Addressing Order Relation Issues with Constrained Radial Basis Functions and Consistent Indicator Variograms

by

Sebastián Ignacio Sánchez Villar

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mining Engineering

Department of Civil and Environmental Engineering
University of Alberta

Abstract

Quantifying uncertainty is a critical task of resource delineation in the mining industry. Uncertainty is used to assess risk in economic evaluation and for classification in resource reporting. The inference of local distributions from conditioning data is key to quantifying uncertainty. Multiple indicator Kriging (MIK) is a well-established non-parametric local distribution inference technique that does not assume a prior distribution. The local conditional cumulative distribution functions (CCDF) are estimated directly from indicators defined from thresholds. MIK is flexible since allows the addition of soft data and accounts for different spatial correlations related to different thresholds. These advantages can be eclipsed by the fact that MIK CCDFs almost always present order relation issues, that is, probabilities below zero, above 1 and decreasing for increasing thresholds. The aim of this work is to understand the origin of order relations issues and to find ways to reduce or eliminate them. It explores their relation with the negative weights of Kriging. It also focuses on analytical and practical techniques that help to avoid order relations issues. One technique explores the internal consistency of the indicator variograms used in MIK and its connection to order relation issues. The other proposes an interpolation methodology using Radial Basis Functions (RBF) with inequality constraints to produce CCDFs without order relation issues.

There are four main contributions of this research. First, it presents different tests and examples that show how negative weights influence order relation issues and their distributions. It also shows the relation between negative weights and indicator variograms. Second, it explains the internal consistency of indicator variograms related to a bivariate distribution and shows novel equations for the calculation of probabilities of the internal bivariate distribution. Additionally, it proposes a workflow to use the equations as a tool to aid the indicator variogram modeling process. Third, it proposes a new methodology of MIK that uses the RBF framework to add inequality constraints to the estimator. The constraints force the MIK estimates to comply with a licit CCDF. Finally, the equations to calculate bivariate probabilities and the RBF methodology are combined into a single workflow for MIK. The RBF framework is proven to work with a similar performance to classic MIK.

These contributions aid to understand the different modeling processes involved in MIK, from the intrinsic characteristics of Kriging and RBF to the internal consistency of indicator variogram models. The understanding of this helps to build more consistent models in the future.

Dedication

To Marlene, Nelson, Martin, Carolina and don Goyo

Acknowledgments

First, I would like to thank my supervisor Dr. Clayton Deutsch for his support, his guidance, and for being approachable while I undertook this research project. It was a true honor to be mentored by you. Second, I would like to thank my co-supervisor Dr. Jeff Boisvert for his unique ideas, his ability to keep me on track, and the systematic approach he took to guide me through the various stages of this thesis. Third, I would like to extend my gratitude to the CCG members for the financial support that made my research, and this thesis, a reality. Fourth, I would like to thank the whole CCG team I had the privilege of getting to know. The countless geostatistics debates and after-hours gatherings will always be remembered. I am deeply grateful to my parents, Nelson and Marlene, and the rest of my family, for their lifelong support and encouragement when I decided to move to a different country and pursue my master's. Last, I want to thank my girlfriend Breanna and the new friends I have made in Edmonton.

Table of Contents

List of Tables

List of Figures

List of Symbols

Symbol	**Description**
A	Set of locations
$\mathcal{A}$	NxL matrix of RBF values
a	Range of a variogram
ax	Range of anisotropy on direction x
ay	Range of anisotropy on direction y
az	Range of anisotropy on direction z
B	MxL matrix of RBF values
$B(\ )$	Basis function
b	Constant that represents the probability of error in determination of bins
$\mathcal{C}(\ )$	Normalization function
$Cov[\ ]$	Covariance
$Cov(\mathbf{h})$	Covariance of random variables separated by a vector **h**
$Cov_k(\mathbf{h}, z_k)$	Covariance of the indicator random variable for threshold z_k separated by a vector **h**
$C_{Y_iY_j}$	Covariance between the Gaussian random variables Y_i and Y_j
c_0	Nugget effect
c_j	Structure sill contribution
$\mathbf{C}$	Matrix of covariances
c	Vector of covariances
D	Threshold number used for the division in a log-ratio transform
d	Vector of weights
$E[\ ]$	Expected value operator
F	Solution vector with $f(\mathbf{u})$ values
$F(z_k)$	Cumulative density function for indicator $I(\mathbf{u}, z_k)$
$F^*(\mathbf{u}, z_k)$	Cumulative density function estimate for indicator $I(\mathbf{u}, z_k)$ at location **u**
$f(\mathbf{u})$	Sample value function at location **u**
$f_{Y_1 \dots Y_n}(y_1, \dots y_n)$	Multi-variate Gaussian probability density function
$GK(\ )$	Gaussian Kernel function
$\mathcal{H}$	Hilbert space
H	LxL of RBF values
$I(\mathbf{u}_i, z_k)$	Indicator random variable at location $\mathbf{u}_i$ for threshold z_k

Symbol	**Description**
$I^*_{RBF}(\mathbf{u}, z_k)$	Indicator random estimate using RBF at location $\mathbf{u}$ for threshold z_k
$I^*_{RBFC}(\mathbf{u}, z_k)$	Indicator random estimate using RBFC at location $\mathbf{u}$ for threshold z_k
J	Number of structures in a variogram
$j(\mathbf{u}_i, z_k)$	Indicator random variable assigned from the bin related to threshold z_k
$j'(\mathbf{u}_i, z_k)$	Log ratio transform of indicator random variable assigned from a bin
$j^*(\mathbf{u}_i, z_k)$	Indicator random variable assigned from the bin estimate
K	Number of thresholds
k	Threshold number
$K_{kk'}(\mathbf{h}; z_k, z'_k)$	Non-centered indicator cross-covariance for thresholds z_k and z'_k at a lag $\mathbf{h}$
L	Number of samples plus the number of inequality constraints
m	Mean
M	Number of constraints
N	Number of samples
$N(\mathbf{h})$	Number of samples separated by a vector $\mathbf{h}$
P_{ij}	Bivariate probability
$p^*(\mathbf{u}, z_k)$	Probability density function estimate
p, p'	Probabilities
p_{z_k}	Indicator proportion for threshold z_k
q	Vector of inequality values
r_k	Central value of the bin related to the threshold z_k
$r(\mathbf{u}_i, \mathbf{u}_j)$, $r_{i,j}$	Distance function between locations $\mathbf{u}_i$ and $\mathbf{u}_j$
$s(\mathbf{u})$	Interpolator function at location $\mathbf{u}$
$\mathbf{u}$	Location vector
ux_i	Coordinate of the location $\mathbf{u}_i$
uy_i	Coordinate of the location $\mathbf{u}_i$
uz_i	Coordinate of the location $\mathbf{u}_i$
$\mathbf{u}_i$	Sample location
$\mathbf{v}$	Location vector
$Var[\ ]$	Variance operator
$X(\mathbf{u})$	Random variable at location $\mathbf{u}$
Y	Gaussian random variable
$\mathbf{y}$	Vector of Gaussian variables
y_i	Outcome of a Gaussian random distribution

Symbol	**Description**
$Z(\mathbf{u})$	Random variable at location $\mathbf{u}$
$z(\mathbf{u})$	Outcome of the random variable $Z(\mathbf{u})$
$Z^*(\mathbf{u})$	Estimate at location $\mathbf{u}$
z_k	Threshold number k
σ^2	Variance
σ_Y^2	Variance of a Gaussian random variable
$\gamma(\mathbf{h})$	Variograms for lag $\mathbf{h}$
$\gamma_k(\mathbf{h}, z_k)$	Indicator variogram for threshold z_k at lag $\mathbf{h}$
$\gamma_{kk'}(\mathbf{h}; z_k, z_k')$	Indicator cross-variogram for threshold z_k and z_k' at lag $\mathbf{h}$
λ_0	Kriging estimator constant
λ_i	Weights
λ	Vector of weights
$\mathbf{\Sigma}$	Covariance matrix of Gaussian random functions
μ	Vector of means
$\varphi(\ \ )$	Normal scores transformation function
α	Range of a Gaussian kernel
$\phi(\ )$	Radial basis function
ϵ	Constant of a radial basis function
$\mathcal{N}(0, 1)$	Gaussian distribution with zero mean and variance of one
$Sph(h)$	Spherical variogram function
$Exp(h)$	Exponential variogram function
$exp(\ \ )$	Exponential function
$Gaus(h)$	Gaussian variogram function
$Prob\{\ \ \}$	Probability operator

List of Abbreviations

Abbreviation	**Description**
CCG	Centre for computational geostatistics
CDF	Cumulative distribution function
CCDF	Conditional cumulative distribution function
CDIK	MIK with a Gaussian kernel
CPDF	Conditional probability density function
GSLIB	Geostatistical software library
SIK	Simplicial Kriging
MGK	MultiGaussian Kriging
MIK	Multiple indicator Kriging
MIKRBF	Multiple indicator Kriging with RBF with constraints
OK	Ordinary Kriging
PDF	Probability density function
RBF	Radial basis function
RBFC	Radial basis function interpolation with constraints
RMSE	Root mean squared error
SOR	Slope of regression

Chapter 1

Introduction

Since the mid-20th century, with the work of Matheron (Matheron, 1963), geostatistics has focused on the problem of interpolation of sparse data and delineation of resources. The idea of considering sparse data as stationary random variables led to the development of Kriging, a linear interpolator optimized to have the lowest squared error (Journel, 1984). Kriging uses the covariance or variogram matrix between samples to calculate weights. Different types of Kriging have been developed: simple Kriging is the best-unbiased linear estimator (Chiles & Delfiner, 2009); Ordinary Kriging does not assume a prior mean (Isaaks & Srivastava, 1989); co-Kriging considers secondary data (Myers, 1982, 1984); Kriging with a trend assumes a parametric form to the mean (Rossi & Deutsch, 2013); Kriging with an external drift assumes an arbitrary form to the mean (Goovaerts et al., 1997); and Kriging with locally varying anisotropy allows non-stationary directions of continuity (Boisvert, Manchuk, & Deutsch, 2009).

Radial Basis Functions (RBF) interpolation was developed almost in parallel to the formulation of Kriging (Cowan et al., 2003). RBF solves the problem of interpolation of sparse data by considering a weighted linear combination of basis functions as the interpolator. The weights are calculated by forcing the interpolator to reproduce the sample grades at their respective locations (Fasshauer, 2007). Due to its versatility, RBF has been widely used for implicit geological modeling in combination with signed distance functions (Cowan et al., 2003; Osher & Fedkiw, 2003). One of its features is that the RBF framework allows for the addition of inequality constraints in its system which can be solved using quadratic programming (Hillier, Schetselaar, de Kemp, & Perron, 2014). One of the differences between RBF and Kriging is that Kriging uses covariances for the calculation of weights, while RBF can use any positive definite function. Although, as covariances are usually modeled with positive definite functions, they are suitable to be used in the RBF framework. In that case, RBF has the same formulation as dual Kriging (Stewart, de Lacey, Hodkiewicz, & Lane, 2014).

In the mining industry, uncertainty can be used to assess risk in economic evaluation and as a tool for classification and resource reporting (Rossi & Deutsch, 2013). Both Kriging and RBF interpolation produce smooth estimates with less variability than the variable being estimated and give a unique estimate. Kriging also gives a measure of uncertainty, the Kriging variance. However, the Kriging variance must be used carefully since it only depends on the spatial configuration and not the grade values (Isaaks & Srivastava, 1989). Uncertainty in grade estimation comes from several sources: a limited amount of sampling, sampling methodologies, intrinsic variability of data, and modeling decisions. Estimating the local conditional distributions is helpful in quantifying uncertainty. The most common methods to build local distributions rely on the assumption of a multivariate or bivariate Gaussian distribution. These methods include (1) multi-Gaussian Kriging (MGK) (Verly, 1983), which assumes a multivariate Gaussian distribution and then uses simple Kriging to estimate the mean and variance of the local distributions. (2) uniform conditioning uses the discrete Gaussian model as a change of support model to estimate the distribution of selective mining units (SMU) within a panel (Remacre, 1987; Roth & Deraisme, 2000; Vann & Guibal, 1998) (3) disjunctive Kriging assumes a bi-Gaussian distribution to estimate Hermite polynomials fitted from the cumulative distribution function (CDF) using simple Kriging. The local distributions are then determined from the estimated Hermite values (Armstrong & Matheron, 1986; Matheron, 1976). These methodologies are also called parametric since they estimate the parameters of an a priori distribution. The principal limitation of these methodologies is that they are only reliable if the models that describe the samples are suitable (Rossi & Deutsch, 2013).

Not all methodologies for estimating local distributions are parametric or dependent on a Gaussian assumption. Non-parametric methods do not make any assumptions on the prior distribution of the samples, instead, probabilities of the conditional cumulative distribution function (CCDF) at any location are calculated by the conditional expectation of the probability of samples values $z(\mathbf{u})$ being below a threshold z_k (Emery & Ortiz, 2004; Rossi & Deutsch, 2013). Multiple indicator Kriging (MIK) is the most used non-parametric local distributions estimator (Journel, 1983). The methodology consists of (1) choosing thresholds that discretize the CDF of the samples $z(\mathbf{u})$, (2) assigning one indicator variable per threshold which takes the value of 0 if the sample value is below a threshold z_k or 1 is if it above, and (3) estimate the indicator variables at each location. The indicator estimates represent the local CCDF probabilities for the thresholds z_k. Probabilities for values between the thresholds are interpolated and the tails are extrapolated to obtain a complete CCDF (Deutsch & Journel, 1998).

1.1 Problem motivation

The CCDF probabilities obtained from MIK should follow order relations to constitute a licit CCDF. Order relations imply that CCDF probabilities must be above 0, below 1 and nondecreasing for increasing thresholds. Since Kriging is a nonconvex estimator that can produce negative weights (Deutsch, 1996), the indicator approach will almost always present order relations issues (Deutsch & Journel, 1998; Goovaerts, 1994; Journel, 1983, 1984; Journel & Posa, 1990; Suro-Perez & Journel, 1991). These order relation issues are classically fixed by averaging downward and upward corrections and resetting values below zero to zero and above one to one (Deutsch & Journel, 1998) (Figure 1.1). Three specific problems can be identified related to order relation issues. First, most literature claims that order relation issues come from a lack of data, inconsistent variogram modeling and negative weights (Deutsch & Journel, 1998). However, no study shows practically how these reasons work together to produce order relation issues. A deeper study seems necessary to understand the nature of order relations issues and their connection with negative weights.

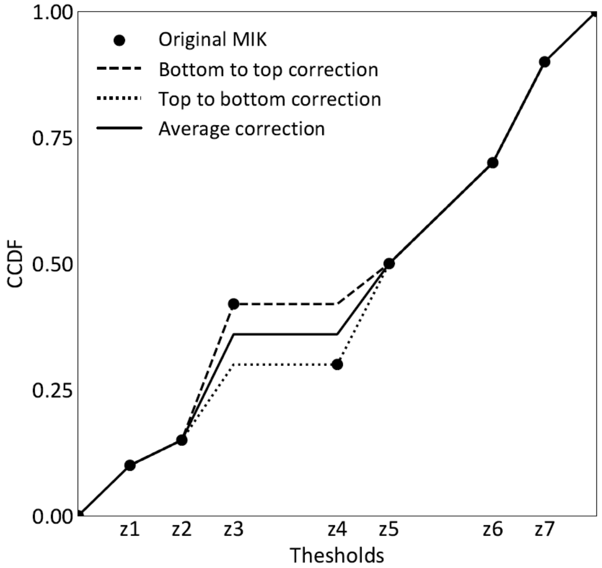


Figure 1.1: Order relation issue correction for a CCDF

Second, MIK calls for the estimation of several indicator variables defined from the continuous data related to thresholds to define local CCDF probabilities. This entails the fitting of each experimental indicator variogram by positive definite functions. A special attribute of experimental indicator variograms is that each of them can be calculated from the bivariate distribution formed by the pairs of samples separated by a distance-vector **h** (Journel & Posa, 1990; Sullivan, 1984) (Figure 1.2). The relation between indicator variograms and a bivariate distribution should not be ignored while modeling the indicator variograms. This characteristic is highlighted by Journel and Posa (1990). They define order relations that the modeled indicator variograms have to follow to comply with an underlying bivariate distribution. Furthermore, the probabilities of the underlying bivariate distribution must sum to 1 and be non-negative. This leads to two main questions: How can the probabilities of the bivariate distribution be derived from modeled variograms? and how can these probabilities be used to help model indicator variograms? The purpose of having consistent models should be to have better estimates. A test is necessary to check the relation between the consistency of modeled indicator variograms with an internal bivariate distribution, order relation issues and local distribution inference performance.

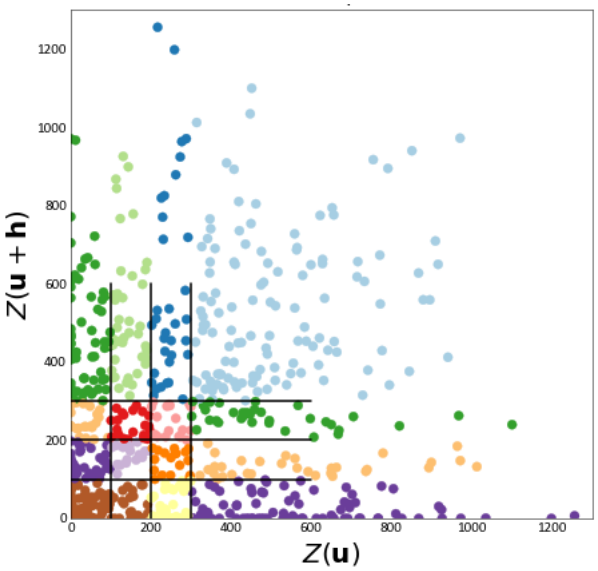


Figure 1.2: Bivariate distribution formed by the pairs $Z(\mathbf{u})$ and $Z(\mathbf{u}+\mathbf{h})$ for three thresholds

Finally, even after accounting for the consistency of models in MIK, order relation issues may arise. Multiple attempts have tried to prevent order relations issues in MIK. The use of a single variogram for all indicators to reduce order relation issues (Deutsch & Journel, 1998), constraints included in the Kriging system to not allow for negative weights (Barnes & Johnson, 1984; Deutsch, 1996), probability Kriging which uses additional information from the bivariate distribution (Carr & Mao, 1993; Journel, 1984; Sullivan, 1984), compositional data methodologies which consider the indicator values as part of a whole (Tolosana-Delgado, Mueller, & van den Boogaart, 2019; Tolosana-Delgado, Pawlowsky-Glahn, & Egozcue, 2008) and the addition of constraints into the Kriging system to force the estimates to respect order relations (Soltani-Mohammadi & Tercan, 2012). Despite all these efforts, the classic post-process to correct order relation issues is still the most widely used methodology. The modification of indicator estimates may produce a poor reproduction of variogram models in sequential indicator simulation (Deutsch & Journel, 1998). Hence, there is a need to produce a practical algorithm to account for order relation issues that do not involve a post-correction.

1.2 Thesis statement

There are three main statements in this thesis. (1) Order relations issues are produced by negative weights when changing variograms from one indicator to another, changing indicator values and the combination of both. This is seen in the case of median MIK which produces fewer order relation issues than median MIK. On the other hand, dissimilar variograms produce more order relation issues. (2) Equations can be derived from modeled indicator variograms and proportions to calculate the probabilities of an internal bivariate distribution. The probabilities will depart from the experimental bivariate distribution when fewer samples are used for their calculation. This is proven by a multi-Gaussian test and a sample density test. The probabilities can be used to aid the indicator variogram modeling. The methodology is tested in several real datasets. (3) Finally, this research proposes a new methodology to estimate local conditional distributions without order relation issues using MIK combined with RBF with constraints. The methodology is proven to have a similar performance to MIK and outperforms compositional data methodologies.

1.3 Thesis outline

The thesis is outlined as follows. Chapter 2 presents all the theoretical frameworks used in the thesis. It reviews the concept of regionalized variables and random function, spatial correlation modeling, simple and dual Kriging and parametric and non-parametric local distribution inference techniques. Chapter 3 presents studies to understand order relations issues in MIK. It maps the weights of a MIK estimation in a single location. It explores the distributions of order relation issues for MIK and median MIK. The relation between variogram dissimilarity and order relation issues is also explored. Chapter 4 reviews the internal bivariate distribution of indicator variograms and shows equations to calculate bivariate probabilities. The equations are tested in a multivariate Gaussian distribution. Tests with different data densities are performed. A performance comparison in multiple datasets is also explored. Chapter 5 introduces a new methodology to avoid order relation issues using MIK. The proposed methodology is compared with classic MIK, multi-Gaussian Kriging and two compositional data methodologies. Chapter 6 shows a demonstration of the use of bivariate probabilities for indicator variogram modeling plus the RBF with constraints methodology for MIK in a single workflow.

Chapter 2

Theoretical background

This chapter summarizes the theoretical framework used throughout the thesis. This includes the notion of regionalized variables and random functions, spatial correlation modeling, simple Kriging and RBF interpolation, local distribution inference and MIK methodologies.

2.1 Regionalized variables and random functions

Delineation of resources is a process that is included in several steps of a life of a mine. The idea is to have an estimation of the grade at any point in a space A. Since the natural processes that produce mineralization are too complex to be described deterministically, the forces that control the mineralization are assumed to be probabilistic (Isaaks & Srivastava, 1989). The probabilistic model allows the estimation of resources and quantification of uncertainty. The model uses the concepts of random variables. A random variable $Z(\mathbf{u})$ is defined by a set of possible outcomes $\{z_1(\mathbf{u}), z_2(\mathbf{u}), ..., z_k(\mathbf{u})\}$, each of them having a probability of occurrence. The model consists of considering the samples as a regionalized variable $z(\mathbf{u})$ and a realization of a random variable $Z(\mathbf{u})$. All random variables in space A, $\{Z(\mathbf{u}) : \mathbf{u} \ \in A\}$ form a random function. In a deposit, mineralization is not independent; depending on the genetic process involved, the deposits will have different shapes. For example, banded iron deposits are more continuous in the horizontal axis, while gold nugget deposits present almost no spatial correlation. This phenomenon is translated to the probability model as the spatial distribution of the random function. As with any other random variable in statistics, they are determined by the mean $m = E[Z(\mathbf{u})]$, also called the drift of $Z(\mathbf{u})$, the variance $\sigma^2 = Var[Z(\mathbf{u})] = E[Z^2(\mathbf{u}) - m]$, the covariance $Cov(\mathbf{u}, \mathbf{v}) = E[\{Z(\mathbf{u}) - m(\mathbf{u})\}\{Z(\mathbf{v}) - m(\mathbf{v})\}]$, and the variogram $2\gamma(\mathbf{u}, \mathbf{v}) = E[\{Z(\mathbf{u}) - Z(\mathbf{v})\}^2]$ for $\mathbf{u}, \mathbf{v} \ \in A$. Aside from the use of probabilistic models, another assumption is stationarity, that is, moments obey the same probability rules independent of their location. This assumption allows the covariance and variogram to be dependent only on the separation $\mathbf{u} - \mathbf{v} = \mathbf{h}$. Then, the covariance and variogram can be written only in terms of a lag $\mathbf{h}$: $Cov(\mathbf{h}) = E[\{Z(\mathbf{u}) - m\}\{Z(\mathbf{u}+\mathbf{h}) - m\}]$, and the variogram $2\gamma(\mathbf{h}) = E[\{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\}^2]$. These conditions are called second-order stationarity (Chiles & Delfiner, 2009).

2.2 Spatial correlation modeling

By assuming stationarity, the covariance and the variogram only depend on the separation vector $\mathbf{h}$ of the samples. For spatial correlation analysis, the variogram is historically preferred against the covariance since the former does not make use of the mean for its calculation (Chiles & Delfiner, 2009). The experimental variogram is calculated from the samples by:

$$2\gamma(\mathbf{h}) \simeq \frac{1}{N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \left[z\left(\mathbf{u}_i\right) - z\left(\mathbf{u}_i + \mathbf{h}\right) \right]^2$$

Where $N(\mathbf{h})$ is the number of pairs at a certain lag $\mathbf{h}$. The variogram represents the mean of the square differences between pairs of samples separated by a lag $\mathbf{h}$. Intuitively, closer samples are expected to be similar in value, so the variogram should be small. On the other side, samples that are far away are not spatially correlated, and their variogram value is the variance of the samples $Var[Z(\mathbf{u})] = C(0)$. The distance a where the variograms reach the sill or variance is called the range of the variogram. Estimation and simulation geostatistical tools make use of the variogram model for their calculations. To get variogram values for any lag $\mathbf{h}$ in any direction, the experimental variograms are fitted, manually or automatically, with positive definite functions. The most widely used functions are the following (Table 2.1):

Table 2.1: Positive definite variogram model functions (Rossi & Deutsch, 2013)

Name	Equation
Spherical	$Sph(h) = \begin{cases} 1.5(h/a) - 0.5(h/a)^3, h \leq a \\ 1, \text{ otherwise} \end{cases}$
Exponential	$Exp(h) = 1 - \exp(-3h/a)$
Gaussian	$Gaus(h) = 1 - \exp\left(-3(h/a)^2\right)$

These functions are also called nested structures. The variograms may present different behavior for different ranges, thus, it is convenient to use several functions to fit one variogram. The modeled variograms are composed of a nugget effect c_0, which represents the variability at a very short range, and nested structures that add a contribution c_j to the sill with $j \in \{1, \dots, J\}$ and J the number of structures (Rossi & Deutsch, 2013).

$$\gamma(\mathbf{h}) = c_0 + c_1 \cdot Sph(\mathbf{h}, a_1) + c_2 \cdot Sph(\mathbf{h}, a_2) + \dots$$

If the lag $\mathbf{h}$ does not take into account the direction, the variogram is called omnidirectional. In practice, deposits have directions of increased spatial continuity. This phenomenon is called anisotropy. Usually, three perpendicular ranges are defined to model the anisotropy.

2.3 Simple Kriging and dual Kriging

Consider a set of unsampled locations $\mathbf{u} \in A$. The goal is to get a grade estimation at those locations conditioned to a set of samples $z(\mathbf{u_i})$. A linear combination of weighted samples can be considered as the estimator:

$$Z^*(\mathbf{u}) = \sum_{i=1}^{N} \lambda_i z(\mathbf{u}_i) + \lambda_0$$

Where N is the number of samples. The weight values λ_i are chosen to minimize the square error $E[Z^*(\mathbf{u}) - Z(\mathbf{u})]^2$. In the case of λ_0 the value chosen is $\lambda_0 = m - \sum_{i=1}^{N} \lambda_i m$ to make the estimator unbiased. The mean m is assumed known and stationary. Then, the previous expression is written as:

$$Z^*(\mathbf{u}) = m + \sum_{i=1}^{N} \lambda_i z(\mathbf{u}_i - m)$$

Which is the same as working with $X(\mathbf{u}) = Z(\mathbf{u}) - m$:

$$X^*(\mathbf{u}) = \sum_{i=1}^{N} \lambda_i X(\mathbf{u}_i)$$

Notice that $E[X(\mathbf{u})] = 0$. The error variance of this new variable can be expanded in covariances terms:

$$E[\{X^*(\mathbf{u}) - X(\mathbf{u})\}^2] = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j Cov(\mathbf{u}_i, \mathbf{u}_j) - 2 \cdot \sum_{i=1}^{N} \lambda_i Cov(\mathbf{u}, \mathbf{u}_j) + Cov(0)$$

To minimize the previous expression, partial derivatives are applied and the results are set to 0. This leads to a system of equations:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_j Cov(\mathbf{u}_i, \mathbf{u}_j) = Cov(\mathbf{u}, \mathbf{u}_i) \qquad \text{for } i = 1, \dots, N$$

This problem is equivalent to solving a matrix multiplication of the form $\mathbf{C}\lambda = c$ where $\mathbf{C}$ is a NxN matrix of covariances $Cov(\mathbf{u}_i, \mathbf{u}_j)$, λ is a N vector of the weights, and c is a N vector of covariances $Cov(\mathbf{u}, \mathbf{u}_j)$. The system of equations can also be written as $\lambda^T = c^T \mathbf{C}^{-1}$. This expression can be replaced in the interpolator to obtain:

$$
\begin{aligned}
X^*(\mathbf{u}) &= c^T\mathbf{C}^{-1}X \\
X^*(\mathbf{u}) &= Y^T\mathbf{C}^{-1}c \\
X^*(\mathbf{u}) &= d^Tc
\end{aligned}
$$

The first expression is called the primal form of Kriging and the last expression is named the dual form of Kriging. The weights d^T are calculated by:

$$
\begin{aligned}
d^T &= X^T\mathbf{C}^{-1} \\
d &= \mathbf{C}^{-1}X \\
\mathbf{C}d &= X
\end{aligned}
$$

In both primal and dual forms, the covariances between locations are used for the calculation of weights; the covariances can be obtained from the modeled variogram by the expression:

$$Var(\mathbf{h}) = Cov(0) - Cov(\mathbf{h})$$

The expressions shown are assumed to use all samples for the estimation. In the case of the primal form, the samples can be cut off to a local neighborhood to avoid long estimation times. In the case of the dual form, it must use all the samples for the weight calculation but the process is done once.

2.4 Local distributions inference

Kriging produces smooth estimates with less variance than the original variable and gives only one representation of reality. Usually, regionalized variables present uncertainty due to a lack of information and the intrinsic variability of the data, hence, only one realization may not be suitable to get a complete picture of a deposit. Estimating local distributions is a common practice to quantify uncertainty (Deutsch & Journel, 1998). Two methods of estimation of local distribution are reviewed: Multi-Gaussian Kriging and Multiple Indicator Kriging.

