the minimum horizontal range and the vertical range.

The last two parameter lines must be repeated for the number of nested structures specified by **nst.** An example of this block of parameters is shown in Figure A-8.

```
START OF GLOBAL VARIOGRAM MODEL SPECIFICATION:
0.123   0.439                 - nugget effect, exp
2  0.877  0.0  0.0  0.0       - it,cc,ang1,ang2,ang3
           7.6  7.6  7.6       - a_hmax, a_hmin, a_vert
```

Figure A-8: An example of the variogram model block of parameters for `globfit`.

The parameter lines below define the neighbourhood of anchor points for comparing the local variogram model parameter. Figure A-9 shows an example of these parameters.

- **nap**: the number of anchor points where the experimental local measures of spatial continuity were calculated.
- **napmax**: the maximum number of anchor points to be considered in a neighbourhood.
- **radius**, **radius1**, and **radius2**: the radii of the ellipsoid used for searching the neighbouring anchor points.
- **sang1**, **sang2**, and **sang3**: angles that define the orientation of the search ellipsoid.

```
START OF ANCHOR POINTS DEFINITION
400                           - number of anchor points
4                             - maximum number of anchor points for parameter comparison
20.0  20.0  20.0              - maximum search radii
 0.0   0.0   0.0              - angles for search ellipsoid
```

Figure A-9: An example of the anchor points block of parameters for `globfit`.

The advanced options specify the constraints on the local variogram model parameters to be fitted at the multiple anchor point locations. The corresponding block in the parameter file is described below, and an example of it is presented in Figure A-10.

- **izonal**, **izonal1**, and **izonal2**: if any of these indicator parameters is set to 1 the program will automatically add one structure for zonal anisotropy modelling.
- **cychmax**, **cychmin**, and **cycvert**: if any of these indicator parameters is set to 1 the program will automatically add one structure for modelling the cyclicity in the corresponding direction. Only one direction is allowed for cyclicity modelling.
- **flexcc**, **ccmin**, and **ccmax**: indicator for allowing a variable initial sill value and the lower and upper limits for local sill values. If **flexcc** is set to 0 the initial sill value is taken from the global variogram definition. If it is set to 1, the initial local sill will be

calculated as the average of the local experimental variogram points at the three longest lag distances minus the initial local nugget effect. If **ccmin** and **ccmax** have the same value, the total sill is fixed to that value during optimization for all local variograms, with any value of **flexcc**.

- **flexc0, c0min**, and **c0max**: the indicator for allowing a variable initial nugget effect and the lower and upper limits for the local nugget effect. If **flexc0** is set to 0 the initial nugget effect value is taken from the global variogram definition. If it is set to 1, this will be calculated as the projection to the origin of the local experimental variogram points at the three shortest lag distances. If c0min and c0max have the same value, the total nugget effect is fixed to that value during optimization for all local variograms, with any value of **flexcc**.
- **nfixit**: the number of structures to fix during optimization. The same number of lines is required below for specifying the order and type of the fixed structure.
- **fixit**, and **it**: the structure number and structure type code. The codes for the structure types are: 1 for Spherical, 2 for Exponential, 3 for Gaussian, 5 for Hole Effect, and 6 for the Stable Model. The last structure type is allowed only for the first structure.
- **alfamin**, **alfamax**: the lower and upper limits for the local exponent in the stable variogram models. If these parameter values are the same the exponent of the stable structure is fixed for all anchor point locations during optimization.
- **nhmax**: the number of maximum horizontal ranges to constrain or fix during optimization for all anchor points. The same number of lines must be included beneath, specifying the ranges to be controlled.
- **fixhmax**, **hmaxmin**, and **hmaxmax**: the structure number for the maximum horizontal range to be controlled or fixed, and the lower and upper limits for the range value. If **hmaxmin** is equal to **hmaxmax**, the maximum horizontal range will be fixed to that value for all anchor points.
- **nhmin**: the number of minimum horizontal ranges to constrain or fix during optimization for all anchor points. The same number of lines must be included beneath, specifying the ranges to be controlled.
- **fixhmin**, **hminmin**, and **hminmax**: the structure number for the minimum horizontal range to be controlled or fixed, and the lower and upper limits for the range value. If **hminmin** is equal to **hminmax**, the minimum horizontal range will be fixed to that value for all anchor points.

