

University of Alberta

Block LU simulation with an Approximate Model of Coregionalization

by

Tong Wang

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

Master of Science
in
Mining Engineering

Department of Civil and Environmental Engineering

©Tong Wang
Fall 2009
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Examining Committee:

Dr. Clayton V. Deutsch (Supervisor). Department of Civil and Environmental Engineering.

Dr. Oy Leuangthong (Examiner). Department of Civil and Environmental Engineering.

Dr. Qi Liu (Examiner). Department of Chemical and Materials Engineering.

Never, ever give up

Abstract

Geostatistical techniques are used to estimate recoverable reserves at unsampled locations and to quantify uncertainty. Several variables are often measured and important for reserve evaluation. Using more variables improves the quality of modeling, but quantifying the relationships between the variables is difficult. The traditional linear model of coregionalization has been used to quantify the relationship between multiple variables, but ensuring the mathematical validity of the model is cumbersome. This research proposes an approximate method that improves the speed and practicality of the numerical modeling process by easily modeling multiple regionalized variables. The proposed algorithm is based on block LU simulation and takes local transformation into consideration. Application to a nickel deposit demonstrates the proposed methodology.

Acknowledgements

I would like to extend my thanks and appreciation to the members of the Centre for Computational Geostatistics at the University of Alberta. I would especially like to thank my advisor, Dr. Clayton V. Deutsch. It would not be possible to complete this thesis without the support and insightful guidance.

I dedicate this thesis to my family. To my father, mother and my sister for their encouragement and love that keep me move towards my goal. To my friend Dong, I thank you for being the special person you are.

Finally, I would like to thank all others who have helped in my life.

Table of Contents

Abstract	
Acknowledgements	
Chapter 1 Introduction.....	1
1.1 Problem setting	1
1.2 Dissertation outline.....	3
Chapter 2 Background	4
2.1 Random function concept.....	4
2.2 Linear model of coregionalization.....	5
2.3 Kriging	7
2.4 Techniques for multivariable block LU simulation.....	8
2.4.1 Unconditional LU simulation.....	9
2.4.2 Conditional LU simulation.....	10
2.4.3 Computational aspects of LU simulation.....	12
Chapter 3 Improving LU simulation methodology.....	14
3.1 Approximate model of coregionalization	14
3.1 Transformation.....	16
Chapter 4 Case study.....	19
4.1 Data.....	19
4.2 Variogram analysis.....	22
4.3 Comparative simulation with AMC and LMC.....	25
4.4 Sensitivity analysis.....	35
4.5 CPU-time study.....	37
4.6 Transformation study.....	39
Chapter 5 Conclusions and future work.....	42
Bibliography	
Appendix: A program blusim_AMC.....	A-1

List of Tables

Table 4.1 Summary statistics for Ni, Fe and MgO.....	21
Table 4.2 Correlation between variables.....	22
Table 4.3 Summary simulation statistics for Ni, Fe and MgO.....	27
Table 4.4 Summary variance statistics Ni, Fe and MgO.....	27

List of Figures

Figure 3.1 Block discretization with search data.....	15
Figure 3.2 Double search radiuses using the local transformation	17
Figure 3.3 Locally transformation illustrations.....	18
Figure 4.1 Location map of drillhole data.....	20
Figure 4.2 Histogram for Ni, Fe and MgO.....	20
Figure 4.3 Declustered mean versus Cell size for variable Ni.....	21
Figure 4.4 Histograms for Ni, Fe and MgO by using declustering weights.....	21
Figure 4.5 Scatter plot of Ni, Fe and MgO.....	22
Figure 4.6 Ni, Fe and MgO variogram and cross variogram fitted by LMC.....	24
Figure 4.7 Ni, Fe and MgO variogram model fitted separately.....	25
Figure 4.8 Comparative cross variogram model between LMC and AMC.....	25
Figure 4.9 Block averaged simulation histograms using LMC.....	28
Figure 4.10 Block averaged simulation histograms using AMC.....	28
Figure 4.11 Block global variance using LMC.....	29
Figure 4.12 Block global variance using AMC.....	29
Figure 4.13 Scatter plot of block averaged simulation using AMC and using LMC...30	
Figure 4.14 Block averaged simulation plots using LMC.....	30
Figure 4.15 Block averaged simulation plots using AMC.....	31
Figure 4.16 Block global variance plots using LMC.....	31
Figure 4.17 Block global variance plots using AMC.....	32
Figure 4.18 Ni mean grades and proportion of grid above cutoffs using LMC.....	33
Figure 4.19 Ni mean grades and proportion of grid above cutoffs using AMC.....	34
Figure 4.20 Grade tonnage cure of Ni, Fe and MgO using different block size.....	34
Figure 4.21 Std.dev of Ni, Fe and MgO versus the data used in the search	36
Figure 4.22 Std.dev of Ni, Fe and MgO results versus discretization points used	37
Figure 4.23 A case for two blocks using the same discretization.....	38
Figure 4.24 CPU-time study on improved algorithm.....	38
Figure 4.25 Block averaged simulation histogram using local transformation.....	39
Figure 4.26 Block averaged simulation plots using local transformation	40
Figure 4.27 Block global variance histograms using local transformation	40
Figure 4.28 Block global variance plots using local transformation	41

List of Symbols

u	Location in space
h	Distance vector
b_{ij}	Contribution to each variogram structure
m	mean value
$F(z)$	Cumulative distribution function of Z
$\gamma(\mathbf{h})$	Variogram function at lag h
$\gamma_{ZY}(\mathbf{h})$	Cross variogram between the two variables
$C(0)$	Covariance function at lag h=0
$C(h)$	Covariance function at lag h
$E\{\}$	Expected values
$Y(\mathbf{u})$	Residual data
$Z(\mathbf{u})$	A random variable
σ^2	Variance
σ^2_{sk}	Simple kriging Variance
σ^2_{sck}	Cokriging Variance
AMC	Approximate model of coregionalization
BLUE	Best linear unbiased estimate
PDF	Probability density function
CDF	Cumulative distribution function
GSLIB	Geostatistical software library
LU	Lower and Upper decomposition
LMC	Linear model of coregionalization
RV	Random variable
SK	Simple kriging
Spd	A spherical variogram structure
Exp	A exponential variogram structure
SGS	Sequential Gaussian simulation
BLUSIM	Block Lower and Upper decomposition simulation

Chapter 1

Introduction

1.1 Problem setting.

Geostatistics is a relatively new branch of applied statistics, geology and computer science. It has been applied to many spatial problems including reservoir characterization and ore body modeling. Geostatistical techniques are used to estimate recoverable reserves at unsampled locations and also to quantify uncertainty. For the mining engineer, as well as the metallurgist and chemist, it is often essential to be able to predict the variations in the characteristics of the ore at various stages in the operation (Journel and Huijbregts, 1978). The ore is represented by several correlated variables. Quantifying the relationship between many variables is difficult. This thesis is focused on improving numerical modeling by simultaneous modeling of multiple regionalized variables.

Assessment of uncertainty is also important for numerical models of mineral deposits. Uncertainty and variability must be taken into consideration in decision making. Geostatistical simulation generates realizations that depict the fluctuations of the variable. The realizations honor the data and reproduce spatial correlation of the regionalized variables. The variability between realizations is a measure of spatial uncertainty.

The geostatistical realizations are drawn from local conditional distributions that are determined by using kriging techniques. The kriging system accounts for data redundancy, closeness of the data to what is being predicted and spatial/multivariate correlations.

A model of coregionalization specifies the spatial relationship of each variable and the cross correlation between all pairs of variables. The traditional linear model of coregionalization is cumbersome and the practical application of this model is difficult.

There are several geostatistical methods to estimate and model uncertainty. Most methods are based on the multivariate Gaussian distribution after transforming each variable to a Gaussian histogram. The common algorithm for generating realizations from a multivariate Gaussian distribution is Sequential Gaussian simulation.

Often, estimates are required at a practical mining scale and not the scale of the drill core data. The basic idea is to discretize each block into many points. Each point is simulated and the final value is given to that block by averaging the discretized points. This could be very CPU expensive if there are millions of blocks to be simulated. The cost for many variables becomes even greater.

The Lower and Upper (LU) decomposition technique was first proposed in geostatistical simulation by Davis (1987). When the total number of conditioning data plus the number of nodes to be simulated is small and a large number of realizations is requested, simulation through LU decomposition of the covariance matrix provides the fastest solution (Deutsch, 1998).

All variables must be accounted for simultaneously. Multivariate geostatistics gives a tool to integrate multiple variable information, which will improve the quality of reservoir estimates. The collocated cosimulation builds on a Markov-type hypothesis (Almeida, 1994). Another variant of cokriging (Babak, 2008) considers only the single most correlated or most relevant secondary variable.

The objective of this research is to improve the accounting for multiple variables. It is aimed at using all relevant variables at the same time. The development will extend cosimulation through the LU algorithm. An approximate model of coregionalization will be developed to infer the correlation between multiple variables. Another development is the transformation algorithm. A local search for relevant data will be considered. It takes local variations into consideration. The data used in transformation is different at each location. Theoretical and practical implementation details will be explored.

1.2 Dissertation Outline

Chapter 2 discusses the background required in this thesis. Basic geostatistical concepts and methods are described. Chapter 3 focuses on improving the Block LU simulation methodology. Chapter 4 documents a case study with a real 2D nickel deposit using the proposed methodology. Chapter 5 gives conclusions and future work.

Chapter 2

Background

This chapter presents an overview of geostatistical background for mineral resources. Many geostatistical techniques have evolved over the last fifty years, but the basic statistical and mathematical theories remain the same and are essential to these techniques. Based on the random function concept and stationarity, geostatistics deals with predictions at unsampled locations. This requires a spatial correlation model based on the available information. Simulation techniques sample a specified random function model.

2.1 Random Function.

The sample data are viewed as outcomes of a random function. A random variable (RV) Z is a variable that can take a variety of outcome values according to some probability density function (pdf) or equivalently a cumulative distribution function (cdf). The location-dependence of a RV is denoted by $Z(\mathbf{u})$, where \mathbf{u} is a location within the domain of interest A . A specific outcome at \mathbf{u} is denoted by $z(\mathbf{u})$. The cumulative distribution function (cdf) of a RV $Z(\mathbf{u})$ is:

$$F_Z(\mathbf{u}; z) = Prob\{Z(\mathbf{u}) \leq z\} \quad (2.1)$$

If there are K dependent random variables, the random function is defined as follows (Deutsch 1988):

$$F(\mathbf{u}_1, \dots, \mathbf{u}_k; z_1, \dots, z_k) = Prob\{Z(\mathbf{u}_1) \leq z_1, \dots, Z(\mathbf{u}_k) \leq z_k\} \quad (2.2)$$

A decision of stationarity is made in order to substitute the need for repetitive realizations at all locations for scattered sampling at sample locations (Journel and Huijbregts 1978). The assumption of stationarity is required to make predictions at unsampled locations. The first order of stationarity assumes that the mean of the variable is constant in the domain. The second order of stationarity assumes that the

variance and covariance of the data are constant in the domain.

The spatial relationship is measured by the covariance:

$$C(\mathbf{h}) = E\{Z(\mathbf{u} + \mathbf{h})Z(\mathbf{h})\} - E\{Z(\mathbf{u})\}E\{Z(\mathbf{u} + \mathbf{h})\} \quad (2.3)$$

where \mathbf{h} is a distance vector. The covariance is required for all practical \mathbf{h} values within a search radius. If \mathbf{h} is zero, the covariance becomes the variance under stationary assumption.

$$C(0) = E\{Z(\mathbf{u})^2\} - E\{Z(\mathbf{u})\}^2 = Var\{Z\} = \sigma^2 \quad (2.4)$$

The variogram is a geostatistical tool to quantify spatial variability. It is defined as the expected squared difference between two sample values separated by a distance vector:

$$\gamma(\mathbf{h}) = \frac{1}{2} E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2\} \quad (2.5)$$

Under the stationary assumptions, the relationship between variogram and covariance is $\gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$ (Leuangthong, Khan and Deutsch, 2008).

In multivariate geostatistics, under the same assumption, the spatial relationship between different variables is defined by the cross variogram:

$$\gamma_{ZY}(\mathbf{h}) = \frac{1}{2} E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})][Y(\mathbf{u}) - Y(\mathbf{u} + \mathbf{h})]\} : \quad (2.6)$$

If the variables are normalized, the above is equal to:

$$\begin{aligned} \gamma_{ZY}(\mathbf{h}) &= C_{ZY}(0) - \frac{1}{2}[C_{ZY}(\mathbf{h}) + C_{YZ}(\mathbf{h})] \\ &= \rho_{ZY}(0) - \rho_{ZY}(\mathbf{h}) \end{aligned} \quad (2.7)$$

Where ρ_{ZY} is the correlation coefficient for the pairs of variables, C_{ZY} is the cross covariance between Z and Y, C_{YZ} is the cross covariance between Y and Z.

2.2 Linear model of coregionalization (LMC)

The linear model of coregionalization provides a method for modeling the auto- and cross-variograms (Isaaks and Srivastava, 1989). It is useful for the analysis and interpretation of multivariate spatial information.

Consider P stationary random functions, $\mathbf{Z} = \{Z_1, \dots, Z_P\}$. The particularity of the

LMC is that all $P(P+1)/2$ variograms are modeled as linear combinations of the set of L basic variogram functions. The LMC can be written as a multivariate nested covariance model:

$$C_{i,j}(\mathbf{h}) = \sum_{l=0}^L b_{i,j}^l c_l(\mathbf{h}) \quad i, j = 1 \dots P \quad (2.8)$$

Where $b_{i,j}$ is the contribution to each structure. The conditions sufficient for the matrix of functions $C_{i,j}(\mathbf{h})$ to be a permissible model of coregionalization are: (1) the functions $c_l(\mathbf{h})$ are permissible covariance models, and (2) the (L+1) coregionalization matrices B_l are all positive definite, which requires that $b_{ii}^l \cdot b_{jj}^l \geq b_{ij}^l \cdot b_{ji}^l, \forall i, j, l$

We can define the Linear model of coregionalization by direct and cross variogram model as follows:

$$\gamma_{i,j}(\mathbf{h}) = \sum_{l=0}^L b_{i,j}^l g_l(\mathbf{h}) \quad i, j = 1, \dots, P \quad (2.9)$$

Where the $g_l(\mathbf{h})$ is the same structure for all direct and cross variogram. Equation 2.9 can be expressed in matrix notation as:

$$\begin{bmatrix} \gamma_{11}(\mathbf{h}) & \cdots & \gamma_{1p}(\mathbf{h}) \\ \vdots & \ddots & \vdots \\ \gamma_{p1}(\mathbf{h}) & \cdots & \gamma_{pp}(\mathbf{h}) \end{bmatrix} = \begin{bmatrix} b_{11}^0 & \cdots & b_{1p}^0 \\ \vdots & \ddots & \vdots \\ b_{p1}^0 & \cdots & b_{pp}^0 \end{bmatrix} g_0(\mathbf{h}) + \cdots + \begin{bmatrix} b_{11}^l & \cdots & b_{1p}^l \\ \vdots & \ddots & \vdots \\ b_{p1}^l & \cdots & b_{pp}^l \end{bmatrix} g_l(\mathbf{h}) \quad (2.10)$$

The above procedure will be explained in greater detail for the two variable U and V case. The direct and cross-variogram models of U and V are constructed using the same variogram models (Isaaks and Srivastava, 1989).

$$\begin{aligned} \gamma_U(\mathbf{h}) &= u_0 \gamma_0(\mathbf{h}) + u_1 \gamma_1(\mathbf{h}) + \cdots + u_l \gamma_l(\mathbf{h}) \\ \gamma_V(\mathbf{h}) &= v_0 \gamma_0(\mathbf{h}) + v_1 \gamma_1(\mathbf{h}) + \cdots + v_l \gamma_l(\mathbf{h}) \\ \gamma_{UV}(\mathbf{h}) &= w_0 \gamma_0(\mathbf{h}) + w_1 \gamma_1(\mathbf{h}) + \cdots + w_l \gamma_l(\mathbf{h}) \end{aligned} \quad (2.11)$$

Where, $\gamma_U(\mathbf{h})$ and $\gamma_{UV}(\mathbf{h})$ are the auto- and cross-variogram models for U and V, respectively; the basic variogram models are given by $\gamma_0(\mathbf{h}), \gamma_1(\mathbf{h}), \dots, \gamma_L(\mathbf{h})$; $u, v,$

and w are coefficients, possibly negative.