2.4.1 Multi-Gaussian Kriging (MGK)

The multi-Gaussian Kriging is reviewed from Ortiz (2019). The multivariate Gaussian probability density function is a joint parametric density function of n Gaussian random variables Y_i and it is completely defined by the mean μ_i, the variance $\sigma^2_{Y_i}$ and the covariances $C_{Y_iY_j}$:

$$f_{Y_1 \dots Y_n}(y_1, \dots y_n) = \frac{exp(-\frac{1}{2}(\mathbf{y}-\mu)^T \mathbf{\Sigma}^{-1}(\mathbf{y}-\mu))}{\sqrt{(2\pi)^n |\mathbf{\Sigma}|}}$$

The mean and variance vectors and the covariance matrix are defined as:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \qquad \mathbf{\Sigma} = \begin{pmatrix} \sigma^2_{Y_1} & C_{Y_1Y_2} & \cdots & C_{Y_1Y_n} \\ C_{Y_2Y_1} & \sigma^2_{Y_2} & \cdots & C_{Y_2Y_n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{Y_nY_1} & C_{Y_nY_2} & \cdots & \sigma^2_{Y_n} \end{pmatrix}$$

The mean and the variance of a single Gaussian random variable Y_0 conditioned to a set of random variables Y_i for $i = 1, \dots, n$ can be obtained using Bayes' law:

$$\mu_{0|1} = m + \mathbf{\Sigma}_{01}\mathbf{\Sigma}^{-1}(\mathrm{y}_1 - m)$$
$$\sigma^2_{0|1} = \sigma^2_0 - \mathbf{\Sigma}_{01}\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_{10}$$

The set of Gaussian variables Y_i is considered to be the samples, and, by assuming stationarity, the mean m is the same for all Y_i. $\mathbf{\Sigma}_{01}\mathbf{\Sigma}^{-1}$ are written as:

$$\mathbf{\Sigma}_{01} = \begin{pmatrix} C_{Y_0Y_1} \\ C_{Y_0Y_2} \\ \vdots \\ C_{Y_0Y_n} \end{pmatrix} \qquad \mathbf{\Sigma} = \begin{pmatrix} \sigma^2_{Y_1} & C_{Y_1Y_2} & \cdots & C_{Y_1Y_n} \\ C_{Y_2Y_1} & \sigma^2_{Y_2} & \cdots & C_{Y_2Y_n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{Y_nY_1} & C_{Y_nY_2} & \cdots & \sigma^2_{Y_n} \end{pmatrix}$$

Notice that the equations described above are the same as the estimate and the Kriging variance of simple Kriging. Thus, the mean and variance of a multivariate Gaussian random variable distribution at any location conditioned to a set of multivariate Gaussian random variables can be obtained by simple Kriging.

In practice, the distribution of samples $z(\mathbf{u})$ from naturally occurring phenomena is rarely multi-Gaussian. To get a Gaussian distribution, the original $Z(\mathbf{u})$ is transformed to a random variable Y with a Gaussian distribution of mean 0 and variance 1. A Gaussian percentile value is assigned to each percentile of the original $Z(\mathbf{u})$ distribution (Figure 2.1):

$$Z = \varphi(Y)$$
$$Y = \varphi^{-1}(Z)$$
$$Y \sim \mathcal{N}(0,1)$$

Since the distribution of the samples is discrete, values are interpolated to fill the gaps and the tails are extrapolated to complete the distribution. This process is usually named normal score transformation.

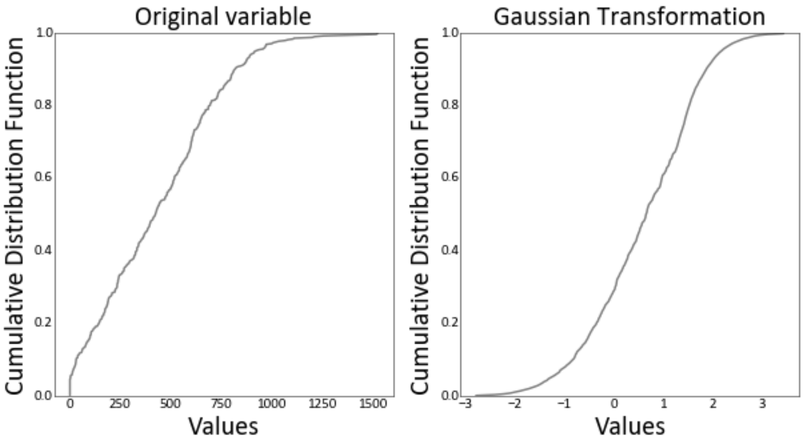


Figure 2.1: Normal score transformation

To use the normal score samples their distribution has to be multi-Gaussian or at least bi-Gaussian (Emery, 2005). The normal score transformation does not ensure a multi-Gaussian or bi-Gaussian distribution. Some of the checks to test the bi-Gaussianity of the samples are (1) the use of h-scattergrams $\{Y(\mathbf{u}), Y(\mathbf{u}+\mathbf{h})\}$, which should look similar to elliptic isodensity curves (Marechal, 1976), (2) the use of normalized Hermite polynomials by checking their relation to the experimental regression curves of the h-scattergrams, and (3) the comparison between experimental indicator variograms and analytical indicator variograms (Emery, 2005; Xiao, 1985).

Finally, the local conditioned distributions at any location can be estimated by transforming the data to a Gaussian distribution from the original data, checking if it is multi-Gaussian or bi-Gaussian, modeling variograms of the Gaussian values to get the covariance matrix and estimating the mean and Kriging variance using simple Kriging. To get a complete distribution in original values, any quantile can be back-transformed to original values.

2.4.2 Multiple Indicator Kriging (MIK)

To get local distributions using multi-Gaussian Kriging, the distribution of the samples must be multi-Gaussian. As seen before, a common method is transforming the data to a Gaussian distribution, but this does not always ensure multi-Gaussianity (Emery, 2005). Journel (1983) develop a non-parametric methodology that does not assume any prior distribution for the conditioning data. The methodology is based on indicators defined from thresholds. Then, the CCDF probabilities at each location are calculated using simple Kriging. The indicator function is a step function related to a threshold z_k. Indicator functions are assigned at N sample locations considering K thresholds:

$$I(\mathbf{u}_i, z_k) = \begin{cases} 1, \text{ if } Z(\mathbf{u}_i) \leqslant z_k \\ 0, \text{ otherwise} \end{cases}, \; k = 1, \dots, K, \; \imath = 1, \dots, N$$

where $I(\mathbf{u}_i, z_k)$ is an indicator function at location $\mathbf{u}_i$ for threshold z_k. Thresholds z_k are defined manually to discretize the sample distribution. They have to be chosen with care, too few thresholds will lose information and too many will make the calculations unnecessarily tedious (Deutsch & Journel, 1998). The expected value of an indicator random variable gives the cumulative density function value for threshold z_k:

$$\begin{aligned} E[I(\mathbf{u}, z_k)] = F(z_k) &= 1 \cdot \text{Prob}\{Z(\mathbf{u}) \leqslant z_k\} + 0 \cdot \text{Prob}\{Z(\mathbf{u}) > z_k\} \\ &= \text{Prob}\{Z(\mathbf{u}) \leqslant z_k\} \end{aligned}, \; k = 1, \dots, K$$

Two points statistics can be calculated from the bivariate distribution $Z(\mathbf{u} + \mathbf{h}, z_k)$ and $Z(\mathbf{u}, z_k)$. The covariance is:

$$\begin{aligned} Cov_k(\mathbf{h}, z_k) &= E\{I(\mathbf{u} + \mathbf{h}, z_k) \cdot I(\mathbf{u}, z_k)\} - E\{I(\mathbf{u} + \mathbf{h}, z_k)\} \cdot E\{I(\mathbf{u}, z_k)\} \\ &= \text{Prob}\{Z(\mathbf{u} + \mathbf{h}) \leqslant z_k \; \text{ and } Z(\mathbf{u}) \leqslant z_k\} - F(z_k)^2 \end{aligned}$$

The semi-variogram:

$$\gamma_k(\mathbf{h}, z_k) = \frac{1}{2} E\{[I(\mathbf{u} + \mathbf{h}, z_k) - I(\mathbf{u}, z_k)]^2\}$$
$$= \frac{1}{2}[\text{Prob}\{Z(\mathbf{u} + \mathbf{h}, z_k) \leqslant z_k \text{ and } Z(\mathbf{u}, z_k) > z_k\} + \text{Prob}\{Z(\mathbf{u} + \mathbf{h}, z_k) > z_k \text{ and } Z(\mathbf{u}, z_k) \leqslant z_k\}]$$

A linear combination of the indicator random variables yields an estimate of the probability for a specific threshold (Journel, 1983). The probabilities describe a local CCDF at any location $\mathbf{u}$ for threshold z_k. In the case of MIK, the simple Kriging algorithm is used:

$$\begin{aligned} F^*(\mathbf{u}, z_k) - F(z_k) &= \sum_{i=1}^{N} \lambda(\mathbf{u}_i, z_k)[I(\mathbf{u}_i, z_k) - F(z_k)] \\ F^*(\mathbf{u}, z_k) &= \sum_{i=1}^{N} \lambda(\mathbf{u}_i, z_k) I(\mathbf{u}_i, z_k) + \left[1 - \lambda(\mathbf{u}_i, z_k)\right] \cdot F(z_k) \end{aligned}, \; k = 1, \dots, K, \; \mathbf{u} \in A$$

where $\lambda(\mathbf{u}_i, z_k)$ are the simple Kriging weights and $F^*(\mathbf{u}, z_k)$ is the probability estimate. To obtain the complete local distribution, the probability estimates are interpolated and tails extrapolated. A standard method is to scale the global experimental CDF into the local CCDF gaps to maintain the samples distribution shape (Deutsch & Journel, 1998). Due to indicators being estimated independently, the estimates do not always satisfy order relations needed to correctly describe a CCDF:

$$\begin{matrix} F^*(\mathbf{u}, z_k) \in [0,1], \\ F^*(\mathbf{u}, z_k) \leqslant F^*(\mathbf{u}, z_{k'}) \end{matrix} \quad \text{for all } k < k';\, k, k' = 1, \ldots, K,\, \mathbf{u} \in A$$

Probabilities must be between 0 and 1 and incremental. The classic methodology to fix order relation violations is to reset the values to $F^*(\mathbf{u}, z_k) = F^*(\mathbf{u}, z_{k'})$ upwards and downwards and average them. Values outside below zero and above one are reset to zero and one respectively (Deutsch & Journel, 1998) (Figure 2.2). Alternatively, the CCDF values can be estimated by least squares with constraints (Sullivan, 1984). Attempts have been developed to account for order relations issues. (1) Using the median threshold indicator variogram to estimate each indicator reduces the number of order relation issues (Deutsch & Journel, 1998). (2) Negative weights are one of the main sources of order relation issues; constraints can be included in the Kriging system to force only nonnegative weights (Barnes & Johnson, 1984; Rao & Journel, 1997). (3) By considering indicators as part of a whole, that is, they have to be nonnegative and have to sum to 1, compositional data techniques can be used to get a consistent CCDF without order relation issues (Tolosana-Delgado et al., 2019, 2008). (4) Another method is to solve a Kriging system with inequality constraints using quadratic programming. (Soltani-Mohammadi & Tercan, 2012).

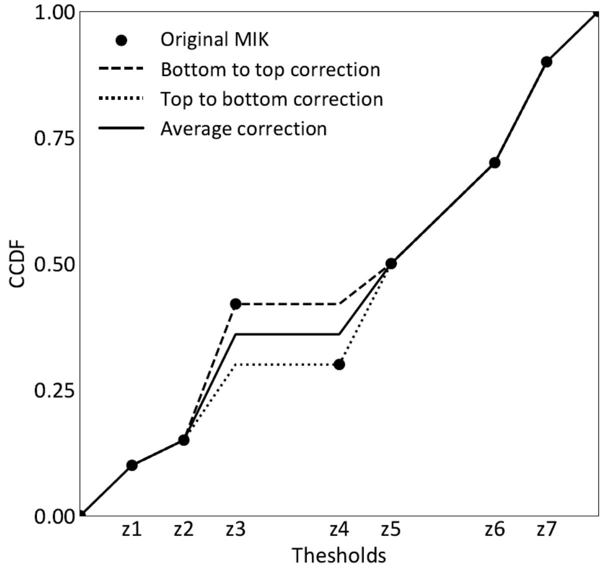


Figure 2.2: Order relation issue corrected by averaging upward and downward corrections

2.4.3 Compositional data with Gaussian kernel (CDIK)

Compositional data with Gaussian kernel is reviewed from the work of Hadavand and Deutsch (2021). As seen before, since the indicators are estimated using Kriging in MIK, the local distributions will almost always have order relation problems. Considering indicators assigned from bins $j(\mathbf{u}, z_k)$ as compositional data, that is, they must sum to one and be nonnegative, is an alternative approach to avoid order relation problems:

$$\sum_{k=1}^{K+1} j(\mathbf{u}, z_k) = 1, \;\; j(\mathbf{u}, z_k) \geqslant 0 \;\; , \mathbf{u} \in A$$

Instead of assigning the indicators as 1 if values are below a threshold and 0 otherwise, an indicator value of 1 is assigned if the value falls into a bin. Indicators are considered probabilities to be in a certain bin. To preserve the compositional data relations, that is, positiveness and summation to 1, the compositional data approach uses the relative magnitude between variables by transforming the variable values into ratios. One common transformation is the additive log-ratio transformation (Aitchison, Barceló-Vidal, Martín-Fernández, & Pawlowsky-Glahn, 2000):

$$j'(\mathbf{u}, z_k) = log\left(\frac{j(\mathbf{u}, z_k)}{j(\mathbf{u}, z_D)}\right), \quad k = 1, \dots, K+1$$

$$j(\mathbf{u}, z_k) = \frac{exp(j'(\mathbf{u}, z_k))}{\sum_{k=1}^{K+1} exp(j'(\mathbf{u}, z_k)) + 1}, \quad k = 1, \dots, K+1$$

$j(\mathbf{u}, z_k)$ are the original variables, $j'(\mathbf{u}, z_k)$ are the transformed variables and K+1 is the number of bins related to K number of thresholds. The first equation is the forward transformation, the transformed values $j'(\mathbf{u}, z_k)$ are then used for geostatistical analysis. The result can be back-transformed to original values. The transformation requires dividing the variables $j(\mathbf{u}, z_k)$ by one of them $j(\mathbf{u}, z_D)$. To be consistent the dividing variable has to be the same for each transformation. Since indicators have only values of 0 and 1, which can make the division undetermined, the indicator values are replaced by values from a Gaussian kernel with a short range. As the Gaussian kernel depends on the absolute difference between the values and the center of each bin, the new values will be close to 1 in the corresponding bin and close to 0 for the rest:

$$GK(z(\mathbf{u}_i), r_k) = exp(-(\alpha|r_k - z(\mathbf{u}_i)|)^2) \;\; , k = 1, \dots, K+1$$

where α is the range of the Gaussian kernel and r_k is the center value of the corresponding bin. Then, the values are normalized to sum 1. Similar to a Gaussian distribution, more values will fall in the center bin, making this bin top have higher Gaussian kernel values. Then, the center bin is used for the log-ratio transform. For the data to behave well and avoid extreme values that naturally occur after the Gaussian kernel and the log-ratio transformation, the data is transformed to normal scores. To get the distribution, the normal score values are interpolated and then back-transformed and log-ratio back-transformed. The values obtained will correctly describe a conditional probability density function (CPDF). A CCDF can be obtained from the CPDF.

2.4.4 Simplicial indicator Kriging SIK

The methodology is presented by Tolosana-Delgado et al. (2008). As in the compositional data with a Gaussian kernel approach, indicators $j(\mathbf{u}_i, z_k)$ are assigned from bins and estimated at any location $j^*(\mathbf{u}, z_k)$. Then, the following formula is used to get the final CPDF values:

$$p^*(\mathbf{u}, z_k) = \mathscr{C}(exp(\beta \cdot j^*(\mathbf{u}, z_k)) \ , k = 1, \dots, K+1, \mathbf{u} \in A$$

where $\beta = log((1-b)((K+1)-1)/b)$, b is a chosen value that represents the probability of error in the determination of bins and $\mathscr{C}$ is the normalization function. The equation is a simplification of using indicators as vector coordinates on an orthonormal basis of the simplex. Usually, this method gives better results if the indicators values are estimated using co-Kriging, simple Kriging can also be used but with less accuracy.

2.5 RBF interpolation

In a scattered data problem, a weighted linear combination of functions can be used to infer values at unknown locations. The weighted linear combination of functions is assumed to reproduce the sample values at their respective locations or, in other words, the function must be fitted to the samples. After getting the weights, the use of the function for getting location values is called interpolation. Samples are not exactly taken on a grid; thus this process is called scattered data interpolation (Fasshauer, 2007).

2.5.1 Interpolation

RBF interpolation is a widely used interpolation method. To get an interpolated value at a unknown location $\mathbf{u}$ from a set of data $f(\mathbf{u}_i)$, $i = 1, \dots, N$, a weighted linear combination of basis functions B_i is considered (Fasshauer, 2007):

$$s(\mathbf{u}) = \sum_{i=1}^{N} \lambda_i B(\mathbf{u}_i)$$

Where s($\mathbf{u}$) is the interpolated value, and λ_i are the weights. The interpolator is assumed to give back the sample values at their respective locations $s(\mathbf{u}_i) = f(\mathbf{u}_i)$ leading to a system of linear equations:

$$\begin{bmatrix} B_1(\mathbf{u}_1) & B_2(\mathbf{u}_1) & \cdots & B_N(\mathbf{u}_1) \\ B_1(\mathbf{u}_2) & B_2(\mathbf{u}_2) & \cdots & B_N(\mathbf{u}_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(\mathbf{u}_N) & B_2(\mathbf{u}_N) & \cdots & B_N(\mathbf{u}_N) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} f(\mathbf{u}_1) \\ f(\mathbf{u}_2) \\ \vdots \\ f(\mathbf{u}_N) \end{bmatrix}$$

A univariate function of the distance between the unknown location and the data can be used $\phi(r(\mathbf{u}, \mathbf{u}_i))$ as the basis function. The distance function can be either the Euclidian distance:

$$r(\mathbf{u}_i, \mathbf{u}_j) = \sqrt{(ux_i - ux_j)^2 + (uy_i - uy_j)^2 + (uz_i - uz_j)^2}$$

or the anisotropic distance to include different directions of continuity into the interpolation:

$$r(\mathbf{u}_i, \mathbf{u}_j) = \sqrt{\left(\frac{ux_i - ux_j}{ax}\right)^2 + \left(\frac{uy_i - uy_j}{ay}\right)^2 + \left(\frac{uz_i - uz_j}{az}\right)^2}$$

Where ux_i, uy_i and uz_i are the coordinates of location $\mathbf{u}_i$ and ax, ay, and az are the ranges of anisotropy. To include the directions of anisotropy, the coordinates are rotated in the directions of continuity (Rossi & Deutsch, 2013). The basis functions of the form $\phi(r(\mathbf{u}, \mathbf{u}_i))$ are called radial basis functions since they give the same value for a distance r independently of the direction, hence, they are radially symmetric with respect a center $\mathbf{u}$ (Fasshauer, 2007). Some examples of RBFs can be seen in the table below (Table 2.2):

Table 2.2: Positive definite RBFs (Fasshauer, 2007)

Name	Equation
Gaussian	$\phi(r) = e^{-\epsilon^2 r^2}$
Spherical	$\phi(r) = 1.5\epsilon r - 0.5(\epsilon r)^3$
Exponential	$\phi(r) = e^{-3r/\epsilon}$
Multiquadratic	$\phi(r) = \sqrt{1 + (\epsilon r)^2}$
Linear	$\phi(r) = r$

By replacing the RBF into the basis functions system, it changes to:

$$\begin{bmatrix} \phi(r_{1,1}) & \phi(r_{1,2}) & \cdots & \phi(r_{1,N}) \\ \phi(r_{2,1}) & \phi(r_{2,2}) & \cdots & \phi(r_{2,N}) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(r_{N,1}) & \phi(r_{N,2}) & \cdots & \phi(r_{N,N}) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} f(\mathbf{u}_1) \\ f(\mathbf{u}_2) \\ \vdots \\ f(\mathbf{u}_N) \end{bmatrix}$$

After solving the system and getting the weights the values can be interpolated by:

$$s(\mathbf{u}) = \sum_{i=1}^{N} \lambda_i \phi(r(\mathbf{u}_, \mathbf{u}_i))$$

Notice how the RBF system is the same as the simple Kriging dual form with a 0 mean (Stewart et al., 2014). RBF does not need the use of covariances for interpolation, although, since the covariances are modeled with positive definite functions, they are suitable to use in RBF interpolation. As dual Kriging, RBF interpolation uses all samples to calculate the weights. This is advantageous for computational efficiency if the data do not exceed 10,000 samples.

2.5.2 Inequality constraints

A benefit of RBF formulations is the possibility of adding inequality constraints to the interpolations. To include inequality constraints, the quadratic form of the interpolator is minimized adding all the constraints (Hillier et al., 2014), leaving a quadratic optimization problem to calculate the weights. Consider N number of samples, M number of locations with inequality constraints and $L = N + M$:

$$\begin{aligned} &\text{minimize } \|s\|^2_{\mathcal{H}} = \lambda^T H \lambda \\ &\text{subject to } \mathcal{A}\lambda = F, \; B\lambda \leqslant q \end{aligned}$$

H is LxL matrix with $\phi(r_{i,j})$ for $i, j = 1, \ldots, L$ where L is the number of locations of the samples plus the inequality constraints locations, $\mathcal{A}$ is NxL matrix with $\phi(r_{i,j})$ for $i = 1, \ldots, N$, $j = 1, \ldots, L$, F is the solution vector of length N with $f(\mathbf{u}_i)$ for $i = 1, \ldots, N$, B is a MxL matrix with $\phi(r_{i,j})$ for $i = N + 1, \ldots, L$ and q is the inequalities vector with length M.

If the original estimate is less than the inequality value, the RBF interpolation with inequality constraints (RBFC) will force the estimate to be equal to the inequality value at that location. If the original estimate satisfies the inequality, the RBFC will maintain the original estimate (Figure 2.3):

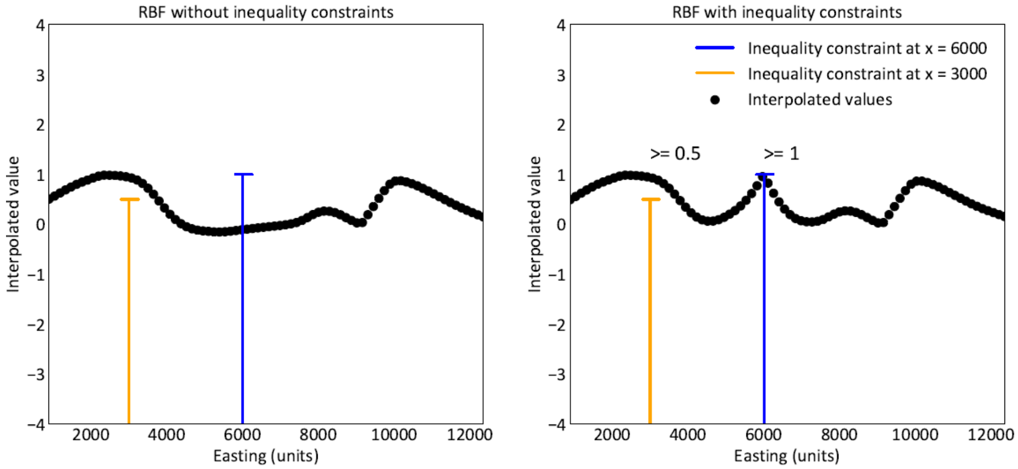


Figure 2.3: Example of RBF with and without constraints along the x-axis

In the example, inequality constraints of $>=$ 1 at location $x = 6000$ and $>=$ 0.5 at location $x = 3000$ are added to the RBF estimation. Notice how the estimates only change where the inequality constraint is not satisfied.

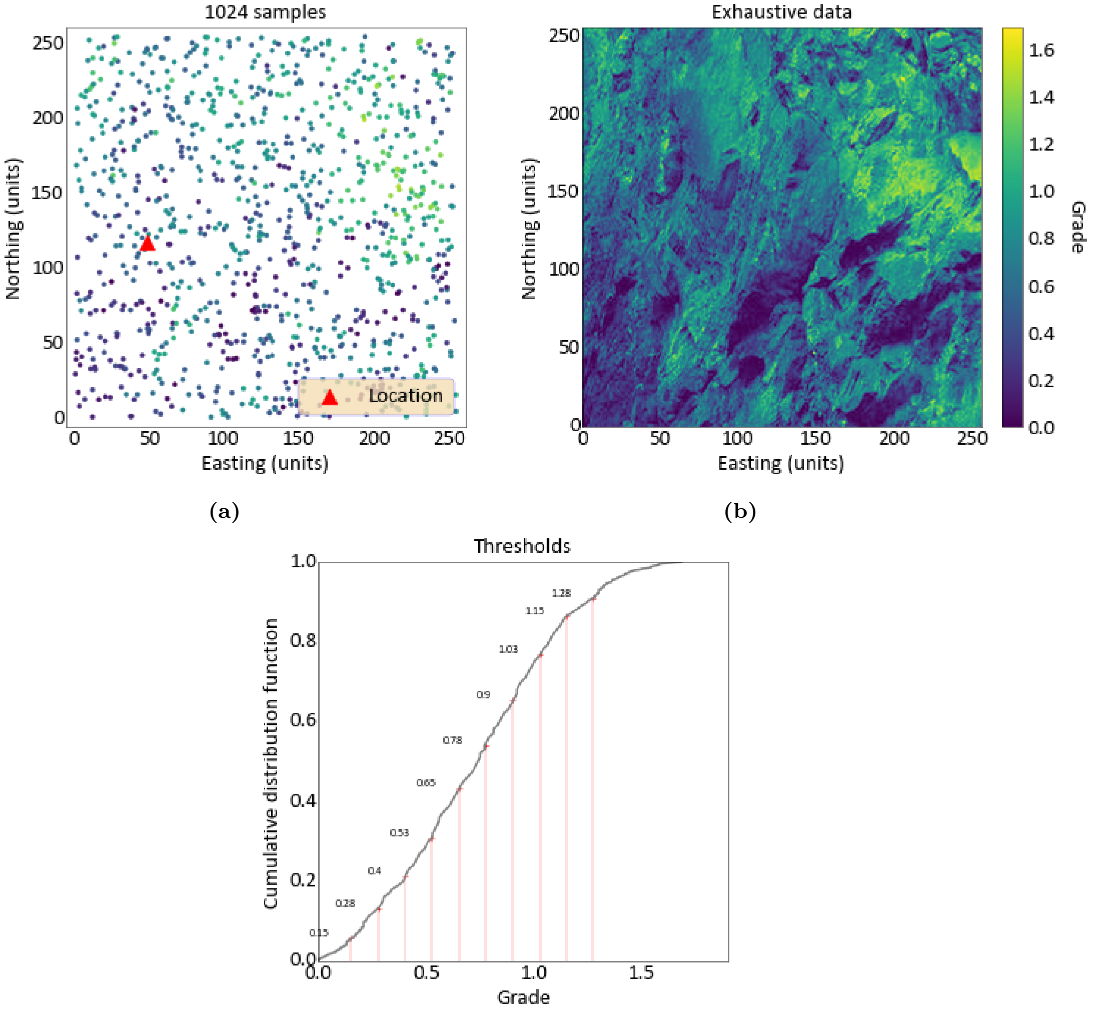
Chapter 3

Order relation issues and negative weights

Order relation issues are a drawback of MIK. Multiple attempts have been made to correct, avoid or eliminate order relations issues. Computing the estimation of all indicators using the median indicator variogram can reduce the number of order relations issues (Deutsch & Journel, 1998). Constraints have been added to the Kriging system to not allow negative weights (Barnes & Johnson, 1984; Rao & Journel, 1997). Compositional data techniques have been used to avoid order relation issues (Tolosana-Delgado et al., 2019, 2008). Other methods involve using quadratic programming to correct the CCDF estimation using constraints in the Kriging system (Soltani-Mohammadi & Tercan, 2012). Although the main reasons for order relations issues described in the literature are the lack of variogram consistency, lack of samples and negative weights, there is no study that goes deep into how all these reasons work together to produce order relation issues. In this chapter, several MIK tests are performed to get insight into how order relations are produced and which factors are the most important to take into account when estimating local distributions. The tests consider estimation at a single location. The distribution of the number of order relations issues at each indicator is investigated. A comparison with median MIK and the influence of inconsistent indicator variograms are also investigated.

3.1 Single location MIK

Order relations issues occur due to different factors related to the use of Kriging to estimate distribution probabilities. These factors include negative weights, different variograms, the spatial distribution of samples, and the prior mean if simple Kriging is used. To have a better understanding of how these factors influence order relations, MIK is performed at a single location conditional to a set of samples (Figure 3.1a). The samples are obtained from exhaustive data from the CCG data validation project (Figure 3.1b) (Mokdad et al., 2022). The dataset is chosen due to the stability of its indicator variograms. Ten thresholds are chosen to discretize the distribution of the samples (Figure 3.1c). The thresholds are selected so the number of samples between each threshold is more than 4% of the total (Table 3.1). Variograms are modeled for all indicators (Figure 3.2). Each indicator is estimated using simple Kriging with a search radius of 100 units and a maximum of 40 samples. The CCDF obtained at the chosen location has a decreasing order relation issue between indicators 3 and 4 (Figure 3.3). Figure 3.4 shows the changes in the weights and indicator values from indicators 3 and 4 on the samples used for the estimation of the location. Notice how the weights do not change very much since the variograms for those indicators are similar. In the case of indicator values, there are three samples (red circles) that change values and two of them present negative weights (Table 3.2). The negative weights get lower for indicator 4, and due to the change of indicator value from 0 to 1, the estimated value gets smaller than the estimate for indicator 3. Even though there is one sample with a positive weight that changes from 0 to 1, its weight is not positive enough to counter the negative weights and avoid the order relation issue. The decreasing values of indicators 3 and 4 estimates are due to a combination of negative weights and changes in the value of indicators. To explore how much the weights and the change of indicators influence the order relation issue, the experiment is carried out using only the variograms of indicator 3 for both estimates. Then, the same experiment is performed but using the same indicator values with different variograms.