- **nhver**: the number of vertical ranges to constrain or fix during optimization for all anchor points. The same number of lines must be included beneath, specifying the ranges to be controlled.
- **fixhver**, **hvermin**, and **hvermax**: the structure number for the vertical range to be controlled or fixed, and the lower and upper limits for the range value. If **hvermin** is equal to **hvermax**, the vertical range will be fixed to that value for all anchor points.
- **angfl**: a file containing the previously defined local anisotropy angles. The data order in this file must be the same as the order in which the location-dependent experimental variograms were calculated in **`gamvlocal`**.
- **iang1**, **iang2**, and **iang3**: the column numbers in file **angfl** for the previously defined azimuth, dip and plunge angles of the anisotropy ellipsoid.
- **nang1**: number of anisotropic structures with a controlled or fixed azimuth angle. The next line must be repeated nang1 times.
- **ivar1**, **ang1min**, and **ang1max**: the number of the structure to be controlled or fixed, and the minimum and maximum azimuth angles. If ang1min and ang1max are equal, the corresponding azimuth angle is fixed.
- **nang2**: the number of anisotropic structures with a controlled or fixed dip angle. The next line must be repeated **nang2** times.
- **ivar2**, **ang2min**, and **ang2max**: the number of the structures to be controlled or fixed, and the minimum and maximum dip angles. If ang1min and ang1max are equal, the corresponding dip angle is fixed.
- **nang3**: the number of anisotropic structures with a controlled or fixed plunge angle. The next line must be repeated **nang3** times.
- **ivar3**, **ang3min**, and **ang3max**: the number of the structures to be controlled or fixed, minimum and maximum plunge angles. If ang1min and ang1max are equal, the corresponding dip plunge is fixed.
- **npref**: the number of variogram directions with preference for the minimization of the mean square error.
- **ipref**, and **rpref**: the number of the preferred variogram and preference weighting factor.
- **nhvanis**: the number of structures with fa ixed or controlled anisotropic ratio between the maximum horizontal and vertical ranges.

- **ihvanis**, **hvanismin**, and **hvanismax**: the number of the structure to be fixed or controlled, and lower and upper limits for the controlled anisotropy ratio. If **hvanismin** value is equal to **hvanismax** the ratio between the maximum horizontal and vertical ranges is fixed.
- **nhhanis**: the number of structures with a fixed or controlled anisotropic ratio between the maximum and minimum horizontal ranges.
- **ihhanis**, **hhanismin**, and **hhanismax**: the number of the structure to be fixed or controlled, and lower and upper limits for the controlled anisotropy ratio. If **hhanismin** value is equal to **hhanismax** the ratio between the maximum and minimum horizontal ranges is fixed.

```
START OF ADVANCED OPTIONS:
0    0    0          - zonal Anis: Hmax, Hmin, Vert (0=no, 1=yes)
0    0    0          - cyclicity:  Hmax, Hmin, Vert (0=no, 1=yes)
0  1.0  1.0          - Variable initial Sill (0= no, 1 = yes), lower and upper tolerance
0  0   0.4           - Variable initial nugget effect (0=no, 1=yes), lower and upper limits
1                    - number of structure types to fix
1  2                 - structure number and structure type (6=stable variogram model)
0.5  2.0             - Lower and upper limits for stable variogram exponent
1                    - number of Hmax ranges to control/fix
1  10  180           - structure number, lower and upper limits
1                    - number of Hmin ranges to fix
1  5   100           - structure number, lower and upper limits
1                    - number of Vert ranges to fix
1  2.5  80           - structure number, lower and upper limits
ackt3d_ang.out       - file with prior local anisotropic angles
1 2 3                - columns for local ang1, ang2 and ang3
1                    - number of azimuth angles to control/fix
1   0.0  180.0       - structure number, lower and upper limits
1                    - number of dip angles to control/fix
1   0.0  90.0        - structure number, lower and upper limits
1                    - number of plunge angles to control/fix
1  -90.0  0.0        - structure number, lower and upper limits
0                    - number of variogram preferences
1                    - number of Hmax/Vert anis. to control/fix
1   0.5    1.0       - structure number, lower and upper limits
1                    - number of Hmin/Hmax anis. to control/fix
1   0.5   1.0        - structure number, lower and upper limits
```