Combinations of the first basic model, $\gamma_0(h)$

$$\begin{bmatrix} \gamma_{U,0}(\mathbf{h}) & \gamma_{UV,0}(\mathbf{h}) \\ \gamma_{VU,0}(\mathbf{h}) & \gamma_{V,0}(\mathbf{h}) \end{bmatrix} = \begin{bmatrix} u_0 & w_0 \\ w_0 & v_0 \end{bmatrix} \cdot \begin{bmatrix} \gamma_0(\mathbf{h}) & 0 \\ 0 & \gamma_0(\mathbf{h}) \end{bmatrix} \quad (2.12)$$

Combinations of the second basic model, $\gamma_1(h)$

$$\begin{bmatrix} \gamma_{U,1}(\mathbf{h}) & \gamma_{UV,1}(\mathbf{h}) \\ \gamma_{VU,1}(\mathbf{h}) & \gamma_{V,1}(\mathbf{h}) \end{bmatrix} = \begin{bmatrix} u_1 & w_1 \\ w_1 & v_1 \end{bmatrix} \cdot \begin{bmatrix} \gamma_1(\mathbf{h}) & 0 \\ 0 & \gamma_1(\mathbf{h}) \end{bmatrix} \quad (2.13)$$

Combinations of the m^{th} basic model, $\gamma_m(h)$

$$\begin{bmatrix} \gamma_{U,i}(\mathbf{h}) & \gamma_{UV,i}(\mathbf{h}) \\ \gamma_{VU,i}(\mathbf{h}) & \gamma_{V,i}(\mathbf{h}) \end{bmatrix} = \begin{bmatrix} u_i & w_i \\ w_i & v_i \end{bmatrix} \cdot \begin{bmatrix} \gamma_i(\mathbf{h}) & 0 \\ 0 & \gamma_i(\mathbf{h}) \end{bmatrix} \quad (2.14)$$

To ensure that linear model given above is positive semidefinite, the following conditions must be fulfilled.

$$\begin{aligned} u_j &\geq 0 \text{ and } v_j \geq 0 \\ u_j v_j &\geq w_j w_j \end{aligned} \quad \text{For all } j=0, \dots, 1 \quad (2.15)$$

This model becomes very difficult to fit as the number of variables increase.

2.3 Kriging

Kriging is a linear regression algorithm to estimate at unsampled location values using the related nearby data. Kriging is often referred as the Best Linear Unbiased Estimate (BLUE). It is based on the minimization of the estimation variance. Each nearby data is assigned a weight in a linear combination of these data. Block kriging amounts to kriging all the unsampled location in the block and then averaging the block values to block kriging estimate. The cokriging algorithm provides the ability to use secondary variables to make a better estimate.

The most common form is Simple Kriging (SK). Other variations are made on SK with constraints. The following is an overview of SK.

Assume the mean is constant and stationary (m). Consider n data values $Z(\mathbf{u}_\alpha)$, $\alpha=1, \dots, n$ and mean m . The SK estimate is written:

$$[Z_{SK}^*(\mathbf{u})-m]=\sum_{\alpha=1}^n \lambda_{\alpha}[Z(\mathbf{u}_{\alpha})-m] \quad (2.16)$$

If define a new variable as the residual of the data $Y=Z-m$. The estimate is:

$$Y^*(\mathbf{u})=\sum_{\alpha=1}^n \lambda_{\alpha}Y(\mathbf{u}_{\alpha}) \quad (2.17)$$

The Kriging variance is defined as.

$$\begin{aligned} \sigma_{SK}^2 &= E\{[Y(\mathbf{u})-Y^*(\mathbf{u})]^2\} \\ &= E\{[Y(\mathbf{u})Y(\mathbf{u})]-2E\{[Y(\mathbf{u})Y^*(\mathbf{u})]\}+E\{[Y^*(\mathbf{u})Y^*(\mathbf{u})]\} \\ &= C(0)-2\sum_{\alpha=1}^n \lambda_{\alpha}C(\mathbf{u},\mathbf{u}_{\alpha})+\sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha}\lambda_{\beta}C(\mathbf{u}_{\alpha},\mathbf{u}_{\beta}) \end{aligned} \quad (2.18)$$

The optimal weights are calculated from the minimization of Equation 2.18:

$$\sum_{\beta=1}^n \lambda_{\beta}C(\mathbf{u}_{\alpha},\mathbf{u}_{\beta})=C(\mathbf{u}_{\alpha},\mathbf{u}_{\beta}) \quad \alpha=1,\dots,n \quad (2.19)$$

Then the minimum error variance is simplified to:

$$\sigma_{SK}^2=C(0)-\sum_{\alpha=1}^n \lambda_{\alpha}C(\mathbf{u},\mathbf{u}_{\alpha}) \quad (2.20)$$

Cokriging is a formalism for multiple data types. Consider P data types and each data type has n_p nearby data for each type. Consider the residual as the variable. The cokriging estimate is written:

$$Y_i^*(\mathbf{u})=\sum_{p=1}^P \sum_{\alpha=1}^{n_p} \lambda_{\alpha p}Y_p(\mathbf{u}_{\alpha p}) \quad (2.21)$$

The weights for primary and secondary data are determined by the cokriging system of equations:

$$\sum_{q=1}^P \sum_{\beta=1}^{n_p} \lambda_{\beta q}C(\mathbf{u}_{\alpha p},\mathbf{u}_{\beta q})=C_{ip}(\mathbf{u},\mathbf{u}_{\alpha p}) \quad \forall p=1,\dots,P; \alpha=1,\dots,n_p \quad (2.22)$$

And the corresponding estimation variance is:

$$\sigma_{SCK}^2=C_{ii}(0)-\sum_{p=1}^P \sum_{\alpha=1}^{n_p} \lambda_{\alpha p}C_{ip}(\mathbf{u},\mathbf{u}_{\alpha p}) \quad (2.23)$$

Where the above equation indicates the minimized variance of cokriging system. All of the required covariances for different and same variable are calculated by the LMC model.

2.4 Techniques for multivariable Block LU simulation

In mining, the decision of what material in a deposit is worth extracting and processing is based on the estimated block values; it is not based on the point scale of data. An estimate of each block is required. The Block LU simulation amounts to discretize each blocks into many points and then average the individual point simulations to get the block values.

2.4.1 Unconditional LU simulation.

Let C be the covariance matrix between the data. It contains the relationship between the data points. C is symmetric and positive-definite and its Lower and Upper decomposition could be obtained (Davis 1987):

$$\begin{aligned} C &= LU \\ L' &= U \end{aligned}$$

Consider the random vector F that is a vector with standard normal distribution $N(0, 1)$ values. The multiplication of Lower matrix by the random vector F is:

$$y = LF$$

The expected value of the $n \times n$ matrix YY' is given by

$$\begin{aligned} E\{YY'\} &= E\{FF'U\} \\ &= LE\{FF'\}U \end{aligned}$$

Because F is a vector of independent standard normal distribution $N(0, 1)$.

$$E\{FF'\} = I$$

So

$$\begin{aligned}
E\{YY'\} &= E\{LFF'U\} \\
&= LIU \\
&= C
\end{aligned}$$

Thus, the simple multiplication provides the unconditional realizations of the random function at data location with covariance matrix C (Alabert, 1987). This method can handle a moderate-sized grid up to 7000 points.

2.4.2 Conditional LU simulation.

Consider the covariance matrix of data locations and the grid to be generated. The matrix is partitioned the matrix as follows:

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \quad (2.24)$$

Where the C_{11} is the covariance and cross-covariance between n data points. C_{12} is the covariance and cross-covariance between the data points and the grid discretization points or grid points. C_{22} is the covariance and cross-covariance matrix between the N unsampled locations. To be specific with one variable:

$$C_{11} = \begin{bmatrix} c(\mathbf{u}_1, \mathbf{u}_1) & \cdots & c(\mathbf{u}_1, \mathbf{u}_n) \\ \vdots & \ddots & \vdots \\ c(\mathbf{u}_n, \mathbf{u}_1) & \cdots & c(\mathbf{u}_n, \mathbf{u}_n) \end{bmatrix} \quad (2.25)$$

$$C_{12} = \begin{bmatrix} c(\mathbf{u}_1, \mathbf{u}_1) & \cdots & c(\mathbf{u}_1, \mathbf{u}_N) \\ \vdots & \ddots & \vdots \\ c(\mathbf{u}_n, \mathbf{u}_1) & \cdots & c(\mathbf{u}_n, \mathbf{u}_N) \end{bmatrix} \quad (2.26)$$

$$C_{22} = \begin{bmatrix} c(\mathbf{u}_1, \mathbf{u}_1) & \cdots & c(\mathbf{u}_1, \mathbf{u}_N) \\ \vdots & \ddots & \vdots \\ c(\mathbf{u}_N, \mathbf{u}_1) & \cdots & c(\mathbf{u}_N, \mathbf{u}_N) \end{bmatrix} \quad (2.27)$$

$$Z_\alpha = \begin{bmatrix} z(\mathbf{u}_1) \\ \vdots \\ z(\mathbf{u}_n) \end{bmatrix} \quad (2.28)$$

The LU approach can consider data of different types or different variables.

Incorporating a multivariate version of block LU simulation allows the modeler to consider data of different types simultaneously. There are many secondary variables that could provide additional information on the variable of interest (Boisvert and Deutsch 2007).

If there are 3 variables u, v and w. Each matrix is extended as follows (n data location and N unsampled locations):

$$C_{11} = \begin{bmatrix} c(\mathbf{u}_1, \mathbf{u}_1) & \cdots & c(\mathbf{u}_1, \mathbf{u}_n) & c(\mathbf{u}_1, \mathbf{v}_1) & \cdots & c(\mathbf{u}_1, \mathbf{v}_n) & c(\mathbf{u}_1, \mathbf{w}_1) & \cdots & c(\mathbf{u}_1, \mathbf{w}_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c(\mathbf{u}_n, \mathbf{u}_1) & \cdots & c(\mathbf{u}_n, \mathbf{u}_n) & c(\mathbf{u}_n, \mathbf{v}_1) & \cdots & c(\mathbf{u}_n, \mathbf{v}_n) & c(\mathbf{u}_n, \mathbf{w}_1) & \cdots & c(\mathbf{u}_n, \mathbf{w}_n) \\ & & & c(\mathbf{v}_1, \mathbf{v}_1) & \cdots & c(\mathbf{v}_1, \mathbf{v}_n) & c(\mathbf{v}_1, \mathbf{w}_1) & \cdots & c(\mathbf{v}_1, \mathbf{w}_n) \\ & & & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & c(\mathbf{v}_n, \mathbf{v}_1) & \cdots & c(\mathbf{v}_n, \mathbf{v}_n) & c(\mathbf{v}_n, \mathbf{w}_1) & \cdots & c(\mathbf{v}_n, \mathbf{w}_n) \\ & & & & & & c(\mathbf{w}_1, \mathbf{w}_1) & \cdots & c(\mathbf{w}_1, \mathbf{w}_n) \\ & & & & & & \vdots & \ddots & \vdots \\ & & & & & & c(\mathbf{w}_n, \mathbf{w}_1) & \cdots & c(\mathbf{w}_n, \mathbf{w}_n) \end{bmatrix} \quad (2.29)$$

$$C_{12} = \begin{bmatrix} c(\mathbf{u}_1, \mathbf{u}_1) & \cdots & c(\mathbf{u}_1, \mathbf{u}_N) & c(\mathbf{u}_1, \mathbf{v}_1) & \cdots & c(\mathbf{u}_1, \mathbf{v}_N) & c(\mathbf{u}_1, \mathbf{w}_1) & \cdots & c(\mathbf{u}_1, \mathbf{w}_N) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c(\mathbf{u}_n, \mathbf{u}_1) & \cdots & c(\mathbf{u}_n, \mathbf{u}_N) & c(\mathbf{u}_n, \mathbf{v}_1) & \cdots & c(\mathbf{u}_n, \mathbf{v}_N) & c(\mathbf{u}_n, \mathbf{w}_1) & \cdots & c(\mathbf{u}_n, \mathbf{w}_N) \\ & & & c(\mathbf{v}_1, \mathbf{v}_1) & \cdots & c(\mathbf{v}_1, \mathbf{v}_N) & c(\mathbf{v}_1, \mathbf{w}_1) & \cdots & c(\mathbf{v}_1, \mathbf{w}_N) \\ & & & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & c(\mathbf{v}_n, \mathbf{v}_1) & \cdots & c(\mathbf{v}_n, \mathbf{v}_N) & c(\mathbf{v}_n, \mathbf{w}_1) & \cdots & c(\mathbf{v}_n, \mathbf{w}_N) \\ & & & & & & c(\mathbf{w}_1, \mathbf{w}_1) & \cdots & c(\mathbf{w}_1, \mathbf{w}_N) \\ & & & & & & \vdots & \ddots & \vdots \\ & & & & & & c(\mathbf{w}_n, \mathbf{w}_1) & \cdots & c(\mathbf{w}_n, \mathbf{w}_N) \end{bmatrix} \quad (2.30)$$

$$C_{22} = \begin{bmatrix} c(\mathbf{u}_1, \mathbf{u}_1) & \cdots & c(\mathbf{u}_1, \mathbf{u}_N) & c(\mathbf{u}_1, \mathbf{v}_1) & \cdots & c(\mathbf{u}_1, \mathbf{v}_N) & c(\mathbf{u}_1, \mathbf{w}_1) & \cdots & c(\mathbf{u}_1, \mathbf{w}_N) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c(\mathbf{u}_N, \mathbf{u}_1) & \cdots & c(\mathbf{u}_N, \mathbf{u}_N) & c(\mathbf{u}_N, \mathbf{v}_1) & \cdots & c(\mathbf{u}_N, \mathbf{v}_N) & c(\mathbf{u}_N, \mathbf{w}_1) & \cdots & c(\mathbf{u}_N, \mathbf{w}_N) \\ & & & c(\mathbf{v}_1, \mathbf{v}_1) & \cdots & c(\mathbf{v}_1, \mathbf{v}_N) & c(\mathbf{v}_1, \mathbf{w}_1) & \cdots & c(\mathbf{v}_1, \mathbf{w}_N) \\ & & & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & c(\mathbf{v}_N, \mathbf{v}_1) & \cdots & c(\mathbf{v}_N, \mathbf{v}_N) & c(\mathbf{v}_N, \mathbf{w}_1) & \cdots & c(\mathbf{v}_N, \mathbf{w}_N) \\ & & & & & & c(\mathbf{w}_1, \mathbf{w}_1) & \cdots & c(\mathbf{w}_1, \mathbf{w}_N) \\ & & & & & & \vdots & \ddots & \vdots \\ & & & & & & c(\mathbf{w}_N, \mathbf{w}_1) & \cdots & c(\mathbf{w}_N, \mathbf{w}_N) \end{bmatrix} \quad (2.31)$$

Because C is symmetric, so $C_{21} = C_{12}'$, Any symmetric, positive definite, matrix C has a unique LU decomposition (Horn and Johnson 1985):

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ A_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & B_{12} \\ 0 & U_{22} \end{bmatrix} \quad (2.32)$$

The SK estimator is written as:

$$\begin{aligned} Z^* &= \lambda' Z_\alpha = C_{21} C_{11}^{-1} Z_\alpha \\ \text{where } C_{11} \lambda &= C_{12} \\ \Rightarrow \lambda &= C_{11}^{-1} C_{12} \\ \Rightarrow \lambda &= C_{21} C_{11}^{-1} \end{aligned} \quad (2.33)$$

The lower matrix multiple a vector \mathbf{F} of N independent random numbers with a standard normal distribution.

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} L_{11} & \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{bmatrix} \quad (2.34)$$

The conditional simulation replaces \mathbf{F}_1 by \mathbf{G}_1 , where \mathbf{G}_1 is the solution to equation

$$L_{11} \mathbf{G}_1 = z_1 \quad (2.35)$$

because the data has been transformed to a standard normal distribution. So

$$E(z_1 z_1') = C_{11}$$

Thus

$$\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^{-1} z_1 \\ \mathbf{F}_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ L_{21} L_{11}^{-1} z_1 + L_{22} \mathbf{F}_2 \end{bmatrix} \quad (2.36)$$

Produces a conditional simulation of the random function, which is $L_{21} L_{11}^{-1} z_1$.