(a)

(b)

(c)

Figure 3.1: (a) 1024 samples and the location estimated (b) exhaustive data (c) CDF with thresholds

Table 3.1: Sample percentage per class

Indicators min - 1 sample %	5.8
Indicators 1-2 sample %	7.4
Indicators 2-3 sample %	8.0
Indicators 3-4 sample %	9.8
Indicators 4-5 sample %	12.3
Indicators 5-6 sample %	10.8
Indicators 6-7 sample %	11.2
Indicators 7-8 sample %	11.6
Indicators 8-9 max sample %	9.4
Indicators 9-10 max sample %	4.6
Indicators 10 max sample %	9.1

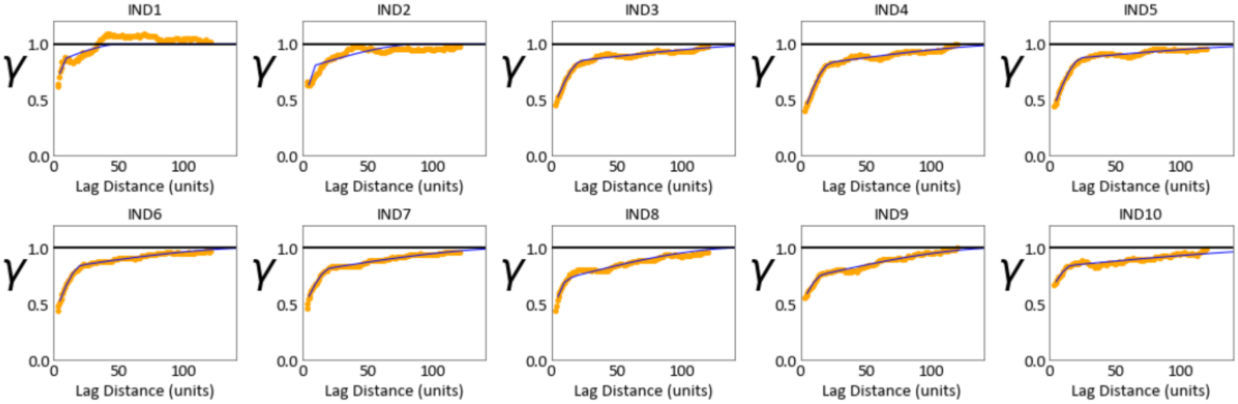


Figure 3.2: Modeled variograms for indicators

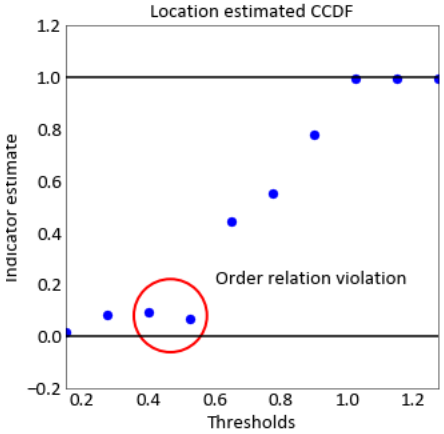


Figure 3.3: Probability estimates for a single location

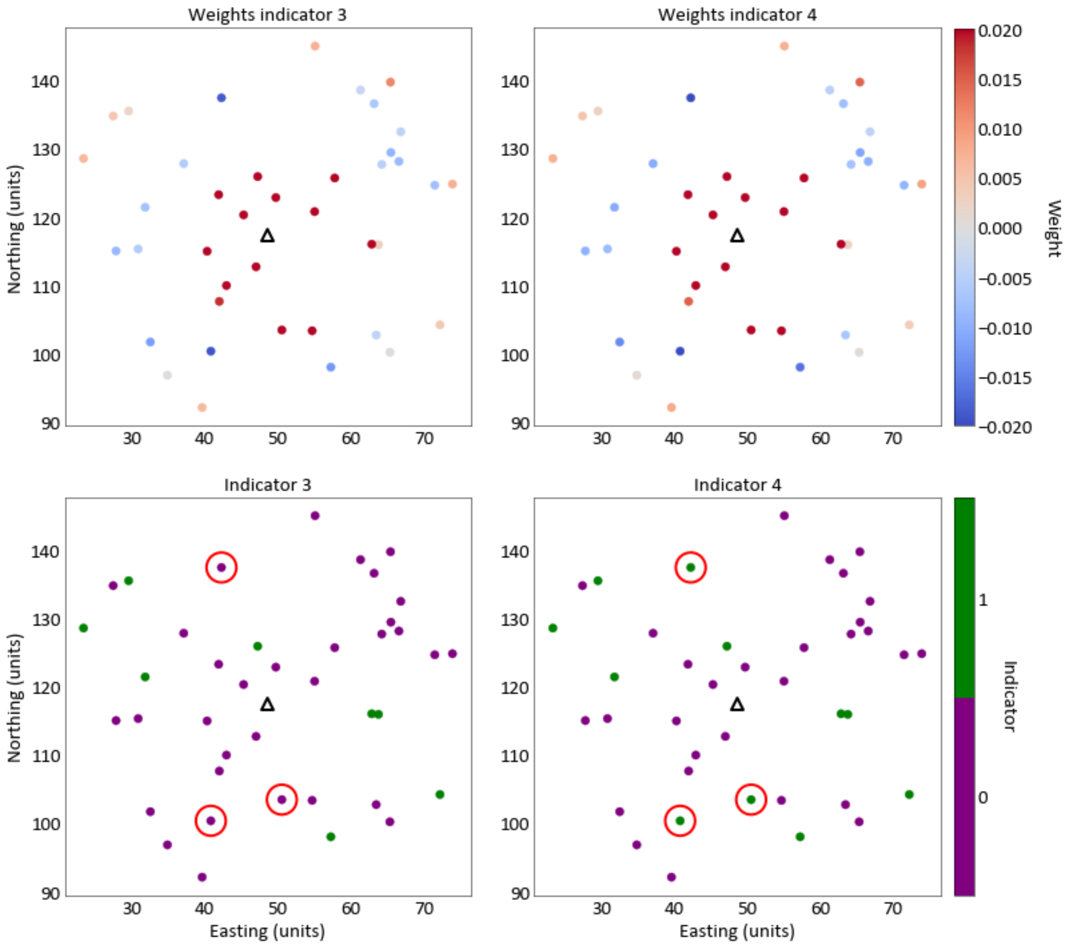


Figure 3.4: Samples (circles) weights and values for the estimation of one location (triangle)

Table 3.2: Samples with different indicator values

X	Y	Ind3	Ind4	Wt3	Wt4
50.56	103.41	0	1	0.031	0.029
40.84	100.31	0	1	-0.019	-0.026
42.29	137.41	0	1	-0.019	-0.023

Table 3.3: Samples with different indicator values and same variograms

X	Y	Estimate 3	Estimate 4	Ind3	Ind4	Wt3	Wt4
50.56	103.41	0.096	0.093	0	1	0.031	0.031
40.84	100.31	0.096	0.093	0	1	-0.019	-0.019
42.29	137.41	0.096	0.093	0	1	-0.019	-0.019

In the case with the same variograms, the weights for both estimates are equal (Table 3.3). The contribution to the estimate from all the samples remains the same except for the samples where there is a change of indicator value from 0 to 1. As before, from the three samples that change their value, two of them have negative values and contribute to the order relation issue. Notice that the change of indicator value only affects the estimation due to the existence of negative weights. In the case with the same indicator values and different variograms, since the indicator values are the same in each estimation, only the weights for samples with a value of 1 are analyzed (Table 3.4). The summation of weights for samples with values of 1 is smaller for indicator 4 than for indicator 3. The negative weights get more negative and some of the positive weights get less positive. In this case, the change in the variograms from indicator 3 to 4 produces smaller weights which relates to the order relation issue.

Indicators 2 and 5, which do not have order relation issues, are also reviewed (Figure 3.5). The variograms change considerably from indicators 2 to 3 but remain similar on indicators 3, 4, and 5. The variogram for indicator 2 produces only positive weights. For the rest of the indicators, the variograms produce negative values between distances of around 15 units and 26 units to the estimated location, which is close to the range distance for the second structure of the variograms with around 80% of the sill contribution. The closest samples are receiving the highest weights, leaving further samples with negative weights. This is the screening effect of Kriging. For indicators 2 to 3 the indicator values that change correspond to samples close to the estimated location, therefore, they receive a high positive weight and the estimate of indicator 3 is higher than the estimate of indicator 2. Something similar happens from indicator 4 to indicator 5, where indicator values change close to the estimation location. Only from indicators 3 to 4, indicator values change at locations where the negative weights are the majority. In that sense, the order relation issue is produced by the combination of the change in indicator values, sample spatial distribution and negative weights. Notice that, if the indicators values do not change considerably from one indicator to the other, but the weights change considerably at those locations, even without negative weights there is a possibility for the estimates to produce an order relation issue.

Table 3.4: Samples with indicator value 1 and different variograms

X	**Y**	**Estimate 3**	**Estimate 4**	**Ind3**	**Ind4**	**Wt3**	**Wt4**
63.79	115.9	0.0962	0.0814	1	1	0.0033	0.0026
72.21	104.16	0.0962	0.0814	1	1	0.0043	0.0037
31.85	121.38	0.0962	0.0814	1	1	-0.0074	-0.0105
57.27	97.97	0.0962	0.0814	1	1	-0.0126	-0.0163
62.87	115.98	0.0962	0.0814	1	1	0.0385	0.0390
47.25	125.88	0.0962	0.0814	1	1	0.0568	0.0488
23.41	128.53	0.0962	0.0814	1	1	0.0067	0.0074
29.58	135.47	0.0962	0.0814	1	1	0.0022	0.0031
					Sum	0.0918	0.0777

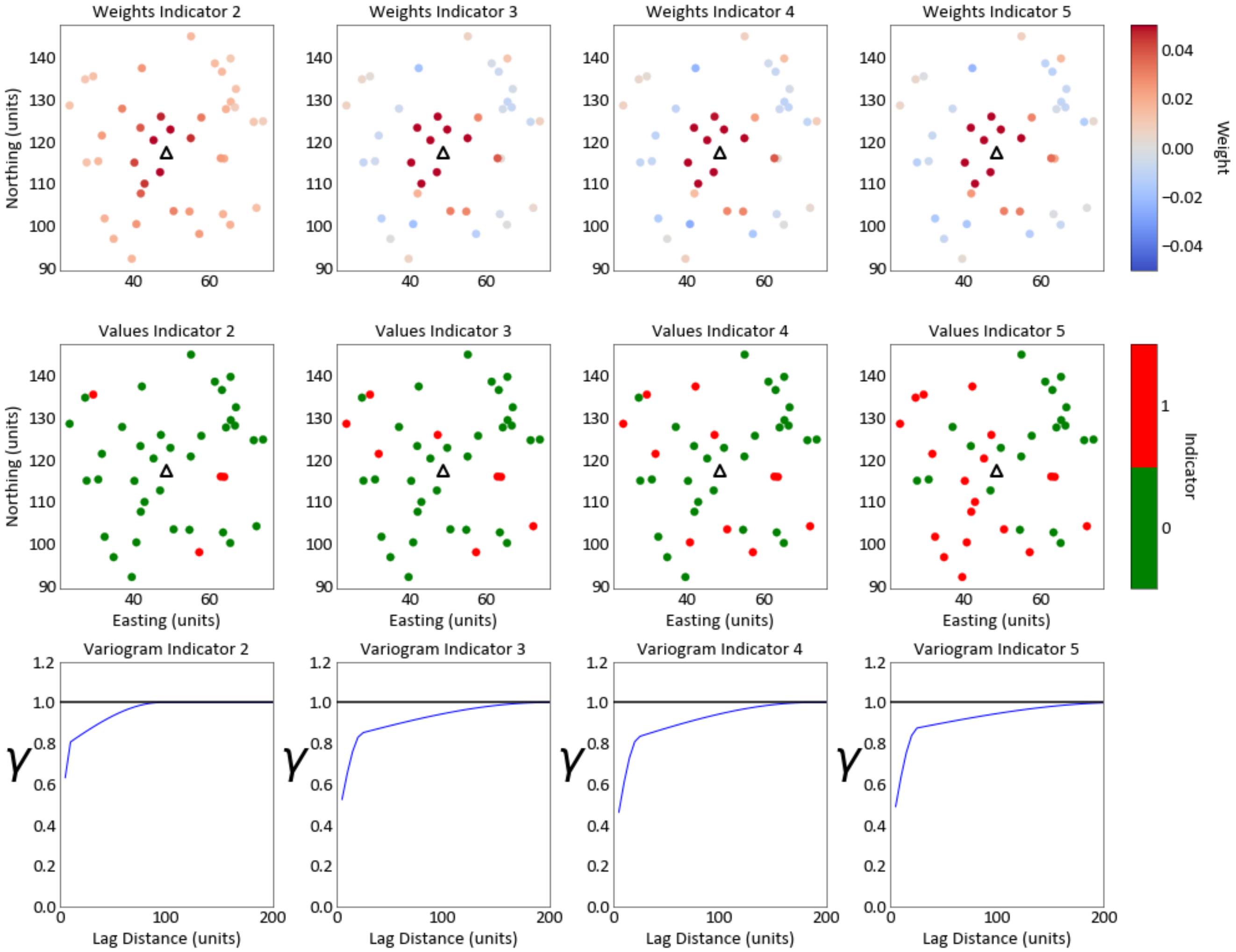


Figure 3.5: Comparison between indicators 2, 3, 4, and 5. Weights, indicators values and variograms are shown.

3.2 Order relations distributions and median MIK

Negative weights are produced by the screen effect of Kriging, which, in combination with the change of indicator values from 0 to 1, leads to order relation issues. As seen in the previous section, weights can get negative or more negative from one indicator to another due to differences between indicator variograms. To highlight this effect, a test is performed using a median variogram. Since all indicators have the same variogram and weights, order relations issues are only affected by the change of indicator values and only if negative weights are involved. It is impossible to produce an order relation issue by changing the values if all weights are the same and positive. The test is performed using the same 1024 samples, thresholds, variograms and Kriging parameters used in the previous section. The 5th indicator variogram is used as the median variogram. Both classic MIK and median MIK order relations are plotted for each indicator number and the type of weights related to the order relations is highlighted (Figure 3.6).

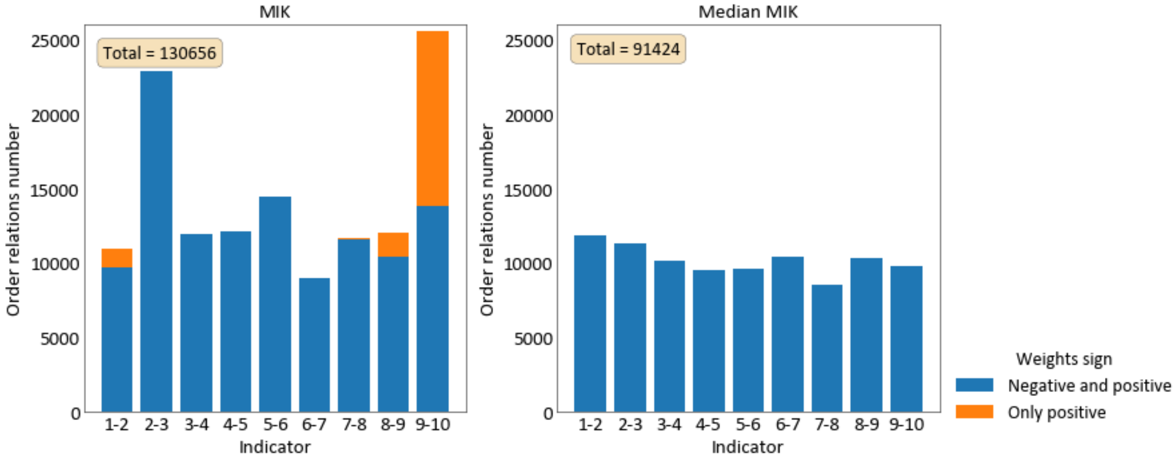


Figure 3.6: Distribution of the number of order relations per indicator

Most order relations are produced between indicators 2 and 3 and between indicators 9 and 10. Indicators from 3 and 8, show a similar order relation number. This is due to the difference between the variograms. Dissimilar variograms have a higher difference between weights, which increases the number of order relations issues. In this case, variograms between indicators 2 and 3 and indicators 9 and 10 are the most dissimilar. In the previous single location example, indicators 2 and 3 did not produce an order relation issue due to the change of indicator values for that specific location. The results also show order relation issues produced by positive weights-only samples. The positive weights-only order relation issues happen between indicators that have a small change in indicator values, which is more likely to happen at high and low indicators. Indeed, positive weights-only order relations issues occur for indicators 1-2, 8-9, and 9-10. In the case of the median variogram MIK, all order relations issues involving only positive weights are erased, and the negative weight order relations are reduced in all indicators. Since the weights are the same for all the indicators, an order relation issue occurs only where an indicator value changes from 0 to 1 on a negative weight sample.

MIK estimation proceeds by estimating the local distributions and then calculating the mean of those distributions using the programs ik3d.exe and postik.exe, respectively (Deutsch & Journel, 1998). The means obtained by MIK are compared with the exhaustive data (Figure 3.7). The comparison between MIK and median MIK shows that both methodologies are globally similar, they present the same mean of 0.723, and a slightly different variance of 0.257 for MIK and 0.268 for median MIK. The mean squared error presents a little difference with 0,249 for MIK and 0.251 for median MIK. The median MIK is less biased with respect to the diagonal with a slope of regression of 1.018 versus 1.070 of MIK. The differences in global statistics are not significant. The estimation maps give a better insight into the local estimation. The median MIK estimation is smoother than the classic MIK estimation. The classic MIK presents a nugget effect and shorter range in low grades and high grades. MIK is able to capture the spatial distribution for different thresholds better than median MIK.

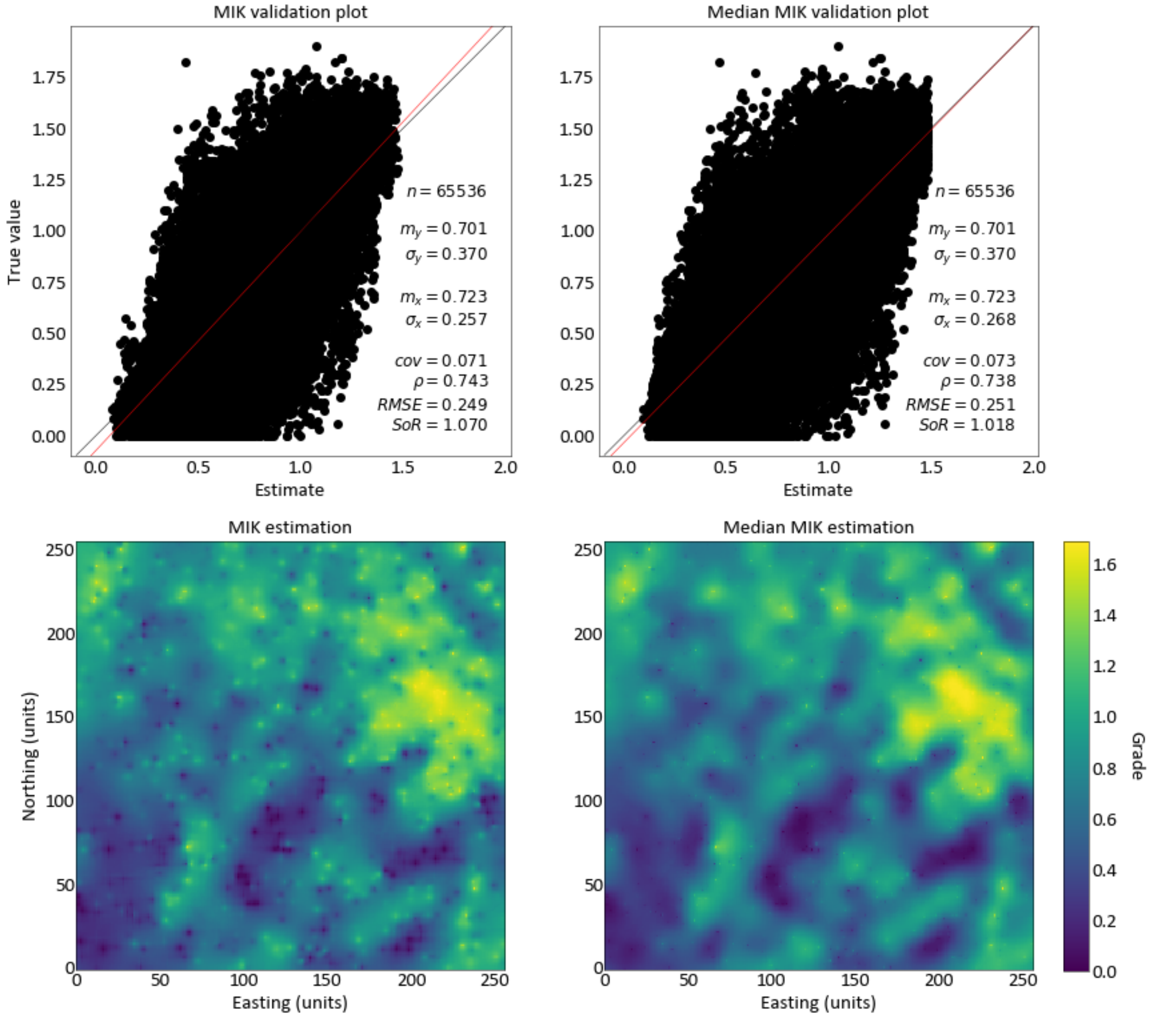


Figure 3.7: Validation and map plots of MIK and median MIK

3.3 Order relations and number of samples

The last section showed that the more different two indicator variograms are, the more likely they are to produce order relation issues. Less stable variograms from fewer samples should produce dissimilar variograms and more order relation issues. A sensitivity analysis is performed to check how the instability of indicator variograms affects the number of order relations. A set of samples from the exhaustive data with 256 samples is used to model variograms (Figure 3.8). Then, MIK is performed using the original 1024 samples with both the original variograms (Figure 3.9) and the variograms modeled from 256 samples (Figure 3.10). As expected, less stable variograms estimation produces more order relations (Figure 3.11). The number of locations where order relation issues happen also increases. However, the distribution of order relations is different than the original MIK. In the estimation with unstable variograms, there are more order relations between indicators 9 and 10 than in the previous original estimate. Also, there is a reduction in order relations between indicators 2 and 3. In both cases, the differences are related to the variograms. In the case of indicators 2 and 3, indicator variograms for 256 samples are less continuous than for 1024 samples which may produce fewer negative weights and along with that fewer order relation issues. For indicators 9 and 10, the variograms seem more dissimilar and more continuous which may produce more negative weights and more order relation issues. In the central indicators, the unstable variograms have less continuity for short ranges but more continuity at long ranges, they are also quite different. That could produce different weights and stronger negative weights in the short range of the variograms, producing more order relation issues.

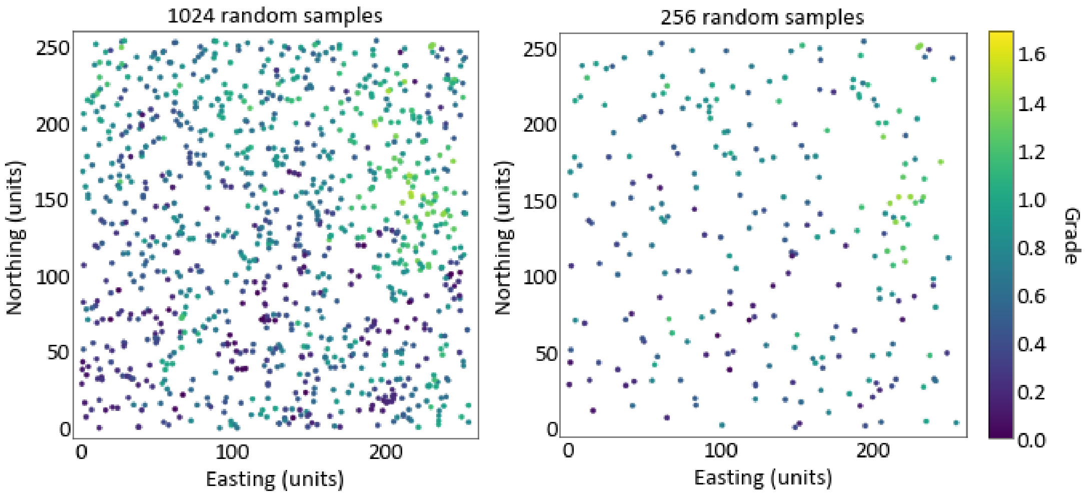


Figure 3.8: 1024 and 256 random samples for variograms destabilization

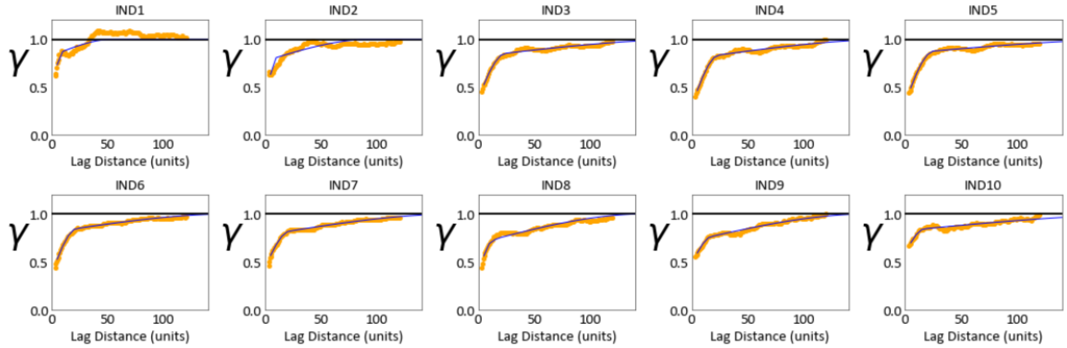


Figure 3.9: Indicator variograms for 1024 samples

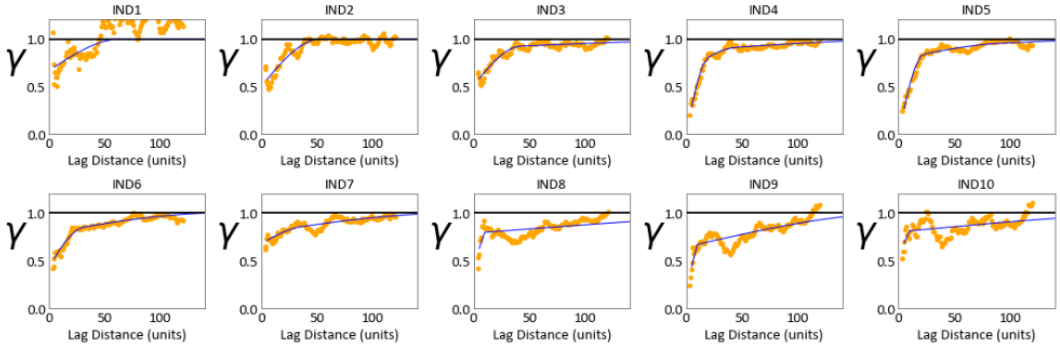


Figure 3.10: Indicator variograms for 256 samples

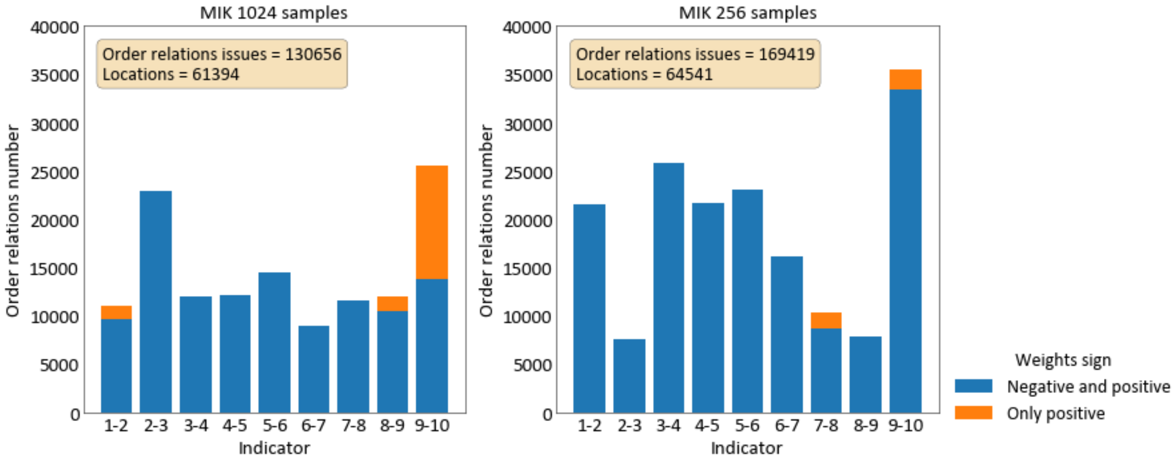


Figure 3.11: MIK order relations issues for 1024 samples at 65536 locations with both sets of variograms

3.4 Summary of order relation issues and negative weights

Negative weights are produced by the screen effect of Kriging. Continuous variograms produce negative weights. Negative weights can produce order relation issues (1) in combination with the change of indicator values from one threshold to the next one, (2) without any change in indicator values, or (3) with the same weights and different indicator values. Also, order relation issues can be produced by samples without negative weights. This usually happens when indicator values of two consecutive thresholds do not change. The median MIK makes all weights for each indicator the same. This reduces the number of order relation issues. Order relation issues with only positive weights are erased in median MIK. More dissimilar variograms mean more order relation issues. MIK and median MIK have an almost equal global performance despite the reduction in order relation issues. Median MIK gives smoother estimates but classic MIK respects the spatial distribution for different grades. Finally, fewer samples will produce more unstable and dissimilar variograms which leads to an increase in order relation issues and locations where this happens. The unstable variograms can also alter significantly the distribution of order relation issues.

Chapter 4

Assessing bivariate consistency of indicator variograms from continuous variables

In Kriging for continuous variables, variograms are modeled from experimental variograms which are calculated by pairing samples separated by a distance **h** (Chiles & Delfiner, 2009; Isaaks & Srivastava, 1989; Rossi & Deutsch, 2013). Indicators are a step function dependent on a continuous variable (Journel, 1983) and are related to each other by the variable. A bivariate distribution can be constructed using the pairs made from the continuous variable (Figure 4.1). The bivariate distribution can be divided by the thresholds to form probabilities. The probability areas are positive and sum to 1. Experimental variograms and proportions of indicators can be calculated from the bivariate distribution. In that sense, modeled indicator variograms should be related to each other by a consistent bivariate distribution. Chapter 3 reviewed how negative weights affect the number of order relation issues in MIK. It explained that one of the main contributors to negative weights and order relation issues are variograms. Indicator variograms are usually calculated from experimental variograms, without considering the bivariate distribution behind them. This chapter explores how the consistency of modeled indicator variograms with a bivariate distribution affects order relation problems and MIK estimations. Past methodologies on how to calculate bivariate probabilities from indicator variograms are reviewed. Novel equations to calculate bivariate probabilities from modeled indicator variograms are derived and tested on a multi-variate Gaussian case. This test includes a sensitivity analysis of sample density. A workflow is presented on how to use the bivariate probabilities for variogram modeling. The workflow is carried out in multiple datasets to explore the effectiveness of the methodology.