Figure A-10: Example of the advanced options block of parameters for `globfit`.

A.4. Spatial Prediction with Location-Dependent Statistics

The estimation and simulation programs for spatial prediction with location-dependent statistics are modified versions of previously available programs for kriging and sequential simulation with global statistics. Among the different techniques that could be adapted for locally stationary estimation, only simple, ordinary kriging and locally stationary multiGaussian kriging have been implemented so far. Locally stationary indicator kriging has not been developed mainly due to the difficulty of modelling the location-dependent indicator variograms at multiple locations and for several cut-off's. A locally stationary co-kriging program is also pending. Two programs are available for

locally stationary simulation: **SGSim_loc** for continuous variables and **SISim_loc** for categorical variables.

A.4.1. Locally Stationary MultiGaussian Kriging: kt3d_LMG

This program for locally stationary point and block support estimation was developed form the GSLIB program **kt3d** (Deutsch & Journel 1998, pp.96-100). It is able to perform locally stationary simple and ordinary kriging and locally stationary multiGaussian kriging with local normal score transformations. For multiGaussian block support estimation, the program performs change of support using the discrete Gaussian model with local variograms. If none of the location-dependent parameters required for locally stationary estimation is provided, the **kt3d_LMG** will perform exactly as the **kt3d** program with globally stationary parameters. The details of the required parameter file are given below. Figure A-11shows an example of this file.

- **datfl**: This is the original data file, the same as used for **LDWgen**.
- **idhl**, **ixl**, **iyl**, **izl**, **ivr**, and **iextv**: the column numbers for the x, y, and z coordinates, the variable used in estimation and the collocated non-stationary mean.
- **tmin** and **tmax**: all values in **datfl** smaller than **tmin** and greater than **tmax** are ignored.
- **lhercofl**: a file containing the Hermite coefficients interpolated at the resolution of the estimation grid defined below. If this file is not provided then local normal score transformation is not performed and the program works in original units.
- **npol** and **ipol**: the number of expansions used in the Hermite polynomial fitting of the local normal score transformation functions, and the column number for the first local coefficient.
- **acerr**: the acceptable error in the block variance when building the change of support model.
- **zmin** and **zmax**: the minimum and maximum values in the back transformed distributions.
- **koption**: if set to 0, point or block kriging of all nodes in the grid is performed. If set to 1, cross-validation with the data in **datfl** is performed. If set to 2, jackknifing is performed.
- **jackfl**: a file with locations to perform jackknife.

