Titre: Title:	New Developments in Covariance Modeling and Coregionalization for the Study and Simulation of Natural Phenomena
Auteur: Author:	Min Liang
Date:	2015
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Liang, M. (2015). New Developments in Covariance Modeling and Coregionalization for the Study and Simulation of Natural Phenomena [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2008/

UNIVERSITÉ DE MONTRÉAL

NEW DEVELOPMENTS IN COVARIANCE MODELING AND COREGIONALIZATION FOR THE STUDY AND SIMULATION OF NATURAL PHENOMENA

MIN LIANG
DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE MINÉRAL)
DÉCEMBRE 2015

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

NEW DEVELOPMENTS IN COVARIANCE MODELING AND COREGIONALIZATION FOR THE STUDY AND SIMULATION OF NATURAL PHENOMENA

présentée par : LIANG Min
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

M. CHOUTEAU Michel, Ph. D., président
M. MARCOTTE Denis, Ph. D., membre et directeur de recherche
M. GLOAGUEN Erwan, Ph. D., membre
M. EMERY Xavier, Ph. D., membre externe

DEDICATION

To my parents

ACKNOWLEDGMENTS

Working as a Ph.D student of geostatistics was an amazing and challenging experience to me. During the five years, many people helped me in my research project. Without their support, this thesis would not be possible.

First I would like to express the deepest appreciation to my research director Professor Denis Marcotte, for his guidance, encouragement and financial support. Every time I met problem on my project, he always provides me his ideas and suggestions patiently. Without his contribution, I would not be able to complete my research work and this dissertation.

I am deeply thankful to all members of the jury, professor Michel Chouteau, Xavier Emery and Erwan Gloaguen, for reading and evaluating this thesis.

I thank Nicolas Benoit from Gelogical Survey of Canada for providing data and technical supports.

I sincerely acknowledge China Scholarship Council, for offering me scholarship of studying.

I also want to thank my colleagues and friends in the lab, Martine, Véronique, Pejman, Abderrezak, Eric, Alain, Clémence, Louis, Hassan and Raphael. Thank you for all information you shared, research suggestions, good restaurants, planting, baby caring and so on.

Moreover, I am especially grateful for my husband Xiaoyu for his support on my research project (discussion on geostatistics, suggestion and English improvement on my articles) and on my life (cooking delicious food). I also want to thank my parents for their constant encouragement and my daughter Sophie for her birth and smile which give me plenty of energy.

RÉSUMÉ

La géostatistique s'intéresse à la modélisation des phénomènes naturels par des champs aléatoires univariables ou multivariables. La plupart des applications utilisent un modèle stationnaire pour représenter le phénomène étudié. Il est maintenant reconnu que ce modèle n'est pas assez flexible pour représenter adéquatement un phénomène naturel montrant des comportements qui varient considérablement dans l'espace (un exemple simple de cette hétérogénéité est le problème de l'estimation de l'épaisseur du mort-terrain en présence d'affleurements). Pour le cas univariable, quelques modèles non-stationnaires ont été développés récemment. Toutefois, ces modèles n'ont pas un support compact, ce qui limite leur domaine d'application. Il y a un réel besoin d'enrichir la classe des modèles non-stationnaires univariable, le premier objectif poursuivi par cette thèse.

Dans le cas multivariable, en plus du choix stationnaire, la plupart des applications sont limitées à l'utilisation du modèle linéaire de corégionalisation (LMC). Cette limitation est probablement due à 1) la facilité d'évaluer l'admissibilité du LMC, et 2) le manque de méthodes de simulation rapides pour les modèles qui ne sont pas LMC (N-LMC). Des progrès significatifs ont été faits sur le premier point récemment, mais moins sur le second. Par conséquent, le second objectif principal de cette thèse est de fournir une méthode de simulation rapide pour N-LMC.

Cette thèse se compose principalement de trois articles. Le premier article utilise un modèle non-stationnaire univariable existant pour étudier le problème de l'estimation de l'épaisseur du mort-terrain en présence de nombreux affleurements. L'affleurement a une influence locale sur la distribution de l'épaisseur du mort-terrain car l'épaisseur y est nulle par définition. Le modèle non-stationnaire est ici utilisé pour limiter la distance d'influence des affleurements. À l'intérieur de cette distance d'influence, les paramètres de covariance sont supposés être des fonctions régulières simples de la distance à l'affleurement le plus proche. Au-delà de cette distance d'influence, la fonction de covariance de l'épaisseur de mort-terrain est supposée être stationnaire et l'influence des affleurements ne se fait plus sentir. La méthode est testée avec des données réelles. Les résultats montrent que le modèle non-stationnaire améliore la précision et le réalisme de l'estimation, en particulier à proximité des affleurements.

Le deuxième article introduit de nouvelles fonctions non-stationnaires avec support compact, remplissant ainsi le premier but principal de la thèse. Les fonctions développées sont issues du modèle sphérique. Elles sont dérivées par convolution d'hypersphères dont le rayon varie spatialement en $\Re^n$. On applique ensuite une transformation de Radon dont l'ordre contrôle

la continuité de la fonction de covariance résultante. Les expressions explicites des covariances non-stationnaires isotropes sont dérivées pour les modèles sphériques, cubiques, et penta-sphériques. Aussi une méthode de simulation des modèles non stationnaires utilisant la moyenne pondérée d'un bruit blanc gaussien est décrite pour les cas isotrope et anisotrope.

Le troisième article présente une méthode de simulation efficace pour N-LMC basée sur la transformée de Fourier rapide (FFT) et la moyenne mobile (GFFTMA), réalisant ainsi le deuxième but principal de la thèse. Cette méthode permet de simuler des variables avec des structures spatiales différentes. Dans le domaine spectral, les matrices de densités sont décomposées séparément, en valeurs propres-vecteurs propres, à chaque fréquence discrète. Le champ corrélé est obtenu par la multiplication, à chaque fréquence, de la racine carrée de la matrice spectrale par les spectres de bruits blancs gaussiens suivi d'une transformée inverse de Fourier. Ceci permet d'imposer les structures spatiales directes et croisées désirées pour chaque variable. Cette méthode possède une complexité Nlog (N) (N, le nombre de pixels à simuler). La méthode GFFTMA est testée sur des exemples synthétiques à deux ou trois variables et formés de différentes combinaisons de modèles parmi les sept modèles de base disponibles. Tout les cas testés montrent des réalisations dont les variogrammes expérimentaux directs et croisés reproduisent très bien le modèle cible.

Les trois articles utilisent comme étude de cas le problème de l'estimation et la simulation de l'épaisseur du mort-terrain dans les basses terres de Saint-Laurent et de la l'est de la Montérégie, Québec, Canada. Les nouveaux outils développés dans cette thèse permettent de mieux étudier les phénomènes naturels aussi bien pour le cas univariable, où les modèles non stationnaires à support compact offrent plus de flexibilité, que dans le cas multivariable stationnaire, où une méthode de simulation non-LMC efficace basée sur la FFT a été développée. Ensemble, ces outils devraient permettre d'améliorer la modélisation des phénomènes naturels.

ABSTRACT

Geostatistics focus on modeling natural phenomena by univariate or multivariate spatial random fields. Most applications rely on the choice of a stationary model to represent the studied phenomenon. It is now acknowledged that this model is not flexible enough to adequately represent a natural phenomenon showing behaviors that vary substantially in space (a simple example of such heterogeneity is the problem of estimating overburden thickness in the presence of outcrops). For the univariate case, a few non-stationary models were developed recently. However, these models do not have compact support, which limits in practice their range of application. There is a definite need to enlarge the class of univariate non-stationary models, a first goal pursued by this thesis.

In the multivariate case, in addition to the stationary choice, most applications are limited to use of the linear model of coregionalization (LMC). This limitation is probably governed by 1) the ease of assessing the admissibility of the LMC, and 2) the lack of availability of fast simulation methods for models that are not LMC (N-LMC). Significant progress were made on the first point recently, but not much on the second. Hence, the second main goal of this thesis is to provide a fast simulation method for stationary non-LMC.

This thesis is mainly composed of three articles. The first article uses the existing univariate non-stationary model to study the problem of estimating the overburden thickness in presence of many outcrops. The outcrop has local influence on overburden thickness distribution as the thickness value is zero on the outcrop. The non-stationary model is used to restrict the distance of influence of outcrops. Within this distance of influence, covariance parameters are assumed to be simple functions of the distance to the nearest outcrop. Beyond the distance of influence of outcrops, the thickness covariance is assumed stationary. The method is tested with real data. The results show that the non-stationary model improves the precision of estimation and provides realistic map, especially at points close to outcrops.