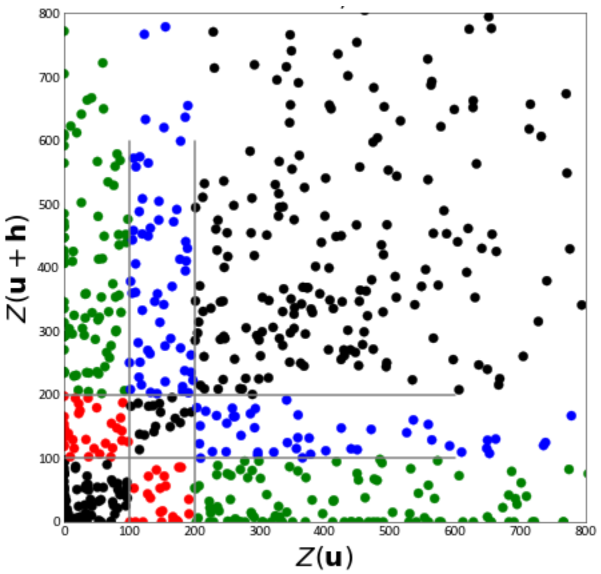


Figure 4.1: Bivariate distribution calculated from pairs at a distance **h**

4.1 Review of calculating bivariate probabilities

The pairs $Z(\mathbf{u}+\mathbf{h})$ and $Z(\mathbf{u})$ describe a bivariate distribution for each lag $\mathbf{h}$ that can be divided into probabilities by thresholds z_k for $k = 1, \dots, K$ with K being the number of thresholds. Since the samples are assumed to be stationary, the probabilities are symmetrical with respect to the diagonal (Figure 4.2). The indicator experimental variograms can be calculated from the bivariate distribution $Z(\mathbf{u}+\mathbf{h})$ and $Z(\mathbf{u})$. As the indicators can only take values of 0 and 1, the pairs making a non-zero contribution to the variograms will be the ones that satisfy the following conditions:

$$\text{or} \begin{array}{c} Z(\mathbf{u}) \leq z_k \text{ and } Z(\mathbf{u}+\mathbf{h}) > z_k \\ Z(\mathbf{u}) > z_k \text{ and } Z(\mathbf{u}+\mathbf{h}) \leq z_k \end{array}, \; k = 1, \dots, K, \; \mathbf{u} \in A$$

Where $\mathbf{u}$ is a location on a space A. One method to obtain direct and cross-variograms is to assume a prior distribution for the samples. For a stationary random variable $Y(\mathbf{u})$ with a multi-Gaussian distribution $G(z)$ with zero mean and unit variance the direct and cross-variograms for thresholds $z_K = G^{-1}(p)$ and $z_{k'} = G^{-1}(p')$ for probabilities p and p' can be calculated by (Journel & Posa, 1990; Xiao, 1985):

$$
\begin{aligned}
2\gamma_{kk'}\left(\mathbf{h}; z_k, z_{k'}\right) =& 2\min\left(p, p'\right) - 2p \cdot p' \\
& - \frac{1}{2\pi}\int_0^{\arcsin C_Y(\mathbf{h})} \exp\left[-\frac{z_k^2 + z_{k'}^2 - 2z_k z_{k'}\sin\theta}{2\cos^2\theta}\right] d\theta \\
& - \frac{1}{2\pi}\int_0^{\arcsin C_Y(-\mathbf{h})} \exp\left[-\frac{z_k^2 + z_{k'}^2 - 2z_k z_{k'}\sin\theta}{2\cos^2\theta}\right] d\theta
\end{aligned}
$$

where $C_Y(\mathbf{h})$ is the covariance for lag $\mathbf{h}$ of the random variable $Y(\mathbf{u})$. In the multi-Gaussian case, the correct variograms can be calculated analytically; therefore, they are consistent with the bivariate Gaussian distribution. If the prior distribution of the samples is unknown, the direct and cross-variograms can be obtained by modeling the experimental variograms. Then, the inverse process of calculating the bivariate probabilities can be performed from the complete matrix of the modeled cross and direct variograms. The probabilities are calculated from the non-centered indicator cross-covariance (Journel, 1983; Journel & Posa, 1990):

$$K_{kk'}(\mathbf{h}; z_k, z_{k'}) = E\{I(\mathbf{u}; z_k) \cdot I(\mathbf{u}+\mathbf{h}; z_{k'})\}$$

The non-centered indicator cross-covariance is related to the cross-variogram by:

$$2\gamma_{kk'}(\mathbf{h}; z_k, z_{k'}) = 2F(\min(z_k, z_k')) - K_{kk'}(\mathbf{h}; z_k, z_{k'}) - K_{k'k}(\mathbf{h}; z_{k'}, z_k)$$

Where $F(z) = E\{I(\mathbf{u}; z)\}$. In practice, sample distributions are not multi-Gaussian, and only the direct variograms are used. One variogram is fitted per threshold (Figure 4.3).

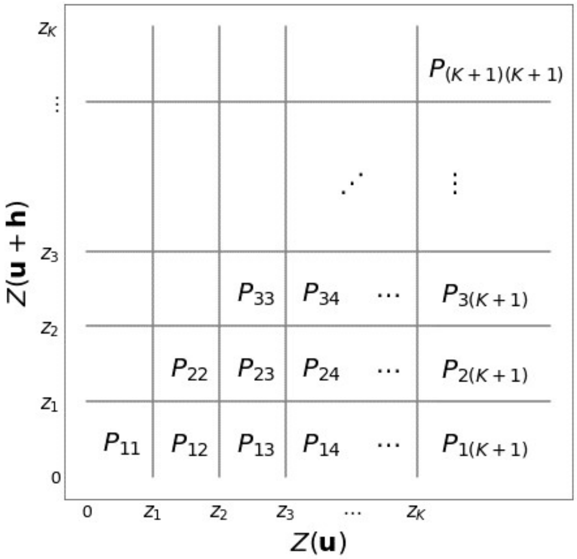

z_K
⋮
z_3
z_2
z_1
0
$Z(\mathbf{u}+\mathbf{h})$
$P_{(K+1)(K+1)}$
⋰
⋮
P_{33} P_{34} ⋯ $P_{3(K+1)}$
P_{22} P_{23} P_{24} ⋯ $P_{2(K+1)}$
P_{11} P_{12} P_{13} P_{14} ⋯ $P_{1(K+1)}$
0 z_1 z_2 z_3 ⋯ z_K
$Z(\mathbf{u})$

Figure 4.2: Probabilities in a bivariate scatter plot for a lag **h**

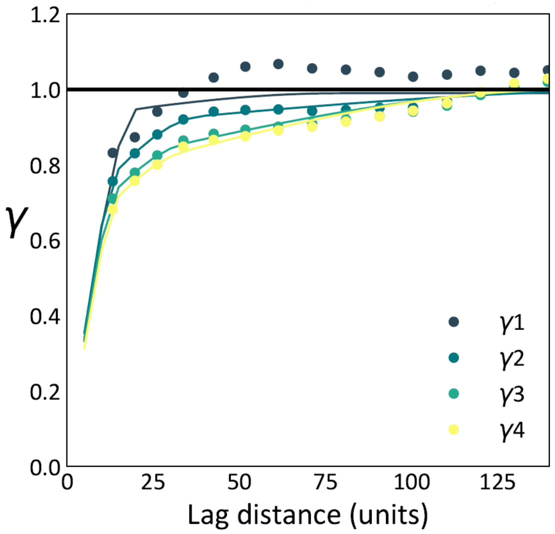


Figure 4.3: Indicator variogram modeling (lines) of experimental indicator variograms (dots) for different thresholds (colors)

Direct variograms are commonly automatically fit to experimental variograms; this only considers the experimental variograms and not the relationship between indicator variograms. This may result in inconsistent direct variograms with a bivariate distribution (Matheron, 1989). To check consistency, probabilities of the internal bivariate distribution can be calculated from the direct variograms and information from the experimental data.

4.2 Derivation of bivariate probabilities equations from indicator variograms

The probabilities for two and three thresholds are derived in the next sections. Then, a generalization for K thresholds is presented.

4.2.1 Case: 2 thresholds

The probabilities in the bivariate distribution at any lag $\mathbf{h}$ for two thresholds are labeled as follows (Figure 4.4).

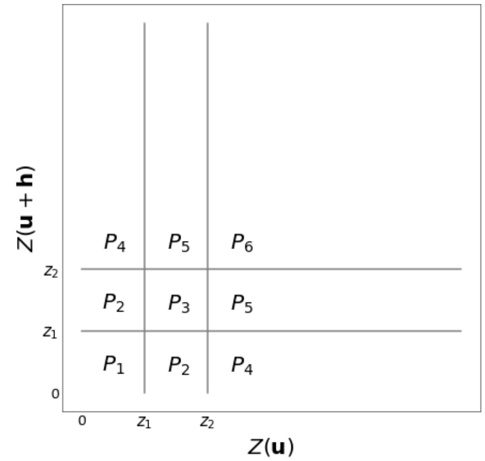


Figure 4.4: Probability labels for 2 thresholds for any lag $\mathbf{h}$

The indicator proportions p_{z_k} can be calculated as a summation of probabilities that are below the corresponding threshold. Similarly, the variograms γ_k can be written as the summation of the probabilities where $I(\mathbf{u}; z_k) \neq I(\mathbf{u}+\mathbf{h}; z_k)$, that is, probability areas where indicators have a different value. Additionally, all the probabilities must sum to 1 and are considered to be symmetric to the diagonal. Then, the equations available using only proportions and modeled variograms are:

$$\begin{aligned}
p_{z_1} &= P_1 + P_2 + P_4 \\
p_{z_2} &= p_{z_1} + P_2 + P_3 + P_5 \\
\gamma_1 &= P_2 + P_4 \\
\gamma_2 &= P_4 + P_5
\end{aligned}$$

$$P_1 + 2P_2 + P_3 + 2P_4 + 2P_5 + P_6 = 1$$

where γ_1 and γ_2 are the modeled indicator variograms values for a certain lag $\mathbf{h}$ for thresholds z_1 and z_2. As there are 5 equations for 6 variables, the system is underdetermined. To solve it, P_3 is assumed and obtained from the samples. By re-arranging the equations, the probabilities are obtained from the indicator proportions, modeled indicator variograms, and P_3:

$$
\begin{aligned}
P_1 &= p_{z_1} - \gamma_1 \\
P_2 &= (p_{z_2} - p_{z_1} + \gamma_1 - \gamma_2 - P_3)/2 \\
P_4 &= (p_{z_1} - p_{z_2} + \gamma_2 + \gamma_1 + P_3)/2 \\
P_5 &= (p_{z_2} - p_{z_1} + \gamma_2 - \gamma_1 - P_3)/2 \\
P_6 &= 1 - p_{z_2} - \gamma_2
\end{aligned}
$$

A similar derivation is considered in the case of three thresholds.

4.2.2 Case: 3 thresholds

The probabilities in the bivariate distribution at any lag **h** for three thresholds are labeled as follows (Figure 4.5):

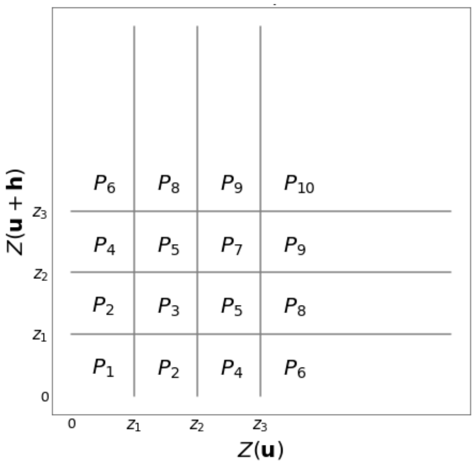


Figure 4.5: Probability labels for 3 thresholds for any lag **h**

The equations for indicator proportions and modeled indicator variograms are calculated as in the case of two thresholds. Then, the equations available are:

$$
\begin{aligned}
p_{z_1} &= P_1 + P_2 + P_4 + P_6 \\
p_{z_2} &= p_{z_1} + P_2 + P_3 + P_5 + P_8 \\
p_{z_3} &= p_{z_2} + P_4 + P_5 + P_7 + P_9 \\
\gamma_1 &= P_2 + P_4 + P_6 \\
\gamma_2 &= P_4 + P_5 + P_6 + P_8 \\
\gamma_3 &= P_6 + P_8 + P_9
\end{aligned}
$$

$$
P_1 + 2P_2 + P_3 + 2P_4 + 2P_5 + 2P_6 + P_7 + 2P_8 + 2P_9 + P10 = 1
$$

Similar to the case of two thresholds, the number of equations is less than the number of variables. Central probabilities P_3, P_5, and P_7 are assumed and calculated from the experimental data. The equations to calculate each probability are:

$$\begin{aligned}
P_1 &= p_{z_1} - \gamma_1 \\
P_2 &= (p_{z_2} - p_{z_1} - P_3 + \gamma_1 - \gamma_2)/2 \\
P_4 &= (\gamma_2 - p_{z_2} - 2P_5 - P_7 + p_{z_3} - \gamma_3)/2 \\
P_6 &= (p_{z_1} + \gamma_1 + P_3 + 2P_5 + P_7 - p_{z_3} + \gamma_3)/2 \\
P_8 &= (p_{z_2} + \gamma_2 - p_{z_1} - \gamma_1 - P_3 - 2P_5)/2 \\
P_9 &= (p_{z_3} - \gamma_3 - p_{z_2} - \gamma_2 - P_7)/2 \\
P_{10} &= 1 - p_{z_3} - \gamma_3
\end{aligned}$$

The procedure for obtaining the probabilities for two and three thresholds can be generalized to define the probabilities for any number of thresholds.

4.2.3 Generalization

The probability labels used are described in Figure 4.2. Let z_k be thresholds with $k = \{1, 2, \cdots, K\}$. $2 \cdot (K+1) - 1$ probabilities at the edges of the symmetric bivariate distribution $Z(\mathbf{u}), Z(\mathbf{u}+\mathbf{h})$ can be obtained using the indicator proportions p_{z_k}, central bivariate probabilities P_{ij} obtained from experimental data, and the modeled indicator variograms γ_k:

$$\begin{aligned}
&P_{11} = p_{z_1} - \gamma_1 \\
&2P_{12} = p_{z_2} - p_{z_1} - \gamma_2 + \gamma_1 - P_{22} \\
&2P_{1i} = p_{z_i} - p_{z_{(i-1)}} - \gamma_i + \gamma_{i-1} - P_{ii} - \sum_{j=2}^{i-1} 2 \cdot P_{ji} \qquad \text{for } i = \{3, \cdots, K\} \\
&2P_{1(K+1)} = p_{z_1} - p_{z_K} + \gamma_1 + \gamma_K + \sum_{j=2}^{K-1} \sum_{m=j+1}^{K} 2 \cdot P_{jm} + \sum_{j=2}^{K} P_{jj} \\
&2P_{i(K+1)} = p_{z_i} - p_{z(i-1)} + \gamma_i - \gamma_{i-1} - P_{ii} - \sum_{j=i+1}^{K} 2 \cdot P_{ij} \qquad \text{for } i = \{2, \cdots, K-1\} \\
&2P_{K(K+1)} = p_{z_K} - p_{z(K-1)} + \gamma_K - \gamma_{K-1} - P_{KK} \\
&P_{(K+1)(K+1)} = 1 - p_{z_K} - \gamma_K
\end{aligned}$$

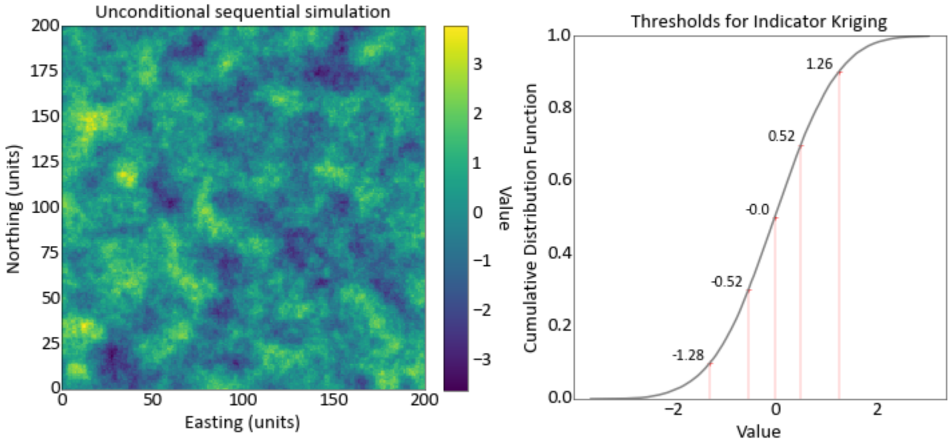
The probabilities obtained from the indicator variogram models should be consistent with a bivariate distribution. The probabilities should sum to one and be non-negative:

$$\begin{aligned}
P_{ij} &\geq 0 \\
\sum_{i=1}^{K+1} \sum_{j=1}^{K+1} P_{ij} &= 1
\end{aligned}$$

Probability P_3 for the case of two thresholds and probabilities P_3, P_5 and P_7 for the case of 3 thresholds are assumed and taken from the experimental probabilities to calculate the rest of the probabilities. Generally, any probability could be assumed to calculate the other ones. The probabilities to be calculated are chosen because they are at the edges of the bivariate distribution. A bivariate distribution shows higher density in the diagonal for lower lags due to closer samples being more correlated. For higher lags, it shows a more spread behavior. Either way, the probabilities at the edges are closer to zero and are more likely to be negative if calculated from modeled variograms than the central probabilities.

4.3 Multi-variate Gaussian example

As shown, indicator variograms of a random variable with a Gaussian distribution can be obtained analytically. The analytically correct variograms can be used to obtain the bivariate distribution probabilities using the equations presented and compare them to the experimental probabilities. To emulate a random variable with a Gaussian distribution, an unconditional numerical simulation is created using sgsim.exe from the GSLIB library (Deutsch & Journel, 1998) (Figure 4.6a). The program simulates values sequentially in a grid with a user-defined spatial correlation. The resulting simulated values have a mean of 0, a variance of 1, and a Gaussian distribution. The spatial correlation is inputted as a variogram; for this case, the variogram is a one-structure omnidirectional spherical variogram with a zero nugget effect, sill of 1, and a range of 20 units. The simulation is run with a 200 units search volume and uses 200 previously simulated data to ensure the correct spatial correlation. The simulation is a 200 x 200 grid with blocks of size 1 unit. The thresholds used are calculated from the proportion values [0.1, 0.3, 0.5, 0.7, 0.9] (Figure 4.6b). The theoretical indicator variograms are calculated using bigaus.exe (Deutsch & Journel, 1998) with the same variogram used for the simulation (Figure 4.7).



(a)

(b)

Figure 4.6: (a) Unconditional Gaussian simulation (b) CDF with thresholds

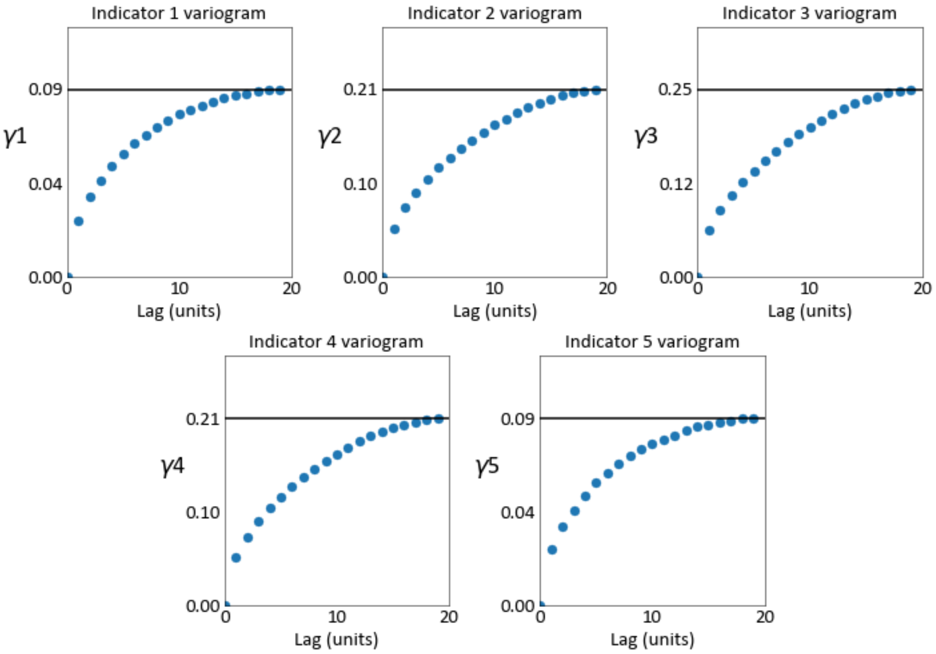


Figure 4.7: Analytically calculated indicator variograms of a Gaussian distribution with omnidirectional anisotropy of 20 units

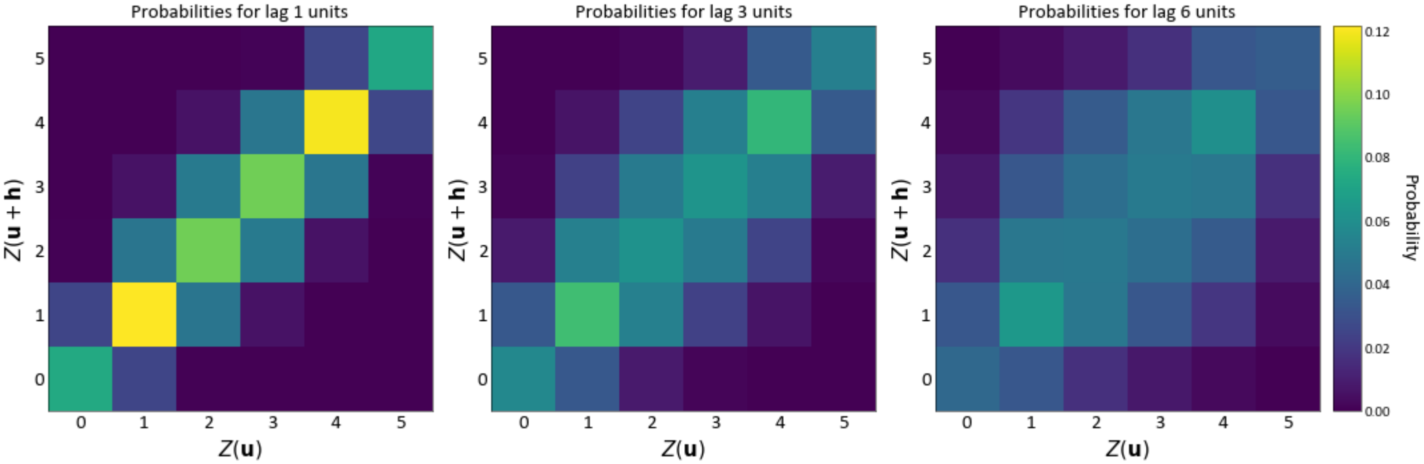


Figure 4.8: Experimental probabilities calculated from the bivariate distribution $Z(\mathbf{u})$, $Z(\mathbf{u}{+}\mathbf{h})$ for different lags $\mathbf{h}$

The intermediate thresholds have more continuity than high and low thresholds which is expected for a Gaussian distribution. A bivariate distribution $Z(\mathbf{u})$, $Z(\mathbf{u}+\mathbf{h})$ can be divided into probability areas by the indicator thresholds (Figure 4.5). To calculate the experimental bivariate probabilities, samples separated by a certain lag $\mathbf{h}$ are paired, then, the probabilities are calculated by dividing the number of pairs that fall in a certain probability area by the total number of pairs. There is one set of probabilities for each lag (Figure 4.8). The probabilities of the bivariate distributions show more correlation at lower lags. The symmetry with respect to the diagonal is due to the stationarity feature of the samples. Once the experimental probabilities are calculated, the equations presented are used to calculate the bivariate probabilities from the analytical variograms. Then, the calculated probabilities are compared with the experimental probabilities. The comparison is made at each lag (Figure 4.9).

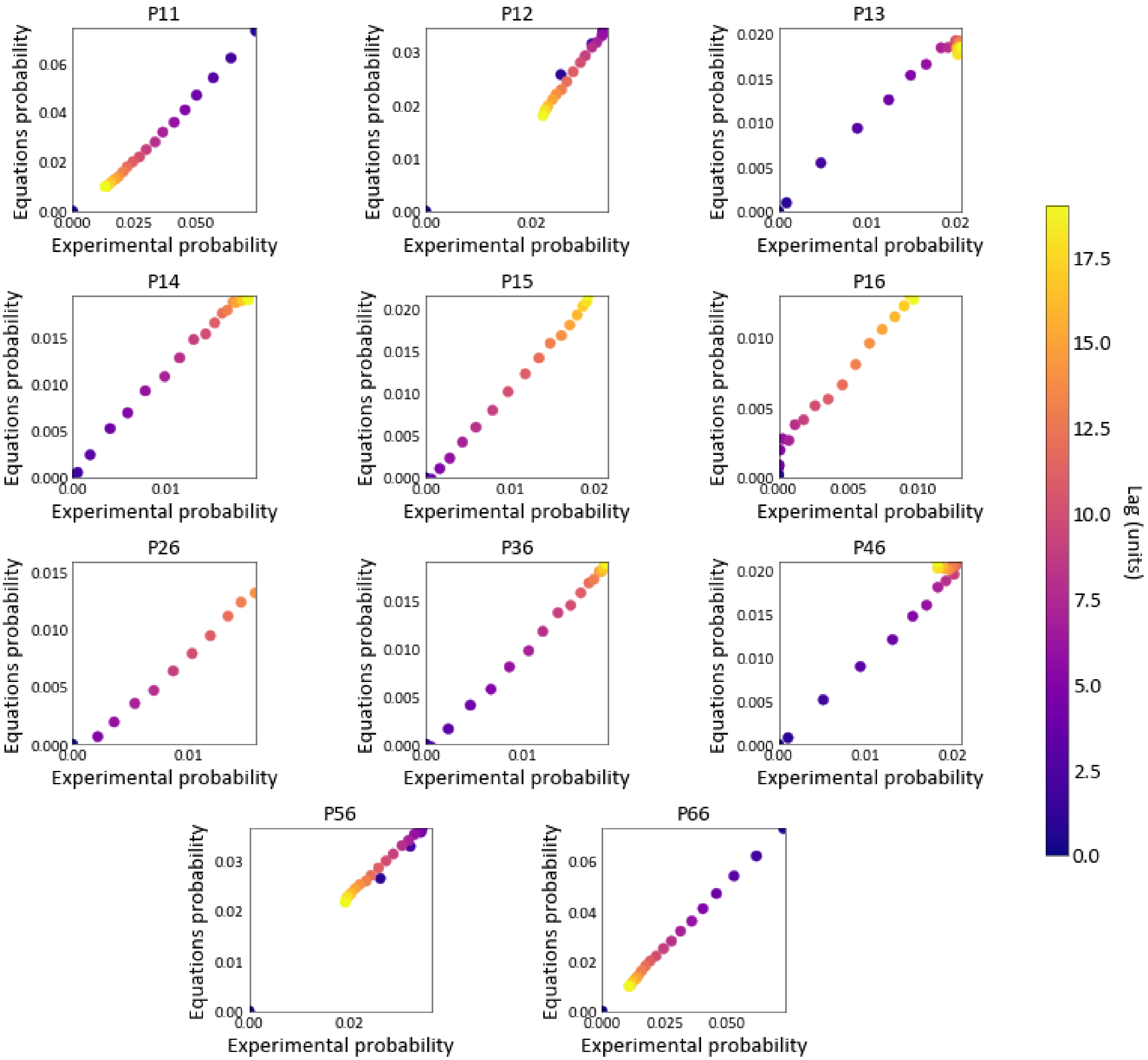


Figure 4.9: Experimental probabilities comparison with equations probabilities at different lags

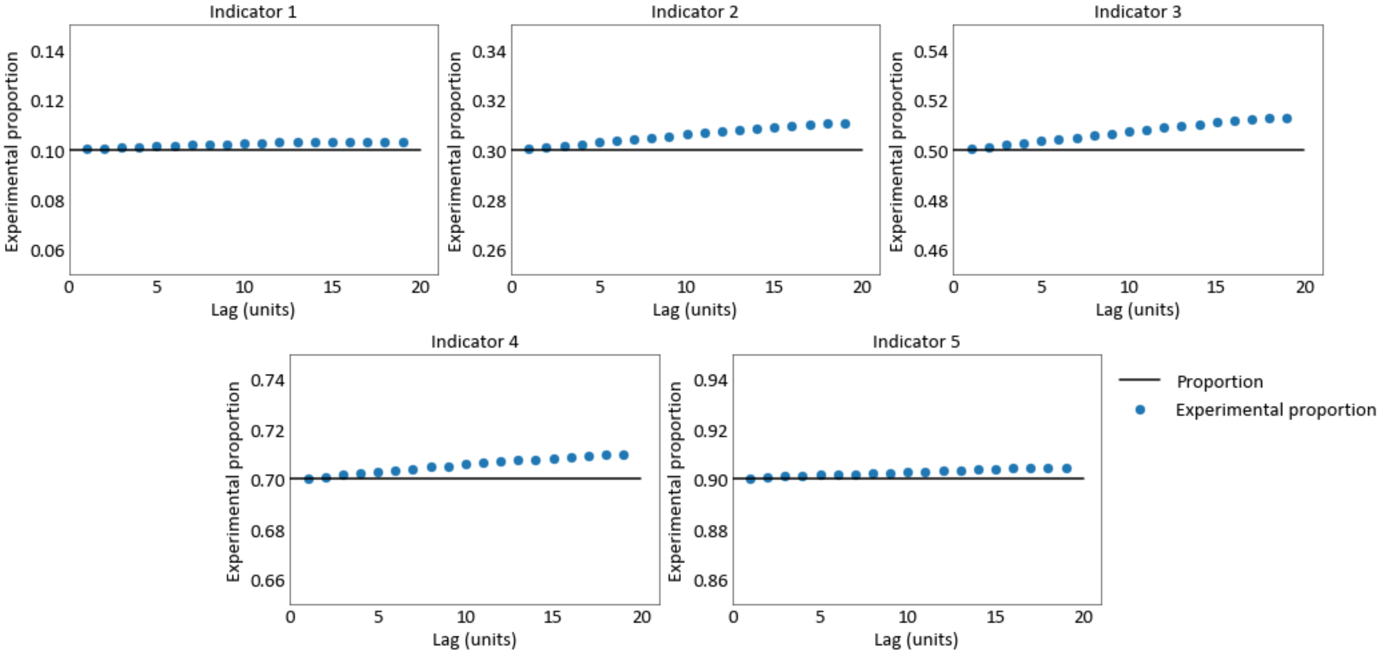


Figure 4.10: Indicator proportion deviation for different lags

The analytically calculated probabilities should correlate with the experimental probabilities, showing experimentally that the equations derived in this paper are correct. Although both probabilities present a satisfactory correlation, some differences can be appreciated. The probabilities obtained from the equations are calculated using fixed proportions for each lag. These proportions come from the original threshold CDF values. However, proportions calculated from the experimental bivariate distributions at each lag slightly differ from the originals (Figure 4.10).