- **icoljx**, **icoljy**, **icoljz**, **icoljvr**, and **icoljsec**: the column numbers for the x, y, and z coordinates, the variable and the collocated non-stationary mean in **jackfl**.
- **idbg**: indicates the debugging output level. Level 0 is for no debugging output. The maximum level, 3, provides the kriging matrices at every estimated location.
- **dbgfl**: the file for debugging output.
- **outfl**: the file for the estimation or crossvalidation output. If **koption** is set to 0, the output is a grid file containing the estimates and estimation variances in original and Gaussian units, the p-value for the threshold specified by cut-off, the local change of support coefficient and the local block support variance. If **koption** is set to 0, the output file contains the sample x, y, and z coordinates, the true and estimated sample values, the estimation error, the percentiles corresponding to the true and estimated values in the local distribution and some significant P-values of the local distribution.
- **nx**, **xmn**, and **xsiz**: the number of blocks, the coordinate of the first block centre and the size of the blocks in the *x* direction.
- **ny**, **ymn**, and **ysiz**: the number of blocks, the coordinate of the first block centre and the size of the blocks in the *y* direction.
- **nz**, **zmn**, and **zsiz**: the number of blocks, the coordinate of the first block centre and the size of the blocks in the *z* direction.
- **nxdis**, **nydis**, and **nzdis**: the number of blocks discretization points parallel to each direction. In all these parameter values are set to 1, point kriging is performed.
- **ndmin** and **ndmax**: the minimum and maximum number of data points within the search neighbourhood used in estimation .
- **noct**: the maximum number of data points to consider within an octant of the search ellipsoid. Octants are not used if this parameter is set to 0.
- **radius**, **radius1**, and **radius**: the radii of the search ellipsoid in the maximum horizontal direction, minimum horizontal direction and vertical direction.
- **sang**, **sang2** and **sang3**: the azimuth, dip and plunge angles describing the orientation of the search ellipsoid.
- **ikrige** and **skmean**: the kriging type and global mean value. If **ikrige** parameter is set to 0, then simple kriging with constant mean specified by **skmean** will be performed. If **ikrige** is set to 1, then ordinary kriging will be performed. An **ikrige** value of 2 is used for non-stationary simple kriging with the local means in **secfl**. If **ikrige** is set to 3, then kriging with external drift will be performed. If **ikrige** is set to 4, locally

stationary simple kriging with local means taken from **secfl** is performed. When performing locally stationary multiGaussian kriging with local normal scores transformations, the **ikrige** value must be set to 0.

- **idrift(i)**, i=1,...,9: the parameters indicating the drift terms that will be used in the trend model (Deutsch & Journel, 1998, p. 99).
- **itrend**: if set to 0, a variable trend is considered, if set to 1, the trend is estimated. A value of 0 must be used if performing locally stationary kriging.
- **secfl**: the file containing the gridded external drift variable or the local means at all the locations to be estimated. This file is required if **ikrige** is 2, 3 or 4.
- **iseccol**: the number of the column in **secfl** containing the external drift or the local mean.
- **cutoff**: the threshold value for reporting its corresponding p-value if locally stationary multiGaussian kriging is performed.
- **nst**, **c0g**, and **alfag**: the number of structures, the nugget effect and the exponent of the stable model for the global variogram model. The number of structures defined for the global variogram model also defines the number of structures of the local variogram models. The next two lines must be repeated **nst** times.
- **itg**, **ccg**, **ang1g**, **ang2g**, and **ang3g**: the global variogram parameters for the type of the structure, sill contribution of the structures, and the angles defining the geometric anisotropy. The types of structures of the global variogram model are the same as those for the local variogram models.
- **aag**, **a2g**, and **a3g**: the global variogram parameters for the maximum horizontal range, the minimum horizontal range and the vertical range.
- **locvarfl**: a grid file containing the location-dependent variogram model parameters. The grid definition of this file must be the same as the grid definition of the estimation.
- **ic0l**, and **ialfal**: the column numbers in **locvarfl** for the local nugget effect and the local exponent if the stable model in the first variogram structure is used. If any of these column numbers is set to 0 the program will take the corresponding parameter value from the definition of the global variogram model.