2.4.3 Computational aspects of LU simulation.

When coming into the computational aspect of the LU simulation, there is no need to do the LU of the big matrix C. From the following we can see (Wackernagel, 1989):

$$\begin{aligned} C &= \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ A_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & B_{12} \\ 0 & U_{22} \end{bmatrix} \\ &= \begin{bmatrix} L_{11} U_{11} & L_{11} B_{12} \\ A_{21} U_{11} & A_{21} B_{12} + L_{22} U_{22} \end{bmatrix} \end{aligned} \quad (2.37)$$

With L_{11}, L_{22} : lower triangular matrices, U_{11}, U_{22} : upper triangular matrices.

$$\begin{aligned}
C_{11} &= L_{11}U_{11} \\
C_{12} &= L_{11}B_{12} \Rightarrow B_{12} = L_{11}^{-1}C_{12} \\
C_{21} &= A_{21}U_{11} \Rightarrow A_{21} = C_{21}U_{11}^{-1} \\
&\Rightarrow A_{21}^1 = (U_{11}^{-1})C_{21} = L_{11}^{-1}C_{12} = B_{12} \quad (2.38) \\
C_{22} &= A_{21}B_{12} + L_{22}U_{22} \\
\text{with } A_{21}B_{12} &= C_{21}U_{11}^{-1}L_{11}^{-1}C_{12} = C_{21}C_{11}^{-1}C_{12} \\
&\Rightarrow L_{22}U_{22} = C_{22} - C_{21}C_{11}^{-1}C_{12}
\end{aligned}$$

So, if given $K_{22} = L_{22}U_{22} = C_{22} - C_{21}C_{11}^{-1}C_{12}$, in the case where there is a single point u to be estimated ($m=1$), the matrix K_{22} reduces to the SK variance:

$$\begin{aligned}
\sigma_{sk}^2(u) &= C(0) - \sum_{\alpha=1}^n \lambda_{\alpha}(u_{\alpha} - u) = C_{22} - \lambda' C_{12} \\
&= C_{22} - C_{21}C_{11}^{-1}C_{12} \geq 0 \quad (2.39)
\end{aligned}$$

In the general case, consider the matrix of error covariance is:

$$\begin{aligned}
&E\{[Z(u_i) - Z^*(u_i)][Z(u_j) - Z^*(u_j)]\} \quad \forall i, j=1, \dots, N \\
&= E\{[Z(u_i) - Z(u_j)]\} - E\{[Z(u_i) - Z^*(u_j)]\} - E\{[Z(u_i) - Z^*(u_j)]\} + E\{[Z^*(u_i) - Z^*(u_j)]\} \\
&= C_{22} - 2\lambda' C_{12} + \lambda' C_{11} \lambda \\
&= C_{22} - \lambda' C_{12} \quad \text{Since } C_{11} \lambda = C_{12} \\
&= K_{22}
\end{aligned}$$

It is an error covariance matrix and it is symmetric and positive definite. Thus the LU decomposition of K_{22} is always possible. From the LU simulation expression (another expression but equal to (2.36));

$$\begin{aligned}
y_{cs} &= y^* + r_s = A_{21}(L_{11}^{-1}Z_{\alpha}) + L_{22}F_2 \\
&= (L_{11}^{-1}C_{12})^{-1}(L_{11}^{-1}Z_{\alpha}) + L_{22}F_2 \quad (2.40)
\end{aligned}$$

Where L_{22} is the LU decomposition of K_{22} .

Once the C_{11} , C_{12} , C_{22} are obtained. Just decompose C_{11} to get L_{11} , then inverse of L_{11} . Then the first part of (2.40) is obtained. At the same time K_{22} is also obtained by (2.39). Then the simulation part is achieved.

Chapter 3

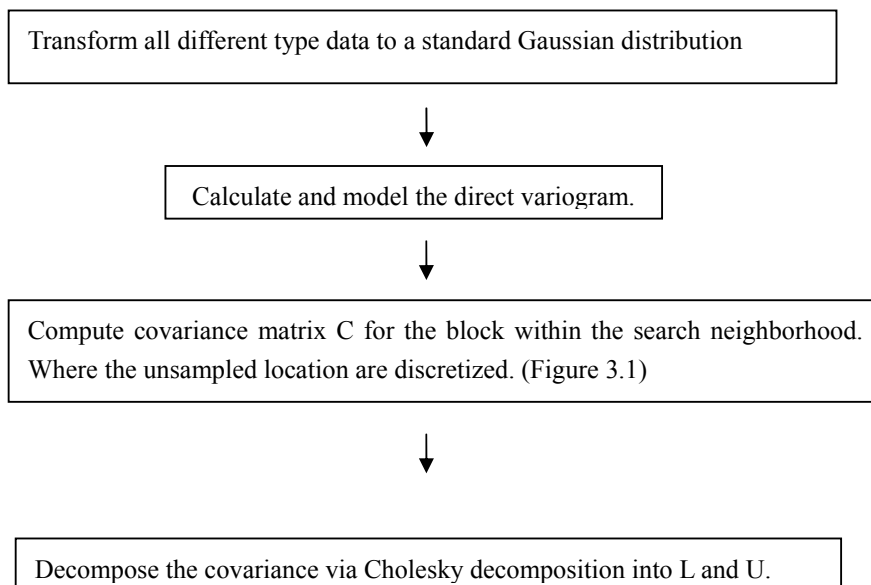
Improving LU simulation methodology.

3.1 Approximate model of coregionalization.

The GSLIB (Deutsch, 1998) like program, Blusim_mv, is used for calculating the block simulated values. The existing LU formalism approach allows us to use multivariate spatial data. The calculation of multivariate simulated results from Blusim requires the linear model of coregionalization (LMC), which is considered a difficult task (Vargas-Guzmán, 2004), although there are some programs to model the variogram and cross variogram. If there are P variables, the method requires $P*(P+1)/2$ variogram models. And the positive definite of the LMC model conditions have to be fulfilled.

In this chapter Block LU simulation with an Approximate model of coregionalization (AMC) is demonstrated. A new program is developed to complete this idea (Appendix A). This method does not require the cross variogram model between variables; it only needs the direct variogram for each variable and a correlation matrix between collocated samples.

The main algorithm is illustrated below:



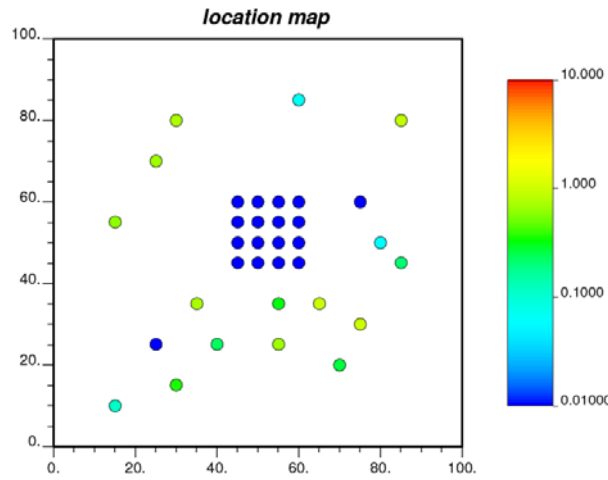
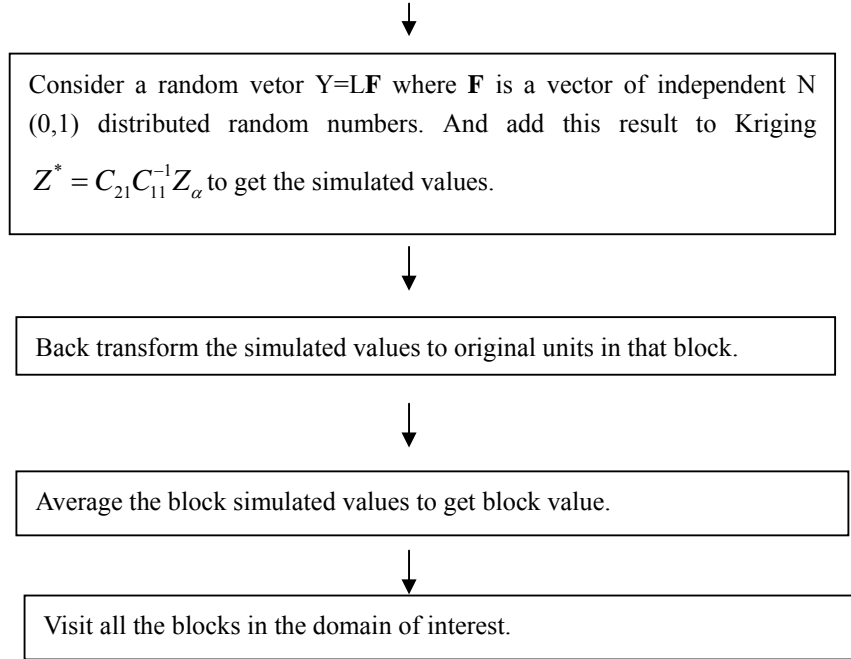


Figure 3.1: Block discretization with search data.

The proposed approach uses the direct variogram and the correlation matrix between variables as the following:

$$\gamma_{p,p'}(\mathbf{h}) = \rho_{p,p'} \sqrt{\gamma_p(\mathbf{h}) * \gamma_{p'}(\mathbf{h})} \quad 3.1$$

The cross-variograms are obtained by using a geometric average of the corresponding directional variogram and scaled to correct correlation between the variables.

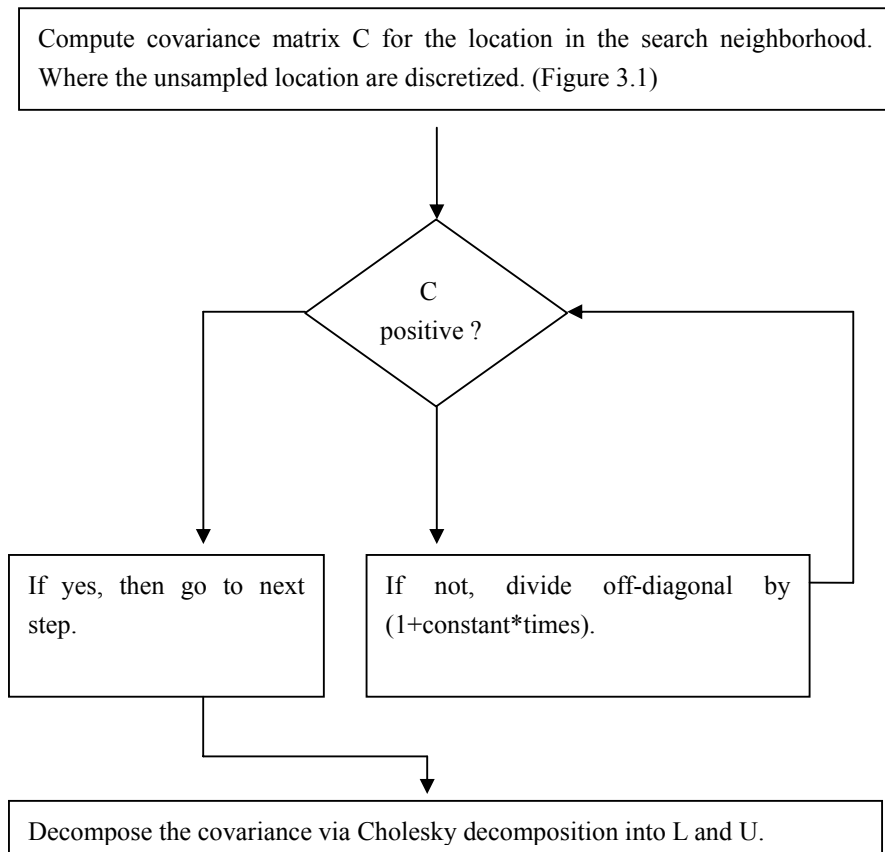
Another optional methodology is using the arithmetic average of the corresponding

directional variogram values and then scale to the correct correlation coefficient between variables.

$$\gamma_{p,p'}(\mathbf{h}) = \rho_{p,p'} \frac{\gamma_p(\mathbf{h}) + \gamma_{p'}(\mathbf{h})}{2} \quad 3.2$$

The geometric average will be considered as our base case because it is always less than or equal to the arithmetic mean and this makes it more stable. The off-diagonal values in the C matrix will be slightly lower and this reduces the chance of C becoming not positive definite.

This methodology does not guarantee that the C matrix is positive definite. It will be likely close to positive definite. A dynamic correction is proposed to make sure it is positive definite if required:



The constant we choose for default is 0.025, which is reasonable because the diagonal of C is 1 when the data are transformed to normal values.

3.1 Transformation.

The algorithm developed requires the variable to be transformed to standard Gaussian variables and the simulated values to be back transformed (Deutsch, 1998). There are often local variations in the variable of interests that are not reflected well by a global transformation. In the improved methodology, the transformation table is built every single local search and then back transformation proceeds with correct local back transformation. The program has two searches, one is a bigger search for the transformation table, and another is a smaller search for the conditioning data (Figure 3.2).

The first step is to build the transformation table, and then back transform the simulated values according to that table (Figure 3.3). One potential problem is that there may not be enough data to build the transform table. If this happens, the global table will be used for that location. Another potential problem in transform is how to deal with the minimum data and maximum data values in the transform table and how it was transformed. In the following chapter, the approximate minimum and maximum is used for transform.

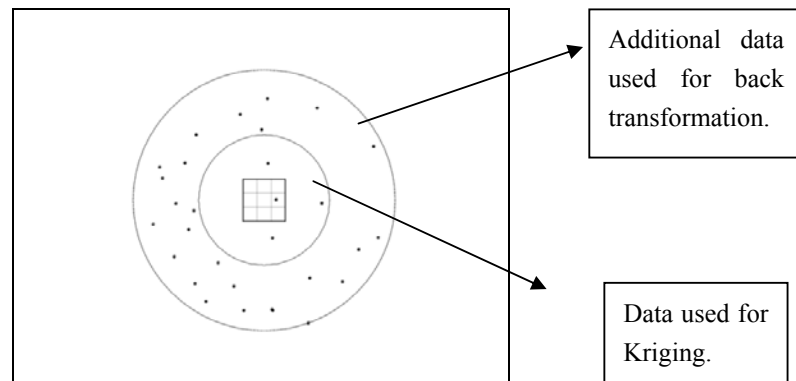


Figure 3.2 Double search radii using the local transformation.

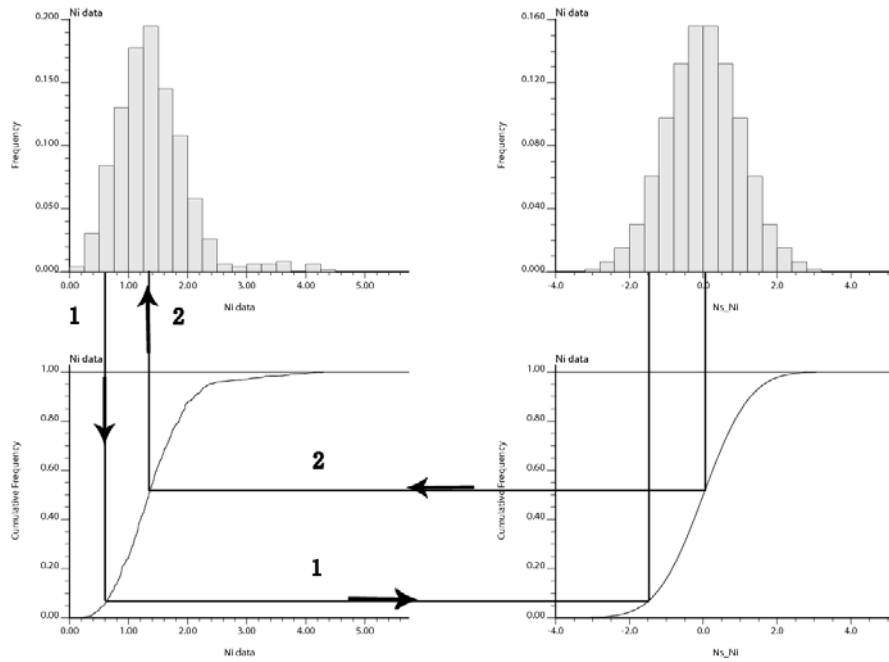


Figure 3.2. Local transformation illustrations. Upper left is the local data histogram. Bottom left is the cdf of local data histogram. Bottom right is the cdf normal distribution. Upper right is the histogram of normal distribution.

Chapter 4

Case study

A case study illustrates the methodology of Block LU simulation with an approximate model of coregionalization (AMC). A comparison with a full Linear model of coregionalization (LMC) is also conducted. The local transformation is also implemented with this data.