4.4 Sensitivity of bivariate probabilities to sampling density

Even using exhaustive data with a Gaussian distribution and analytically correct variograms, the bivariate probabilities deviate (Figure 4.9). In the case of a real deposit with sparse data, samples are paired using tolerances for variogram and probabilities calculation to avoid lags with an insufficient amount of pairs. An example with sparse data is presented to quantify the deviation in the probabilities. To maintain Gaussianity, the samples are first randomly chosen, and then a simulation is run in the samples locations. The number of samples used for the experiment is 1024, 676, and 256 (Figure 4.11):

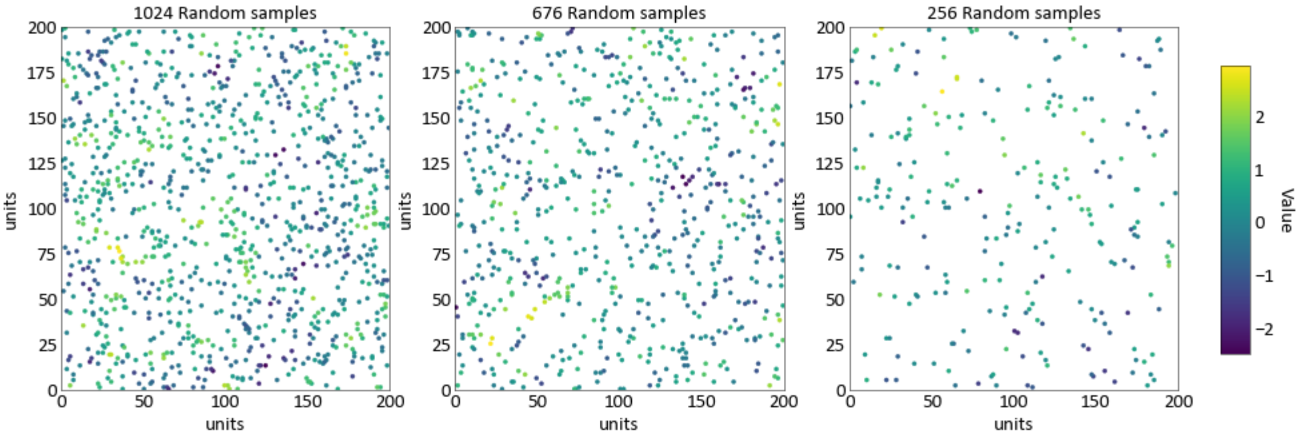


Figure 4.11: Different sample density for sensitivity analysis

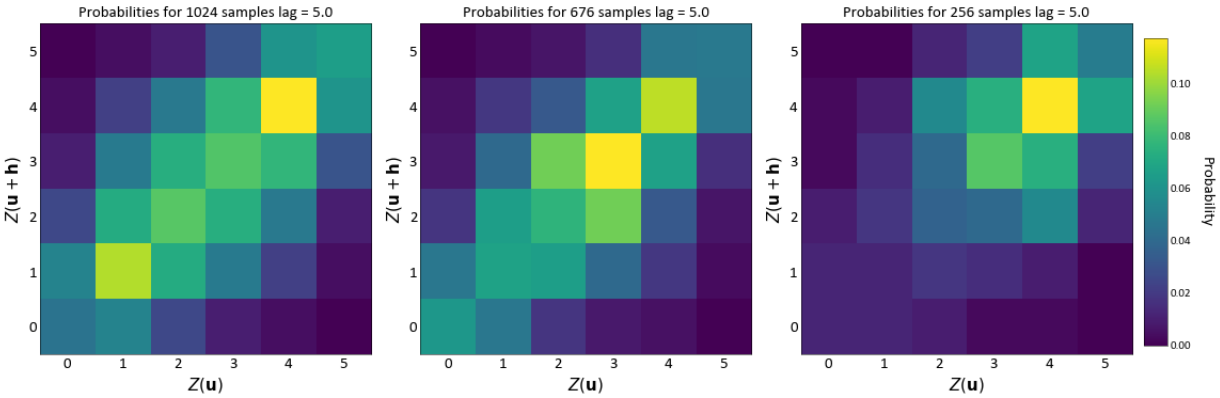


Figure 4.12: Probabilities of different data sets

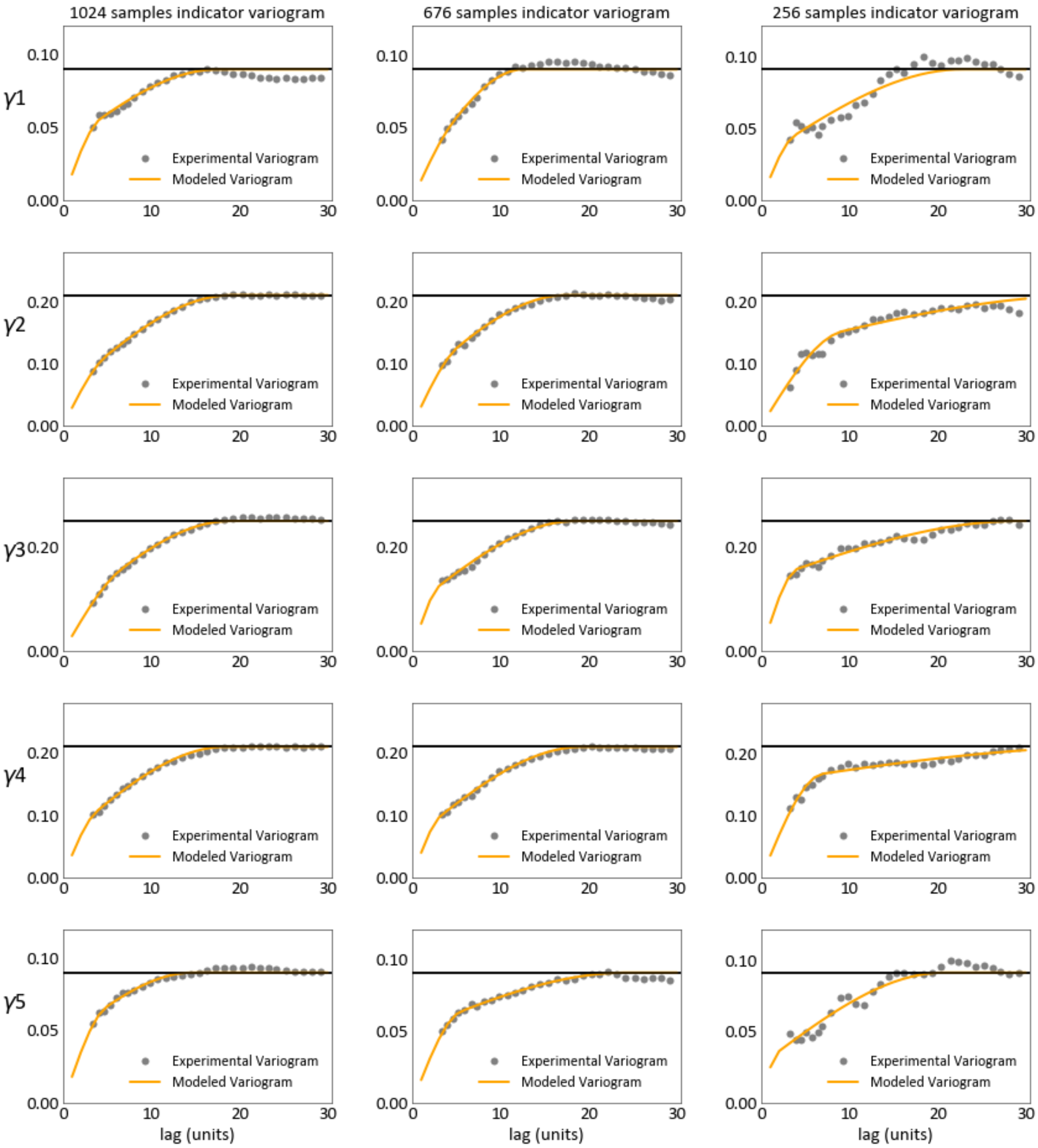


Figure 4.13: Indicator variograms of different data sets

The thresholds remain the same as before [0.1, 0.3, 0.5, 0.7, 0.9]. The distribution values of the thresholds are slightly different for each set of samples. For each sample set, the probabilities of bivariate distributions at each lag are calculated. A tolerance in the lag is added for sample pairing due to samples being sparsely distributed; this reduces correlation at low lags (Figure 4.12). The variograms are modeled to the experimental variograms (Figure 4.13). The two first sets of samples present similar continuity and are related to the original simulation variogram with a range of 20 units (Figure 4.13). The required probabilities are calculated from the modeled variograms. As expected, the correlation between the experimental probabilities and the calculated probabilities decreases (Figure 4.14). This suggests that with more samples, probabilities show more Gaussian behavior and with fewer samples, probabilities deviate from a bivariate Gaussian distribution and are related to a different bivariate distribution; this does not indicate that the variograms are inconsistent.

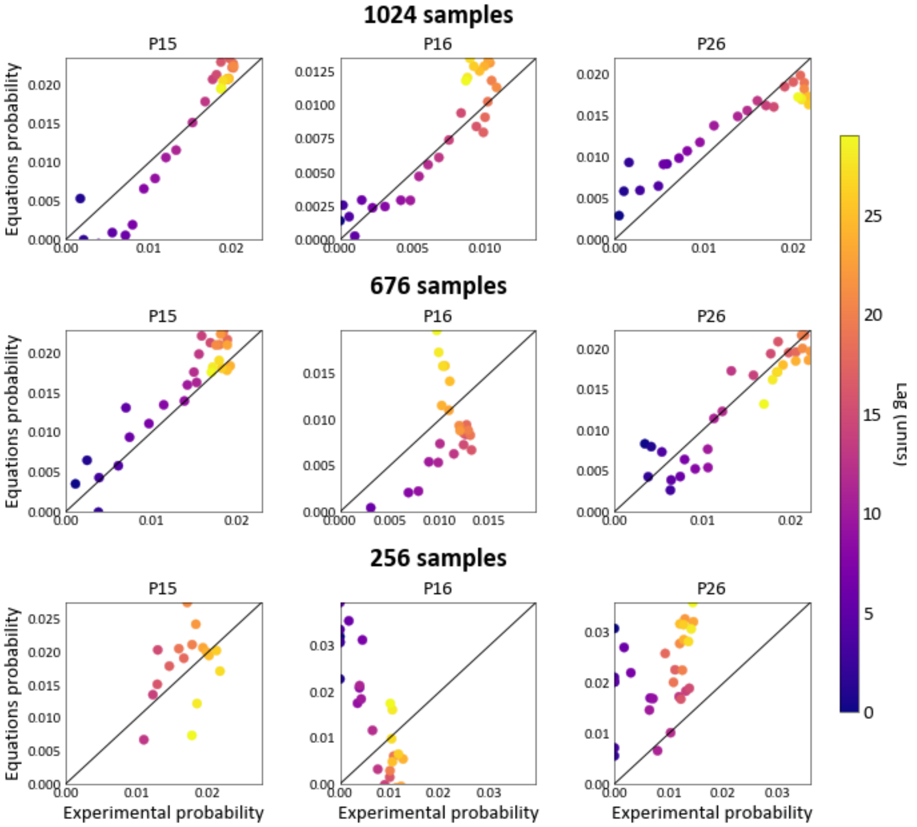


Figure 4.14: Different data sets samples calculated probabilities compared with experimental probabilities

4.5 Metric for assessing the consistency of indicator variograms

The use of probabilities as consistency metrics for indicator variogram modeling is shown on an exhaustive data set from the CCG data validation project (Mokdad et al., 2022). The data is randomly sampled by 1024 samples (Figure 4.15):

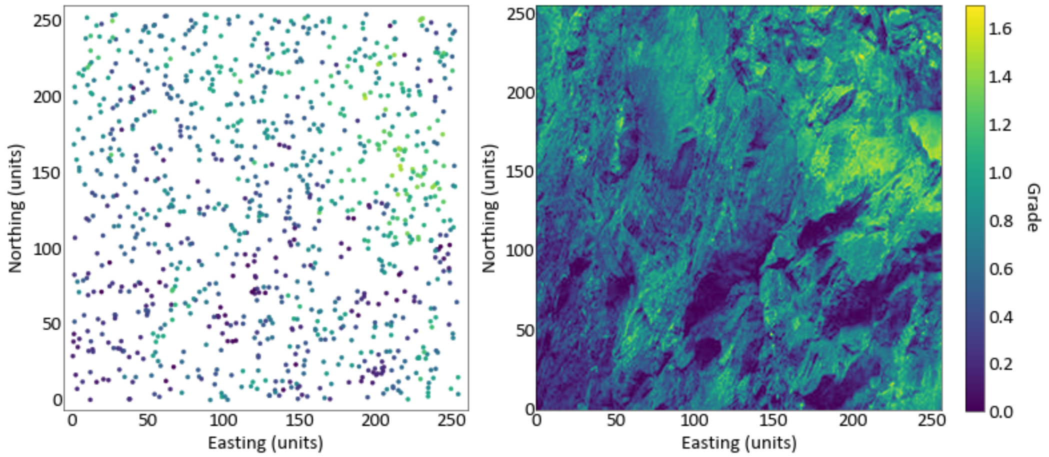
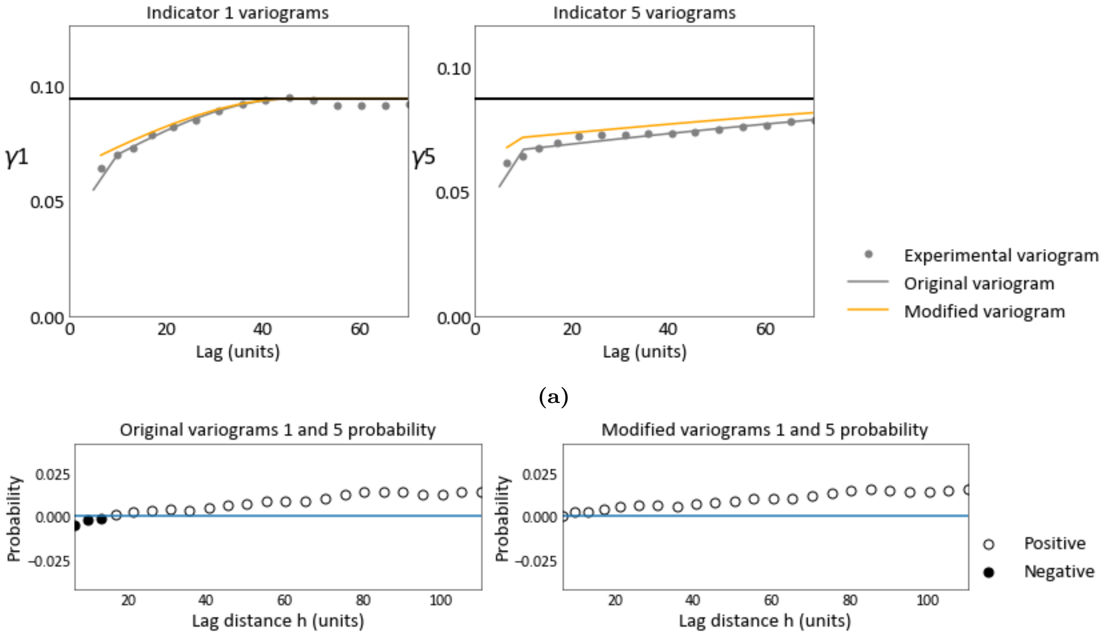


Figure 4.15: Samples and exhaustive data



(a)

(b)

Figure 4.16: (a) Original and modified variograms (b) Probability associated to the original and modified variograms

Table 4.1: Original and modified variograms for 1st indicator. Modified parameters are highlighted in red

Variogram	Structure	Sill	Azm	Dip	hmax	hmin
Original γ_1	Spherical	0.67	0	0	6.74	6.74
	Spherical	0.202	0	0	45.56	45.56
	Spherical	0.128	0	0	47.77	47.77
Modified γ_1	Spherical	0.62	0	0	7.74	7.74
	Spherical	0.252	0	0	45.56	45.56
	Spherical	0.128	0	0	47.77	47.77

Table 4.2: Original and modified variograms for 5th indicator. Modified parameters are highlighted in red

Variogram	Structure	Sill	Azm	Dip	hmax	hmin
Original γ_5	Spherical	0.74	0	0	8.43	8.43
	Spherical	0.236	0	0	141.03	141.03
	Spherical	0.024	0	0	9999	9999
Modified γ_5	Spherical	0.8	0	0	8.04	8.04
	Spherical	0.196	0	0	141.03	141.03
	Spherical	0.004	0	0	9999	9999

Experimental omnidirectional indicator variograms are calculated from the samples and modeled using the GSLIB program varmodel.exe. The variograms have no nugget effect and three spherical structures for each indicator. Central probabilities are obtained from the experimental data. The probabilities calculated from the equations presented are used as a metric to check the consistency of variograms. The probability related to variograms of thresholds 1 and 5 presents negative values for lags less than 20 units. The variograms for thresholds 1 and 5 have a positive sign in the probability equations. Hence, the only option to change probability values from negative to positive is to make the modeled variograms less continuous. This is done by changing the ranges and sill contribution of the variograms (Figure 4.16a) (Tables 4.1 and 4.2). Then, the probability negative values are changed to positive (Figure 4.16b) and the variograms become consistent with a valid bivariate distribution. After fixing the consistency of the variograms, the test is performed again to ensure that the modification did not alter other probabilities.

4.6 Test on multiple data

The previous section showed how the bivariate probabilities calculated from modeled indicator variograms can help to model indicator variograms so they are consistent with a bivariate distribution. This section examines if the consistency of the indicator variogram models improves the number of order relation issues and the results of MIK estimation. A test is carried out on 10 data sets (Figures 4.17 and 4.18) from the CCG database validation project (Mokdad et al., 2022). The datasets are chosen to test different types of spatial correlation. Five thresholds are chosen and indicator variograms are modeled per each dataset. Bivariate probabilities are calculated from the modeled indicator variograms. Then, modeled variograms are modified to remove as many as possible negative values of the bivariate probabilities. The sill and the range of the variograms are modified to make the probabilities positive. MIK using both original and modified variograms is performed and results compare with the exhaustive data. Decreasing order relations issues and global statistics are calculated from the MIK results.

The results (Table 4.3) show that five out of ten datasets present fewer order relations when modifying the variograms. In all datasets, the global mean is practically the same for both MIK with original variograms and with modified variograms. The RMSE changes only in a millesimal order of magnitude and is not related to the number of order relations. The results do not show a consistent improvement in the number of order relations or the global statistics for the MIK with modified variograms. To get a deeper insight, locations where there are order relation issues for MIK with original variograms and no order relation issues for MIK with modified variograms and vice-versa are also explored (Tables 4.4 and 4.5). In general, the mean and the RMSE do not appear to improve on the MIK with modified variograms in a consistent manner for both cases.

4.7 Limitations

To eliminate all negative values from the calculated bivariate probabilities, the modeled variograms have to depart from the experimental spatial correlation considerably, changing the ranges and the sill of the variograms. Also, there is no consistency between the modification of the variograms and improvements in the number of order relation issues and MIK results. An automatic fitting tool could be developed to ensure that variograms respect the spatial correlation of the samples while also maintaining bivariate probabilities as positively as possible. Even so, the departure from the spatial correlation to get consistent bivariate probabilities could mean that the variograms are not enough to capture both features for indicators.

Not all bivariate probabilities are calculated with the equations. Only probabilities that are on the edges of the bivariate distribution are calculated. The equations make use of experimental probabilities to calculate the bivariate probabilities from the modeled variograms. Since these probabilities come from the experimental data they present uncertainty. A test could be made to add uncertainty to the bivariate probabilities calculation and see how this affects the consistency of the bivariate probabilities.

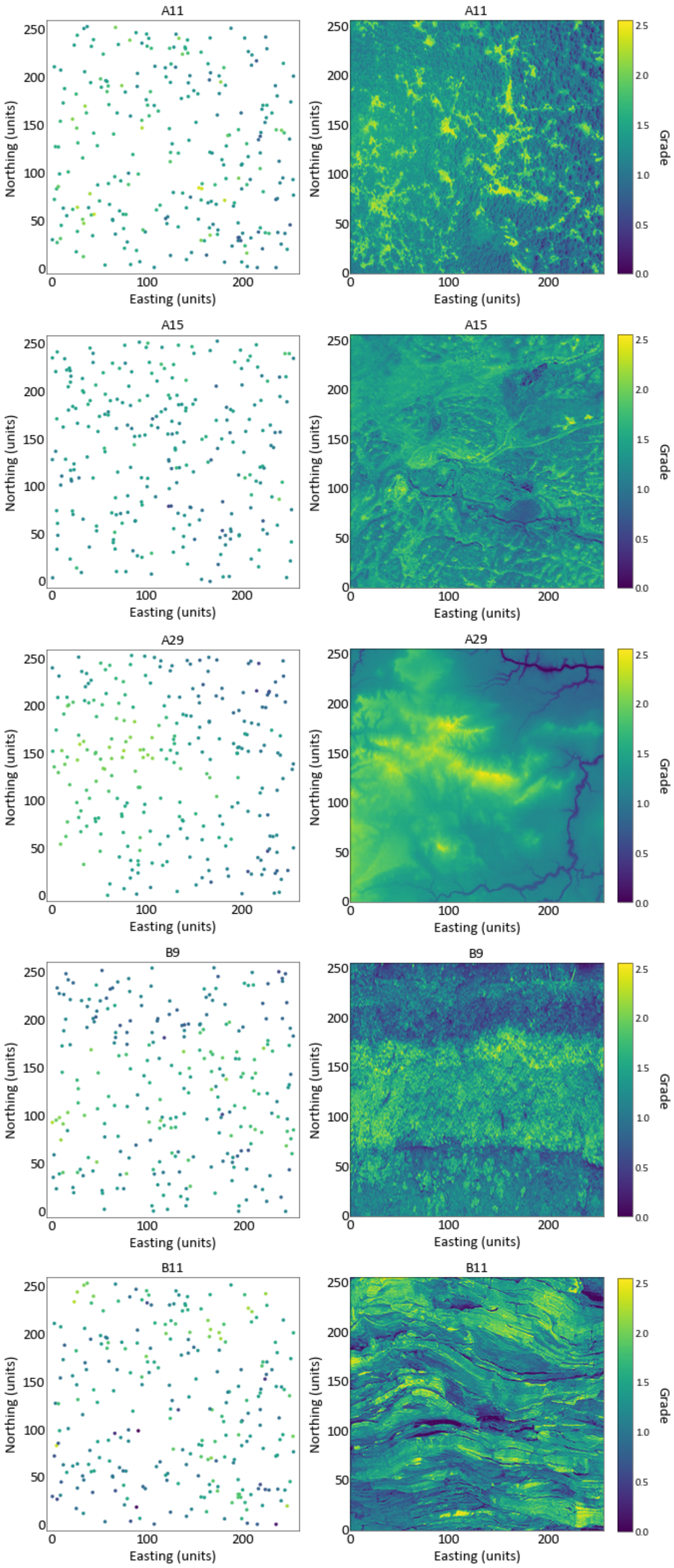


Figure 4.17: Datasets used for the tests

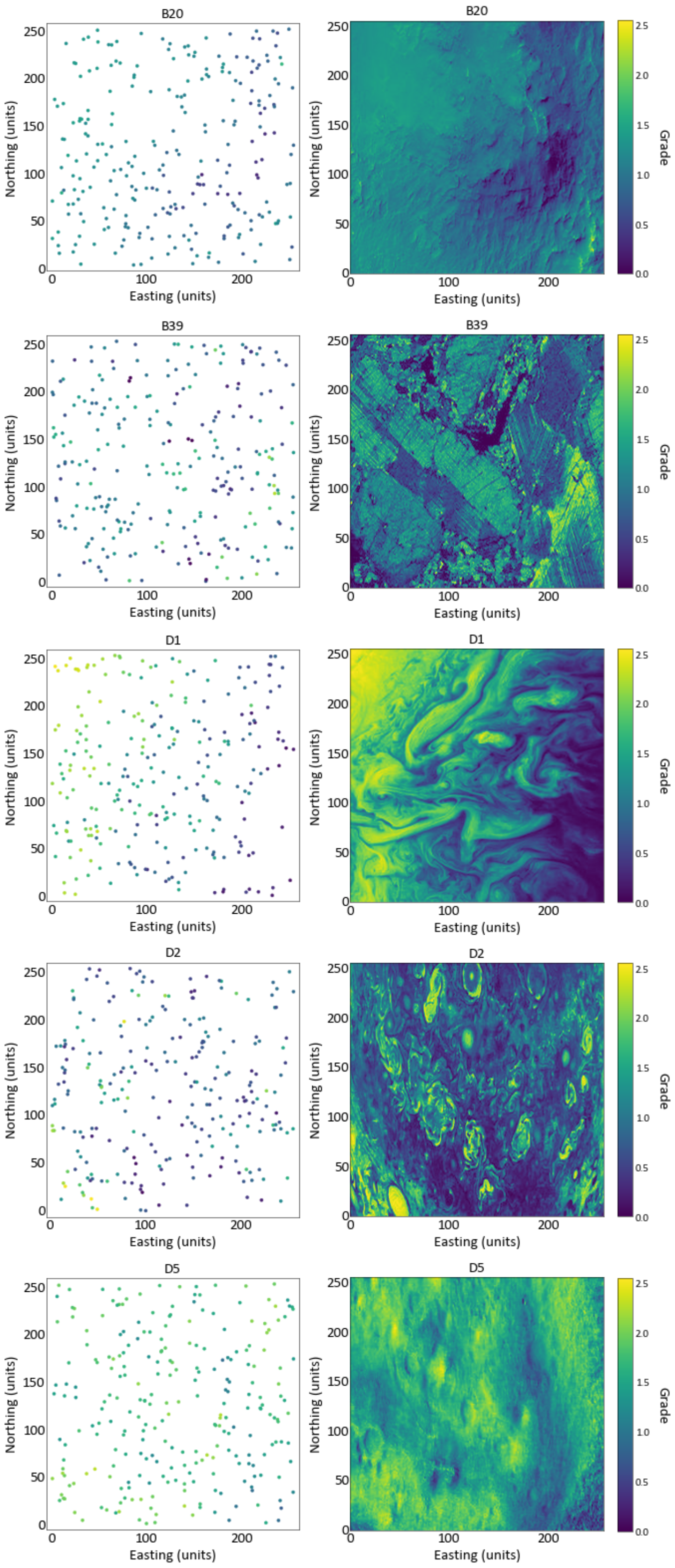


Figure 4.18: Datasets used for the tests

Table 4.3: Order relations and global statistics compared with exhaustive data. Data sets with fewer order relation issues after the variogram modification are highlighted in green. Data sets with smaller RMSE after the variograms modification are highlighted in red

	Order relation number		**Global mean**			**RMSE**	
Dataset	**OV**	**MV**	**Data**	**OV**	**MV**	**OV**	**MV**
A11	8805	10785	1.431	1.456	1.455	0.277	0.277
A15	8731	8592	1.083	1.072	1.072	0.187	0.186
A29	88665	88405	1.411	1.400	1.401	0.150	0.152
B9	28087	33126	1.128	1.136	1.136	0.229	0.229
B11	7706	8455	1.263	1.320	1.321	0.426	0.426
B20	64456	62341	0.869	0.861	0.860	0.146	0.151
B39	29796	26409	0.988	1.011	1.012	0.448	0.447
D1	21492	13707	1.177	1.198	1.210	0.311	0.332
D2	15123	17088	0.877	0.898	0.899	0.439	0.440
D5	46224	46860	1.627	1.587	1.587	0.188	0.190

OV = MIK with original variograms
MV = MIK with modified variograms

Table 4.4: Order relations and global statistics for locations with order relation issues for MIK with original variograms but no order relations issues for MIK with modified variograms

		Global mean			**RMSE**	
Dataset	**Locations**	**Data**	**OV**	**MV**	**OV**	**MV**
A11	1137	1.309	1.226	1.310	0.224	0.288
A15	2026	1.079	1.070	1.145	0.166	0.170
A29	3814	1.736	1.626	1.479	0.229	0.135
B9	421	1.190	1.237	1.271	0.310	0.247
B11	781	1.226	1.270	1.278	0.370	0.327
B20	2449	0.811	0.770	0.763	0.133	0.184
B39	5131	1.031	1.031	1.085	0.447	0.445
D1	13272	0.846	0.885	1.084	0.291	0.355
D2	119	0.812	0.863	0.912	0.330	0.409
D5	1843	1.630	1.650	1.620	0.165	0.190

OV = MIK with original variograms
MV = MIK with modified variograms

Table 4.5: Order relations and global statistics for locations with no order relation issues for MIK with original variograms but with order relations issues for MIK with modified variograms

		Global mean			**RMSE**	
Dataset	**Locations**	**Data**	**OV**	**MV**	**OV**	**MV**
A11	2132	1.671	1.626	1.476	0.237	0.261
A15	1900	1.110	1.120	0.944	0.131	0.198
A29	1704	1.701	1.702	1.552	0.124	0.102
B9	7543	1.099	1.102	1.056	0.213	0.212
B11	1491	1.137	1.153	1.161	0.494	0.405
B20	5086	0.851	0.819	0.951	0.132	0.098
B39	2181	0.986	0.989	1.029	0.431	0.491
D1	4935	0.344	0.496	1.248	0.235	0.305
D2	1209	0.708	0.776	0.788	0.370	0.300
D5	1514	1.680	1.643	1.516	0.188	0.158

OV = MIK with original variograms
MV = MIK with modified variograms

Chapter 5

Data driven RBF interpolation

There are several techniques to avoid order relations problems: replace the indicator variables with values from a Gaussian kernel to smooth the distributions, model indicator variograms with proportional parameters, or use the variogram of the median threshold for all indicators (Deutsch & Journel, 1998). Since indicators are estimated independently and Kriging is a nonconvex estimator that produces negative weights (Deutsch, 1996), these techniques only reduce order relations problems but do not eliminate them completely. The most common methodology to fix order relation issues is a post-process that consists of resetting values to $I^*(\mathbf{u}, z_k) = I^*(\mathbf{u}, z_{k'})$ upwards and downwards and average them. Values below zero and one are reset to zero and 1 respectively (Deutsch & Journel, 1998). A method for MIK without order relations issues and no post-process is proposed in this chapter. This novel methodology uses the RBF formulation with constraints recursively to obtain the final CCDF values without order relation issues. The methodology is compared against classic MIK, MGK, SIK and CDIK.