The following two lines must be repeated for the number of structures indicated in the specification of the global variogram model:

- **iccl**, **iang1l**, **iang2l**, and **iang3l**: the column numbers in **locvarfl** for the contribution in the local sill, and the angles defining the local geometric anisotropy. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.
- **iaal**, **ia2l**, and **ia3l**: the column numbers in locvarfl for the local maximum horizontal range, the local minimum horizontal range, and the local vertical range. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.

```
                  Parameters for KT3D_lMG
                  ***********************

START OF PARAMETERS:
example.dat                       - file with data
0  2  3  4  12  0                 - columns for DH,X,Y,Z,var,sec var
-10.0   1.0e21                    - trimming limits
LocalHerpol.dbg                   - file with local Hermite polynomials (in columns)
36 7                              - number of hermite polynomials and column for phi(0)
0.1                               - acceptable error for block variance
1   26                            - minimum and maximum value in original units
0                                 - option: 0=grid, 1=cross, 2=jackknife
xvk.dat                           - file with jackknife data
1   2   0    3    0               - columns for X,Y,Z,vr and sec var
1                                 - debugging level: 0,1,2,3
MGkt3d-loc-blk.dbg                - file for debugging output
MGkt3d-loc-blk.out                - file for kriged output
100   2.0  4.0                    - nx,xmn,xsiz
100   2.0  4.0                    - ny,ymn,ysiz
100  -398   4.0                   - nz,zmn,zsiz
1    1      2                     - x,y and z block discretization
2    16                           - min, max data for kriging
0                                 - max per octant (0-> not used)
65.0   65.0  65.0                 - maximum search radii
 0.0   0.0   0.0                  - angles for search ellipsoid
0      0.000                      - 0=SK,1=OK,2=non-st SK,3=exdrift,4=locally stationary
0 0 0 0 0 0 0 0 0                 - drift: x,y,z,xx,yy,zz,xy,xz,zy
0                                 - 0, variable; 1, estimate tren
extdrift.dat                      - gridded file with drift/mean
0                                 - column number in gridded file
0                                 - threshold for reporting local p-value
1    0.41  1.0                    - nst, nugget effect
2    0.59  0.0   0.0   0.0        - it,cc,ang1,ang2,ang3
         21.0  21.0  21.0        - a_hmax, a_hmin, a_vert
locvar.out                        - file for variogram parameters (same grid)
5   7                             - columns:   nugget effect, exp
8   9   0   0                     -            cc,ang1,ang2,ang3
12  13   0                        -            a_hmax, a_hmin, a_vert
```

Figure A-11: An example parameter file for `kt3d_LMG`.

A.4.2. Locally Stationary Sequential Gaussian Simulation: `ultimateSGSIM v.2.0`

This new version of **`ultimateSGSIM`** program (Deutsch & Zanon 2002) is able to perform sequential Gaussian simulation with locally normal scores transformed values and local variogram models. A new block of parameters for taking into account the local normal score transformation and the local variogram models has been added at the end of the parameter file. If this block does not exist or the parameters in it are erroneous, then **`ultimateSGSIM`** will perform global normal score transformation and use the global

variogram model parameters. The other blocks of parameters in this file remain the same as those for the previous version. This new version of **`ultimateSGSIM`** accepts only location-dependent direct variograms. The detail of the parameters in this new block for location-dependent statistics and local normal scores transformations is given below. Figure A-12 presents an example for this module.

- **ilds**: if set to 0, global normal scores transformation will be performed and the global variogram model will be used.
- **ivr** and **iv1**: the rock type and variable number for which the location-dependent parameters specified below will be used. This line and the following lines of parameters can be repeated for the combination of all variables and rock types specified in the main parameter block. If any combination is missing, the program will perform the simulation of the missing variable number at the missing rock type with the global parameters indicated in the transformation and variogram blocks.
- **lhercofl**: the file containing the Hermite coefficients interpolated at the resolution of the estimation grid defined in the main block of parameters. If this file is not provided then global normal score transformation is performed according to the parameters in the transformation block.
- **npol** and **ipol**: the number of expansions used in the Hermite polynomial fitting of the local normal score transformation functions, and the column number for the first local coefficient.
- **locvarfl**: a grid file containing the parameters of the location-dependent variogram model. The grid definition of this file must be the same as the grid defined in the main block.
- **ic0l**, and **ialfal**: the column numbers in **locvarfl** for the local nugget effect and the local exponent if the stable model in the first variogram structure is used. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.