The following steps are described: (1) exploratory data analysis, (2) spatial continuity analysis, (3) comparative study between AMC and LMC, (4) a sensitivity study. (5) CPU-time study and (6) a local transformation study.

4.1 Data

The data is a 2D nickel deposit with measurements of the mass fraction of Ni, Fe and MgO. The particulars of the data set will be kept for confidential reason. The exploratory data analysis is important for getting representative statistics for the variables in the deposit. The area of interest is shown in Figure 4.1. It is 1200m east-west, and 700m north-south. From these location maps, one can see that the drill holes are collected in a spatially clustered pattern. The univariate statistics of the three variables are shown in Figure 4.2. The summary of the raw statistics are shown in Table 4.1.

Declustering is needed to obtain representative global statistics. The decision of the appropriate cell size is based on Figure 4.3. A declustering cell size of 85 is chosen for variable Ni. Declustering weights are largely determined on the basis of the geometric configuration of the data, therefore, only one set of declustering weights is calculated in presence of the three variables (Deutsch 2002). The declustering results are shown in Figure 4.4. The mean of the variables decreases after declustering.

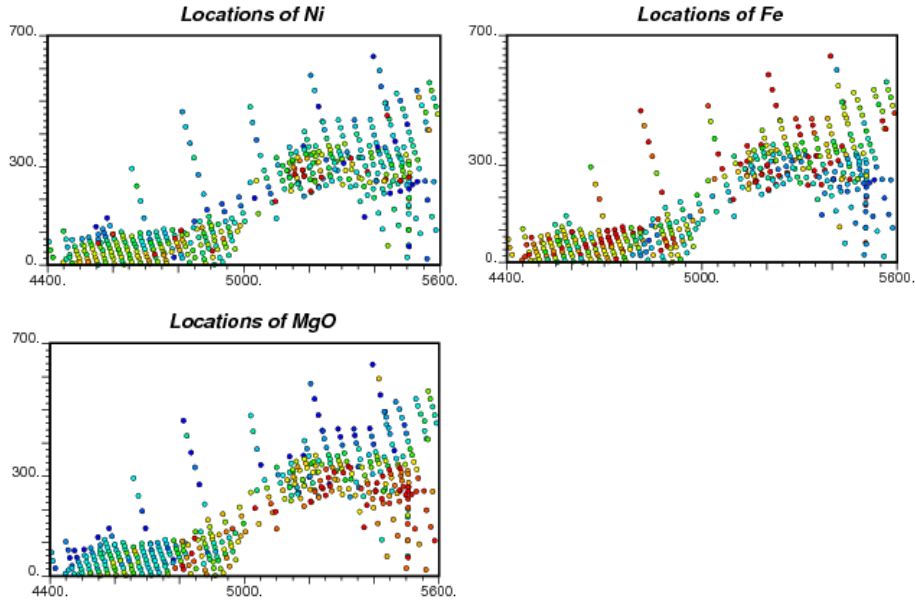


Figure 4.1 Location map of drillhole data.

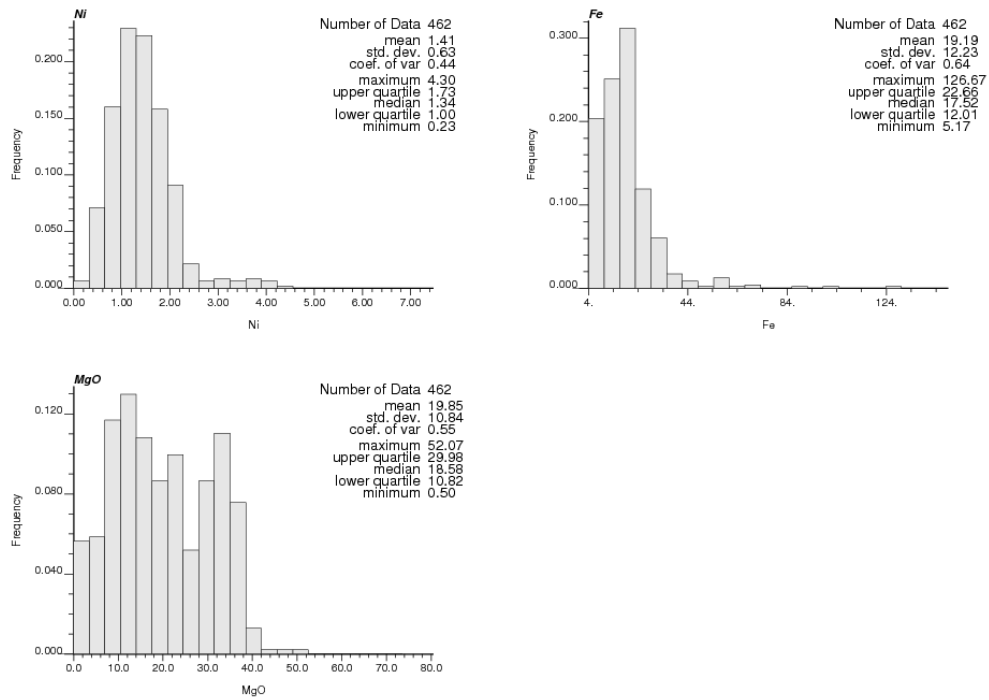


Figure 4.2 Histogram for Ni, Fe and MgO (percent mass fraction).

	Ni(%)	Fe(%)	MgO(%)
Mean	1.41	19.19	19.85
Std.dev	0.63	12.23	10.84
Min	0.23	2.17	0.50
Max	4.30	126.67	52.07

Table 4.1 Summary statistics for Ni, Fe and MgO.

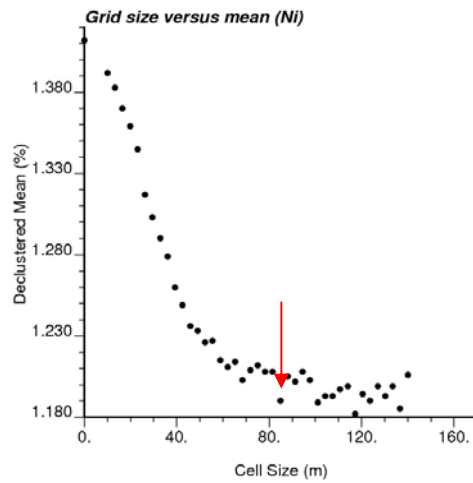


Figure 4.3 Declustered mean versus Cell size for variable Ni.

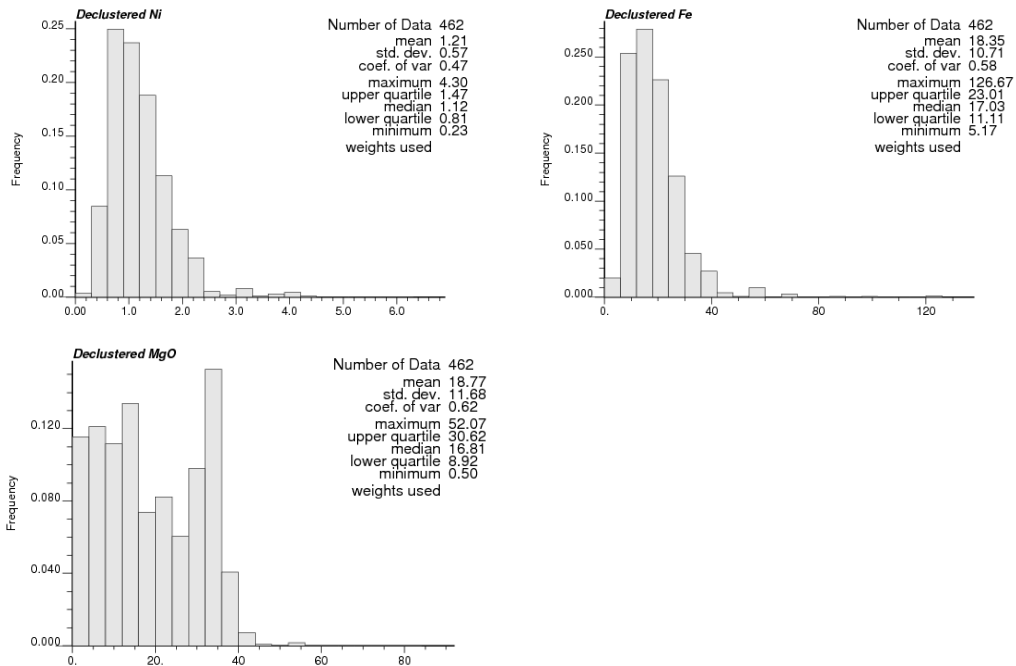


Figure 4.4 Histograms for Ni, Fe and MgO by using declustering weights.

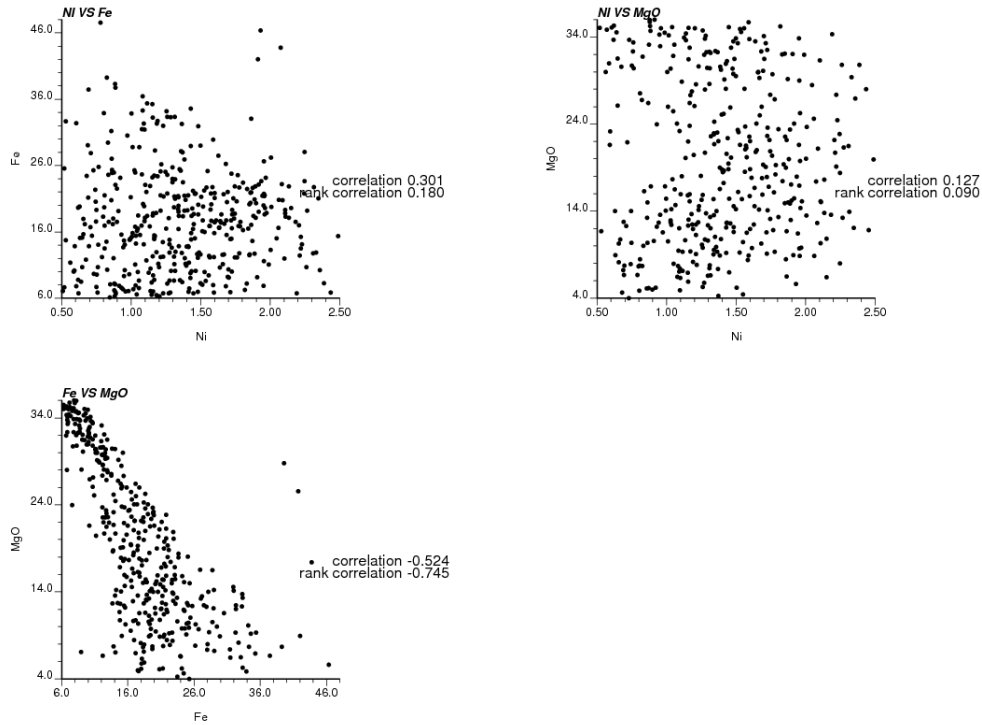


Figure 4.5 Scatter plot of Ni, Fe and MgO.

The bivariate study is shown in Figure 4.5. The correlation between Ni and Fe is 0.301, the correlation between Ni and MgO is 0.127. These two relationships are direct, but not very strong. The correlation between Fe and MgO is -0.524; which shows an inverse correlation. These are summarized in correlation matrix shown in Table 4.2

	Ni	Fe	MgO
Ni	1	0.301	0.127
Fe	0.301	1	-0.524
MgO	0.127	-0.524	1

Table 4.2 Correlation between variables.

4.2 Variogram Analysis.

The omnidirectional horizontal experimental variograms are used for data pattern reasons and they are shown in Figure 4.6. The variogram is sensitive to the lag spacing, tolerance and other parameters. Larger lag spacing, tolerance parameters could result in more stable variograms, but the values would be less precise.

Note that before the variogram is calculated, the data are transformed to Gaussian units. Simulation is based on the Gaussian units. The variograms were modeled (Figure 4.7). GSLIB programs are used to fit these variograms and cross variograms. A positive definite LMC is the following:

Variograms:

$$\gamma_{Ni}(\mathbf{h}) = 0.01 + 0.751Exp_{a=38}(\mathbf{h}) + 0.248Exp_{a=205}(\mathbf{h})$$

$$\gamma_{Fe}(\mathbf{h}) = 0.01 + 0.737Exp_{a=38}(\mathbf{h}) + 0.262Exp_{a=205}(\mathbf{h})$$

$$\gamma_{MgO}(\mathbf{h}) = 0.01 + 0.223Exp_{a=38}(\mathbf{h}) + 0.776Exp_{a=205}(\mathbf{h})$$

Cross variograms:

$$\gamma_{Ni-Fe}(\mathbf{h}) = 0.01 + 0.066Exp_{a=38}(\mathbf{h}) + 0.234Exp_{a=205}(\mathbf{h})$$

$$\gamma_{Ni-MgO}(\mathbf{h}) = 0.01 + 0.084Exp_{a=38}(\mathbf{h}) + 0.042Exp_{a=205}(\mathbf{h})$$

$$\gamma_{Fe-MgO}(\mathbf{h}) = 0.01 - 0.388Exp_{a=38}(\mathbf{h}) - 0.137Exp_{a=205}(\mathbf{h})$$

The sill of all direct variograms is 1.0, and the sill of the cross variogram is correlation coefficient between the pair of variables. The AMC model only requires the direct variograms. The direct variograms are modeled manually (Figure 4.7) as follows:

$$\gamma_{Ni}(\mathbf{h}) = 0.01 + 0.726Exp_{a=35}(\mathbf{h}) + 0.273Sph_{a=200}(\mathbf{h})$$

$$\gamma_{Fe}(\mathbf{h}) = 0.01 + 0.710Exp_{a=30}(\mathbf{h}) + 0.289Exp_{a=200}(\mathbf{h})$$

$$\gamma_{MgO}(\mathbf{h}) = 0.01 + 0.310Exp_{a=35}(\mathbf{h}) + 0.689Exp_{a=200}(\mathbf{h})$$

Figure 4.8 shows the cross variogram model between Fe and MgO by using the two different models.

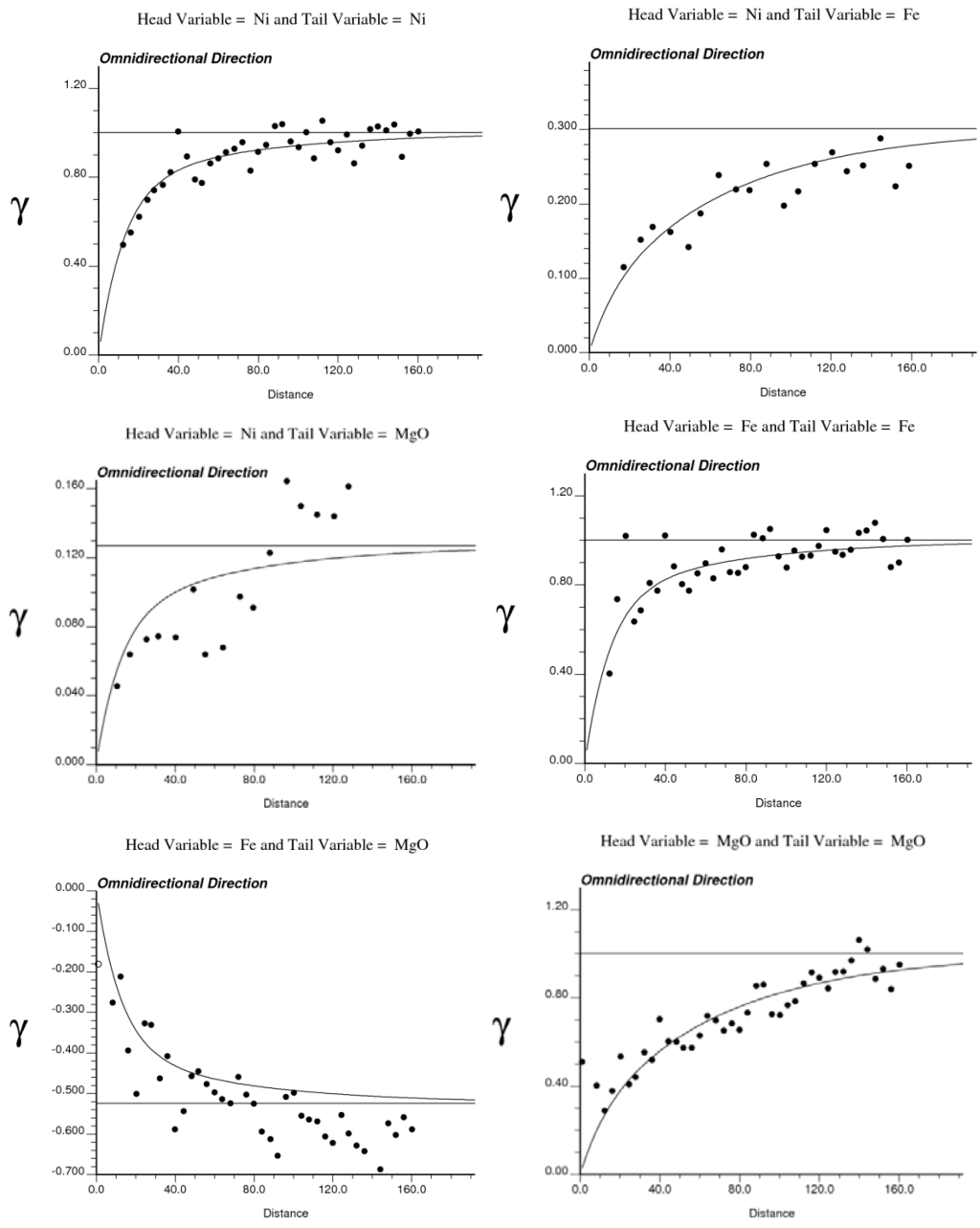


Figure 4.6 Ni, Fe and MgO variogram and cross variogram fitted by LMC.