5.1 Methodology

As seen in Chapter 2, inequality constraints can be added to any location in the RBF formulation by minimizing the quadratic form of the interpolator. The methodology estimates indicators using RBF without constraints for each threshold to identify the locations where order relations issues are occurring; then, RBF with inequality constraints (RBFC) in those locations is performed to get a final result without order relations issues. Since each inequality constraint will modify the estimated value at their respective location, the identification of order relation issues and the estimation of indicator values using RBFC must be sequential. Consider $I^*_{RBF}(\mathbf{u}, z_k)$ and $I^*_{RBFC}(\mathbf{u}, z_k)$ as the estimates at location $\mathbf{u}$ of the indicator variable assigned for the threshold z_k using RBF and RBFC respectively, the sequential methodology proposed follows the next steps:

1. Select thresholds
2. Assign indicators for each threshold
3. Model indicator variograms for each indicator
4. Perform an RBF interpolation without constraints for each indicator
5. Reset values to 0 and 1 where $I^*_{RBF}(\mathbf{u}, z_1) < 0$ and $I^*_{RBF}(\mathbf{u}, z_1) > 1$

6. Identify locations where $I^*_{RBF}(\mathbf{u}, z_1) > I^*_{RBF}(\mathbf{u}, z_2)$ for $z_1 < z_2$ and define inequality constraints at those locations

7. Run the RBFC with constraints on each location identified for the second indicator

8. Reset values to 0 and 1 where $I^*_{RBFC}(\mathbf{u}, z_2) < 0$ and $I^*_{RBFC}(\mathbf{u}, z_2) > 1$

9. Identify locations where $I^*_{RBFC}(\mathbf{u}, z_2) > I^*_{RBF}(\mathbf{u}, z_3)$ for $z_2 < z_3$ and define inequality constraints at those locations

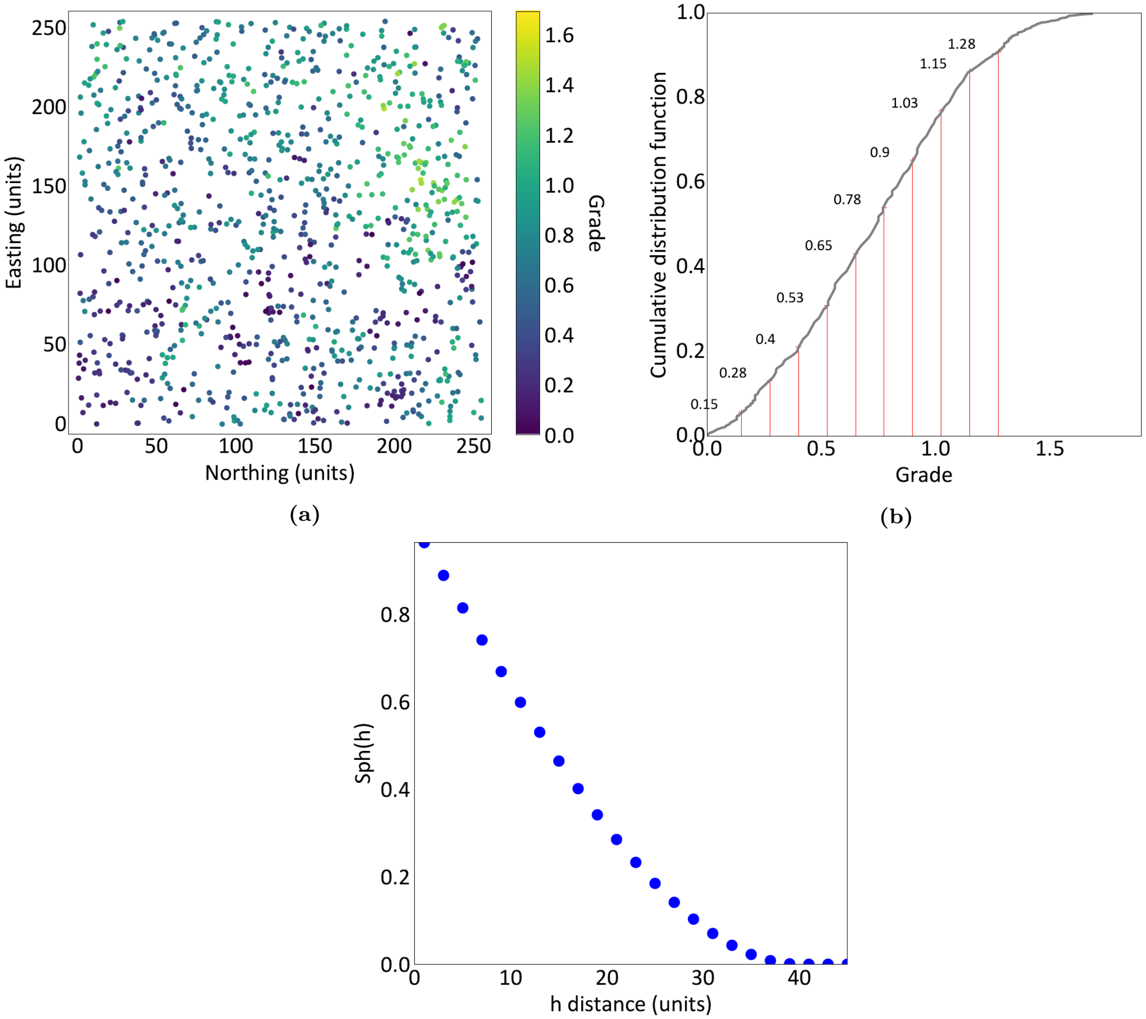
10. Run the RBFC with constraints on each location identified for the third indicator

11. Repeat for the rest of the indicators

Instead of using a sequential methodology, inequality constraints could be added at each location, forcing the weights of the Kriging system to comply with order relations. However, each inequality constraint is added to the quadratic minimization system and, since RBF also uses all the samples available to calculate the weights, including more elements in the system will increase the computational time making it unpractical. The methodology is computationally efficient if $N(\text{number of samples}) + M(\text{number of constraints}) < 10,000$. In that sense, the recursive method is proposed to include inequality constraints only at locations and thresholds where is needed. Values outside [0,1] are reset each time to avoid adding more than one constraint in a single location. Adding more than one constraint in the same location may result in inconsistent estimates and order relation issues. The algorithm can be applied from bottom to top and from top to bottom and average the CCDFs values to obtain a final CCDF. The whole process is called multiple indicator Kriging with RBF with constraints (MIKRBF).

5.2 Example

To test the methodology, 1024 samples from the CCG data validation project (Mokdad et al., 2022) are used for MIKRBF (Figure 5.2a). Ten thresholds are chosen to discretize the CCDF of the samples evenly (Figure 5.2b). Indicators are assigned for each threshold. A spherical kernel with a range of 40 units is used for the interpolation (Figure 5.2c). The same spherical kernel is used for each indicator variable. The indicators are interpolated using RBF and MIKRBF in a 1D array of locations. MIKRBF is applied from bottom to top and from top to bottom. Then, the indicator estimates are plotted along the locations for both methodologies (Figure 5.2). The lighter the shade of the dot, the higher the threshold. Lighter-shaded dots should be above darker-shaded dots to have a licit CCDF. This does not occur for the RBF estimates. There are indicator estimates below 0, above 1 and decreasing for increasing thresholds. MIKRBF does not present any order relation issues. The details of the methodology can be seen in a single location local distribution. Figure 5.3 shows the indicator estimates of each threshold for a single location. Figure 5.3a and 5.3b show the RBF and the MIKRBF estimates from bottom to top and top to bottom respectively. Notice how the indicator estimate is modified to the previous threshold value when an order relation issue arises. Order relation issues related to indicator estimates below zero and above one are reset.

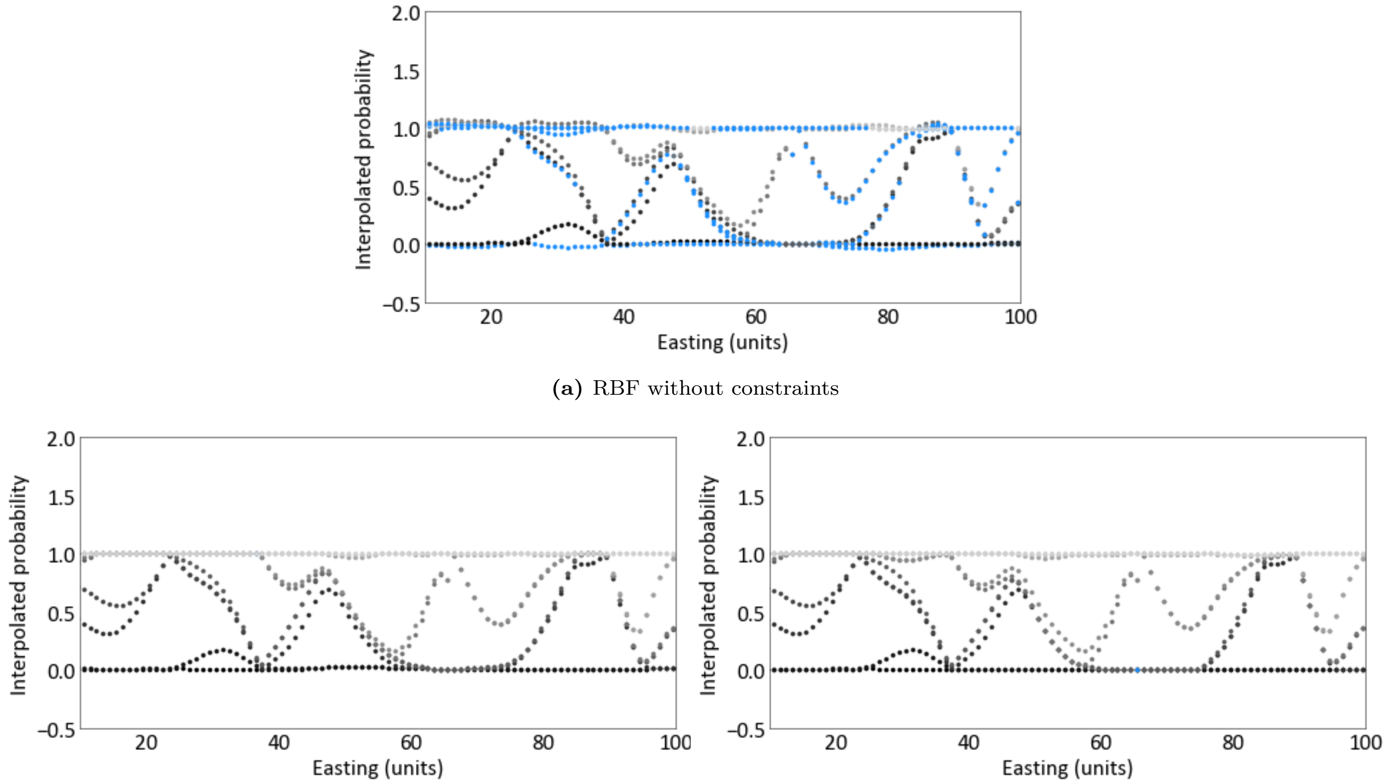


(a)

(b)

(c)

Figure 5.1: (a) 1024 samples for interpolation, (b) CDF of the samples and thresholds for MIK, (c) spherical RBF with 40 units of range

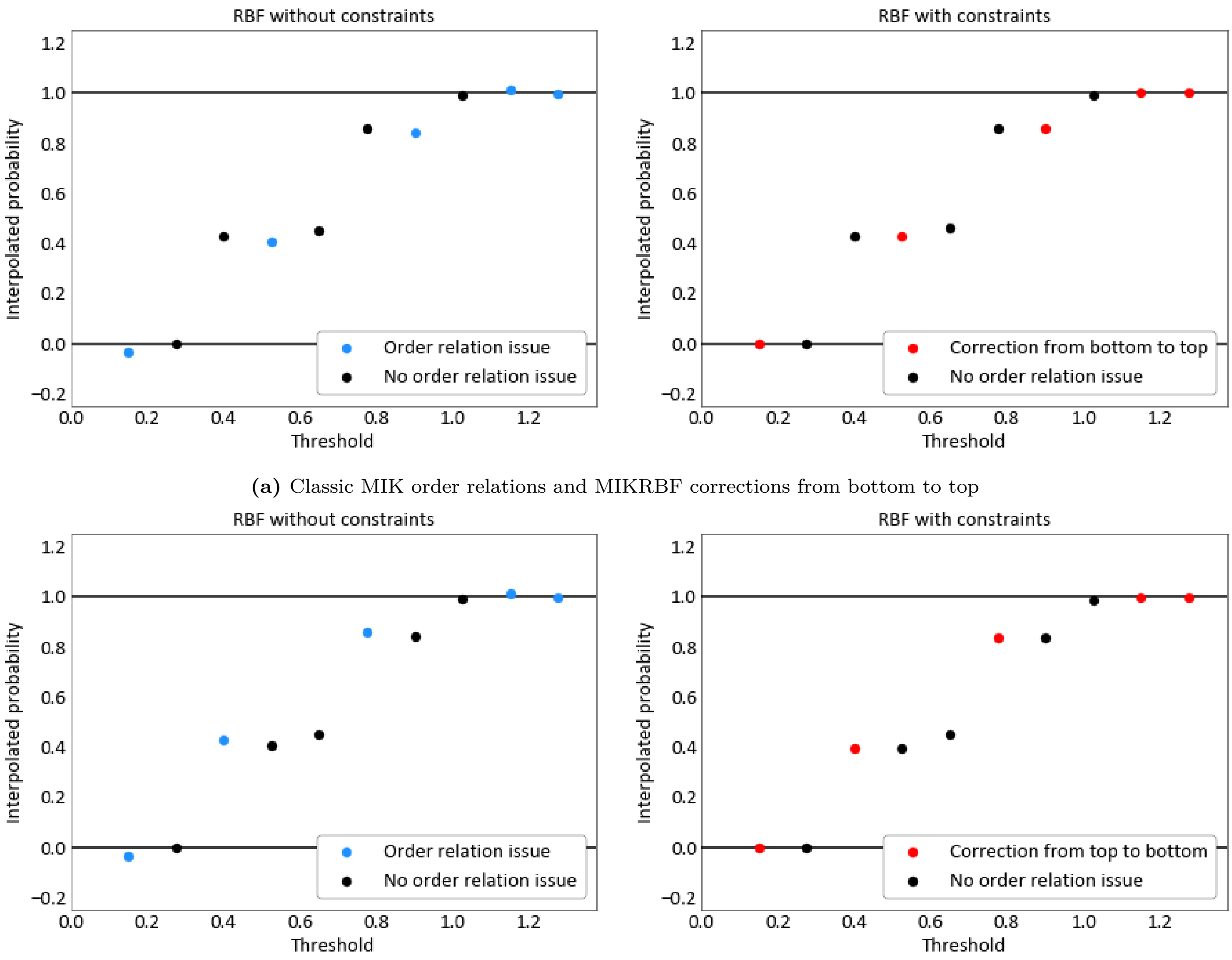


(a) RBF without constraints

(b) MIKRBF from bottom to top

(c) MIKRBF from top to bottom

Figure 5.2: RBF and MIKRBF indicator estimates. Lighter-shaded dots correspond to higher thresholds, darker shaded dots correspond to lower thresholds. Order relation issues are highlighted in blue



(a) Classic MIK order relations and MIKRBF corrections from bottom to top

(b) RBF order relations and MIKRBF corrections from top to bottom

Figure 5.3: Local distribution with corrections from bottom to top and top to bottom

5.3 MIK methodologies comparison

To test the performance of MIKRBF, the new algorithm is compared with classic methodologies for the inference of local distribution such as classic MIK and MGK, and methodologies that use the concept of compositional data to get licit CCDFs. Compositional data describe multiple variables as being part of a whole, such as element compounds of a mineral (Tolosana-Delgado et al., 2019). In the case of MIK with indicators assigned from ranges, indicators can be considered compositional data since they must sum to one and be non-negative. The first of the two compositional data methodologies SIK, a MIK estimation methodology introduced by Tolosana-Delgado et al. (2008). The procedure interpolates indicators without order relation issues by considering the indicator variables assigned from bins as probability vector coordinates on an orthonormal basis of the simplex. After a classic MIK interpolation, a transformation, introduced in Chapter 2, is applied to the estimates to obtain the final CPDF values. The values will be positive and sum to 1, correctly describing a CPDF. In the second compositional methodology (CDIK) developed by Hadavand and Deutsch (2021) based on the work of Tolosana-Delgado et al. (2019, 2008), the indicator variables are transformed to ratios using the additive log-ratio transform. Since compositional data transformations involve the division of the variables by one of them, indicators are replaced by the outcome of a Gaussian kernel to avoid zeros. A normal score transformation is applied to the transformed indicators for data stability. After interpolating the transformed indicators, the values are back-transformed to obtain the final CPDFs.

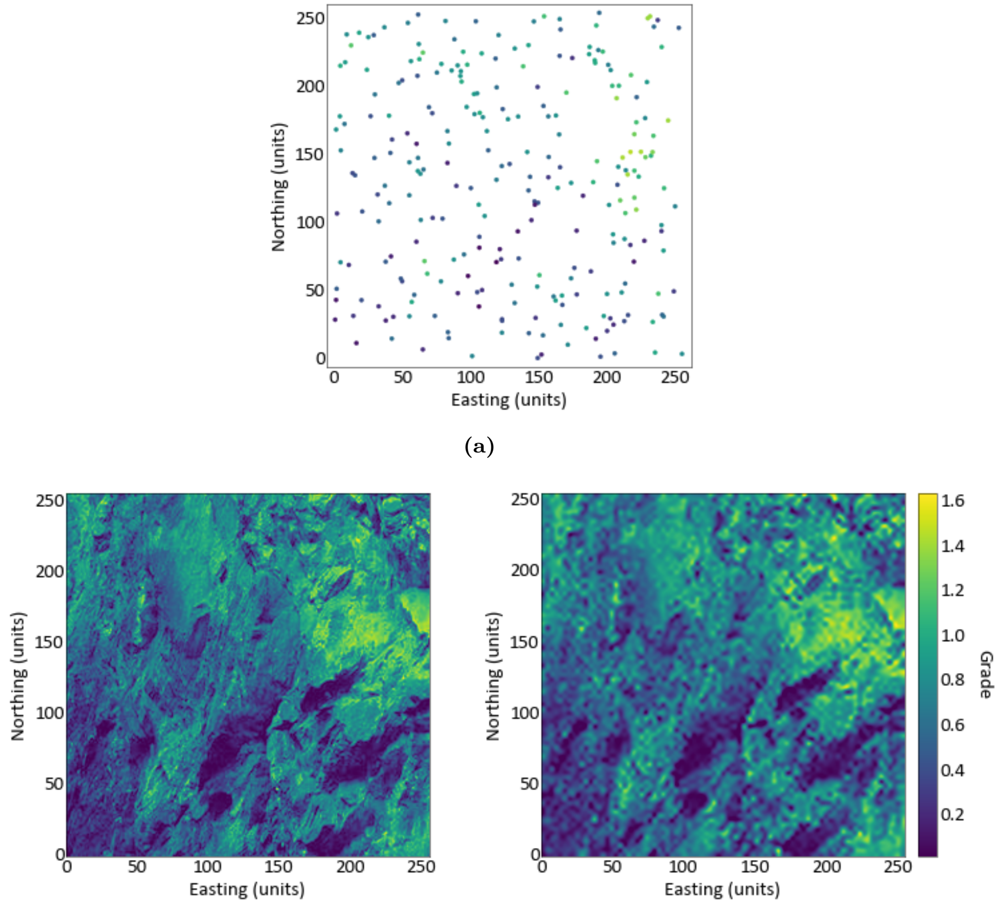
5.3.1 Comparison methodology

The data used in the comparison consist of an exhaustive data set from the CCG validation project (Mokdad et al., 2022). 256 samples are taken randomly from the exhaustive data (Figure 5.4a). The dataset is chosen due to its stability in indicator variograms. The exhaustive data has a dimension of 256x256 which corresponds to 65536 locations separated by 1 unit (Figure 5.4b). Since MIKRBF can only take 10,000 samples plus order relations issues, the number of locations is too high for MIKRBF to work properly. The exhaustive data resolution is downgraded to a grid of 85x85 with 7225 locations separated by 3 units (Figure 5.4c). Simple Kriging for all methodologies uses a search volume of 100 units in all directions and 40 as the maximum amount of samples.

For MGK, the data is declustered using the cell methodology. The cell for declustering has a dimension of 26 units and the mean of the samples changes from 0.74 to 0.71. Then, the samples are normal scores transformed to a Gaussian distribution. Variograms are calculated from the transformed variable for directions 0° and 90° and modeled with three spherical structures (Table 5.1 and Figure 5.5). Simple Kriging is used to obtain the mean and Kriging variance which describes the complete CCDF at each location. The program postmg.exe is used to back-transform the Gaussian distributions and obtain the E-type in original grades.

For classic MIK and MIKRBF, ten thresholds are chosen to equally discretize the CDF of the samples (Figure 5.6). Each class made by the thresholds contains more than 4% of the total samples (Table 5.2). Indicator variables are assigned per threshold. Experimental indicator variograms are calculated and modeled by omnidirectional variograms with three spherical structures (Figure 5.7). For the case of MIK, ik3d.exe is used for the interpolation. The program makes all the order relation issues corrections (Deutsch & Journel, 1998). For the case of MIKRBF, the interpolation is run in Python. The E-types from the local distributions are obtained by using the program postik.exe in both cases.

The compositional data methodologies use the same ten thresholds used for MIK and MIKRBF but, indicator variables are assigned from the probability of being in between thresholds. For SIK, indicator variograms are calculated and modeled with omnidirectional variograms with three spherical structures (Figure 5.8). Some of the indicator variograms are very discontinuous due to the lack of samples. Then, indicator variables are estimated using simple Kriging and transformed to a correct CPDF with the simplicial Kriging algorithm described in Chapter 2. For CDIK, the indicator variables are passed by a Gaussian kernel with a range of 0.15 units. The Gaussian kernel modifies values of one to close to one and values of zero to close to zero. Then, the variables are transformed using the additive log-ratio transform and the normal score transform. Variograms are calculated and modeled with three spherical structures (Figure 5.9). Simple Kriging is applied to the transformed variables and then the estimates are back-transformed to obtain the CPDF. In both cases, the CPDFs are accumulated to obtain the CCDF. The program postik.exe is used to obtain the E-types.



(a)

(b)

(c)

Figure 5.4: (a) Samples from the CCG validation project (Mokdad et al., 2022), (b) 256x256 exhaustive data, (c) 85x85 exhaustive data

Table 5.1: Modeled variograms for MGK

Nugget	Structure	Contribution	hmax	hmin
0.1	Sph	0.233	4.820	10.0
	Sph	0.190	27.03	20.0
	Sph	0.477	109.52	120.0

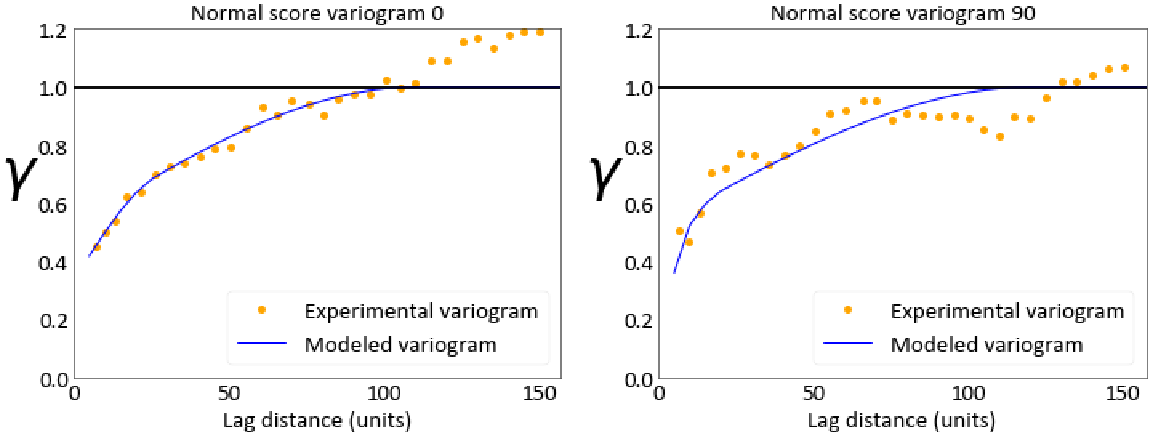


Figure 5.5: Experimental and modeled variograms for MGK

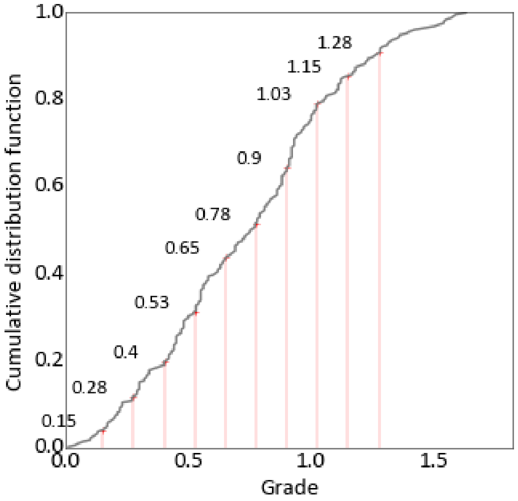


Figure 5.6: Thresholds for MIK and MIKRBF

Table 5.2: Sample percentage per class

Indicators min - 1 sample %	4.3
Indicators 1-2 sample %	7.4
Indicators 2-3 sample %	8.2
Indicators 3-4 sample %	11.3
Indicators 4-5 sample %	12.5
Indicators 5-6 sample %	7.8
Indicators 6-7 sample %	12.9
Indicators 7-8 sample %	14.5
Indicators 8-9 max sample %	6.6
Indicators 9-10 max sample %	5.1
Indicators 10 max sample %	9.4