The following two lines must be repeated for the number of structures indicated in the block of global variogram model parameters for the corresponding variable and rock type.

- **iccl**, **iang1l**, **iang2l**, and **iang3l**: the column numbers in **locvarfl** for the contribution in the local sill, and the angles defining the local geometric anisotropy. If any of these

column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.

- **iaal**, **ia2l**, and **ia3l**: the column numbers in **locvarfl** for the local maximum horizontal range, the local minimum horizontal range, and the local vertical range. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.

```
START OF LDS
1                               - use location dependent statistics? (0=no, 1=yes)
1  1                            - rock type, variable number
Herpol_Rock1-v1.dat             - file with local Hermite polynomials (in columns)
36 1                            - number of hermite polynomials and column for phi(0)
locvar_Rock1-v1.out             - file for variogram parameters (same grid)
5   7                           - columns:   nugget effect, exp
8   9   10   11                 -            cc,ang1,ang2,ang3
12  13  14                      -            a_hmax, a_hmin, a_vert
2  1                            - rock type, variable number
Herpol_Rock2-v1.dat             - file with local Hermite polynomials (in columns)
36 1                            - number of hermite polynomials and column for phi(0)
locvar_Rock2-v1.out             - file for variogram parameters (same grid)
5   7                           - columns:   nugget effect, exp
8   9   10   11                 -            cc,ang1,ang2,ang3
12  13  14                      -            a_hmax, a_hmin, a_vert
```

Figure A-12: An example of the location-dependent statistics block of parameters for `ultimateSGSIM v.2.0.`

A.4.3. Locally Stationary Sequential Indicator Simulation: `sisim_loc`

This program for sequential indicator simulation with local proportions and local variogram models is a modification of the **`sisim_lm`** program (Deutsch & Journel 1998, pp.175-180). The parameters of the new version are almost the same as for **`sisim_lm`**. The only difference is that the new block of parameters for specifying the files containing the location dependent indicator variogram models has been added at the end of the parameter file. An example of the complete parameter file is shown in Figure A-13. The following parameter lines in the block for the location-dependent variogram models must be repeated for the number of categories specified at the second line of the parameter file.

- **locvarfl**: the grid file containing the parameters of the location-dependent variogram model. The grid definition of this file must be the same as the grid defined for simulation. If this file is missing, the global indicator variogram model parameters will be used for the corresponding category.
- **ic0l**, and **ialfal**: the column numbers in **locvarfl** for the local nugget effect and the local exponent if the stable model in the first variogram structure is used. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.