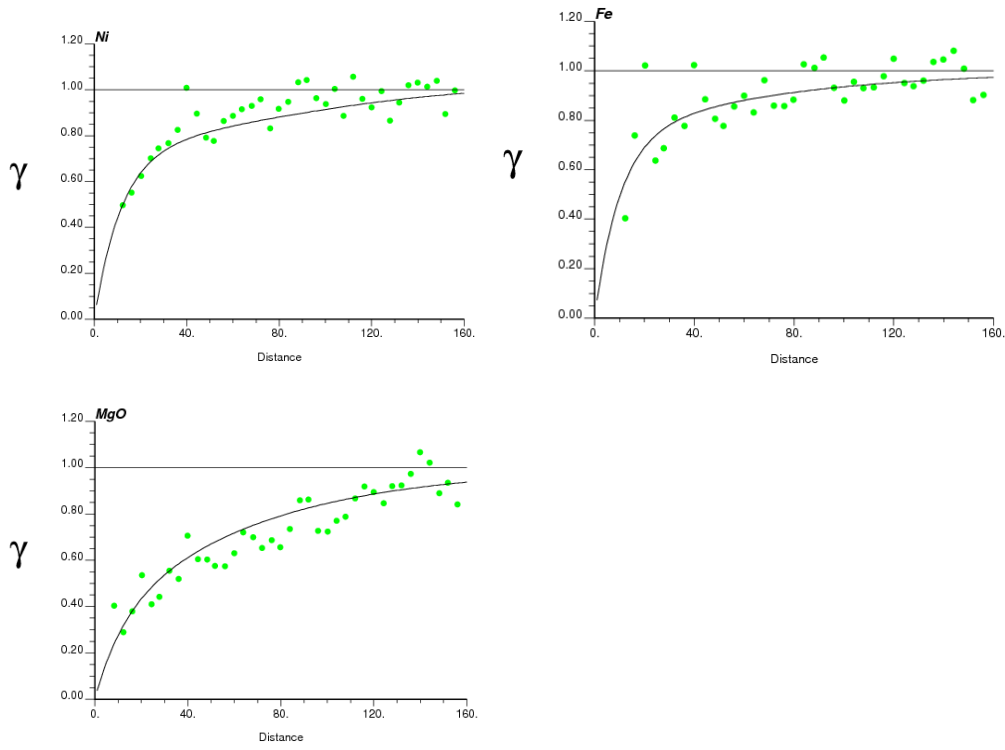


Figure 4.7 Ni, Fe and MgO variogram model fitted separately.

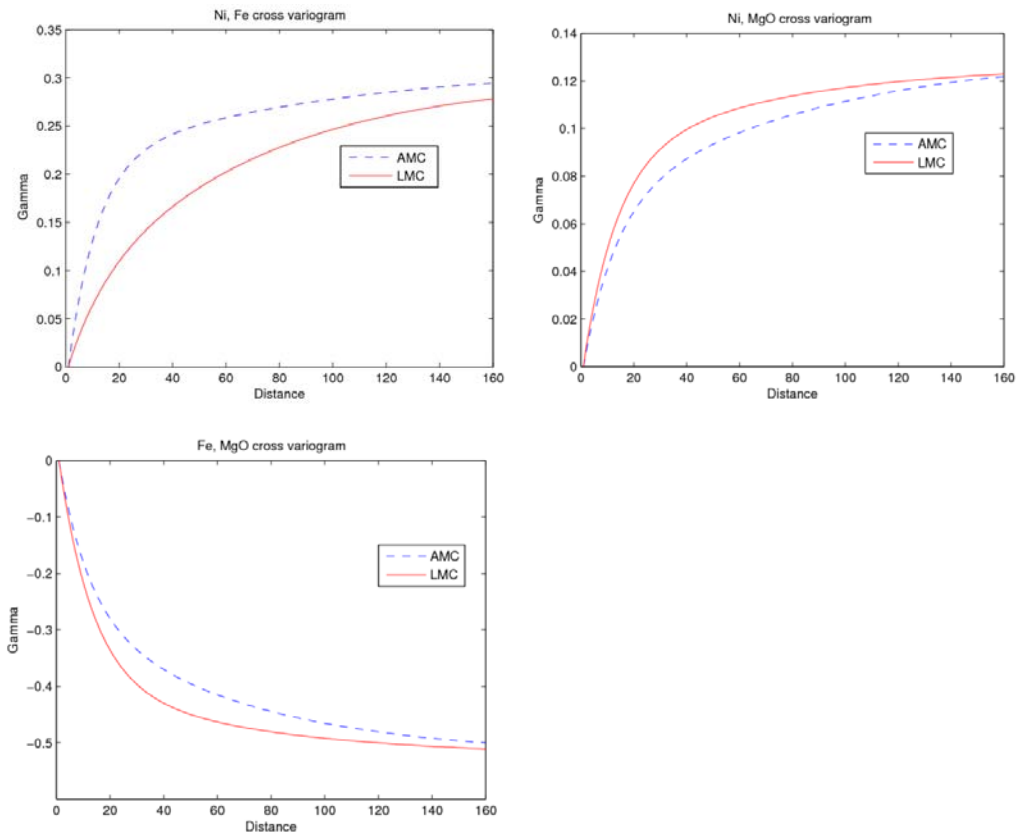


Figure 4.8 Comparative cross variogram model between LMC and AMC

4.3 Comparative simulation with AMC and LMC

A GSLIB FORTRAN program called Blusim_AMC was developed for performing block multivariate LU simulation (Appendix A). An explanation of all parameters is included. This program does not require the LMC model to calculate the cross variogram model but upper diagonal correlation coefficient matrix of variables.

Using the AMC does not assure that the covariance matrix is positive; however, no non-positive definite matrices are found.

The summary statistics of the block averaged simulated value are shown in Table 4.3. 100 realizations are used for both methods. The mean of Fe is slightly bigger using AMC. The mean of Ni are slightly smaller using AMC. The mean of MgO is the same. The AMC also shows slightly lower minimum and slightly higher maximum value (Figure 4.9, 4.10). This is seen in the visualized simulation results in Figure 4.14 and 4.15. From the scatter plot of these two, the high correlation shows that the Blusim with AMC is quite close to Blusim with LMC (Figure 4.13). The computation of mean grade and proportion above a cutoff grade are shown in Figure 4.18 (LMC) in Figure 4.19 (AMC).

The summary statistics of the simulated global variance are shown in Table 4.4, and the histograms are shown in Figures 4.11 and 4.12. The mean of variance for Ni and MgO is quite similar: similar mean, minimum and maximum value. But for Fe, the variance using AMC shows slightly higher variance and slightly higher minimum and maximum value.

Variable \ Model		LMC	AMC
Ni	mean	1.21	1.19
	std.dev	0.23	0.25
	min	0.35	0.37
	max	3.17	3.13

Fe	mean	18.1	18.7
	std.dev	3.58	3.82
	min	5.75	5.82
	max	50.5	46.9
MgO	mean	18.1	18.0
	std.dev	7.10	7.10
	min	0.68	0.86
	max	40.6	40.0

Table 4.3 Summary simulation statistics for Ni, Fe and MgO.

Variable \ Model		LMC	AMC
Ni	mean	0.19	0.18
	std.dev	0.07	0.06
	min	0.00	0.00
	max	0.59	0.58
Fe	mean	62.9	68.6
	std.dev	30.0	35.1
	min	0.16	0.21
	max	324	296
MgO	mean	68.9	69.5
	std.dev	36.0	33.1
	min	0.17	0.38
	max	127	123

Table 4.4 Summary variance statistics Ni, Fe and MgO.

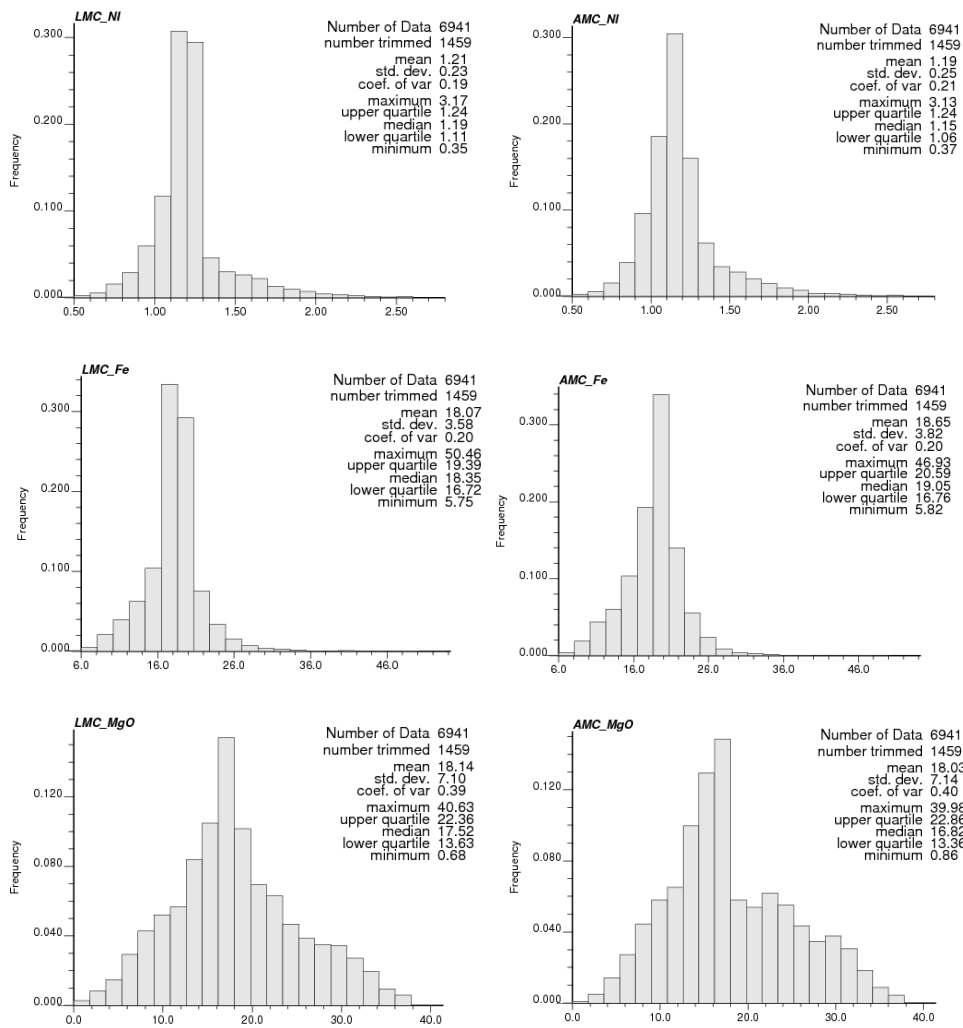


Figure 4.9 Block averaged simulation histograms using LMC (left).

Figure 4.10 Block averaged simulation histograms using AMC (right).

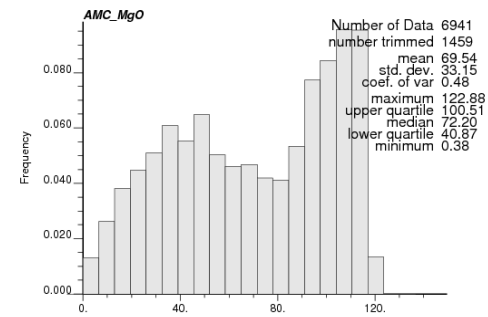
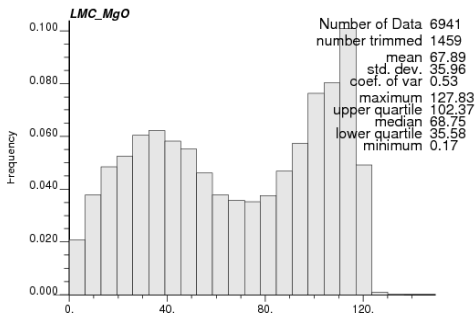
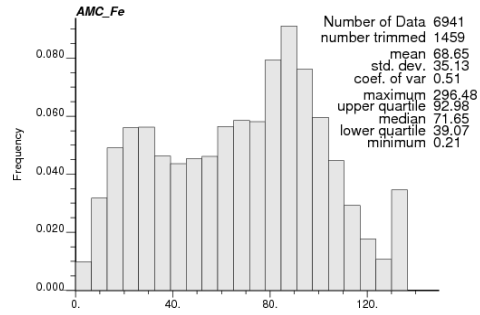
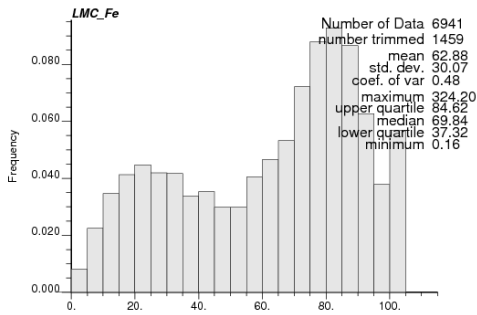
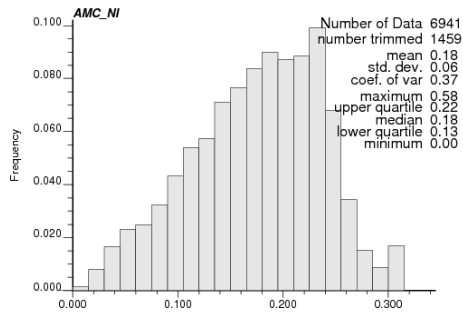
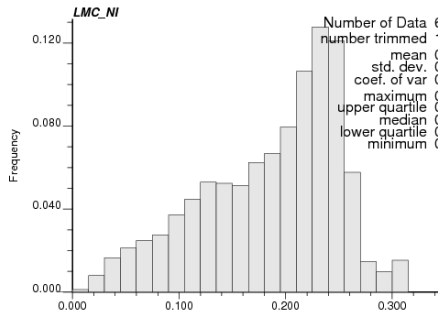


Figure 4.11 Block global variance using LMC (left).
 Figure 4.12 Block global variance using AMC (right).