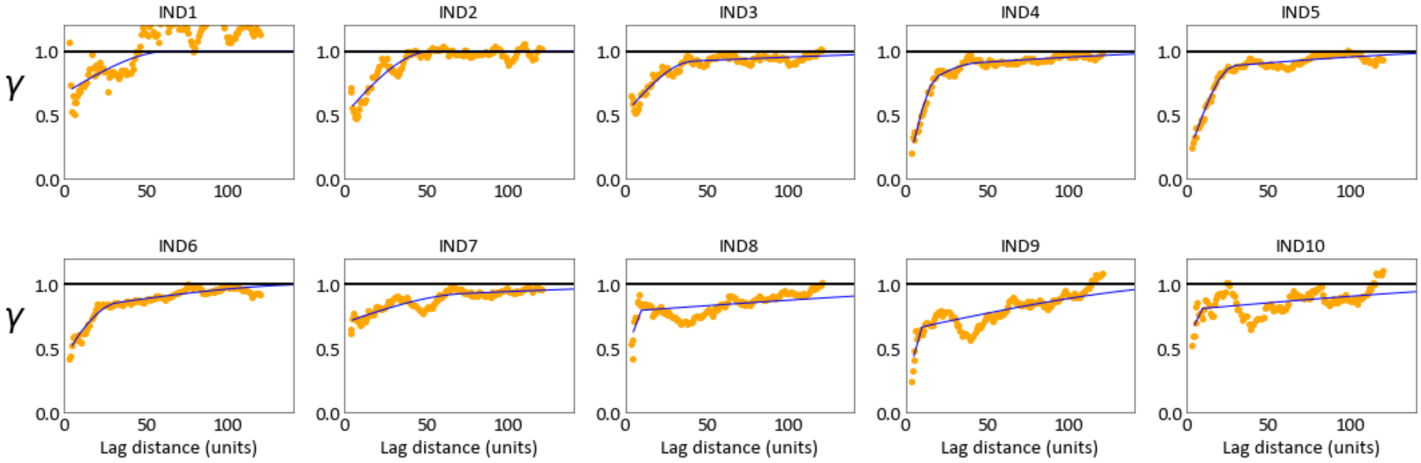


Figure 5.7: Experimental and modeled indicator variograms for MIK and MIKRBF

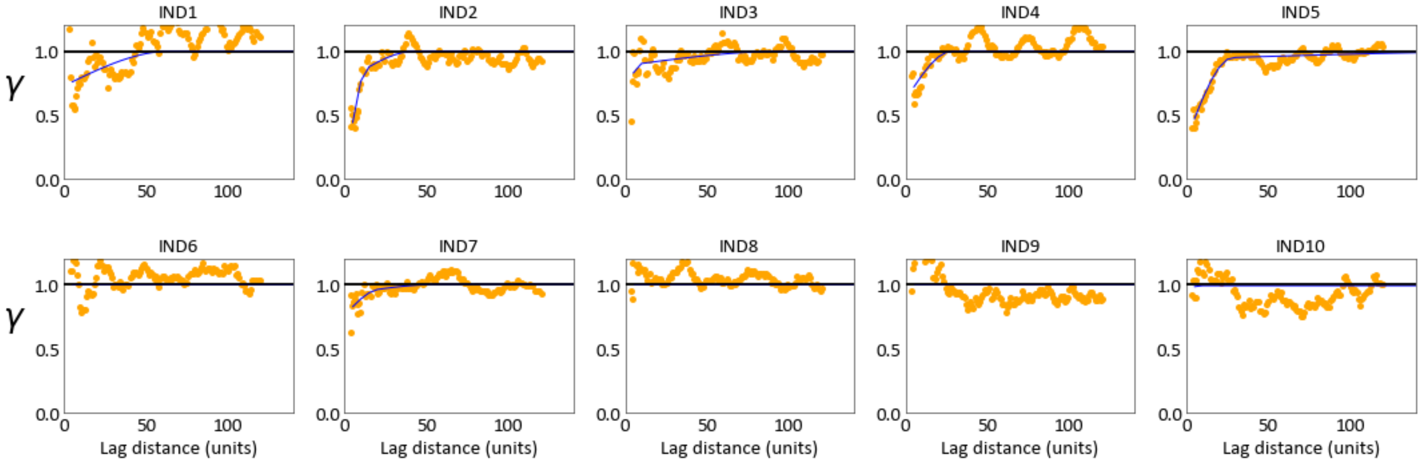


Figure 5.8: Experimental and modeled indicator variograms for SIK

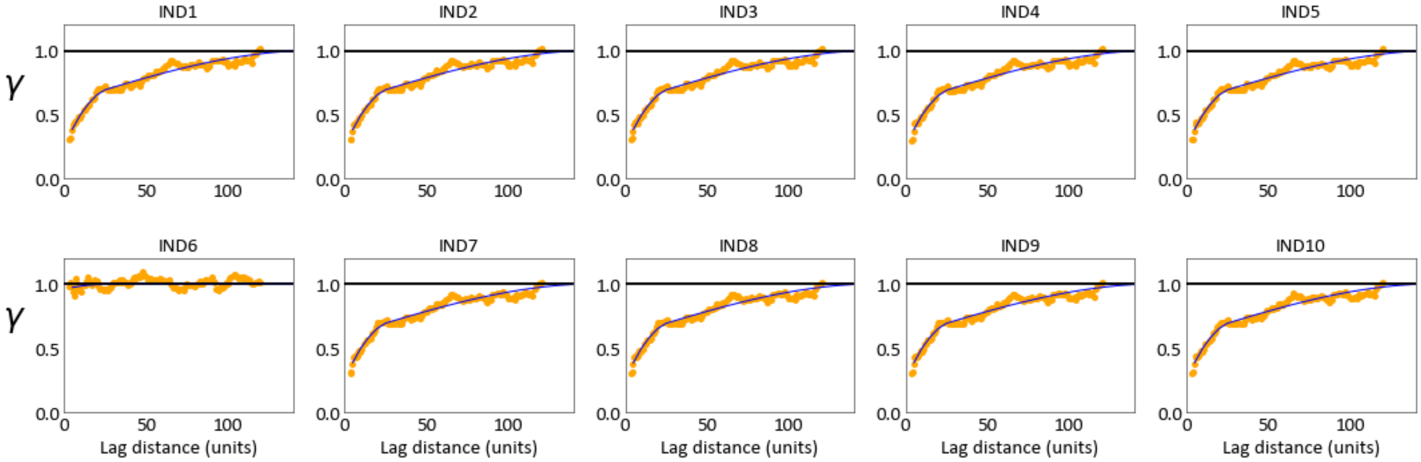


Figure 5.9: Experimental and modeled indicator variograms for CDIK

5.3.2 Results

The E-types of each methodology are compared with the exhaustive data. The E-type statistics comparison (Table 5.3) shows that CDIK has less bias considering the global mean, with an average of 0.711 in comparison with 0.698 of the exhaustive data. It is followed by MGK with 0.715, SIK with 0.719 and MIKRBF with 0.725. MIK is the most biased with a global mean of 0.730. In terms of root mean squared error (RMSE) the methodology with the smallest error is CDIK with an RMSE of 0.286 followed by MGK with 0.285, MIKRBF with 0.294 and CDIK with 0.286. The methodology with the worst error is SIK with 0.319. The slope of regression (SOR) shows that the methodology with the smallest bias is MIK with a SOR of 0.964 followed by MGK with 0.959, MIKRBF with 1.060 and CDIK with 0.907. The worst methodology in terms of SOR is SIK with 1.231. By considering the global statistics, the methodology with the overall best performance is MGK. Even though SIK reproduces the global mean satisfactorily it is the most biased in terms of RMSE and SOR. MIK and MIKRBF present similar results. The results for CDIK and SIK have to be taken carefully. The validation scatter plots which plot E-types and the reference data values (Figure 5.10) show that the estimates of CDIK are discretized by the thresholds used for the interpolation. For SIK, the estimates are similar to the mean. E-type maps show smooth estimates for MGK, MIK and MIKRBF (Figure 5.11). MIK and MIKRBF are the most similar. CDIK presents abrupts changes in grade and SIK shows that most locations have a grade close to the mean. This correlates to the results seen in the scatter validation plots.

Table 5.3: E-types statistics

Method	Mean	St.Dev	RMSE	SOR
MGK	0.715	0.245	0.285	0.959
MIK	0.730	0.234	0.294	0.964
MIKRBF	0.725	0.217	0.291	1.060
CDIK	0.711	0.259	0.286	0.907
SIK	0.719	0.155	0.319	1.231
True Values	0.698	0.369		

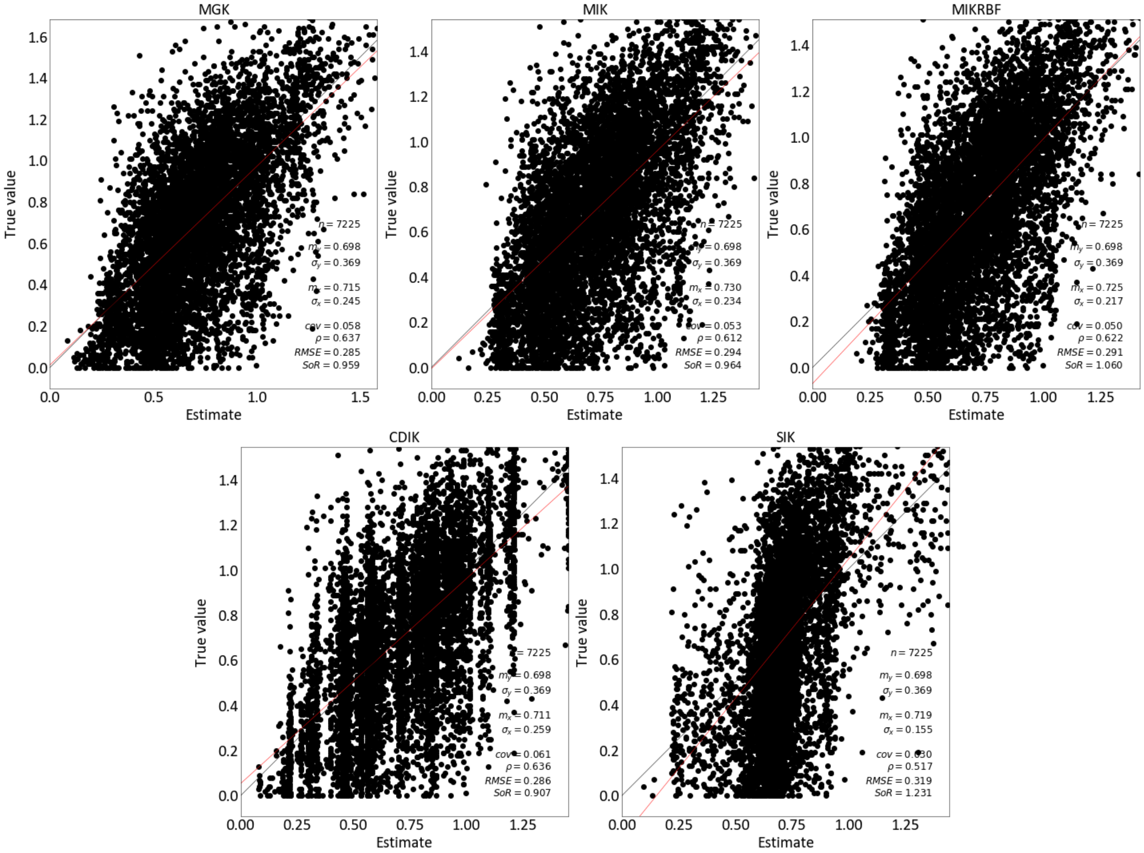


Figure 5.10: Validation scatter plots

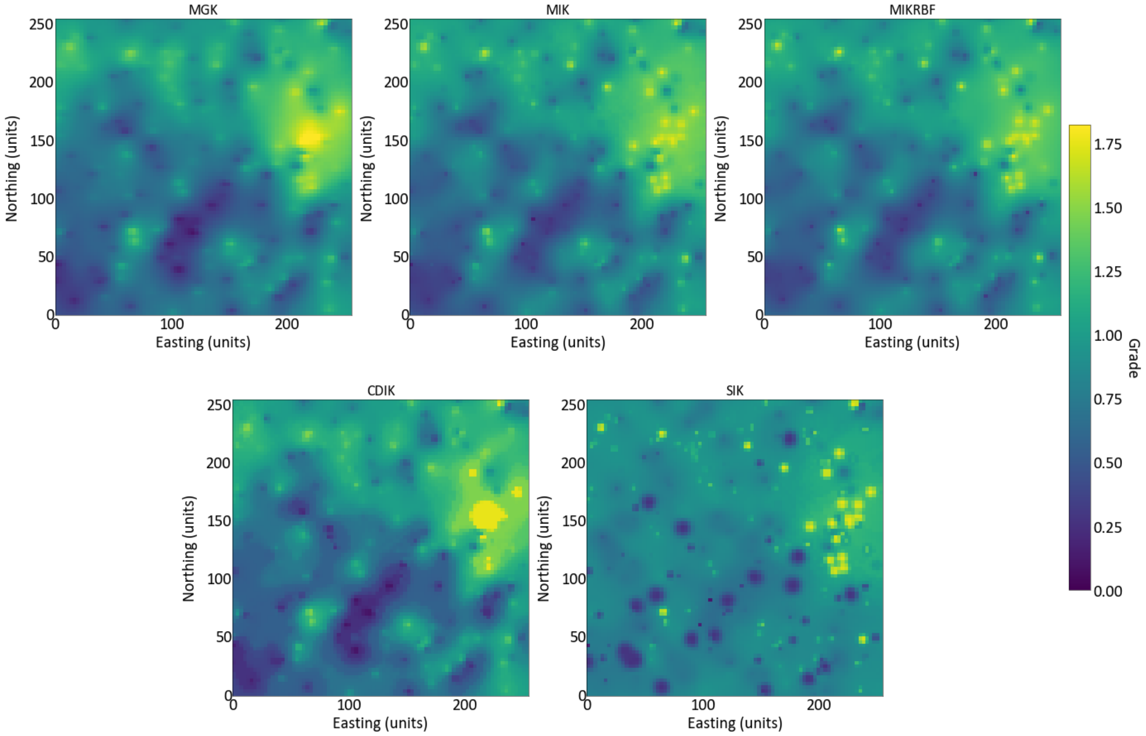


Figure 5.11: Map plots of E-types

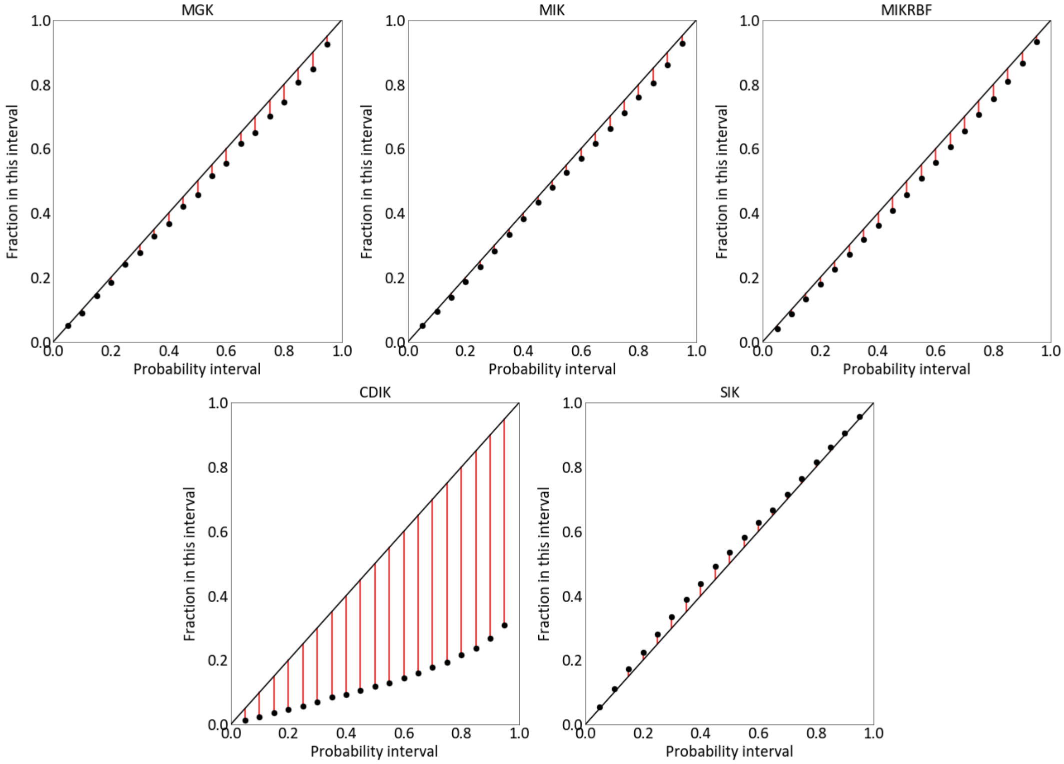


Figure 5.12: Accuracy plots

To check the consistency of the local distributions, accuracy plots are built for each methodology

(Figure 5.12). The accuracy plots show the proportion of times the true value falls in a certain probability interval, thus, the proportion should be similar to the probability interval range. If the proportions fall above the 45° line, then the model is accurate but the uncertainty is too broad. On the other hand, if the proportions are below the 45° line the model is inaccurate and the uncertainty is too narrow. The model is precise and accurate if the proportions fall in the 45° line or close to it (Leuangthong, McLennan, & Deutsch, 2004). Accuracy plots are constructed using the GSLIB program accplt.exe and accplt-sim.exe. Accuracy plots show that MGK, MIK and MIKRBF present proportions below the 45° line but very close to it. Their results are acceptable. SIK presents a good accuracy but its distributions are slightly wider than MGK, MIK and MIKRBF. Its proportions also fall very close to the 45° line, showing good results. CDIK presents a very narrow uncertainty model.

MIKRBF is proven to perform similarly to MIK and MGK. For this dataset, MGK is the methodology with the best overall performance base on the global mean, RMSE and SOR. MIKRBF performs slightly better than classic MIK, however, the computational limitation of MIKRBF is a drawback in comparison with MIK. MIKRBF outperforms the two compositional data methodologies. Even though they both provide licit CPDF, the results show that the compositional transformations are introducing a considerable bias.

Chapter 6

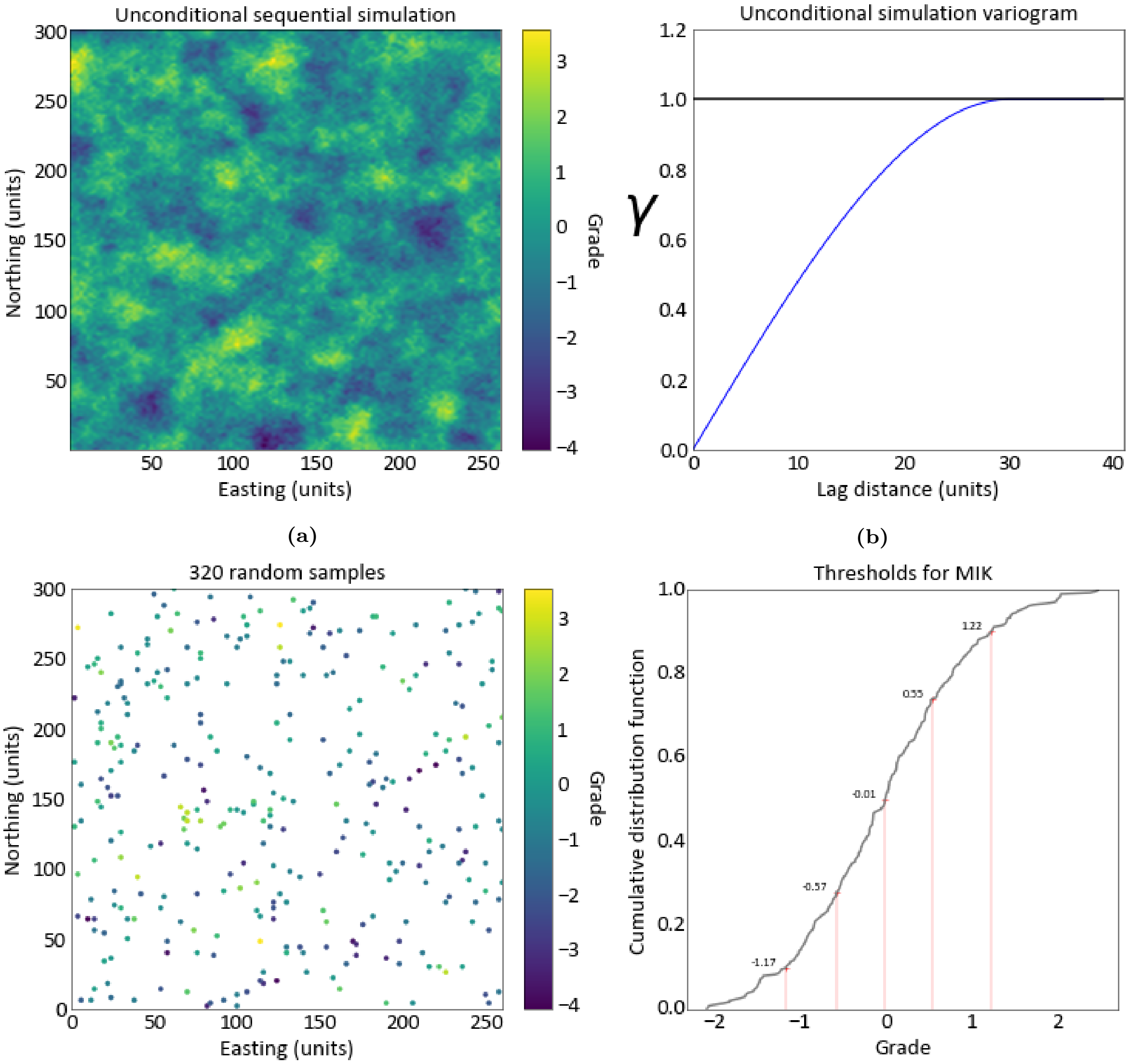
Demonstration

Chapter 4 introduces a methodology to address probabilities consistency for modeled indicator variograms of continuous variables. Chapter 5 shows a quadratic minimization methodology to add constraints to an RBF framework to avoid onder relation issues in MIK. In this chapter, both methodologies are combined into a single MIK workflow. The results are compared with the classic MIK to test the strength and limitations of both methodologies.

6.1 Methodology

The comparison involves two stages in the workflow of MIK: (1) indicator variogram modeling and (2) estimation. The MIKRBF and classic MIK methodologies are tested in an unconditional sequential simulation. A Gaussian simulation on a grid of 130x150 units with points separated by 2 units is used as reference data. The unconditional Gaussian simulation (Figure 6.1a) is obtained using the GSLIB program sgsim.exe (Deutsch & Journel, 1998) considering an omnidirectional variogram with one spherical structure with a range of 30 units (Figure 6.1b). 320 random samples are taken from the reference data to perform the estimation (Figure 6.1c). Five thresholds are chosen to discretize the CDF of the samples evenly (Figure 6.1d). Each class made by the thresholds has more than 10% of the samples (Table 6.1). The number of points in the reference data, the amount of samples and the number of thresholds are chosen to get a reasonable computation time when using the MIKRBF. Since MIKRBF uses all the samples plus the locations where order relations appear to calculate weights, a larger set of samples may be challenging to compute.

Bivariate probabilities are calculated for each lag (Figure 6.2). Since the number of samples used for the calculation is limited, bivariate probabilities will differ from the theoretical Gaussian bivariate probabilities. The bivariate probabilities are then used to calculate the experimental non-standardized variograms of each threshold at each lag. Note that the indicator experimental variogram from each threshold at a certain lag can be obtained from a single set of bivariate probabilities. The variograms are calculated by adding the probability values that fulfill the conditions $Z(\mathbf{u}) \leq z_k$ and $Z(\mathbf{u}+\mathbf{h}) > z_k$ and $Z(\mathbf{u}) > z_k$ and $Z(\mathbf{u}+\mathbf{h}) \leq z_k$. Once the indicator experimental variograms are calculated, they are modeled using the program varmodel.exe (Figure 6.3). Using the equations derived in Chapter 4, the modeled variograms are used to calculate the bivariate implicit probabilities. The underlying bivariate distributions will be consistent if the probabilities calculated from the modeled variograms are all non-negative. If the bivariate probabilities obtained from the modeled variograms are negative, the original variograms are modified. The modeled variograms are modified until there are no negative bivariate probabilities, even if the changes make the variograms deviate from the original spatial correlation. A classic MIK and MIKRBF estimations are performed using both sets of variograms. The number of order relations and their distribution are compared using bar plots. Only decreasing order relation issues between thresholds are evaluated since order relations issues related to 0 and 1 are simply reset in both algorithms. The estimations use a search volume of 100 units and a maximum of 24 samples. CCDFs are obtained at each location and then the mean is calculated. The four mean estimations are compared against the reference Gaussian simulation. The methodologies are also compared with an ordinary Kriging (OK) estimation. The ordinary Kriging estimation uses the same omnidirectional variograms used to simulate the reference data.



(a)

(b)

(c)

(d)

Figure 6.1: (a) Reference data, (b) variogram used for the unconditional sequential Gaussian simulation (c) 320 random samples extracted from the reference data (d) CCDF of the samples with thresholds

Table 6.1: Sample percentage per class

Indicators min - 1 sample %	10.0
Indicators 1-2 sample %	18.1
Indicators 2-3 sample %	21.9
Indicators 3-4 sample %	24.1
Indicators 4-5 sample %	15.9
Indicators 5-max sample %	10.0

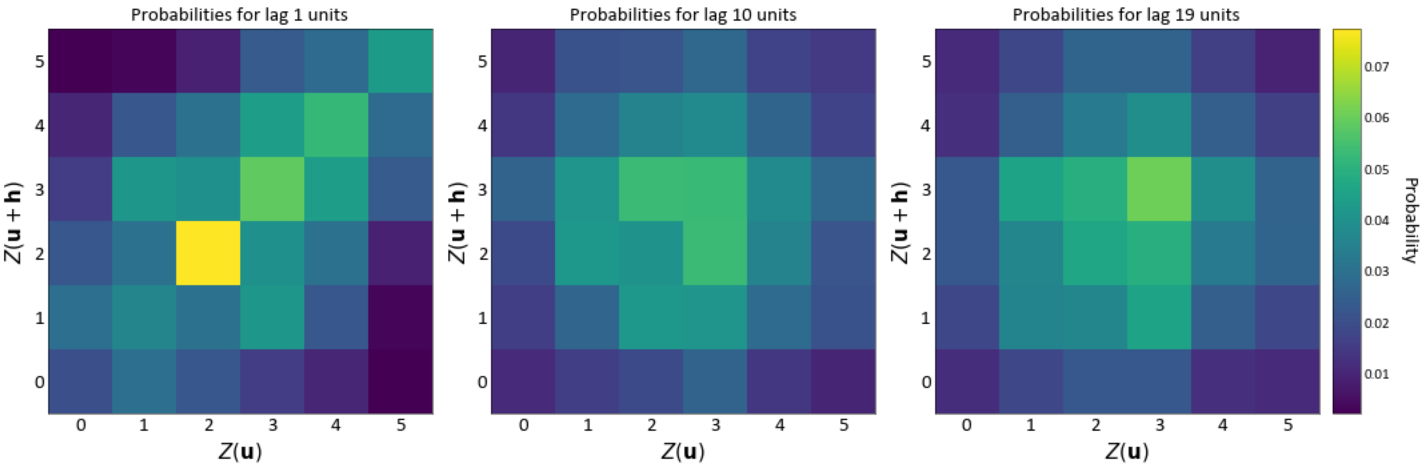


Figure 6.2: Bivariate probabilities calculated for lags 1, 10 and 19

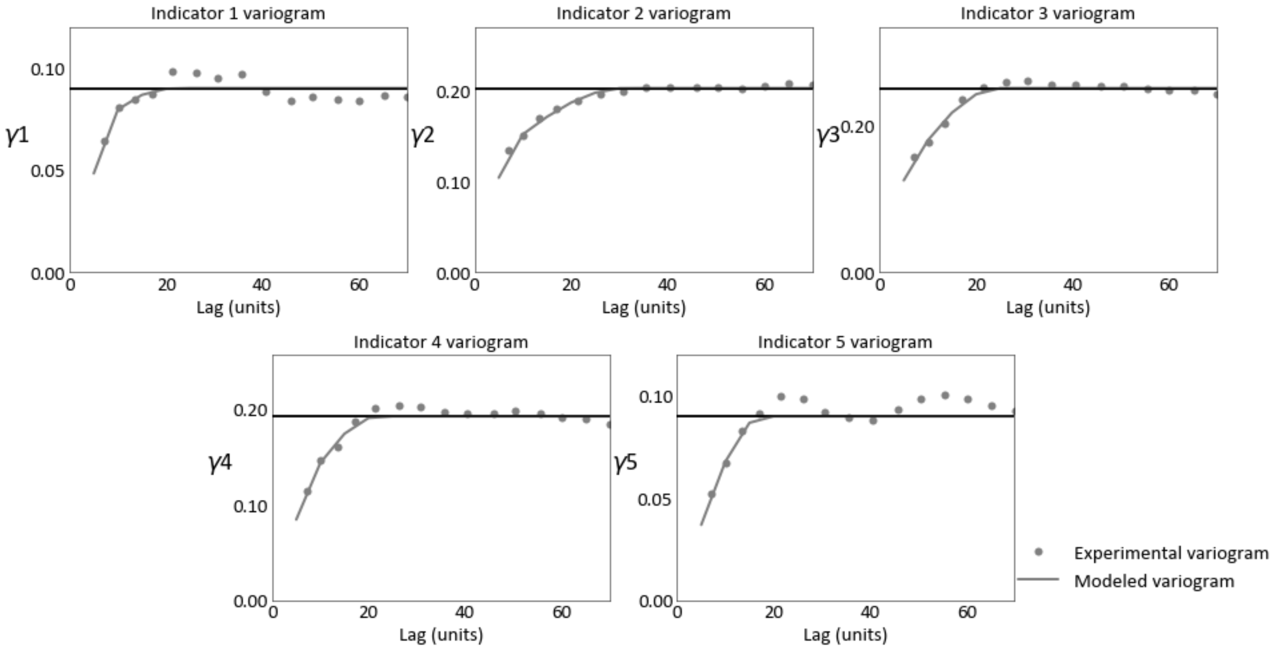
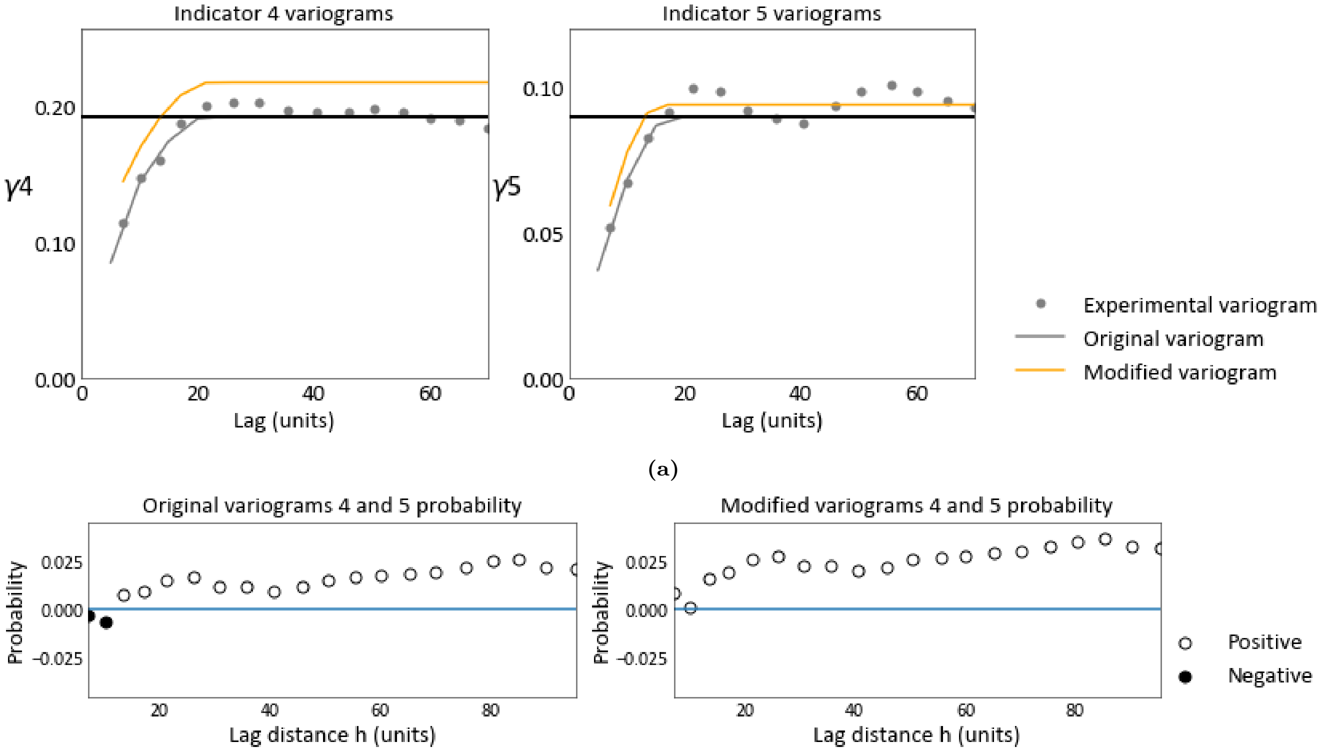