The second article develops new non-stationary functions with compact support, thus fulfilling the first main goal of the thesis. The developed functions include the non-stationary form of the spherical family model. It is derived by convolving hyperspheres (with spatially varying radius) in $\Re^n$ followed by a Radon transform. The order of the Radon transform controls the differentiability of the covariance functions. Closed-form expressions for the isotropic non-stationary covariances are derived for the spherical, cubic, and penta-spherical models. Also a simulation method of the non-stationary models is described by weighted average of independent standard Gaussian variates in both the isotropic and the anisotropic case.

The third article presents a very fast simulation method for N-LMC, the general fast Fourier transform and moving average (GFFTMA), thus realizing the second main goal of the thesis. This method makes available to simulate variables following different spatial structures. In spectral domain, the spectral density matrices are eigen-decomposed separately at each discrete frequency. Correlated spectrum for each variable is produced by the decomposed matrices multiplied with the spectrum of Gaussian white noise. Then taking the inverse Fourier transform, the random field of each variable in spatial domain is created. The CPU- time of this method increases as $Nlog(N)$ (N, the number of pixels to simulate). The GFFTMA is tested in simulation of synthetic examples with two and three variables for different combinations of the seven available models. All the realizations produced fit the desired covariance model well.

The three articles use as illustrative case study the problem of estimating and simulating the overburden thickness in the Saint-Laurence lowlands, Montérégie Est, Québec, Canada. As a whole, this thesis provides new tools to better study the natural phenomena both in the univariate case, where compactly supported non-stationary models offer more flexibility, and in the stationary multivariate case, where an efficient FFT based non-LMC simulation method is derived. Together, they should help improve modeling of natural phenomena.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

FFTMA	Fast Fourier transform moving average
GFFTMA	General fast Fourier transform moving average
LMC	Linear model of coregionalization
N-LMC	Non-linear model of coregionalization
ME	Mean error
MAE	Mean absolute error
RMSE	Root mean square error
K-S	Kriging based on stationary model without outcrop information
K-SO	Kriging based on stationary model incorporating outcrop information
K-NS	Kriging based on non-stationary model incorporating outcrop information
a	Correlation range
d	Distance
f	Function
g	Covariogram
h_{ij}	Distance vector between x_i and x_j
k	Kernel function
$\mathbf{k}$	Covariance vector between observations and aim point
m	Mean
n	Number
p	Proportion
x	Coordinate of a point
A	Amplitude
C	Covariance function
C^S	Stationary covariance function
C^{NS}	Non-stationary covariance function
Cov	Covariance
$\mathcal{D}(f)$	Differential operator
E	Expectation
F	Facies
$\mathcal{F}$	Fourier transform
I	Incomplete beta function
I	Indicator function

$\mathcal{I}(f)$	Integral operator
$\mathbf{K}$	Covariance matrix among observations
$\mathcal{K}_\nu$	Bessel function of order ν
L	Lagrange function
P	Transition kernel
Prob	Probability
R	Correlation
R^S	Stationary correlation function
R^{NS}	Non-stationary correlation function
Rf	Radon transform
S	Power spectral density function
V	Volume
Var	Variance
X	Euclid's hat
$\mathcal{X}_i$	Poission point process
$Y(x)$	Random function
$Z(x)$	Random function
$Z^*(x)$	Estimate of $Z(x)$
Z_{CS}	Conditional simulation
Z_S	Unconditional Simulation
Z_S^*	Estimate using simulation
α	Weight of secondary variable in cokriging
δ	Gaussian random variable
λ	Weight in Kriging
γ	Variogram
μ	Drift
μ_l	Lagrange multiplier
ρ	Correlation function
σ	Standard deviation
ω	Dilution function
φ^{Wu}	Wu's function
φ^{We}	Wendland's function
Σ	Kernel Matrix
$\Re^n$	Real space in n-dimensions
Γ	Gamma function

CHAPTER 1 INTRODUCTION

1.1 Basic concepts

Many natural phenomena can be characterized by spatial variables, such as soil properties, the concentration of pollutants, or the temperature in a region. The studies of spatial variables often involve to determine the value on every point based on imperfect understanding of the phenomena and limited known information. This brings uncertainty into the study. Geostatistical methods can describe the spatial structure hidden by the randomness of spatial variables, such as orientation, smoothness of transition, and so on. To characterize the spatial structure in a simple way, some hypotheses are commonly assumed including typically : 1) the stationarity of the covariance model for univariate study, and 2) the linear model of coregionalization for multiple variables. These assumptions limit the applicability of the methods in a complex real world.

Stationarity is a common assumption in geostatistics. It means that the spatial structure of a variable, typically the mean and the covariance, do not change with locations over the study domain. In fact, many phenomena exhibit local varying structures. For example, variation of the annual precipitation is much greater in mountain than in flat plain (Paciorek and Schervish, 2006). Likewise, the distribution of atmospheric pollutants show large spatial variability depending on the source location, meteorology and the chemical reactions between the source and receptor (Fuentes, 2001). Another typical example is distribution of overburden thickness over a domain. In a plain area, the overburden usually has a uniform thickness. But the presence of outcrops interrupts the continuity, which leads to a locally different distribution of overburden thickness in the vicinity of the outcrops. To better quantify these location dependent structures, non-stationary models emerged and became an appealing solution in recent decades. Despite the powerfulness, computational limitations can happen when data is abundant as the existing non-stationary model do not have compact support. The resulting covariance matrix then becomes large and non-sparse. One has to resort the local neighborhoods to alleviate the problem however that may generate undesirable discontinuities. There is a definite need for non-stationary compactly supported covariance models as an alternative to local neighborhood approaches.

In many instances, more than one variable are studied simultaneously. For example, the geophysical data can reflect the intrinsic properties of overburden and could be considered to better estimate the overburden thickness (Hunter *et al.*, 1984; Caron *et al.*, 2013; Blouin *et al.*, 2013). The linear model of coregionalization (LMC) is commonly used in modeling of

correlated multiple variables. It requires each covariance of variables is a linear combination of the same spatial model. The LMC is practical in limited cases, such as geochemistry and study of animal abundance (Chilès and Delfiner, 1999). In fact, the non-linear model of coregionalization (N-LMC) describes more general and common situations, where variables can follow entirely different structures. For example, in geophysics, the density and magnetic susceptibility are closely correlated while their indicators, gravity and magnetic field data, follow different spatial structures. The gravity field has smoother behavior than the magnetic susceptibility (Maus, 1999). As another example, in geology the spatial field of overburden thickness usually can be described by a model which has linear behavior at the origin. One of its correlated variable, geological domain can be modeled by thresholding a continuous Gaussian variable. To obtain realistic geological domains, it is needed to adopt a variogram with parabolic behavior for the latent variable, hence necessitating a N-LMC to simulate jointly the thickness and the latent variable. To be practical, an efficient method is needed to simulate such fields with N-LMC.

1.2 Research problems

The study of natural phenomena is often limited in the univariate case to methods using stationary models. In the multivariate case, the stationary LMC is by far the most often used model. These models lack flexibility and are not deemed sufficient to represent adequately the complexity of many natural phenomena. There is a definite need to enlarge the class of univariate non-stationary models and to escape from the LMC straitjacket to adequately represent the studied natural phenomena.

More specifically, the following research problems arise :

1. In univariate case, how to model the heterogeneity of a random field with a non-stationary covariance ? How to define non-stationary models with compact support that allows to obtain sparse covariance matrices ? How to simulate efficiently these non-stationary fields ?
2. In multivariate case, how to simulate N-LMC field efficiently ?

1.3 Objectives and contributions of the thesis

The objectives of this thesis are :

1. To develop a non-stationary model to incorporate the outcrop information in overburden thickness estimations.

2. To develop a class of compactly supported non-stationary covariance functions.
3. To develop an approach of simulation based on the non-stationary covariance model.
4. To develop an efficient method of multivariate simulation of N-LMC.

This thesis achieves the following original contributions :

1. A non-stationary model which includes outcrops position in overburden thickness estimation. It produces realistic maps and improves the thickness estimation precision.
2. A class of non-stationary covariance functions with compact support, of which a special case is a non-stationary form of spherical family model. The non-stationary model is applied in overburden thickness estimation incorporating classification of the geological domain and enables a reduction of the estimation error and a more realistic transition between two geological domains.
3. A method for simulating non-stationary models with compact support in both isotropic and anisotropic cases.
4. A general method to simulate multivariate fields with N-LMC based on the fast Fourier transform - moving average method (FFTMA).

1.4 Structure of the thesis

This thesis consists of nine chapters. The first chapter provides a brief overview of the thesis, including basic introduction, research problems, objectives, contributions and structure.