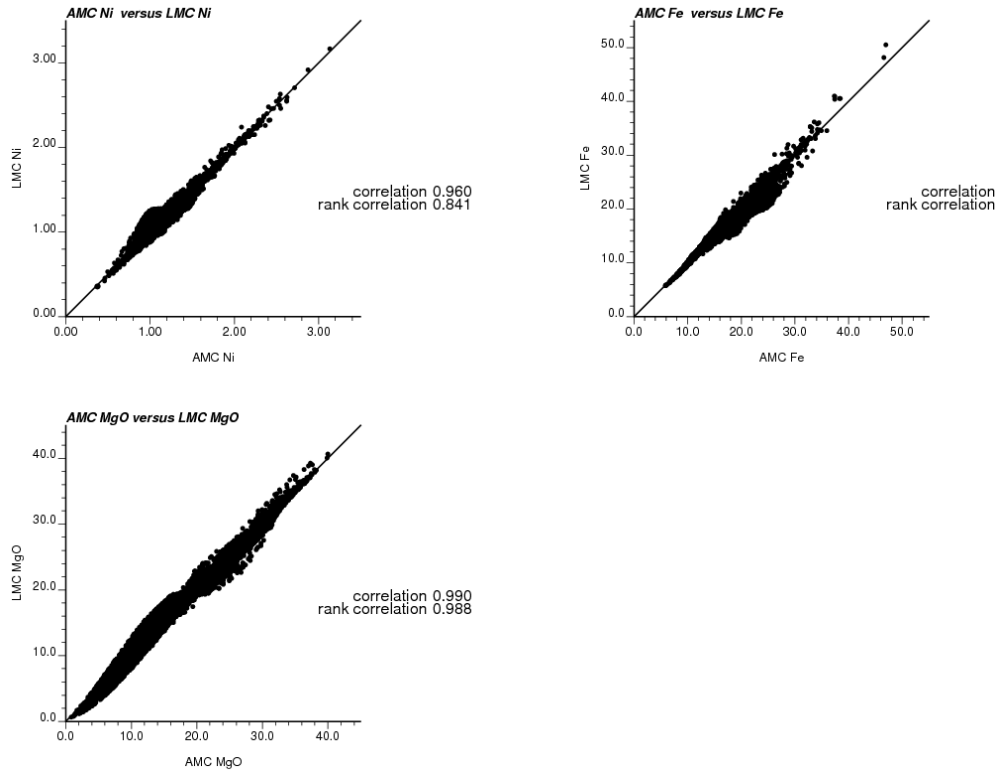
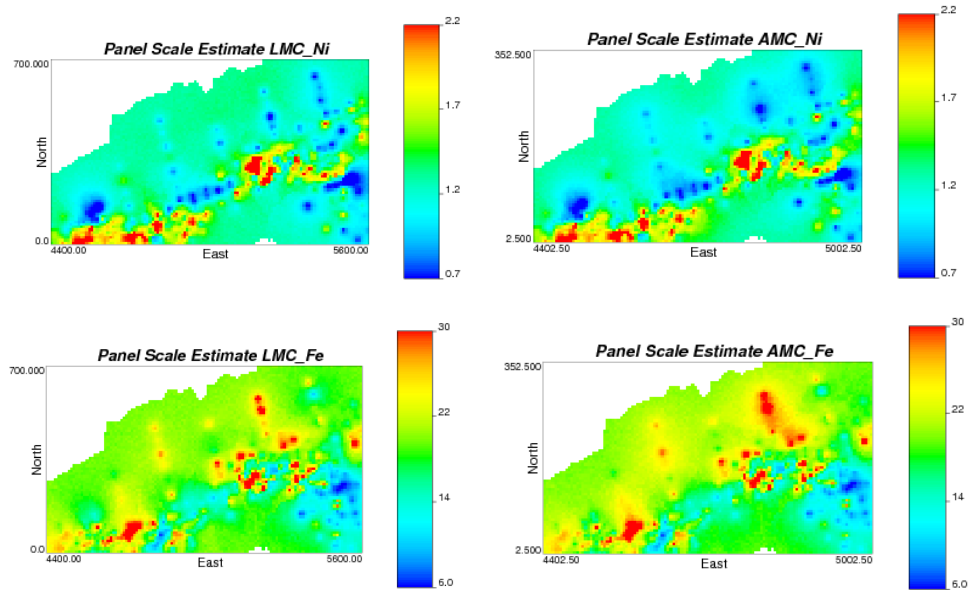


Figure 4.13 Scatter plot of block averaged simulation using AMC and using LMC



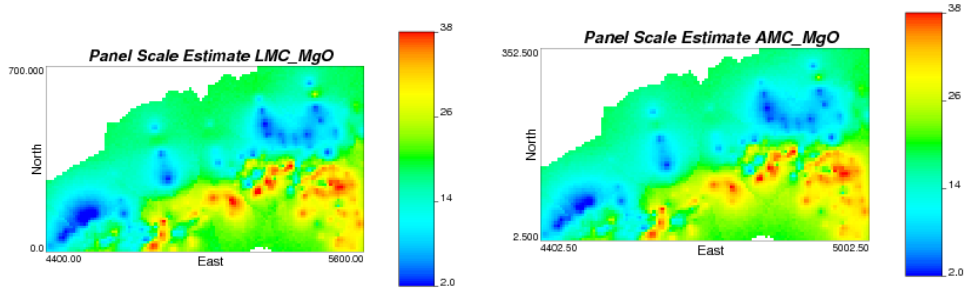


Figure 4.14 Block averaged simulation plots using LMC (left).
 Figure 4.15 Block averaged simulation plots using AMC (right).

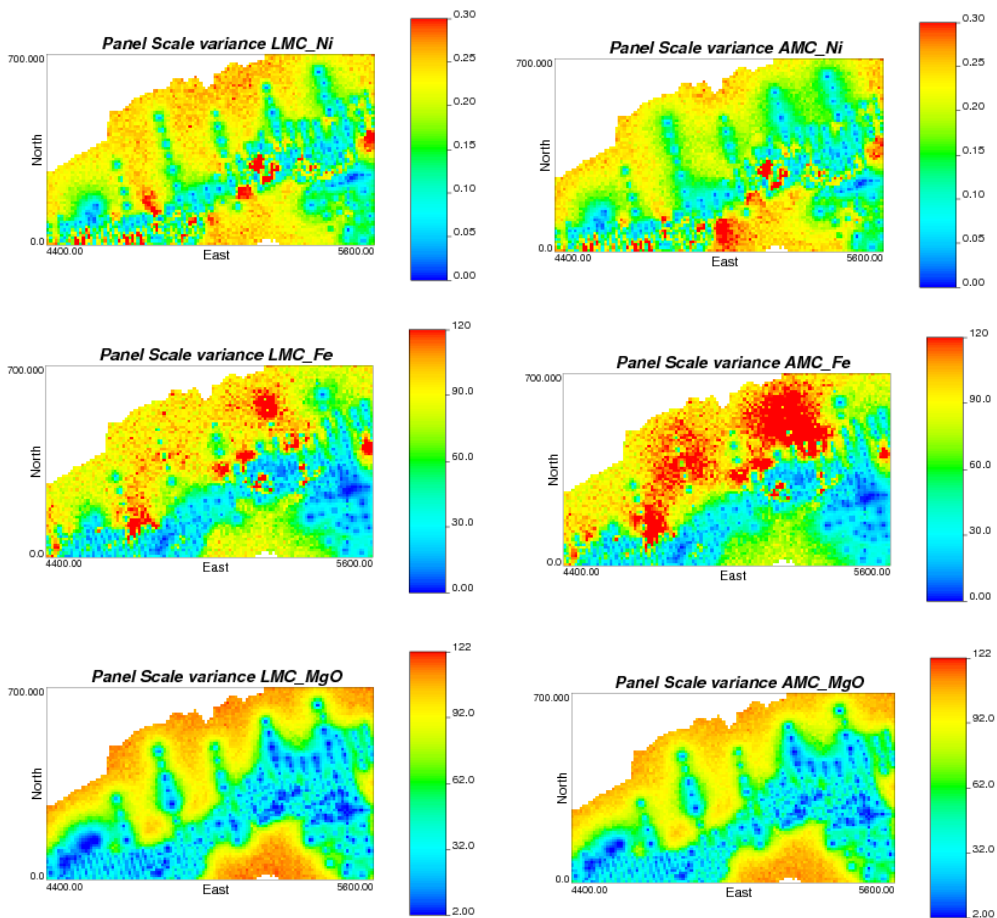


Figure 4.16 Block global variance plots using LMC (left).
 Figure 4.17 Block global variance plots using AMC (right).

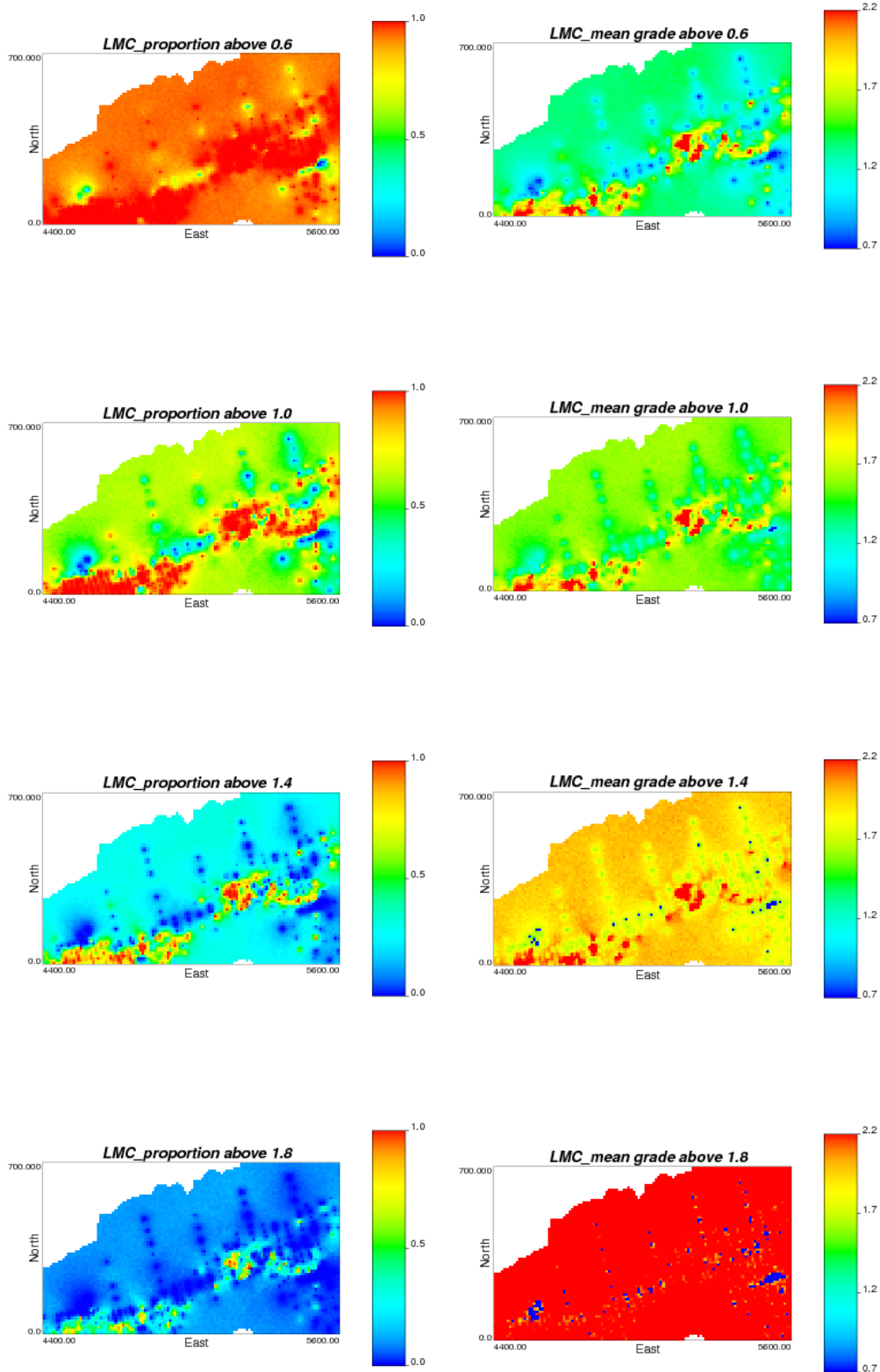


Figure 4.18 Ni mean grades and proportion of grid above cutoffs using LMC.

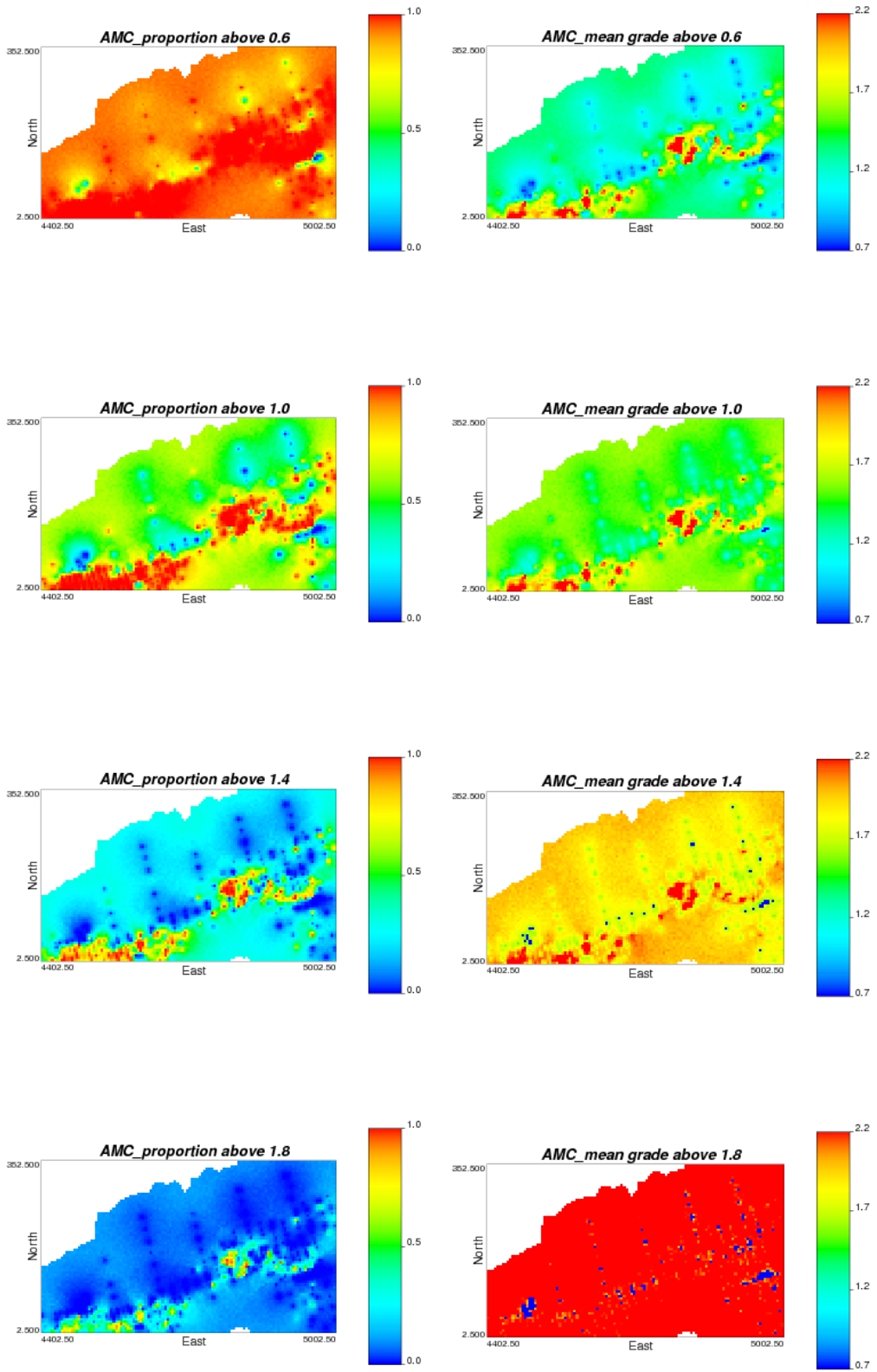


Figure 4.19 Ni mean grades and proportion of grid above cutoffs using AMC.

4.4 Sensitivity analysis.

The sensitivity analysis focuses on the effects of grid size, search strategy used and discretization. To determine the sensitivity of the grid size in this Nickel deposit, 6 different block sizes are chosen. The different increasing cutoff to Ni, Fe and MgO estimate are conducted, and then the tonnage fractions are calculated shown in Figure 4.20. For each block model, the cutoff grade is applied above cut off for SMU sizes 5×5, 10×10, 15×15, 20×20, 25×25, 30×30. Using these maps, the fraction of total tonnage is not so sensitive to the block size.