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

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

START OF LDVs
vfLIC0.sum            - Cat 0: file for variogram parameters (same grid)
5   7                           - columns:   nugget effect, exp
8   9   10  11                  -            cc,ang1,ang2,ang3
12  13  14                      -            a_hmax, a_hmin, a_vert
vfLIC1.sum            - Cat 1: file for variogram parameters (same grid)
5   7                           - columns:   nugget effect, exp
8   9   10  11                  -            cc,ang1,ang2,ang3
12  13  14                      -            a_hmax, a_hmin, a_vert
```

Figure A-13: An example parameter file for `sisim_loc`.

The following two lines must be repeated for the number of structures indicated in the definition of the global indicator variogram model parameters for the corresponding category.

- **iccl**, **iang1l**, **iang2l**, and **iang3l**: the column numbers in **locvarfl** for the contribution in the local sill, and the angles defining the local geometric anisotropy. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.
- **iaal**, **ia2l**, and **ia3l**: the column numbers in **locvarfl** for the local maximum horizontal range, the local minimum horizontal range, and the local vertical range. If any of these column numbers is set to 0, the program will take the corresponding parameter value from the definition of the global variogram model.

Appendix B
Nomenclature

B.1. List of Abbreviations

cdf	cumulative distribution function
ccdf	conditional cumulative distribution function
KT	kriging with a trend model
LDIQR	location-dependent interquantile range
LSMGK	locally stationary multiGaussian kriging
LSOK	locally stationary ordinary kriging
LSSGS	locally stationary sequential Gaussian simulation
LSSIS	locally stationary sequential indicator simulations
LSSK	locally stationary simple kriging
MSE	mean square error
MGK	traditional multiGaussian kriging
RAM	random access memory
OK	traditional ordinary kriging
pdf	probability distribution function
RV	random variable
RF	random function
SK	traditional simple kriging
SGS	traditional sequential Gaussian simulation
SIS	traditional sequential indicator simulation

B.2. List of Most Important Symbols

$C(\mathbf{h})$	global covariance of the RF
$\hat{C}(\mathbf{h})$	experimental global covariance
$\hat{C}(\mathbf{h};\mathbf{o})$	experimental location-dependent covariance
Cov	covariance function
D	a domain in space
$d(\mathbf{u};\mathbf{o})$	Euclidean distance between a sample at location **u** and an anchor point at location **o**.
E	expected value
exp	exponentiation function
ε	Gaussian kernel background value
$F(\mathbf{u};z)$	cumulative distribution function of $Z(\mathbf{u})$ for a cutoff z
$F(\mathbf{u};z \mid n(\mathbf{u}))$	ccdf at a location **u** given $n(\mathbf{u})$
$\hat{F}(\mathbf{u};z;\mathbf{o})$	experimental univariate location-dependent prior cdf for a cutoff z
$\hat{F}(z)$	experimental univariate global cdf for a cutoff z
$G(y)$	standard normal Gaussian cdf
$g\left(\mathbf{u};y \mid (n)\right)$	conditional pdf in Gaussian units
$\gamma(\mathbf{h})$	global variogram of the RF
$\hat{\gamma}(\mathbf{h})$	experimental global variogram
$\gamma(\mathbf{h};\mathbf{o})$	location-dependent variogram model
$\hat{\gamma}(\mathbf{h};\mathbf{o})$	experimental location-dependent variogram
$\mathbf{h}$	2-point distance or lag vector in Euclidean space
$H_q(y)$	q Hermite polynomial
$I(\mathbf{u};z)$	binary indicator for a cutoff z
$\lambda_\alpha(\mathbf{u})$	estimation weight assigned to a sample α in relation to **u**
$\lambda_\alpha^{(LSOK)}(\mathbf{o})$	locally stationary ordinary kriging weights assigned to sample to sample α in relation to location **o**

$\lambda_{\alpha}^{(LSSK)}(\mathbf{o})$	locally stationary simple kriging weights assigned to sample to sample α in relation to location **o**
$\lambda_{\alpha}^{OK}(\mathbf{u})$	ordinary kriging weight assigned to sample α in relation to location **u**
$\lambda_{\alpha}^{SK}(\mathbf{u})$	simple kriging weight assigned to sample α in relation to location **u**
m	global mean of the RF
$\hat{m}$	experimental global mean
$\hat{m}_{-\mathbf{h}}$	experimental global tail mean