Chapter 2 presents a literature review encompassing the origin and development of major topics of the thesis.

Chapter 3 describes the organization of the thesis and the pertinence of the articles with regards to the objectives of the thesis.

Chapter 4 introduces theoretical background including the related theoretical knowledge and techniques in geostatistics.

Chapter 5 presents the article "A comparison of approaches to include outcrop information in overburden thickness estimation" by Min Liang, Denis Marcotte and Nicolas Benoit, published in "Stochastic Environmental Research and Risk Assessment" 28.7 (2014) : 1733-1741, DOI : 10.1007/s00477-013-0835-6. In the article, a non-stationary covariance model is used to describe spatial structure of overburden thickness incorporating outcrops information.

Chapter 6 contains the article "A class of non-stationary covariance functions with compact support" by Min Liang and Denis Marcotte, published in "Stochastic Environmental Research and Risk Assessment" (2015) : 1-15, DOI : 10.1007/s00477-015-1100-y. This article presents

a family of non-stationary covariance functions with compact support. Moreover, a non-stationary simulation method in both isotropic and anisotropic cases is proposed.

Chapter 7 contains an article submitted to the journal of "Computers & Geosciences" at June 6, 2015 with the title of "Simulation of non-linear coregionalization models by FFTMA" by Min Liang, Denis Marcotte and Pejman Shamsipour. A fast simulation method GFFTMA is proposed for multivariate fields of N-LMC.

Chapter 8 is a general discussion of the dissertation.

The last chapter summarizes the contributions and highlights of the thesis, addresses the limitations, and provides suggestions for future researches.

CHAPTER 2 LITERATURE REVIEW

The literature review covers three main parts. First I briefly review the main existing non-stationary models. Secondly, the main contributions on compactly supported covariance functions are described. The last part of the review bears on multivariate geostatistics and available simulation methods.

2.1 Non-stationary model

2.1.1 Stationary model

For a random function $Z(x), x \in \Re^n$, the spatial variation at arbitrary two points x_i and x_j is characterized by its covariance $C(x_i, x_j)$ expressed by

$$C(x_i, x_j) = E[Z(x_i) - m(x_i)][Z(x_j) - m(x_j)]. \tag{2.1}$$

In general, the covariance function $C(x_i, x_j)$ shows how the correlation between two points changes with their locations x_i and x_j. For mathematical convenience, conventional models assume that the random function $Z(x)$ is second-order stationary, which means for any point x and $x + h$ of $\Re^n$

$$E[Z(x)] = m, \tag{2.2}$$

$$E[Z(x) - m][Z(x + h) - m] = C(h). \tag{2.3}$$

Thus, the covariance function $C(x_i, x_j)$ is independent of specific locations x_i and x_j and varies only as a function of vector distance $h_{ij} = x_i - x_j$.

Another tool to analyze the spatial distribution of a random function $Z(x)$ is the variogram. It shows how the differences between $Z(x_i)$ and $Z(x_j)$ evolve with their locations. In the stationary case, similar with the covariance, the variogram γ depends on the vector separation h,

$$\gamma(h) = \frac{1}{2}\mathrm{Var}[Z(x + h) - Z(x)]. \tag{2.4}$$

The commonly used variogram models in geostatistics include spherical family model, exponential family model, Gaussian model, Cauchy class model, Matérn model, $\mid h \mid^{\alpha}$ model, logarithmic model and hole effect model (Chilès and Delfiner, 1999). Note that the $\mid h \mid^{\alpha}$ and the logarithmic model correspond to the case where only the increments $Z(x + h) - Z(x)$ are stationary, not $Z(x)$ itself. This means the class of variogram functions is wider than the

covariance functions. The random function with stationary increments is generalized in the next section.

2.1.2 Non-stationary covariance functions

The non-stationary model describes a more general situation that allows to have location dependent spatial structures. A random function $Z(x)$ can be decomposed as sum of a drift component $m(x)$ and a zero-mean random variable $\delta(x)$:

$$Z(x) = m(x) + \delta(x). \tag{2.5}$$

The non-stationarity can be introduced in the drift component, the stochastic component or both.

To describe the non-stationarity of the drift, Matheron (1973) proposed the intrinsic random functions that characterize the drift as a linear combination of basis functions $f_l(x)$ (for example the monomials, logarithmic or trigonometric functions)

$$m(x) = \sum_l a_l f_l(x). \tag{2.6}$$

An increment of order k is obtained by forming a linear combination $Z(\lambda) = \sum_{i=1}^{n} \lambda_i Z(x_i)$ such that :

$$\sum_i^n \lambda_i f_l(x_i) = 0, \quad \forall l = 0, 1, \cdots, k. \tag{2.7}$$

Such linear combinations are said "authorized". Then, one has $E[Z(\lambda)] = 0$ and $Var(Z(\lambda)) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j K(x_j - x_i)$, where $K(x_j - x_i)$ is called "generalized covariance". The variance of the increment depends only on the separation vectors between points, which means the drift is filtered out. The variance of the increments does not call for the knowledge of the coefficients a_l. Note that $-\gamma(h)$ is the generalized covariance function for increments of order 0. It filters out the influence of a constant unknown mean.

Details and applications of the generalized covariance function are described in the literature Matheron (1973) and Kitanidis (1993). More studies on the non-constant mean $m(x)$, for example universal kriging, differencing or splines, are well documented in Cressie (1993) and Pintore and Holmes (2004).

This thesis focuses on the non-stationarity of random component $\delta(x)$, which is characterized by a non-stationary covariance model. The drift component is assumed constant in the context of the thesis. In recent decades, the main approaches reported in construction of a non-

stationary covariance function are moving window, space deformation and kernel convolution.

Moving window approach

Haas (1990a) proposed a moving window method to model the local feature where stationarity is assumed locally. A moving window is given centered at every point of interest. Then in one window, the spatial covariance structure is estimated by only neighbors in this window. As the window moves through the whole study area, a number of local covariance models with different parameters are obtained that draw the heterogeneity of the field. Developments and applications of this method can be found in Haas (1990b, 1995), Horta *et al.* (2010) and Soares (2010). The local characteristics of the field can be shown in moving windows, however the general covariance structure of the whole area cannot be modeled (Pintore and Holmes, 2004). Moreover, artefact discontinuities between the models and estimations may be introduced with this approach.

Space deformation approach

Another approach estimating non-stationary model is spatial deformation. Based on multidimensional scaling, Sampson and Guttorp (1992) firstly computed a representation of sampling stations. The spatial dispersion was approximated by a monotone function of distances between two points. Then they computed a smooth mapping for the geographic representation of the sampling stations into their multidimensional scaling representation. The anisotropy and nonstationarity of the nature field is included in the composition of the mapping and the monotone function of multidimensional scaling representation. Further applications and extensions of this approach are found in Guttorp *et al.* (1994); Smith (1996); Meiring *et al.* (1997); Damian *et al.* (2001); Schmidt and O'Hagan (2003); Boisvert and Deutsch (2008). This method however concentrates essentially on the model anisotropy directions and ratios. It cannot model non-stationarity of scale, of noise or of differentiability.

Kernel based approach

Priestley (1965) states that if a stationary Gaussian process $Z(x)$ has a correlation function $R(d)$ given by

$$R(d) = \int_{\Re^2} k(x)k(x-d)\mathrm{d}x, \tag{2.8}$$

the process $Z(x)$ can be expressed by a convolution of a Gaussian white noise process $\delta(\cdot)$ and a kernel function $k(\cdot)$ like

$$Z(x) = \int_{\Re^2} k(x-u)\delta(u)\mathrm{d}u. \tag{2.9}$$

Higdon et al. (1999) extended this characterization in non-stationary cases so that the kernel evolves over spatial locations. Then the non-stationary correlation function of a Gaussian process on two points x_i, x_j would be represented by convolution of two kernels at x_i and x_j

$$R^{NS}(x_i, x_j) = \int_{\Re^2} k_{xi}(u)k_{xj}(u)\mathrm{d}u, \tag{2.10}$$

where $k_x(u)$ denotes a kernel centered at location x. Then they focused on the Gaussian kernel that can be represented by an ellipse with local scaling and rotation at a given location. Moreover, parameters of kernel function $k_x(u)$ were discussed that control the smoothness of the kernel.