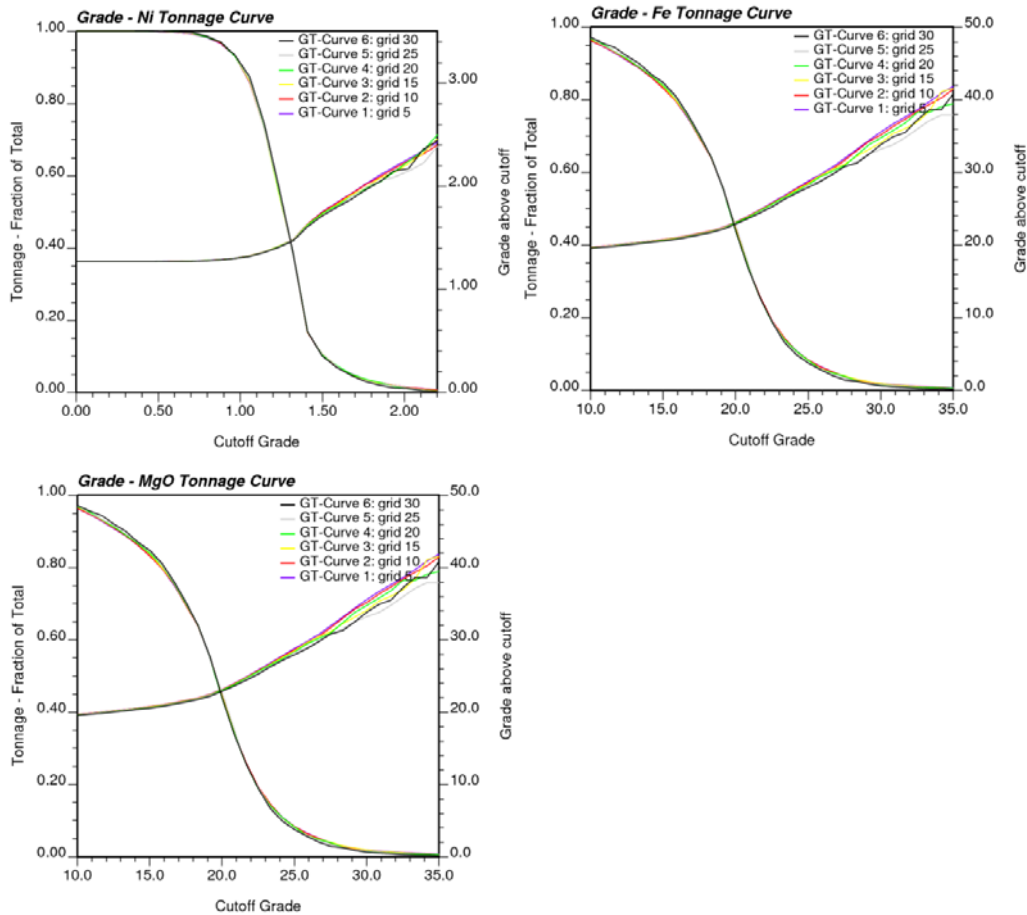


Figure 4.20 Grade tonnage curve of Ni, Fe and MgO using different block size.

A sensitivity study on the effect of the search data is also conducted.

The variance of the simulation is sensitive to the number of data used (Figure 4.21). The standard deviation of the three variables is increasing when the search data increases, and it flattens off at round 20 data used.

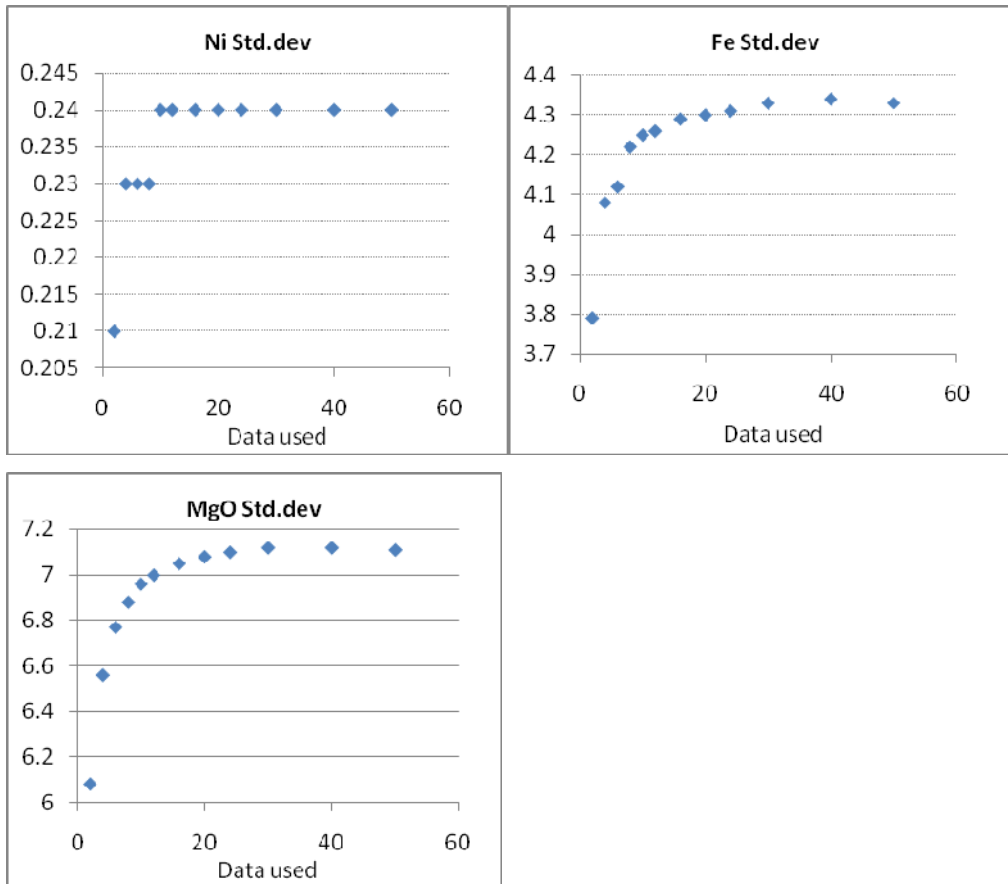


Figure 4.21 Standard deviation of Ni, Fe and MgO versus the number of data used in the search.

A sensitivity study on effect of the discretization points in the grid was conducted. The more discretization points are required, the more points are simulated in the grid, but more computation memory is needed and the computations are slower. In this case, the discretization points are chosen as 2, 4, 6, 8, 10, 12 and 14.

The variance is decreasing as the discretization points is increasing, and reach a stable value (Figure 3.22). Considering CPU-time will increase as discretization points increases, the optimal discretization points should be 4.

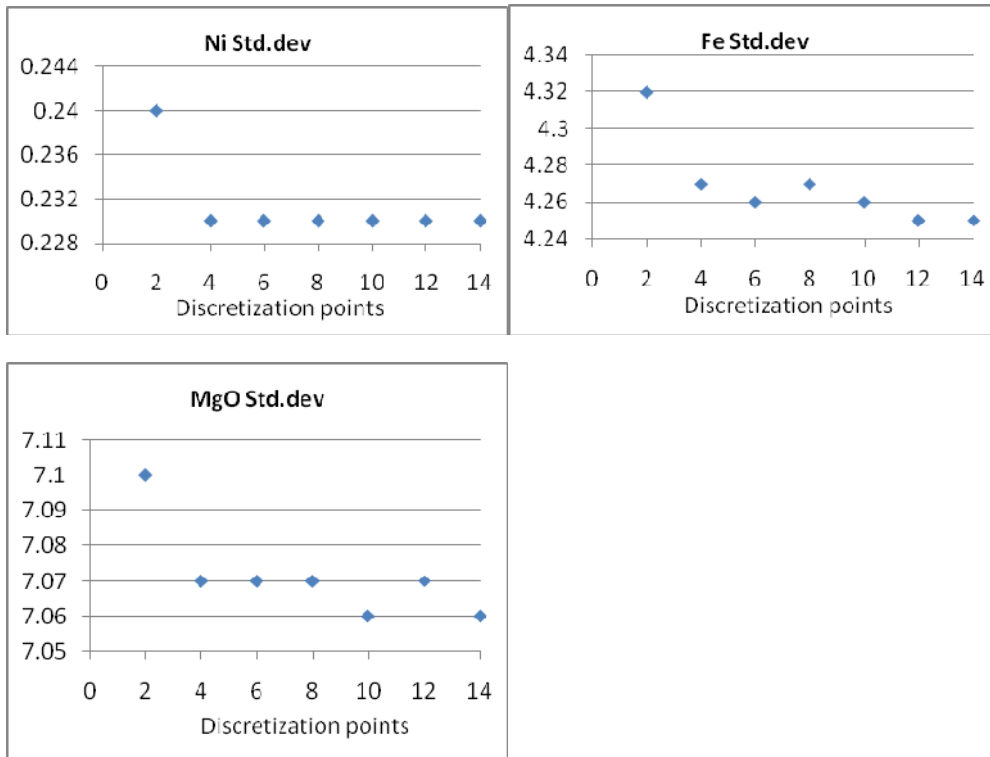


Figure 4.22 Standard deviation of Ni, Fe and MgO results versus the number of discretization points used.

4.5 CPU-time study.

In Figure 4.23, it shows that when doing the block LU simulation with AMC, once the grid size, discretization points and other parameters is set up, the C22 sub-matrix is all the same during algorithm, because C22 sub-matrix is the covariance matrix between discretization points. The improved algorithm is using the same C22 sub-matrix over all the simulation process. Figure 4.24 shows that the compared CPU-time study on improved algorithm and the old one when the simulation number, search data used and discretization points are increasing.

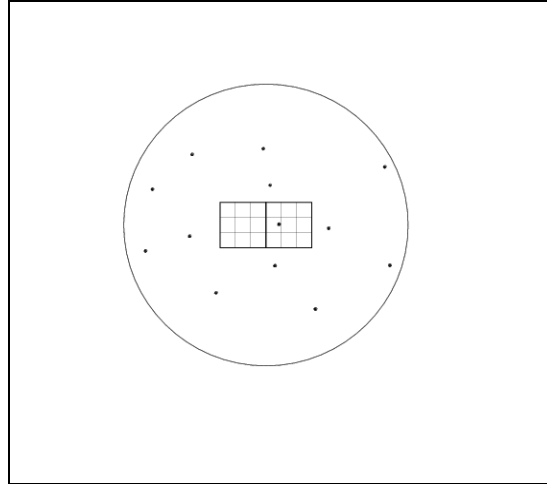


Figure 4.23: A case for two blocks using the same discretization.

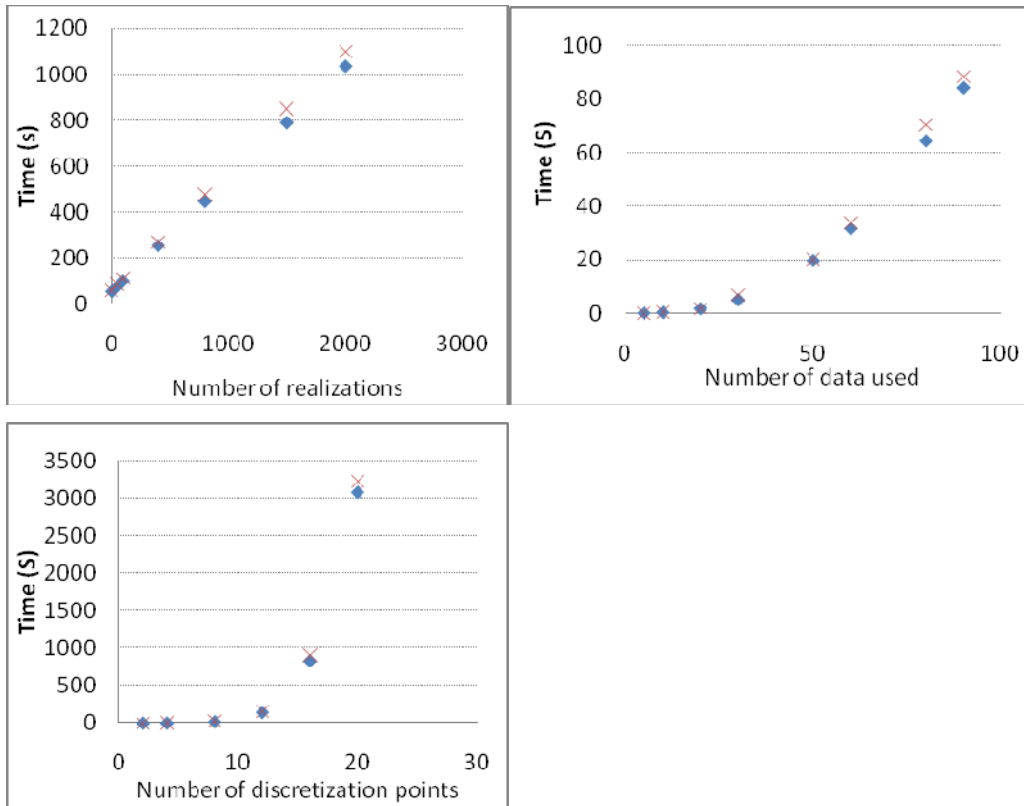


Figure 4.24. CPU-time study on improved algorithm where C22 is calculated once (square) compared with previous where C22 is calculated every time (star).

4.6 Transformation study

The program BLUSIM_AMC allows the data to be transformed locally. The locally data are used to build the transformation table, and back transformed by the results. A larger search will ensure a stable transformation table. Declustering can also be used.

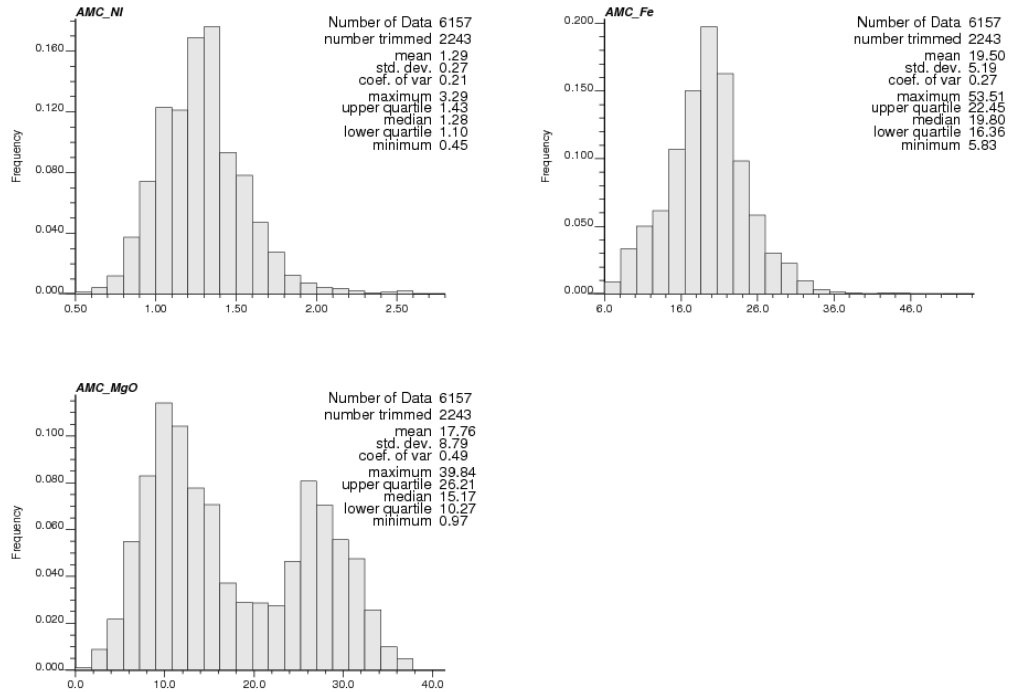


Figure 4.25. Block averaged simulation histograms using local transformation.