$\hat{m}_{+\mathbf{h}}$	experimental global head mean
$m(\mathbf{u})$	local mean
$\hat{m}(\mathbf{o})$	experimental location-dependent mean
$\hat{m}_{-\mathbf{h}}(\mathbf{o})$	experimental location-dependent tail mean
$\hat{m}_{+\mathbf{h}}(\mathbf{o})$	experimental location-dependent head mean
Max	maximum function
max	maximization
Min	minimum function
min	minimization
$N(\mathbf{h})$	number of sample pairs separated by vector **h**
n	total number of samples within a domain
$n(\mathbf{u})$	number of samples within a neighbourhood centered at location **u**
o	a coordinate vector corresponding to an anchor point location in 1-D, 2-D or 3-D Euclidean space
P	number of anchor points within a domain
p	probability value
Prob	probability function
Q	number of Hermite polynomials
$R(\mathbf{u})$	residual RF
$\rho(\mathbf{h})$	global correlogram of the RF
$\hat{\rho}(\mathbf{h})$	experimental global correlograms
$\hat{\rho}(\mathbf{h};\mathbf{o})$	experimental location-dependent correlograms
$\hat{S}(\mathbf{o})$	experimental location-dependent variogram sill

s	Gaussian kernel bandwidth or standard deviation
σ^2	global variance of the RF
$\hat{\sigma}^2$	experimental global variance
$\hat{\sigma}^2(\mathbf{o})$	experimental location-dependent variance
$\hat{\sigma}^2_{-\mathbf{h}}(\mathbf{o})$	experimental location-dependent tail variance
$\hat{\sigma}^2_{+\mathbf{h}}(\mathbf{o})$	experimental location-dependent head variance
$\sigma^2_E(\mathbf{u})$	variance of the estimation error
$\sigma^2_{LSOK}(\mathbf{o})$	locally stationary ordinary kriging variance
$\sigma^2_{LSSK}(\mathbf{o})$	locally stationary simple kriging variance
$\sigma^2_{OK}(\mathbf{u})$	ordinary kriging estimation variance
$\sigma^2_{SK}(\mathbf{u})$	simple kriging estimation variance
t	mixture rule exponential
$\mathbf{u}$	a coordinate vector corresponding to a location in 1-D, 2-D or 3-D Euclidean space
$\overline{\mathbf{uo}}$	vector that joins a sample at location **u** and an anchor point at location **o**
$\varphi_Y(z;\mathbf{o})$	location-dependent normal scores transformation function
$\varphi_Z(y;\mathbf{o})$	location-dependent normal scores back-transformation function
$\phi_q(\mathbf{o})$	q location-dependent Hermite coefficient
w_α	declustering weight assigned to the sample α
Var	variance function
$v(\mathbf{o})$	block centered in **o**
$Y(\mathbf{u})$	a random variable in Gaussian units
$Y^*(\mathbf{u})$	RV estimator in original units
$y(\mathbf{u})$	a Gaussian transformed attribute value at location **u**
$y^*(\mathbf{u})$	estimated value in original units at a location **u**
$Z(\mathbf{u})$	a random variable in original units
$Z^*(\mathbf{u})$	RV estimator in original units
$z(\mathbf{u})$	an attribute value at location **u** in original units

$z^*(\mathbf{u})$	estimated value in original units at a location **u**
$Z^*_{LSMGK}(\mathbf{o})$	locally stationary multiGaussian kriging estimator
$Z^*_{LSOK}(\mathbf{o})$	locally stationary ordinary kriging estimator
$Z^*_{LSSK}(\mathbf{o})$	locally stationary simple kriging estimator
$Z^*_{OK}(\mathbf{u})$	ordinary kriging estimator
$z_p(\mathbf{o})$	location-dependent p-quantile in original units
$\hat{z}_p(\mathbf{o})$	experimental location-dependent p-quantile in original units
$Z^*_{SK}(\mathbf{u})$	simple kriging estimator
$\omega(\mathbf{u};\mathbf{o})$	distance based 1-point weight assigned to a sample at location **u** in relation to the anchor point at location **o**.
$\bar{\omega}(\mathbf{u})$	average of distance weights assigned to a sample at location u in relation to all anchor points
$\hat{\omega}(\mathbf{u};\mathbf{o})$	distance based 1-point weight after correction by declustering weights
$\omega'(\mathbf{u};\mathbf{o})$	distance based 1-point weight after standardization
$\tilde{\omega}_{GK}(\mathbf{u};\mathbf{o})$	distance based 1-point weight obtained with a Gaussian kernel with dynamic bandwidth
$\omega(\mathbf{u},\mathbf{u}+\mathbf{h};\mathbf{o})$	distance based 2-point weights assigned to a sample pair separated by vector **h** and in relation to an anchor point at location **o**.
$\omega'(\mathbf{u},\mathbf{u}+\mathbf{h};\mathbf{o})$	standardized distance based 2-point weights