Paciorek (2003) extended the kernel convolution method of Higdon *et al.* (1999) and produced explicit non-stationary correlation functions with locally varying geometric anisotropies. Then Paciorek and Schervish (2006) developed a class of non-stationary covariance functions including Matérn covariance model and rational quadratic covariance model. Pintore and Holmes (2004) described a framework to construct non-stationary covariance functions of Gaussian processes by evolving the stationary spectrum over space. Two possible spectral decomposition of covariance functions were mentioned, the Fourier and Karhunen-Loève expansions. The non-stationary form of Gaussian and Matérn covariance functions are presented in this article. Stein (2005) proposed an explicit non-stationary covariance function that allows local anisotropy and differentiability varying over space and gave a special form of Matérn model. Then Mateu *et al.* (2013) extended Stein (2005)'s result and gave a general explicit form of non-stationary covariance function. Although the functions contain a compactly supported function, the resulting function does not show the behavior of having a compact support.

Example : Modeling overburden thickness in presence of outcrops with stationary model : Pros and cons

One typical case of non-stationarity occurs naturally in studying overburden thickness with presence of outcrops. Overburden is the material on top of bedrock, usually including soil, sand, sediment, gravel and till (Thrush, 1968). The thickness of overburden is commonly mo-

deled with stationary covariance, which gives simple computation. However, when outcrops occurs in the studied area, the stationary model can result in a few problems.

One intuitive way to treat outcrops is considering them as zero value in data set. The outcrops data can be more abundant than thickness data as it is observed preferentially. The numerous zeros can exaggerate the outcrop sizes thereby leading to a strong bias in the kriging estimation. There are several ways to alleviate the problem by keeping stationary covariance but modifying the kriging method. One way is to add inequality constraints to kriging (Dubrule and Kostov, 1986), and force the overburden thickness being negative on outcrops and positive elsewhere. The main drawback of this method is it involves quadratic programming and therefore it is more difficult to apply the kriging as the solution is obtained iteratively. Another way is to use Gibbs samplers (Geman and Geman, 1984) to add constraints in conditional simulation (Freulon and de Fouquet, 1993). For this method, first a classical conditional simulation is obtained using only observations. Then all simulated points are checked to satisfy the constraint (negative simulated value on outcrops and positive elsewhere). An iterative approach based on the Gibbs sampler is used. A sufficient number of realizations are constructed by repeating this process. Finally, the overburden thickness estimate is taken as the average of these realizations. This approach has precise estimation close to outcrops. However, it is very CPU intensive and requires a Gaussian hypothesis for the Gibbs sampler.

In the first article (Chapter 5), a non-stationary model is proposed for estimation of overburden thickness incorporating outcrop locations. In the non-stationary model, the range of influence of outcrops is defined. Within this scope, the covariance function is non-stationary with parameters (correlation range and nugget) that are functions of distance to the nearest outcrop. Outside the range of influence, the distribution of overburden is not affected by outcrops and assumed stationary.

2.2 Stationary compactly supported covariance functions

The compactly supported function for stationary covariance was well studied in last few decades (Matheron, 1965; Wendland, 1995; Wu, 1995). In the non-stationary case, the compactly supported functions are rarely investigated. In this thesis, I focus on extension of one type of compactly supported functions, which is based on convolution of dilution functions such as Wu's function, into a non-stationary case. First, I introduce approaches of constructing compactly supported functions and the resulting radial functions. Then, some methods are reviewed on treating the non-compactly supported covariance functions to gain more computation efficiency.

2.2.1 Direct construction of compactly supported radial functions

Matheron (1965) proposed two operators to obtain covariogram functions with compact support, 'La montée' and 'La descente', and gave examples in the stationary and isotropic case. 'La montée' is an integral operator, for a positive definite function $f(u), u \geq 0$ in $\Re^n$,

$$\mathcal{I}(f)(r) = \int_r^\infty u f(u) \mathrm{d}u. \tag{2.11}$$

The $\mathcal{I}$ operator is a special case of Radon transform of the second order. The resulting functions $I(f)(r)$ is positive definite in $\Re^{n-2}$ for $n > 3$. They have same compact support but higher smoothness than the original function $f(r)$. Wendland (1995) developed a group of compactly supported radial basis functions by integral operator starting with Askey's power function (Askey, 1973).

'La descente' is a differential operator that preserves the same support as original function but decrease the smoothness. For $u \geq 0$,

$$\mathcal{D}(f)(r) = -\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} f(r) \tag{2.12}$$

which is positive definite in $\Re^{n+2}$. Wu (1995) constructed a class of functions with compact support by differential operator $\mathcal{D}(f)(r)$ based on auto-convolution of a cut-off polynomial function.

Convolution is another operator to create compactly supported covariance functions. The property of positive definiteness or compact support of two functions is kept in their convolution. The stationary spherical model is proposed by Matheron (1965) by auto-convolution of hyperspheres defined in $\Re^n$.

2.2.2 Development of compactly supported covariance functions in geostatistics

Most of popular covariance functions used in geostatistics (such as exponential, Gaussian or Matérn family model) do not have compact support. This means when the distance between two points is much larger than the correlation range, their covariance is still a non-zero value. The very small positive covariance is insignificant, but intensify the computation tasks. Several methods were proposed to create compact support for covariance functions (Moreaux, 2008). One approach is to neglect small correlations below a certain threshold and set these covariances to zero (Rygaard-Hjalsted *et al.*, 1997). However, this approach does not preserve positive definiteness of the covariance function (Furrer *et al.*, 2006). A better approach which can keep positive definiteness of covariance functions is tapering these functions by a

compactly supported function.

Furrer *et al.* (2006) showed how to modify the Matérn model to be compactly supported, by using compactly supported functions (such as spherical or Wendland covariance functions) as tapering functions. Also they gave conditions to ensure positive definiteness of the resulting function. The kriging estimates obtained from the tapered function are asymptotically optimal, but the kriging variances are biased and need rescaling by a factor depending on the data locations. Conditions on tapering functions for other covariance functions than the Matérn class could not be derived by Furrer *et al.* (2006). Besides condition of positive definiteness, we also have to focus on properties of the tapering function, due to the result covariance inherits some of the main properties of the tapering function. For example, a Gaussian covariance tapered by a spherical covariance is no more differentiable at the origin.

Gneiting (2002) presented turning bands operator to construct correlation functions of hole effect model. By transforming a supported function ϕ_n defined in $\Re^n$, the resulting function ϕ_{n-2} still has compact support, same behavior at origin and positive definiteness in $\Re^{n-2}$.

In multivariate case, Porcu *et al.* (2013) developed stationary covariance models with compact support based on scale mixtures of Askey functions (Askey, 1973). Following that, Kleiber and Porcu (2015) obtained locally-stationary covariance models with compact support based on functions of Wendland (1995) and Gneiting (2002) .

Mateu *et al.* (2013) presented a closed-form of a non-stationary covariance function, which is constructed by a complete monotonic function and a compactly supported function. However the property of compact support was not present in the resulting covariance function. Till now, the extension of compactly supported functions to the non-stationary case has not been published. Chilès and Delfiner (1999) mentioned that the non-stationary form of bounded covariance function could be obtained by considering local kernel functions. Based on this idea, in Chapter 6 a class of non-stationary covariance functions with compact support are developed by convolution of local kernel functions.

2.3 Multivariate modeling and simulation

In geological, mining, petroleum and hydrogeological studies, it is common to observe that several properties are spatially correlated to the main property of interest. To better understand and quantify the main property, these secondary information should be taken into account as they provide useful information. Considering a multivariate field with n stationary random functions $Z_i(x), i = 1, 2, \cdots, n$, the full covariance matrix function $\mathbf{C}(h)$ can

be expressed by

$$\begin{pmatrix} C_{1,1}(h) & C_{1,2}(h) & \cdots & C_{1,n}(h) \\ C_{2,1}(h) & C_{2,2}(h) & \cdots & C_{2,n}(h) \\ \vdots & \vdots & \ddots & \vdots \\ C_{n,1}(h) & C_{n,2}(h) & \cdots & C_{n,n}(h) \end{pmatrix} \tag{2.13}$$

where $C_{i,j}(h)$ is cross-covariance function of variable Z_i and Z_j when $i \neq j$. $C_{i,j}(h)$ is defined by

$$C_{i,j}(h) = E[Z_i(x) - m_i(x)][Z_j(x+h) - m_j(x+h)] \tag{2.14}$$

in which $m_i(x)$ is the mean of $Z_i(x)$.