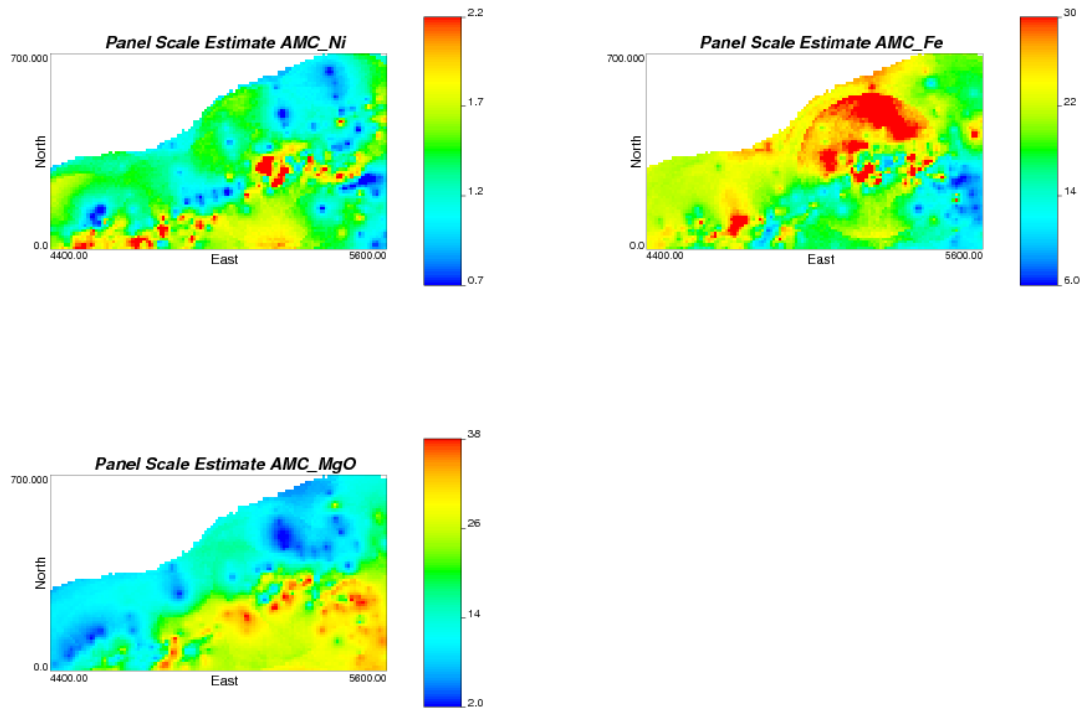


Figure 4.26. Block averaged simulation plots using local transformation.

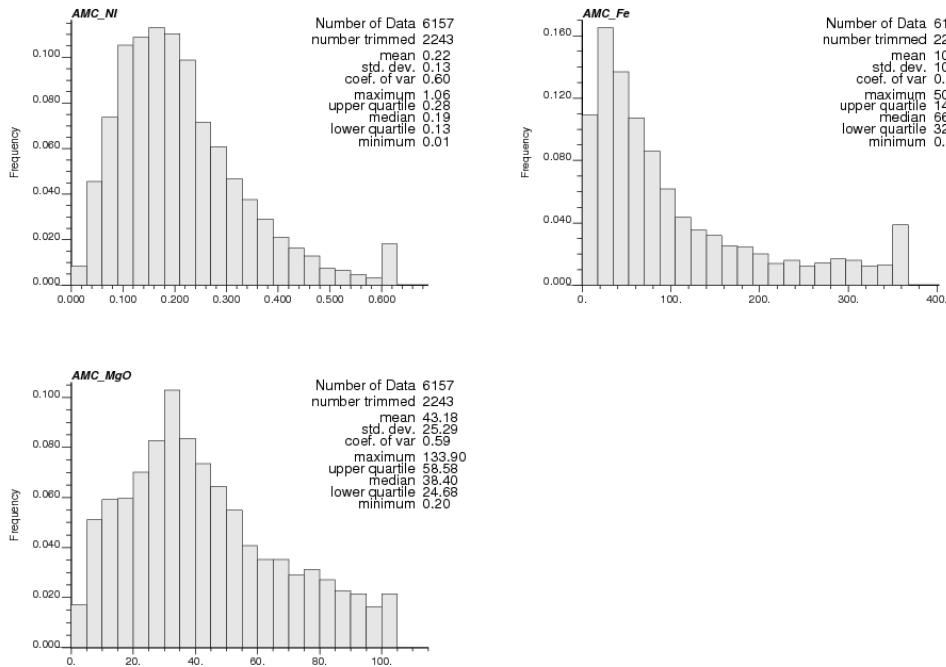


Figure 4.27. Block global variance histograms using local transformation.

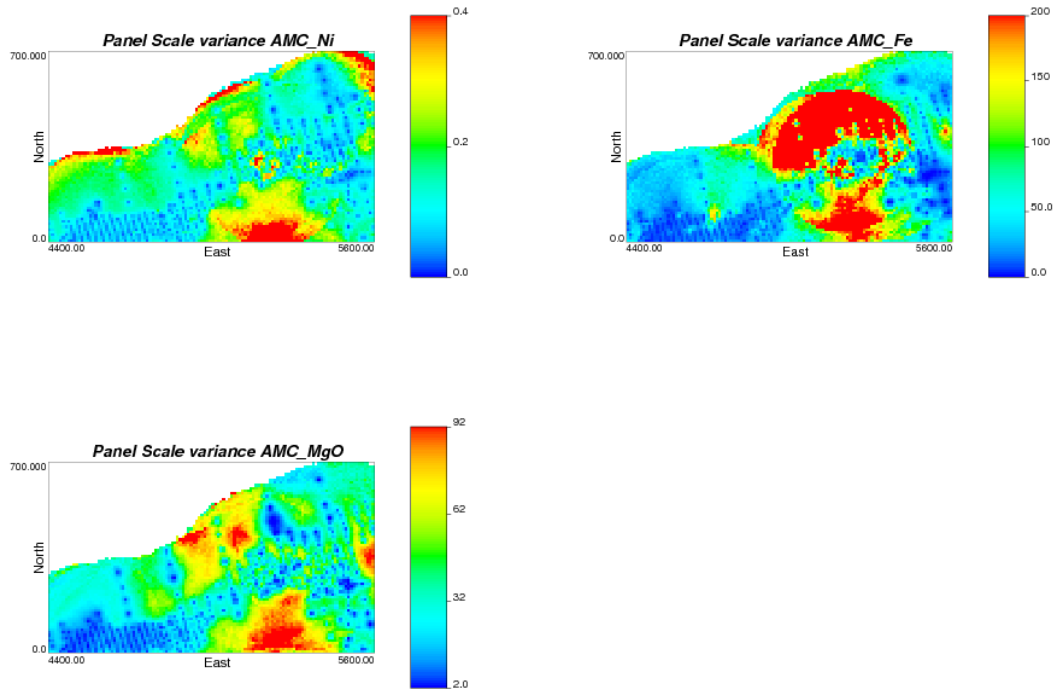


Figure 4.28. Block global variance plots using local transformation.

For this case, 20 data are used to build transformation table. The block averaged simulation histograms are shown in Figures 4.25 and 4.26. From these plots, one can see that the variance is zero at the data points. The local data area are better respected using the local transformation table: the high valued areas have more higher values and the low areas have more low values. For example, for Ni, there are more low values in the upper right corner. The variances are shown in Figure 4.27 and 4.28.

Chapter 5

Conclusions and Future Work

Block simulation using the LU technique in geostatistics has been an active area of research. Accounting for multiple variables improves the quality of geostatistical estimates. The linear model of coregionalization is traditionally required for calculate the relationship between pairs of variable, which is considered a difficult task.

This thesis developed an approximate model of coregionalization in the framework of the block LU simulation technique. The main idea is to model the direct variogram model separately to get the relationship between the same variable and then use the correlation matrix and previous direct variogram to get the relationship to different variables. The covariance matrix is built using these relationships between pairs of variable. To ensure the covariance matrix to be positive definite so that the LU could be implemented, the detection and iterative correction could be applied.

A FORTRAN code called blusim_AMC is developed to apply the methodology. The methodology is applied to a real Ni deposit. Two other variables Fe and MgO are used to get the approximate model coregionalization. The results are compared with the linear model of coregionalization. Similar simulation results for all variables were obtained with both techniques.

Sensitivity analysis based on the approximate model of coregionalization was also conducted. The results show that the optimum grid size, data used for kriging and discretization points. The CPU-time on improved covariance matrix calculation shows some improvement.

Another development in the program is to take local variations into consideration. Two search radii are used, the inner search radius is for kriging and the outer radius search is for transformation. This is explored in the Ni deposit. The result of local transformation shows a closer reproduction of the local data.

Future Work

Regarding future work, the iterative correction of the C22 matrix calculated must be documented more completely. The current case study happened to not need any correction. Further development on this correction scheme is still subject to testing.

The program developed in this thesis deal with the collocated data. This program should be extended to the non-collocated case, that is, when not all variables are at the same location.

Some minor visual artifacts with using the double search radius scheme, the transformation using the local data requires further development and exploration.

Bibliography

- F. Alabert. 1987. The Practice of Fast Conditional Simulations Through the LU Decomposition of the Covariance matrix, *Mathematical Geology*, Vol.19, No.5.
- A. Almeida and A. Journel 1994. Joint simulation of multiple variables with a Markov-type coregionalization model. *Mathematical Geology*. Vol. 26, 565-588.
- O. Babak and C. Deutsch. 2008. Collocated Cokriging Based on Merged Secondary Attributes. *Mathematical Geosciences*. Vol. 42, 1874-8961.
- J. Boisvert and C. Deutsch. 2007. Multivariate Block LU Simulation, CCG annual report 9, 305.
- M. Davis. 1987. Production of Conditional Simulations via the LU Triangular Decomposition of the Covariance Matrix *Mathematical Geology*, Vol. 19, No. 2.
- M. Davis. 1987. Generating large stochastic simulations-The matrix polynomial approximation method. *Mathematical Geology*, Vol. 19(2), 99-107.
- C. Deutsch and A. Journel. 1998. *GSLIB - Geostatistical software library and users guide*. Oxford University press, 2nd Edition p12-13, p139.
- C. Deutsch. 2002. *Geostatistical Reservoir Modeling*. Oxford University Press, p 58.
- C. Dietrich and G. Newsam. 1995. Efficient generation of conditional simulations by Chebyshev matrix polynomial approximation to the symmetric square root of the covariance matrix. *Mathematical Geology*, Vol. 27(2), 207-228.
- P. Dowd and C.Sara. 1993. An extension of the LU decomposition method of simulation, in Armstrong. M. and Dowd, P. A eds., *Geostatistical simulations*: Kluwer Academic, London, p 23–36.
- I. Glacken. 1997. Change of Support and Use of Economic Parameters for Block Selection. in *Geostatistics Wollongong '96*, E. Y. Baafi and N. A. Schofield (eds.), volume 2, p 811-821, Kluwer, 1997.
- R. Horn and C. Johnson.1985. *Matrix Analysis*. Section 7.2. Cambridge University Press. ISBN 0-521-38632-2.
- E. Isaaks and R. Srivastava. 1989. *An introduction to applied geostatistics*. Oxford University Press, Oxford, p 323.

A. Journel and C. Huijbregts CJ, 1978. Mining geostatistics. London, Academic Press. p1-3, p30.

O. Leuangthong, D. Khan, C. Deutsch. 2008. Solved Problems in Geostatistics. Wiley, John & Sons. p60

J. Ortiz and C. Deutsch. 2007. Local Recoverable Reserves Prediction with Block LU Simulation. CCG annual report 9, 301.

J. Vargas-Guzmán. 2004. Fast modeling of cross-covariances in the LMC: A tool for data integration, Vol.18, No. 2

H. Wackernagel, P. Petitgas and Y. YoutTalt, 1989. Overview of methods for coregionaliration. in Armstrong. M, ed. Geostatistics, v. I: Kluwer Acad. Publ. Co. Dordrecht. Netherlands, p 409-420.

E. Xavier. 2008. Change of Support for estimating local Block Grade Distributions. Mathematical Geology, Vol. 40, 671-688

Appendix A

FORTRAN Code: Block LU simulation with AMC

A GSLIB-like program, called blusim_AMC, is developed to calculate the resource reservoir simulation with multi-variable. The program does not require the LMC model. To calculate the cross variogram between variable, it need the upper diagonal correlation coefficient matrix. The parameters required for this program are:

```
1
2
3           Parameters for BLUSIM_AMC
4           *****
5
6 START OF PARAMETERS:
7 small_data_positive_ws.dat - file with data
8 2 3 0 6 5 0 - columns for X, Y, Z, varID, var, wt (var ID runs from 1 to nvar)
9 3 3 - number of variables, number of variables to simulate
10 1 2 3 - ID of variables to simulate (ID s correspond with ID s in data file)
11 0.301 0.127 -0.524 - upper diagonal correlation matrix. (eg. correlation 1,2 correlation 1,3. correlation 2,3 .... )
12 0 1.0e21 - trimming limits
13 2 20 - transform the data (0=no, 1=yes 2=locally transform, and how many data used)
14 blusim.trn - file for output trans table
15 0 - consider ref. dist (0=no, 1=yes)
16 histsmth.out - file with ref. dist distribution (file contains ref dist for ALL variables)
17 1 2 3 4 5 6 7 8 - columns for vr1 and wt1, vr2 and wt2, vr3 and wt3 ...
18 0.2 6 4 126 0 70 - zmin, zmax(tail extrapolation) (nvar*2 numbers, info for var1, var1, var2, var2 ...)
19 1 0.22 1 1.47 1 0 - lower tail option, parameter ( need nvar*2 numbers on this line)
20 1 6.18 1 126 1 70 - upper tail option, parameter ( need nvar*2 numbers on this line)
21 0 - debugging level: 0, 1, 2, 3
22 blusim.dbg - file for debugging output
23 blusim.out - file for kriged output
24 2 - number of realizations to generate
25 24 4425 50 - nx, xmn, xsiz - Pannel Size
26 14 25 50 - ny, ymn, ysiz
27 1 0.5 1 - nz, zmn, zsiz
28 1 1 1 - nbx, nby, nbz- SMUs per Pannel
29 2 2 1 - x, y and z block discretization
30 0 - number of cutoffs for reporting
31 0.4 - cutoffs
32 0 - option to consider an equivalent grade for all variable (0->not used)
33 1 0.8 0.6 - for each variable SIMULATED need a factor to multiply (only used with equivalent grade)
34 0 - option to output all SMU realizations (0=no)
35 SMUreals.out - file for the SMU realizations
36 67515 - random number seed
37 3 10 3 10 3 10 - min, max data for kriging (need nsin*2 numbers, IF THIRD NUMBER=-1 DONT CONSIDER DIFFERENT MIN/MAX)
38 0 - max per octant (0-> not used) (cannot specify data type, uses total of all available data)
39 250 250 250 - maximum search radii
40 00 0.0 0.0 - angles for search ellipsoid
41 1 1 - Semivariogram for 1 and 1
42 2 0.001 - nst, nugget effect
43 2 0.726 00.0 0.0 0.0 - it, cc, ang1, ang2, ang3
44 35 35 35 - a_hmax, a_hmin, a_vert
45 1 0.273 35.0 0.0 0.0 - it, cc, ang1, ang2, ang3
46 200 200 200 - a_hmax, a_hmin, a_vert
47 2 2 - Semivariogram for 2 and 2
48 2 0.001 - nst, nugget effect
49 2 0.710 00.0 0.0 0.0 - it, cc, ang1, ang2, ang3
50 30 30 30 - hmax, a_hmin, a_vert
51 2 0.289 00.0 0.0 0.0 - it, cc, ang1, ang2, ang3
52 200 200 200 - hmax, a_hmin, a_vert
53 3 3 - Semivariogram for 3 and 3
54 2 0.001 - nst, nugget effect
55 2 0.510 00.0 0.0 0.0 - it, cc, ang1, ang2, ang3
56 35 35 35 - hmax, a_hmin, a_vert
57 2 0.689 00.0 0.0 0.0 - it, cc, ang1, ang2, ang3
58 200 200 200 - hmax, a_hmin, a_vert
59
60
```

- Information about the data (line 6-10). Name of the data file, columns for the coordinates X, Y, Z and the variable ID, variable and the weights. The variable ID must be indicated in the data file. The number of variable for line 8 must equal to the highest ID of the variable in the data file.
- Correlation information and trimming options (line 10-11). Line 10 is the upper diagonal correlation matrix of among the variables. Line 11 is the trimming options

- Transformation options (line 12-19). Line 12 and 13 is the option weather to transform the data and the output file. Line 14 to 16 is weather to consider a transformation table. This is in one file with each column representing a different variable.
- Debugging options and output file (line 20-22). This option allows the user to check the calculation, but if set higher debugging level; the debugging file could be large. Line 22 is the name of output file.
- Parameters for simulation grid and discretization (line 23-28). Line 23 is number of realizations, Line 24-26 is the gird definition. Line 27 is number of discretization points in the panel; Line 28 is the number of Selective Mining Unit in the panel. Note that number of discretization points divided by number of SMU should be a whole number.
- Parameters for cutoff report (line 29- 34). Line 29 and 30 is number to cutoffs and value of cutoffs. Line 31 is the option to consider equivalent grade, which requires a factor for each data simulated. Line 32 is the option weather to write out all SMU realizations, which could be large if the realization is large.
- Seed for random number (line 35)
- Search information (line 36-39). Line 36 is the min and max data used in simulation. Line 37 is whether to consider octant search. Line 38 is the search radii.
- Variogram model parameters (line 40-end). These are only the direct variogram models for the data.