Figure 6.3: Experimental and modeled variograms of each indicator

6.2 Results

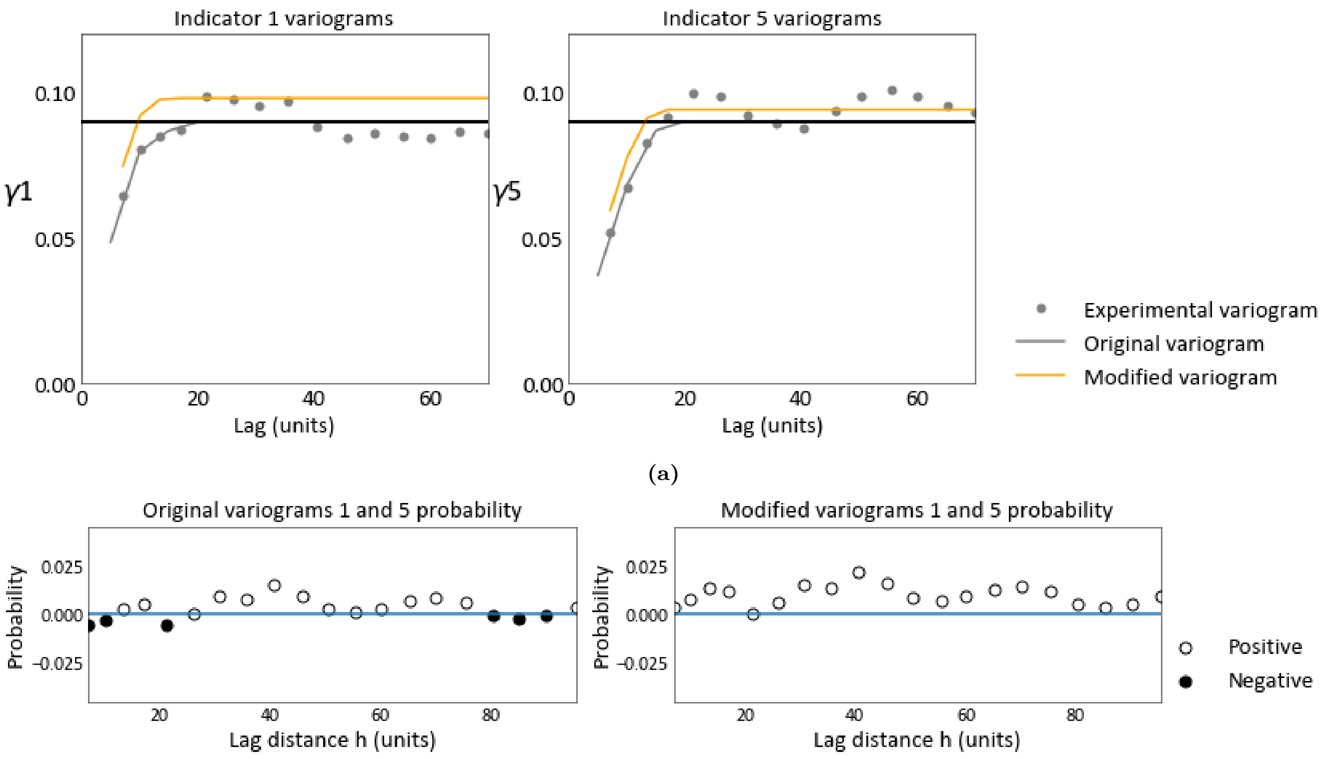
To eliminate any negative values from the bivariate distributions, the modification of the variograms must allow the deviation of variograms from the experimental values. The approach taken here to get consistent bivariate distributions without negative values is to modify the range and the sill of the variograms. The modified variograms correspond to the thresholds 1, 4 and 5 (Figures 6.4a and 6.5a). For probabilities related to indicators 1 and 5, their variograms have a positive sign in their respective probability equation. Thus, the only method to change the probability related to indicators 1 and 5 from negative to positive is to reduce the range and/or increase the sill. Similarly, in the case of the probability related to indicators 4 and 5, the variogram of indicator 4 has a positive sign in the equation, producing the same effect. Indeed, all modified variograms have a reduction in their ranges and an increase in their sills (Tables 6.2, 6.3 and 6.4). Notice that indicators 1 and 5 are also related to a probability on their own in which the variograms have a negative sign in the equations, therefore, there is a limit on how much these variograms can be modified. Indicator 1 range goes from 22.01 to 15.01 in its second structure. Indicator 4 range goes from 10.32 to 8.32 in its first structure. Indicator 5 range goes from 17.77 to 15.77 for its first structure and from 17.83 to 15.83 for its second structure. Their sill increases from 0.9 to 0.98 for indicator 1, from 0.1921 to 0.2171 for indicator 4 and from 0.09 to 0.94 for indicator 5. Increasing the sill of the variogram is considered in this study to check the behavior of MIK with consistent variograms but it is not recommended for real cases. After modifying the variograms, all negative probabilities are converted to positive (Figures 6.4b and 6.5b). The comparison shows a decrease in order relation issues, along with a change in their distribution (Figure 6.6). Order relation issues between 3-4 and 4-5 decrease whereas order relation issues between 1-2 increase. Order relation issues in the interval 2-3 remain the same due to those variograms not being modified.



(a)

(b)

Figure 6.4: (a) Original and modified variograms (b) probabilities related to variograms 4 and 5 before and after the modification



(a)

(b)

Figure 6.5: (a) Original and modified variograms (b) probabilities related to variograms 1 and 5 before and after the modification

Table 6.2: Original and modified variogram parameters for indicator 1. The modified parameters are highlighted in red

Indicator	Type	Nugget	Structure	Contribution	Range	Total Sill
1	**Original**	0.0	Sph	0.066	11.52	
			Sph	0.024	22.01	0.9
1	**Modified**	0.0	Sph	0.66	11.52	
			Sph	0.032	15.01	0.098

Table 6.3: Original and modified variogram parameters for indicator 4. The modified parameters are highlighted in red

Indicator	Type	Nugget	Structure	Contribution	Range	Total Sill
4	**Original**	0.0	Sph	0.06	10.32	
			Sph	0.1321	22.08	0.1921
4	**Modified**	0.0	Sph	0.085	8.32	
			Sph	0.1321	22.08	0.2171

Table 6.4: Original and modified variogram parameters for indicator 5. The modified parameters are highlighted in red

Indicator	Type	Nugget	Structure	Contribution	Range	Total Sill
5	**Original**	0.0	Sph	0.041	17.77	
			Sph	0.049	17.83	0.9
5	**Modified**	0.0	Sph	0.045	15.77	
			Sph	0.049	15.83	0.094

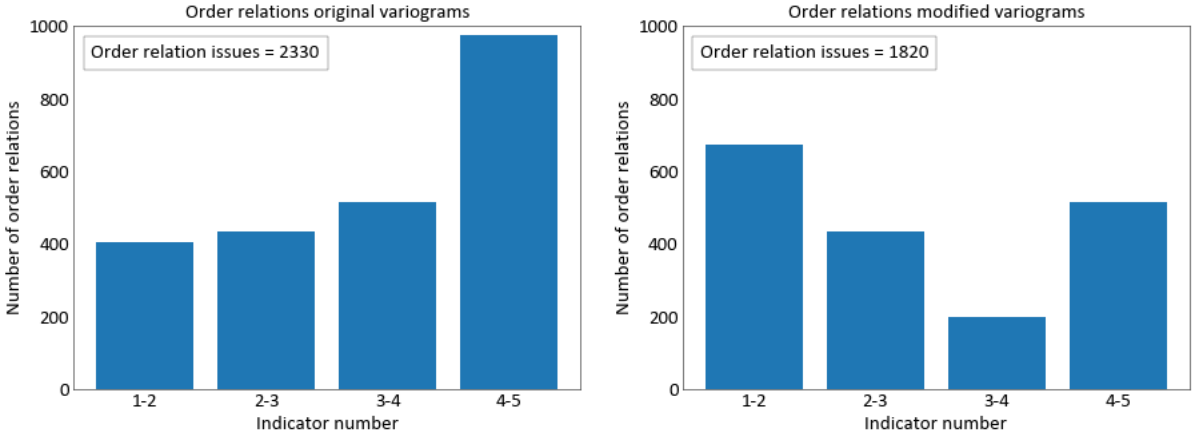


Figure 6.6: Number of order relation issues for original and modified variograms

Table 6.5: E-types statistics

Method	Variograms	Mean	RMSE	SOR
MIK	**Original**	0.002	0.779	1.276
MIK	**Modified**	0.001	0.787	1.326
MIKRBF	**Original**	-0.29	0.778	1.320
MIKRBF	**Modified**	-0.30	0.786	1.368
OK		-0.30	0.705	1.017
True Values		-0.067		

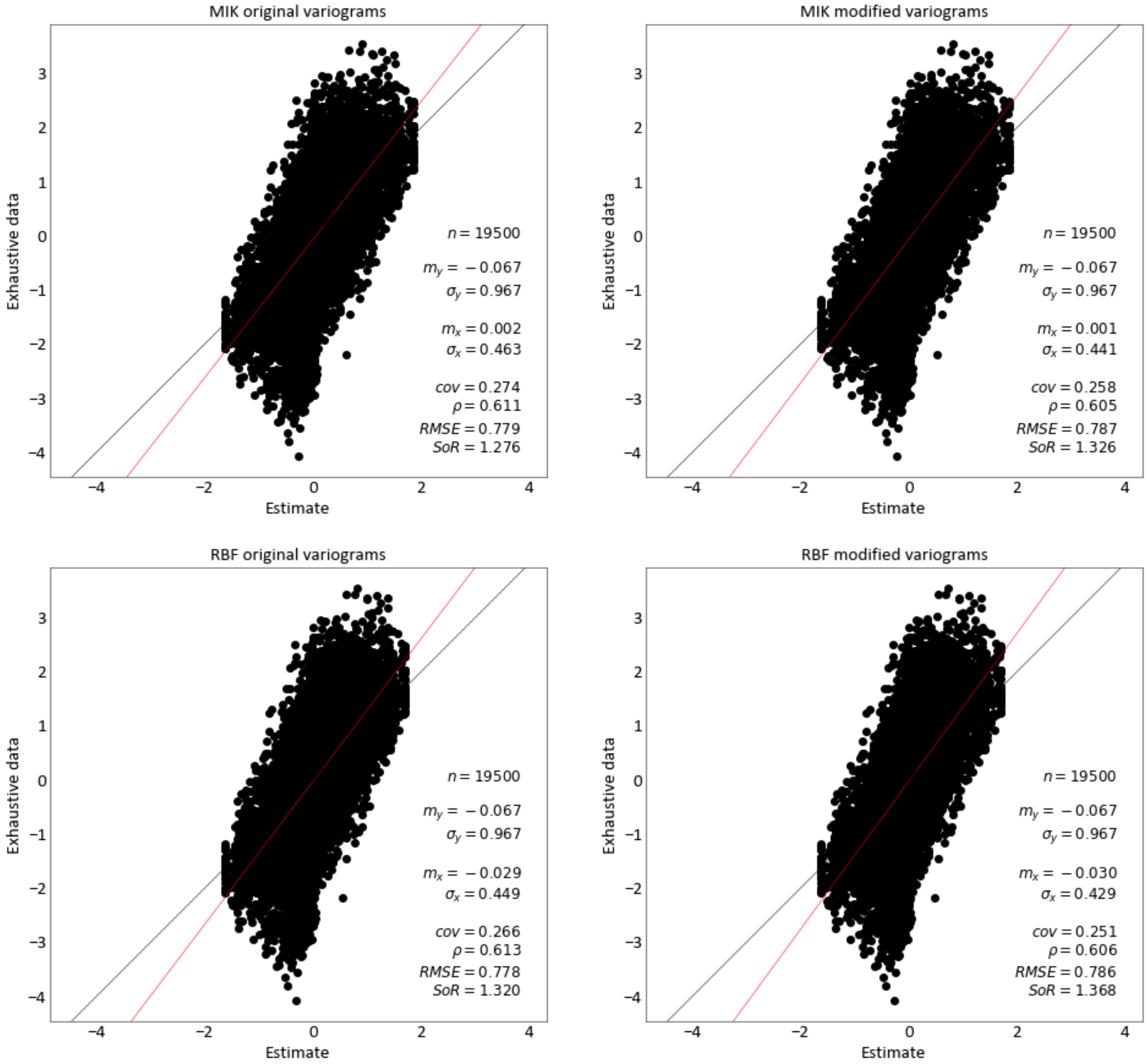


Figure 6.7: Validation plot and statistics of classic MIK and the new methodology in comparison with the reference data

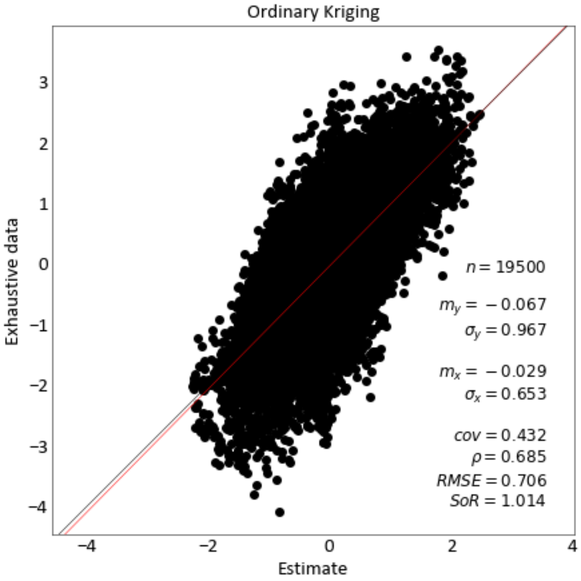


Figure 6.8: Validation plot and statistics of OK in comparison with the reference data

The estimates using the 2 sets of variograms and the two MIK methodologies and the OK estimates are compared with the reference data (Figure 6.7, Figure 6.8 and Table 6.5). The global mean for classic MIK goes from 0.002 to 0.001 and for MIKRBF from -0.029 to -0.030 for original and modified variograms respectively. The RMSE for classic MIK goes from 0.779 to 0.787 and for MIKRBF from 0.778 to 0.786 for original and modified variograms respectively. The slope of regression for classic MIK goes from 1.276 to 1.326 and for MIKRBF from 1.320 to 1.368 for original and modified variograms respectively. In both methodologies, the modification of the variograms makes the global mean closer to the reference mean of -0.067 but the RMSE gets slightly worse. By comparing the methodologies with the same set of variograms, MIKRBF gets a mean closer to the reference data and almost the same RMSE. MIK has the best slope of regression. The results do not show a significant difference between MIK and MIKRBF. On the other hand, OK has a smaller RMSE and a SOR close to one. The two MIK methodologies estimates are over smooth in comparison to OK.

The deviation of the variograms from the experimental values causes the estimates to have a larger RMSE, even with fewer order relation issues. Thus, a two points statistics variogram model might not be suitable to describe the spatial correlation of indicators since it fails to comply with a consistent bivariate distribution. An optimization program could consider both the experimental variogram values and the probabilities calculated from the equations to optimize models that fit the spatial correlation but also are closer to a consistent bivariate distribution. In addition, the experimental probabilities used in the equations might be biased with respect to the real data. In that sense, an uncertainty model could be built for those probabilities instead of using the experimental probabilities.

Chapter 7

Conclusions

This thesis work reviewed three topics regarding MIK and order relation issues: (1) The relation between order relation issues and negative weights, (2) indicator variograms consistency related to a bivariate distribution, (3) the use of RBF with constraints to avoid order relation issues. This chapter summarizes the main contributions and limitations of this research. Future work is also proposed.

7.1 Summary of contributions

This research made several contributions. First, it presented original tests to investigate the behavior of order relation issues in MIK related to negative weights and indicator variograms. Second, novel equations were derived to calculate probabilities of a bivariate distribution from modeled indicator variograms and proportions. Third, an indicator variogram modeling workflow was presented to help practitioners model consistent indicator variograms. Fourth, a methodology using RBF with constraints was developed to avoid order relation issues in MIK.

7.1.1 Order relation issues and negative weights

Order relation issues in MIK are controlled by weights, the spatial distribution of samples and the change of indicator values from one indicator to another. Variograms with better continuity produce zones of negative weights due to the screening effect of Kriging. Order relation issues are related to negative weights in any of the following settings. (1) Identical indicator variograms but different indicators: If an indicator variable value changes from 0 to 1 from one indicator to the next one on a sample with negative weight, the estimate will be smaller than the previous one producing an order relation issue. (2) Identical indicator values and different variograms: If an indicator variogram has a stronger screening effect than the previous one, it will produce more negative weights, increasing the chances of an order relation issue. (3) Different indicators and different variograms: This is a combination of the two previous settings. Order relation issues are not only produced by the presence of negative weights. If consecutive threshold indicator values do not change considerably, which is common on low and high-grade thresholds, order relation issues could arise related to only positive weights. Negative weights are still the main reason for order relation issues.

The median MIK methodology is a way of reducing order relations considerably. Since all indicator estimates use the same variogram, all indicators will have the same weights. This only produces order relation issues of type (1) which decreases the probability of them happening. The use of the same weights also changes the distribution of the order relation per threshold. It makes all thresholds have a similar number of order relation issues. The use of the median variogram also eliminates all order relation issues related to only positive weights. It is impossible to have order relation issues only with positive weights if the weights are the same. The median MIK produces slightly better estimates than the classic MIK for the dataset used in this thesis but it fails to consider all the spatial correlation information for different thresholds.

Dissimilar variograms produce more order relation issues and change the distribution of order relation issues between the thresholds. The more dissimilar variograms are the more order the relation issues will be. The lack of samples usually produces less consistent and dissimilar variograms which will lead to more order relations issues.

7.1.2 Internal bivariate probabilities of modeled indicator variograms

This work provided novel equations for the calculation of probabilities of the internal bivariate distribution of a set of modeled indicator variograms. The inputs to derive the equations are the modeled variograms and the expected values or proportions of the indicators. The first and the last probability can be calculated using only one modeled variogram. The rest of the probabilities compromises two consecutive modeled variograms. The equations are proven to work in a multi-Gaussian case. They satisfactorily reproduce the probabilities of the experimental bivariate distribution when enough density of samples is used. The sensitivity analysis shows that the fewer samples, the more deviation the probabilities have from the experimental probabilities. This characteristic does not mean that the modeled indicator variograms are inconsistent. It means that the modeled indicator variograms are related to another bivariate distribution. The probabilities derived from modeled indicator variograms have to be negative to be inconsistent. In that sense, the calculated probabilities can be used to help model the variograms so they maintain consistency with a bivariate distribution. A methodology is proposed for this purpose. Indicator variograms are fitted to the experimental indicator variograms. Probabilities of the bivariate distribution are calculated using the modeled indicator variograms, proportions and experimental probabilities. Each probability is checked to identify lags where they are negative. Variograms are modified until there are no negative probabilities. Then, the variograms will be consistent with a bivariate distribution. The practitioner must take into consideration that the modification of one variogram can alter the value of multiple probabilities.

7.1.3 RBF interpolation with inequality constraints

Local distribution inference by MIK has the advantage of capturing spatial features at different cut-off grades and it does not need any assumption in the prior distribution of the samples. The main drawback in the methodology is order relation issues. Indicators are estimated independently using Kriging, which can produce negative weights. This means that order relations are common. The framework versatility of RBF interpolation allows for the addition of inequality constraints in the calculation of the weights. The constraints do not affect the estimation if the constraints are not infringed. On the other hand, if the constraints are not satisfied, the value of the estimate will be modified to the value of the constraint. Locations in a radius equal to the range of the RBF are also affected. The ability to include constraints in RBF is beneficial for the MIK formulation since the estimated probabilities have to follow inequality constraints to constitute a licit CCDF. A sequential RBF interpolation with inequality constraints was developed to account for order relations issues and obtain local licit CCDF. The methodology only considers decreasing order relation issues. It first estimates the bottom threshold by RBF without constraints and fixes the 0 and 1 order relation issues manually. Then, it estimates the second threshold without constraints. It identifies all locations with decreasing order relations issues and reruns the estimation using RBF with constraints in those locations. The order relation issues related to 0 and 1 are reset. This process is repeated sequentially until the last threshold. The methodology can be applied from bottom to top and top to bottom and averaged to get a final result. The novel methodology presents a similar performance to MIK and MGK and outperforms compositional data methodologies for the dataset used in this thesis.

7.2 Limitations and future work

The methodology to maintain consistently modeled indicator variograms was tested in multiple real datasets. For the real datasets cases, the variograms must be modified considerably from the experimental variograms to ensure for positiveness in the probabilities of a bivariate distribution. This deviation includes great changes in the ranges and sill of the variograms. The amount of changes is determined by the probability with negative values. Most probability equations contain two modeled variograms with opposite signs to calculate the probability, although, making the variogram model with the positive sign less continuous or adding a bigger sill to it is the easiest way to convert negative probabilities to positive ones. The deviation of the original correlation of the samples makes the estimation mostly worse in comparison to classic MIK. Even though modified variograms produce fewer order relation issues in some cases, the performance of the estimation does not seem related to this. In that sense, the modeled variograms are not able to capture the spatial correlation between the indicators and the consistency of bivariate distributions. A solution to this problem would be to accept some degree of inconsistency in the probabilities. An optimization algorithm could be developed to fit the experimental indicator variograms while maintaining the probabilities to be as positive as possible.

For this work, the central probabilities of the experimental bivariate distribution were chosen to calculate the rest of the probabilities. This is not a necessary condition to prove the consistency of modeled variograms. In general, any probability could be assumed for the equations. In that sense, an uncertainty model could be used for these probabilities. Then, the probability of a set of variograms being consistent could be determined. This approach may also reduce the degree of changes to the modeled variograms.

In the case of MIKRBF, even though order relation issues are avoided by using RBF with constraints, order relations issues related to estimated probabilities below 0 and above 1 are reset in a post-process. Adding more than one constraint in a single location led to inconsistent results and order relation issues. The ideal scenario would be to find a method to add these constraints to the system so that all order relation issues are avoided directly from the estimation. Assigning the constraints in slightly different locations may be a good solution.

As RBF with constraints is a global interpolation methodology, it uses all the samples to calculate the weights. Moreover, each inequality constraint adds a new element to the weights calculation system. In that sense, the methodology is practical only if the number of samples plus the number of order relations issues is less than 10,000. This is not feasible for real-life 3D modeling problems. The methodology could be performed in combination with other global interpolation techniques that account for the computational limitation like the partition of unity (De Rossi & Perracchione, 2017).

7.3 Final remarks

This work aimed to understand different MIK characteristics and solve classic problems related to the methodology. Even though new algorithms are a crucial part of the development of geostatistics, re-visiting classic methodologies and tackling lifetime problems is valuable in that it helps in building consistent models in the future. This research contributes in that sense by expanding the knowledge on negative weights, addressing the consistency of modeled indicator variograms, and proposing another view on how to avoid order relation problems in MIK.

References

Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J. A., & Pawlowsky-Glahn, V. (2000). Logratio analysis and compositional distance. *Mathematical geology*, *32*(3), 271–275.

Armstrong, M., & Matheron, G. (1986). Disjunctive kriging revisited: Part i. *Mathematical Geology*, *18*, 711–728.

Barnes, R. J., & Johnson, T. B. (1984). Positive kriging. In *Geostatistics for natural resources characterization: Part 1* (pp. 231–244). Springer.

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

Carr, J. R., & Mao, N.-h. (1993). A general form of probability kriging for estimation of the indicator and uniform transforms. *Mathematical geology*, *25*, 425–438.

Chiles, J.-P., & Delfiner, P. (2009). *Geostatistics: modeling spatial uncertainty* (Vol. 497). John Wiley & Sons.

Cowan, E., Beatson, R. K., Ross, H. J., Fright, W. R., McLennan, T., Evans, T. R., … others (2003). Practical implicit geological modelling. In *5th international mining geology conference* (Vol. 8, pp. 89–99).

De Rossi, A., & Perracchione, E. (2017). Positive constrained approximation via rbf-based partition of unity method. *Journal of Computational and Applied Mathematics*, *319*, 338–351.

Deutsch, C. V. (1996). Correcting for negative weights in ordinary kriging. *Computers & Geosciences*, *22*(7), 765–773.

Deutsch, C. V., & Journel, A. G. (1998). Gslib: Geostatistical software library and user's guide. *Oxford University Press*, p. 369.

Emery, X. (2005). Variograms of order ω: a tool to validate a bivariate distribution model. *Mathematical Geology*, *37*(2), 163–181.

Emery, X., & Ortiz, J. M. (2004). Shortcomings of multiple indicator kriging for assessing local distributions. *Applied Earth Science*, *113*(4), 249–259.

Fasshauer, G. E. (2007). *Meshfree approximation methods with matlab* (Vol. 6). World Scientific.

Goovaerts, P. (1994). Comparative performance of indicator algorithms for modeling conditional probability distribution functions. *Mathematical Geology*, *26*, 389–411.

Goovaerts, P., et al. (1997). *Geostatistics for natural resources evaluation.* Oxford University Press on Demand.

Hadavand, Z., & Deutsch, C. (2021). *Probabilistic prediction with machine learning and a compositional approach to probabilities.* Retrieved from `http://www.ccgalberta.com`

Hillier, M. J., Schetselaar, E. M., de Kemp, E. A., & Perron, G. (2014). Three-dimensional modelling of geological surfaces using generalized interpolation with radial basis functions. *Mathematical*

Geosciences, *46*, 931-953.

Isaaks, E. H., & Srivastava, R. M. (1989). *Applied geostatistics* (Vol. 561). Oxford University Press, New York.

Journel, A. G. (1983). Nonparametric estimation of spatial distributions. *Journal of the International Association for Mathematical Geology*, *15*, 445-468.

Journel, A. G. (1984). The place of non-parametric geostatistics. In *Geostatistics for natural resources characterization* (pp. 307–335). Springer.

Journel, A. G., & Posa, D. (1990). Characteristic behavior and order relations for indicator variograms. *Mathematical Geology*, *22*, 1011-1025.

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

Marechal, A. (1976). The practice of transfer functions: Numerical methods and their application. In *Advanced geostatistics in the mining industry: Proceedings of the nato advanced study institute held at the istituto di geologia applicata of the university of rome, italy, 13–25 october 1975* (pp. 253–276).

Matheron, G. (1963). Principles of geostatistics. *Economic geology*, *58*(8), 1246–1266.

Matheron, G. (1976). A simple substitute for conditional expectation: the disjunctive kriging. In *Advanced geostatistics in the mining industry: Proceedings of the nato advanced study institute held at the istituto di geologia applicata of the university of rome, italy, 13–25 october 1975* (pp. 221–236).

Matheron, G. (1989). The internal consistency of models in geostatistics. In *Geostatistics* (pp. 21–38). Springer.

Mokdad, K., Binakaj, D., & Boisvert, J. B. (2022). *Data validation project: Validation of 114 spatial 2d datasets nonsynthetic data.* Retrieved from `http://www.ccgalberta.com`

Myers, D. E. (1982). Matrix formulation of co-kriging. *Journal of the International Association for Mathematical Geology*, *14*(3), 249–257.

Myers, D. E. (1984). Co-kriging—new developments. In *Geostatistics for natural resources characterization* (pp. 295–305). Springer.

Ortiz, J. M. (2019). Multigaussian kriging: a review. *Queen's University.*

Osher, S., & Fedkiw, R. (2003). Signed distance functions. In *Level set methods and dynamic implicit surfaces* (pp. 17–22). Springer.

Rao, S. E., & Journel, A. G. (1997). Deriving conditional distributions from ordinary kriging. *Geostatistics Wollongong*, *96*, 92–102.

Remacre, A. Z. (1987). Conditioning by the panel grade for recovery estimation of non-homogeneous orebodies. In *Geostatistical case studies* (pp. 135–148). Springer.

Rossi, M. E., & Deutsch, C. V. (2013). *Mineral resource estimation.* Springer Science & Business Media.

Roth, C., & Deraisme, J. (2000). The information effect and estimating recoverable reserves. In *Kleingeld, wj, krige dg (eds), proceedings of the sixth international geostatistics congress* (pp. 776–787).

Soltani-Mohammadi, S., & Tercan, A. E. (2012). Constrained multiple indicator kriging using sequential quadratic programming. *Computers & Geosciences*, *48*, 211–219.

Stewart, M., de Lacey, J., Hodkiewicz, P. F., & Lane, R. (2014). Grade estimation from radial basis functions–how does it compare with conventional geostatistical estimation. In (Vol. 129, p. 139).

Sullivan, J. (1984). Conditional recovery estimation through probability kriging—theory and practice. *Geostatistics for natural resources characterization*, 365-384.

Suro-Perez, V., & Journel, A. (1991). Indicator principal component kriging. *Mathematical Geology*, *23*, 759–788.

Tolosana-Delgado, R., Mueller, U., & van den Boogaart, K. G. (2019). Geostatistics for compositional data: an overview. *Mathematical geosciences*, *51*, 485-526.

Tolosana-Delgado, R., Pawlowsky-Glahn, V., & Egozcue, J.-J. (2008). Indicator kriging without order relation violations. *Mathematical geosciences*, *40*, 327-347.

Vann, J., & Guibal, D. (1998). Beyond ordinary kriging–an overview of non-linear estimation. In *Proceedings of a one day symposium: Beyond ordinary kriging* (p. 32).

Verly, G. (1983). The multigaussian approach and its applications to the estimation of local reserves. *Journal of the International Association for Mathematical Geology*, *15*, 259-286.

Xiao, H. (1985). A description of the behavior of indicator variograms for a bivariate normal distribution. *Master's thesis, Stanford University.*