2.3.1 The linear model of coregionalization

The linear model of coregionalization (LMC) describe the coregionalization as the sum of proportional covariance models (Chilès and Delfiner, 1999; Myers, 1983; Marcotte, 1991; Journel and Huijbregts, 1978; Wackernagel, 2003),

$$\mathbf{C}(h) = \sum_{k=1}^{s} \mathbf{B_k} C_k(h) \tag{2.15}$$

in which $C_k(h)$ is the basic structure indexed by $k, k = 1, 2, \cdots, s$. It requires that all direct and cross covariances are linear combinations of $C_k(h)$,

$$C_{i,j}(h) = \sum_{k=1}^{s} b_k(i,j) C_k(h) \tag{2.16}$$

A sufficient (but not necessary) condition of the LMC to be admissible is the coefficient matrix $\mathbf{B_k}$ is positive semi-definite for each k. The verification of this condition is easily done by computing the eigenvalues of each coefficient matrix (Journel and Huijbregts, 1978).

In fitting a model for experimental covariances, first the basic structures $C_k(h)$ are determined from the direct covariance. Then the coefficient matrix $\mathbf{B_k}$ is approximated under the constraint

$$\mid b_k(i,j) \mid \leq \sqrt{b_k(i,i) b_k(j,j)}. \tag{2.17}$$

for any i and j and each k due to its positive assumption (Chilès and Delfiner, 1999).

One advantage of the LMC is it is easy to ensure admissibility. Goulard (1989) and Goulard and Voltz (1992) provided an iterative algorithm to fit the coefficients by least squares under the positivity constraint of $\mathbf{B_k}$. Moreover, because the LMC can be simulated by linear

combination of independent univariate simulations, it is possible to use a large variety of simulation algorithms, including the efficient turning bands method (Matheron, 1973; Chilès and Delfiner, 1999; Emery, 2008b) and FFTMA (Le Ravalec-Dupin *et al.*, 2000). The main disadvantage of LMC is that it forces all the variables to have the same spatial structures. In fact, it is common that the secondary variable corresponds to a much larger support than the main variable, then its covariance usually has a smoother behavior close to the origin than the main variable. For example, the geophysical variables, density, electrical resistivity and magnetic susceptibility, are representative of a larger domain and consequently their variograms have more continuous behavior at the origin than the grade measured on cores. In cases that correlated variables have different spatial structures, the LMC cannot fully incorporate this information.

2.3.2 The non-linear model of coregionalization and its simulation

Development of N-LMC and simulation

In the non-linear model of coregionalization, for different variables, their direct and cross covariance C_{ij} in Equation 2.16 have different covariance structure $C_k(h), k = 1, 2, \cdots, s$. One way to test the validity of the model is to verify that frequency dependent spectral matrices are positive definite at every frequency (Chilès and Delfiner, 1999). Yao and Journel (1998) proposed to generate smooth covariance or cross-covariance tables by transforming the experimental covariance into spectral density tables using FFT. However this method cannot be used in case of sampling on irregular grid or when many data occur at distances smaller than the grid used for the FFT. Oliver (2003) constructed a positive definite cross-covariance by involving the square roots of auto-covariance functions of two variables. Marcotte (2015b) derived the spectral density of seven common covariance models (Exponential, spherical, cubic, penta-spherical, Gaussian, Cauchy and K-Bessel models) in the 3D anisotropic case and provided a program for verification of admissibility for non-LMC models with symmetrical cross-covariances.

For the simulation of a N-LMC, a few methods are available. Shinozuka (1971) developed a continuous spectral method to simulate a multivariate homogeneous process by a series of cosine functions. He proposed to use Choleski decomposition of the spectral density matrix, at selected frequencies. Then, the study was extended to non-homogeneous oscillatory process characterized by an evolutionary power spectrum (Shinozuka and Jan, 1972). Mejía and Rodríguez-Iturbe (1974) focused on discussing connection of correlation and spectrum of a random function and provided a simulation method by sampling from the spectral density functions. Mantoglou and Wilson (1982) extended the turning bands method into spectral

domain. Mantoglou (1987) used the spectral turning bands method to simulate multivariate, anisotropic, two- or three-dimensional stochastic processes. In the anisotropic case, the covariance has to be calculated on each line therefore costs much computer time. In addition, in the case of covariance with long range, the calculation in the spectral domain requires large CPU time. Emery *et al.* (2015) improved the spectral turning bands approach through coupling with importance sampling techniques in multivariate simulation of Gaussian random functions. It requires that all direct and cross covariance functions are stationary and the closed-form expressions of their spectral density functions must be known. This approach is fast and has a low memory storage requirement. Marcotte (2015a) proposed a spatial multivariate turning bands method that allows to simulate anisotropic non-LMC directly in the spatial domain. First the line joint covariances to simulate is determined from the expressions for the line spectral densities. Then it is simulated on each line in the spatial domain. Finally the simulation on desired points in $\Re^3$ is generated by combining the simulated values on lines.

Oliver (2003) proposed an approach of multivariate simulation by extending the matrix decomposition method. The cross-covariance function was created by square roots of auto-covariance of two variables. In cases that the square root of covariance function cannot be obtained directly, the square root of Fourier transform of the covariance model is calculated first, then inverse Fourier transform is conducted. Another attempt in simulation of coregionalization is an extension of FFTMA by Le Ravalec-Dupin and Da Veiga (2011) in cosimulation of two variables. This method was based on the Markov-Bayes approximation on two variables, so it did not allow full control and generality of the simulated cross-covariances.

Cosimulation of overburden thickness and latent variable of geological domain

In Chapter 7, a general fast Fourier transform and moving average method (GFFTMA) for multivariate simulation of N-LMC is proposed. In addition, synthetic examples are illustrated for simulations of N-LMC with two and three variables for different combinations of seven available models (Gaussian, exponential, spherical, cubic, penta-spherical, Cauchy and K-Bessel model). Moreover, the GFFTMA is used in a real application of overburden thickness simulation incorporating a secondary variable, geological domain, over an area of Montérégie Est, in Québec.

A number of approaches can be used in the simulation of the categorical variable, such as sequential indicator simulation (Journel and Isaaks, 1984; Deutsch and Journel, 1998; Emery, 2004a), truncated Gaussian simulation (Matheron *et al.*, 1987) and multiple point statistics (Guardiano and Srivastava, 1993; Ortiz and Deutsch, 2004). The truncated Gaussian

simulation is proposed by Matheron *et al.* (1987) to simulate ordered categories with locally varying proportions. The truncated Gaussian simulation was generalized to multiple Gaussian distribution (truncated pluri-Gaussian simulation) by Galli *et al.* (1994). Programs of pluri-Gaussian simulation were published by Xu *et al.* (2006), Emery (2007b), Emery and Silva (2009) and Chopin (2011). Relevant researches and application of truncated Gaussian or pluri-Gaussian simulation are referred to Le Loc'h and Galli (1997); Armstrong *et al.* (2011); Mariethoz *et al.* (2009); Deutsch and Deutsch (2014); Mannseth (2014); Rimstad and Omre (2014).

In simulation of Gaussian fields, one of the commonly used approaches is Gibbs sampler (Geman and Geman, 1984). It is an iterative method based on Markov chain. The general principle of Gibbs sampler is to begin with a simulated field that does not satisfy all the requirements on spatial variability, and then locally update it step by step until all the requirements are met. In practice, when the difference between simulated field and desired distribution are acceptable, the iteration is stopped. The difference is reflected by the transition kernel (Lantuéjoul, 2002), the covariance or the variogram and the histogram in geostatistical simulation. Emery (2008a) proposed to detect the validation of simulation algorithms by some statistical tests. Strategies to improve rate of convergence of Gibbs sampler are discussed by Roberts and Sahu (1997) and Galli and Gao (2001). Applications of Gibbs sampler in simulation of a Gaussian field or Gaussian-based field can be found in Emery (2007c) and Lantuéjoul and Desassis (2012). In Gibbs sampler, it is easy to incorporate inequality constraints in simulation. Freulon and de Fouquet (1993) first proposed to simulate a random function conditional on the hard and soft data with Gibbs sampler. Most program of truncated pluri-Gaussian simulation (Emery, 2007b; Emery and Silva, 2009; Chopin, 2011) are based on this algorithm. Emery *et al.* (2014) improved algorithm of Gibbs sampler by simulated annealing or restricting the transition matrix of the iteration so that it can be used in simulation of large Gaussian random field with inequality constraints.

CHAPTER 3 THESIS ORGANIZATION

This thesis is based on three articles presented in Chapter 5 to 7. In this chapter, the objectives, approach and main contributions of each article are described. Then the link between articles and their consistency related to the research objectives of the thesis are presented.

3.1 The first article

Title : A comparison of approaches to include outcrop information in overburden thickness estimation

Authors : Min Liang, Denis Marcotte and Nicolas Benoit

Published on : Stochastic Environmental Research and Risk Assessment, Volume 28, Issue 7, page 1733-1741

Summary : This article focuses on the problem of overburden thickness estimation in presence of outcrops. The methods based on stationary covariance model (discussed in Section 2.1.2) failed to sufficiently reflect the local spatial structure of overburden thickness near outcrops. In this article, a non-stationary model is developed to incorporate the outcrops information in overburden thickness estimations. Compared to kriging based on a stationary model (with or without involving the outcrop information), the non-stationary model provides more precise estimation, especially at points close to an outcrop. The thickness map obtained with the non-stationary covariance model is more realistic, since the algorithm forced the estimates smoothly reducing to zero close to outcrops without the bias incurred. Highlights of this article are :

- Outcrops information can help improve thickness estimation ;
- A model of non-stationary covariance is able to integrate the outcrop information ;
- Non-stationary proves to be more efficient than two forms of stationary kriging ;
- The non-stationary covariance is modeled by a single additional parameter.

3.2 The second article

One weakness of the model presented in the first paper is the use of the exponential model of covariance. Because the sill is reached only asymptotically, the covariances never vanish.

As a consequence, the covariance matrix used in the non-stationary kriging could become prohibitively large for its use in a global setting. This article seeks to palliate this limitation by developing a family of compactly supported covariance functions in the non-stationary case.

Title : A class of non-stationary covariance functions with compact support

Authors : Min Liang and Denis Marcotte

Published on : Stochastic Environmental Research and Risk Assessment, Published online, page 1-15

Summary : In geostatistics, the popular spherical family models (spherical, cubic and penta models) have compact support, but are limited to stationary cases. The non-stationary form of spherical family model was suggested by Chilès and Delfiner (1999), but never tested and published. Based on the kernel convolution theory, this article proposed a class of non-stationary covariance functions with compact support. The differentiability of functions can be controlled by order of Radon transform. Moreover, a method on simulation of this class of non-stationary models is described by weighted average of independent standard Gaussian variates in both the isotropic and the anisotropic cases. In the end, a real application of the proposed model is provided. Highlights of this article are :

- The closed-form expression of the non-stationary isotropic spherical model in $\Re^n$, and cubic, penta model in $\Re^3$ are derived ;
- Non-stationary anisotropic compactly supported functions can be evaluated numerically ;
- All non-stationary models defined are admissible ;
- A method for simulation based on moving averages is developed ;
- The non-stationary model in estimation of overburden thickness is more precise at the contact of two geological zones.

3.3 The third article

While the two previous articles dealt with the non-stationary univariate case, this paper deals rather with a development of a new efficient simulation method in the multivariate stationary case for the N-LMC model. Indeed, it is frequent that abundant secondary information is

available that could help significantly improve the estimation and simulation of the main variable. As an example for the overburden thickness estimation (or simulation) problem, the surficial geology is known everywhere. This information can be helpful to better estimate the overburden thickness as the complete overburden sedimentary sequence in the Saint-Lawrence is : bedrock - glacial till - clay - sand. Knowing for example that till is observed on surface should typically correspond to a small overburden thickness. One popular method to represent the geology is the truncated Gaussian method, especially appropriate here where the geological facies are naturally ordered. The variogram of the overburden thickness has a linear behavior at the origin. Using a (LMC) would force to represent the latent Gaussian variable with the same variogram, but this would produce unrealistic facies simulation and would not represent adequately the geology indicator variograms and cross-variograms. Hence, a N-LMC model is required.

Title : Simulation of non-linear coregionalization model by FFTMA

Authors : Min Liang, Denis Marcotte and Pejman Shamsipour

Submitted to : Computers & Geosciences in June, 2015

Summary : In this article, a fast and efficient method, GFFTMA, is described to simulate multivariate fields of N-LMC by generalizing the FFTMA method. It allows direct and cross covariance of multiple variables having different structures. Synthetic and real examples illustrated that the simulations of N-LMC by GFFTMA fit the target models well. Highlights of this article are :

- A FFT based method is proposed to simulate the N-LMC ;
- Realizations fit very well the target N-LCM ;
- Computation is very fast and the computing time shows a $Nlog(N)$ relation with simulated size N ;
- An approach based on Cholesky decomposition is suggested to get simulated values at sample points.

3.4 Consistency among the papers

The developments presented in the three papers were all initially motivated by the need to provide better estimation or simulation of overburden thickness in the Saint-Lawrence

lowlands. However, the methods presented are general and can apply to the study of any natural phenomena. The first article illustrates the gain that can be achieved with a non-stationary model compared to a stationary one. The second paper follows from the first one as it fills an important gap in available non-stationary models by providing a simple method to construct and simulate non-stationary models with compact support. Finally, the third paper presents a very efficient simulation method for the multivariate stationary case. In addition to the common application on overburden thickness, the third paper is linked to the first two papers by the observation that the secondary variable can often account for non-stationarity of the mean in the main variable. Admittedly, multivariate non-stationary models in the covariance however are still to be developed and tested.

The first article enables to meet the Objective 1 presented in Section 1.3. The second article enables to fulfill the Objectives 2 and 3, whereas the third article satisfies the Objective 4.

CHAPTER 4 THEORETICAL BACKGROUND

The purpose of this chapter is to provide the theoretical framework used in or related to this thesis. This chapter is organized into three main sections. The first section introduces methods of estimation and simulation in geostatistics. The second section presents the non-stationary covariance functions developed recently. The last section outlines how to construct a function with compact support and some popular compactly support functions.

4.1 Estimation and simulation techniques

4.1.1 Kriging

In the study of a random function, an important task is to estimate its values on unsampled points based on limited observations. Kriging is an estimation approach building on statistical model of the random function (Matheron, 1963). It is trying to find the linear unbiased estimator minimizing the estimation error. A covariance or variogram model is the most important parameter for kriging, which is determined by fitting with the observed data. The input covariance model must be valid, that is positive definite, but not limited to be stationary (Chilès and Delfiner, 1999). For the variable $Z(x)$, kriging estimates the value at unknown point x_0 by the weighted average of known values $Z(x_i)$ at points x_i, $(i = 1, \cdots, N)$. According to properties of the mean, the kriging has three main forms (Chilès and Delfiner, 1999), 1) simple kriging, used in case with known mean ; 2) ordinary kriging, used in case with constant and unknown mean ; and 3) universal kriging, used in case with varying and unknown mean. In the whole thesis, the mean of studied variable is assumed constant therefore only simple kriging and ordinary kriging are reviewed here.

Simple kriging In simple kriging, the mean of a random function $Z(x)$ is known $E[Z(x)] = m(x)$. Therefore one can directly estimate $Z^*(x_0) - m(x_0)$ by the weighted average of $Z(x_i) - m(x_i)$ without any additional constraint on weights λ_i :

$$Z^*(x_0) = \sum_{i=1}^{N} \lambda_i (Z(x_i) - m(x_i)) + m(x_0). \tag{4.1}$$

The estimation variance of kriging is

$$\mathrm{Var}[Z^*(x_0) - Z(x_0)] = \mathrm{Var}[Z(x_0)] + \sum_{i=1}^{N}\sum_{i=1}^{N}\lambda_i\lambda_i\mathrm{Cov}[Z(x_i), Z(x_i)] - 2\sum_{i=1}^{N}\lambda_i\mathrm{Cov}[Z(x_0), Z(x_i)] \tag{4.2}$$

By canceling the partial derivatives with respect to λ_i,the minimum of estimation variance can be obtained. Then the weights λ_i will be written as

$$\mathbf{K}\boldsymbol{\lambda} = \mathbf{k}. \tag{4.3}$$

where $\mathbf{K} = \mathrm{Cov}[Z(x_i), Z(x_i)]$ is a $N \times N$ covariance matrix of observations, $\boldsymbol{\lambda} = [\lambda_i]$ is N vector of weights and $\mathbf{k} = \mathrm{Cov}[Z(x_0), Z(x_i)]$ is a N vector of covariance between observations and points of interest.

Then the estimation variance of simple kriging is

$$\mathrm{Var}(Z^*(x_0) - Z_0) = \mathrm{Var}(Z(x_0)) - \boldsymbol{\lambda}'\mathbf{k} \tag{4.4}$$

in which $\boldsymbol{\lambda}'$ is the transpose of vector $\boldsymbol{\lambda}$.

Ordinary kriging In ordinary kriging, the mean of the variable $Z(x)$, $m(x) = c_0$ in which c_0 is a unknown constant. The estimate $Z^*(x_0)$ on an unknown point x_0 is the weighted average of known values $Z(x_i)$,

$$Z^*(x_0) = \sum_{i=1}^{N}\lambda_i Z(x_i) \tag{4.5}$$

To keep the estimation unbiased, a constraint on the weights must be imposed as

$$\sum_{i=1}^{N}\lambda_i = 1. \tag{4.6}$$

Subject to this constraint on λ_i, the method of Lagrange multipliers is used to minimize the estimation variance (Equation 4.2). The Lagrange function is

$$L(\lambda_i, \mu_l) = \mathrm{Var}(Z^*(x_0) - Z_0) + 2\mu_l(\sum_{i=1}^{N}\lambda_i - 1) \tag{4.7}$$

where μ_l is Lagrange multiplier. By making the partial derivatives of L to zero, the ordinary

kriging system is expressed by

$$\begin{bmatrix} \mathbf{K} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \mu_l \end{bmatrix} = \begin{bmatrix} \mathbf{k} \\ \mathbf{1} \end{bmatrix} \tag{4.8}$$

in which $\mathbf{1}$ is the vector of ones.

The estimation variance of ordinary kriging is expressed as

$$\mathrm{Var}(Z^*(x_0) - Z_0) = \mathrm{Var}(Z(x_0)) - \boldsymbol{\lambda}'\mathbf{k} - \mu_l. \tag{4.9}$$

4.1.2 CoKriging

Cokriging is the kriging for multivariate cases. Here cokriging is illustrated with only two variables for simplicity. Denote the main variable by $Z(x)$ with nz observations and second variable by $Y(x)$ with ny observations. When the means of $Z(x)$ and $Y(x)$ are both known, the simple cokriging estimator can be written as :

$$Z^*(x_0) = m_z + \sum_{i=1}^{nz} \lambda_i (Z_i - m_Z) + \sum_{j=1}^{ny} \alpha_j (Y_j - m_Y) \tag{4.10}$$

where m_Z and m_Y are the means of $Z(x)$ and $Y(x)$, λ is weights of $Z(x)$ and α is weights of $Y(x)$. The weights λ_i and α_j are obtained following

$$\begin{bmatrix} \mathbf{K}_{zz} & \mathbf{K}_{zy} \\ \mathbf{K}_{yz} & \mathbf{K}_{yy} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{k}_{zz} \\ \mathbf{k}_{yz} \end{bmatrix} \tag{4.11}$$

where $\mathbf{K}_{zz}$, $\mathbf{K}_{yy}$ and $\mathbf{K}_{zy}(\mathbf{K}_{yz})$ are covariance matrices between observations of the main variable $Z(x)$, secondary variables $Y(x)$ and cross covariance between $Z(x)$ and $Y(x)$. $\mathbf{k}_{zz}$ (or $\mathbf{k}_{yz}$) are covariance vectors between interested point and observed points of the main variable $Z(x)$ (or the secondary variable $Y(x)$). The variance of estimation in simple cokriging is

$$\mathrm{Var}(Z^*(x_0) - Z_0) = \mathrm{Var}(Z_0) - \boldsymbol{\lambda}\mathbf{k}_{zz} - \boldsymbol{\alpha}\mathbf{k}_{yz} \tag{4.12}$$

When both means are unknown, the cokriging estimator has the following form :

$$Z^*(x_0) = \sum_{i=1}^{nz} \lambda_i Z_i + \sum_{j=1}^{ny} \alpha_j Y_j \tag{4.13}$$

For the estimator to be unbiased, it is sufficient to impose that the weight $\sum_{i=1}^{nz} \lambda_i = 1$ and

$\sum_{j=1}^{ny} \alpha_j = 0$ respectively. The weight can be obtained from

$$\begin{bmatrix} \mathbf{K}_{zz} & \mathbf{K}_{zy} & \mathbf{1} & \mathbf{0} \\ \mathbf{K}_{yz} & \mathbf{K}_{yy} & \mathbf{0} & \mathbf{1} \\ \mathbf{1}' & \mathbf{0}' & 0 & 0 \\ \mathbf{0}' & \mathbf{1}' & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\alpha} \\ \mu_z \\ \mu_y \end{bmatrix} = \begin{bmatrix} \mathbf{k}_{zz} \\ \mathbf{k}_{yz} \\ 1 \\ 0 \end{bmatrix} \tag{4.14}$$

The estimation variance of ordinary cokriging is

$$\mathrm{Var}(Z^*(x_0) - Z_0) = \mathrm{Var}(Z_0) - \boldsymbol{\lambda}\mathbf{k}_{zz} - \boldsymbol{\alpha}\mathbf{k}_{yz} - \mu_z \tag{4.15}$$

4.1.3 Simulations

Stochastic simulation is a representation of a possible spatial distribution of a variable which meets the particular covariance model. By conditional simulation, a series of realizations can be generated, conditional upon observations. Methods of conditional simulation include sequential Gaussian simulation and matrix decomposition. For unconditional simulation, in addition to these methods, we can use among many other approaches, the turning bands, moving average algorithm, dilution functions and continuous or discrete spectral method (Chilès and Delfiner, 1999; Lantuéjoul, 2002). Post conditioning method can calibrate an unconditional simulation to observed data without losing features of the former structure (Chilès and Delfiner, 1999). The post conditioning process is derived by the sum of kriging estimator and simulation of the kriging error. In kriging, the estimator $Z^*(x)$ and the true value $Z(x)$ have relation as

$$Z(x) = Z^*(x) + [Z(x) - Z^*(x)] \tag{4.16}$$

where the $Z(x) - Z^*(x)$ is kriging error. Considering a simulation $Z_S(x)$ in kriging,

$$Z_S(x) = Z_S^*(x) + [Z_S(x) - Z_S^*(x)] \tag{4.17}$$

in which $Z_S^*(x)$ is kriging estimator of $Z_S(x)$. Replacing the unknown kriging error in Equation 4.16 by its simulation $Z_S(x) - Z_S^*(x)$, the conditional simulation is given by

$$Z_{CS}(x) = Z^\star(x) + [Z_S(x) - Z_S^\star(x)] \tag{4.18}$$

In multivariate case, the $Z^\star(x)$ is obtained by cokriging. After conditioning, the simulated field $Z_{CS}(x)$ preserves the required covariance of simulation. Details on conditional simulation

are referred to Chilès and Delfiner (1999).

Simulation of continuous variable

The dilution method Chilès and Delfiner (1999) describes a dilution method by a fixed function to simulate a stationary isotropic field. The dilution function $\omega(x)$ must be square integrable. Corresponding to a random function $Z(x)$, the covariogram $g(h)$ between two points with distance h is expressed by

$$g(h) = (\omega * \breve{\omega})(h) = \int \omega(x)\omega(x+h)dx \tag{4.19}$$

where $\breve{\omega}(h) = \omega(-h)$. At a point x, the random function $Z(x)$ can be expressed by

$$Z(x) = \mu \sum_i \omega(x - \mathcal{X}_i) \tag{4.20}$$

where μ is the mean and $\mathcal{X}_i$ is the point of a Poisson point process. From a practical point, the dilution method works well in cases where the covariance is proportional to the geometric covariogram, such as the field of spherical family model. For other models, Chilès and Delfiner (2012) listed the dilution function for stationary exponential and Gaussian covariance model in $\Re^n$. However, for the non-stationary case, the closed-form expressions for the dilution function have not been provided.

For spherical family models, the indicator function of spheres in $\Re^n$ is used as dilution function to generate the covariogram. It can be extended to the non-stationary case by using indicator function of spheres with different sizes depending on locations. Thus, it is easy to simulate a non-stationary field of spherical family model by the dilution method. In Chapter 6, I adopt a simpler approach to simulate the non-stationary fields of spherical family covariance functions in both isotropic and anisotropic cases. Instead of using dilution functions, equivalent weighting functions are defined. The weights are normalized so as to provide a unit variance for the simulated field at all locations. For the anisotropic case (the kernel is a hyper-ellipsoid in $\Re^n$), the indicator function is transformed to be a hypersphere by rotating and scaling the local grid at each point. Then the non-stationary anisotropy in the original grids can be simulated. Besides, the sequential simulation and matrix decomposition also can be used in simulation of a non-stationary variable. However, they are not applicable in simulations with large size.

Continuous spectral method According to Bochner (1937) theorem, the covariance of a stationary random function $Z(x)$ can be expressed as

$$C(h) = \int_{-\infty}^{+\infty} S(f)\exp(2i\pi f \cdot h)\,\mathrm{d}f \tag{4.21}$$

in which $S(f)$ is the power spectral density function. The covariance function is the inverse Fourier transform of power spectral density function, while the spectral density function can be obtained by Fourier transform on the covariance function,

$$S(f) = \int_{-\infty}^{+\infty} C(h)\exp(-2i\pi f \cdot h)\,\mathrm{d}h. \tag{4.22}$$

It indicates that the power spectral density function $S(f)$ has the same spatial information as the covariance function $C(h)$. Shinozuka (1971) proposed to simulate the stationary random function $Z(x)$ which has the spectral representation by

$$Z_S(x) = \sigma\sqrt{2}\cos(2\pi\langle U, x\rangle + \Phi) \tag{4.23}$$

where

$$\sigma = \sqrt{\int_{-\infty}^{+\infty} S(f)\mathrm{d}f} \tag{4.24}$$

is the standard deviation of $Z(x)$; U is a random vector of $\Re^n$ with probability distribution $S(f)/\sigma^2$; $\langle U, x\rangle = U_1x_1 + U_2x_2 + \cdots + U_Nx_N$; and Φ is a random variable with uniform distribution over $[0, 2\pi]$ independent of U. The realization $Z_S(x)$ has the covariance in form of

$$C(h) = \int \cos(2\pi\langle f, h\rangle)S(f)\mathrm{d}f. \tag{4.25}$$

The realizations simulated by Equation 4.23 is not ergodic (Chilès and Delfiner, 1999). To get ergodic realizations of $Z(x)$ with desired covariance model, one must consider the average of a large number of simulations $Z_S(x)$ being the final realization of $Z(x)$. In practice, the ergodicity of the simulation is slow to reach, especially in the case with widely spread spectrum. Shinozuka and Jan (1972) developed the algorithm of spectral method by discretizing the spectral distribution function. Lantuéjoul (2002), Emery and Lantuéjoul (2006) and Emery *et al.* (2015) described efficient implementations of the continuous approach.

FFTMA The moving average method was proposed by Journel (1974) for simulation of a one-dimensional Gaussian random field. Then Oliver (1995) extended it to simulate Gaussian

fields in two and three dimensions. By moving average, a Gaussian random function $Z(x)$ with mean m and covariance C can be expressed by

$$Z(x) = m + g * \delta \tag{4.26}$$

where g is a convolution root of the covariance C, δ is a Gaussian white noise.

Based on Bochner (1937) theorem (Equation 4.21 and 4.22), Le Ravalec-Dupin *et al.* (2000) proposed to simulate the randomness component $g*\delta$ of Equation 4.26 in the spectral domain by the fast Fourier transform. It is a combination of fast Fourier transform and moving average method therefore it is called FFTMA. Simulations are done in following steps :

1. Calculate the covariance $C(x)$ on all grids to simulate.
2. Calculate the spectral density function $S(f)$ by fast Fourier transform following Equation 4.22.
3. Based on the property of convolution on Fourier transform $\mathcal{F}(g * h) = \mathcal{F}(g) \cdot \mathcal{F}(h)$, obtain the spectral densities $G(f)$ of g by $G(f) = \sqrt{S(f)}$.
4. Simulate the randomness component by $G(f)\mathcal{F}(\delta)$.
5. Apply inverse Fourier transform on $G(f)\mathcal{F}(\delta)$ to get $g\delta$.
6. Add the mean m to get the final simulation as Equation 4.26.

Note that the size of simulated grids must be larger than or equal to the sum of size of grids of interest and correlation range. After simulation, the points beyond grids of interest can be neglected. Applications of the FFTMA can be found in Le Ravalec-Dupin *et al.* (2001); Gloaguen *et al.* (2005b,a); Shamsipour *et al.* (2011).

Turning bands The turning bands method was developed by Matheron (1973). It transforms an n-dimensional simulation to a series of uni-dimensional simulations by projection. In the field to simulate, each point can be projected to lines generated from an arbitrary origin. Denote u_i is the unit vector along each line. A random function $Z(x)$ in $\Re^n$ can be simulated by

$$Z_S(x) = \frac{1}{\sqrt{n_D}} \sum_{i=1}^{n_D} S_i(x \cdot u_i) \tag{4.27}$$

in which S_i is the simulation on each line following a uni-dimensional covariance function that corresponds to the desired model, n_D is the number of lines, and $x \cdot u_i$ represents the inner product of the vectors x and u_i.

In the field given by Equation 4.27, the covariance of two points x_1 and x_2 is expressed by

$$C_S(x_1, x_2) = E[Z_S(x_1)Z_S(x_2)] = \frac{1}{n_D}\sum_{i=1}^{n_D}\sum_{j=1}^{n_D} E[Z_i(x_1 \cdot u_i)Z_j(x_2 \cdot u_j)]. \tag{4.28}$$

Because the simulations along lines are independent, the expectation $E[Z_i(x_1 \cdot u_i)Z_j(x_2 \cdot u_j)]$ is zero unless $i = j$. Thus the covariance of simulation can be written as

$$C_S(x_1, x_2) = \frac{1}{n_D}\sum_{i=1}^{n_D} E[Z_i(x_1 \cdot u_i)Z_i(x_2 \cdot u_i)]. \tag{4.29}$$

Assuming the uni-dimensional realization is second-order stationary with covariance C_1

$$C_S(x_1, x_2) = C_S(h) = \frac{1}{n_D}\sum_{i=1}^{n_D} C_1(h \cdot u_i). \tag{4.30}$$

According to the law of large numbers, this expression can be written as

$$C_S(r) = \lim_{n_D \to \infty}\left\{\frac{1}{n_D}\sum_{i=1}^{n_D} C_1(h \cdot u_i)\right\} = \int_c C_1(h \cdot u) f(u)\mathrm{d}u \tag{4.31}$$

where $f(u)$ is the probability function of u; and c represents the unit circle in 2D and unit sphere in 3D. Respectively, the covariances of simulation are expressed by

$$C_S(r) = \frac{2}{\pi}\int_0^r \frac{C_1(\zeta)}{(r^2 - \zeta^2)^{1/2}}\mathrm{d}\zeta \tag{4.32}$$

in the two-dimensional field and

$$C_S(r) = \frac{1}{r}\int_0^r C_1(\zeta)\mathrm{d}\zeta \tag{4.33}$$

in the three-dimensional field. The uni-dimensional covariance functions for the usually used model can be found in listed in Chilès and Delfiner (1999); Gneiting (1998); Marcotte (2015a).

In the two-dimensional simulation, it is complicated to obtain the uni-dimensional covariance from the Equation 4.32. Mantoglou and Wilson (1982) proposed to use the turning bands method in the spectral domain and listed the corresponding spectral density functions in one dimension for common models. Improvements of the spectral turning bands method are introduced by (Mantoglou, 1987; Emery *et al.*, 2015).

Gibbs sampler The Gibbs sampler (Geman and Geman, 1984) is an iterative method based on Markov chain. It can incorporate inequality constraints in simulation. The general principle is to begin with a simulated field that does not satisfy all the requirements on spatial variability, and then locally update it step by step until the spatial structure is reproduced. Consider a random function $Z(x)$, and the point of simulation contains N points : $x_1, \ldots, x_N$. The iteration of Gibbs sampler is performed as follows :

1. Assign random Gaussian values at simulation points as initial values that respect the known constraints.
2. Select a point on simulated grid, x_i, either randomly or systematically at all nodes.
3. Choose a value for the node from the conditional distribution of $Z(x_i)$ given other values $Z(x_k)$, $k = 1, \ldots, i-1, i+1, \ldots N$. Keep this value if it respects the constraints at this point, otherwise reject it.
4. Repeat steps 2 and 3 until all points are visited.
5. Repeat steps 2-4 finite times until the simulation $Z_S(x)$ converges, as explained below.

In simulation of a Gaussian field, the simulated value in step 2 can be simply obtained by randomly drawing a value from normal distribution of $(Z_{KS}(x_i), \sigma^2_{KS})$, where the $Z_{KS}(x_i)$ is the kriging estimator at x_i, σ^2_{KS} is the variance. With increment of iteration times, the spatial property of simulation, covariance and histogram, will be close to the expected model. When the iteration times K is large enough, although the variograms (and the histogram) calculated from the resulting simulations fluctuate around the target model, their variations are stabilized. Then the iteration is stopped and the simulation is convergent.

Simulation of categorical variables

A few approaches can be used in the simulation of a categorical variable, such as sequential indicator simulation(Journel and Isaaks, 1984; Deutsch and Journel, 1998), truncated Gaussian simulation (Matheron *et al.*, 1987; Armstrong *et al.*, 2011) and multiple point statistics (Guardiano and Srivastava, 1993; Strebelle, 2002; Ortiz and Deutsch, 2004). Truncated Gaussian simulation is proposed by Matheron *et al.* (1987) to simulate ordered categories with locally varying proportions. A categorical random field with N facies $F_1, F_2, \cdots, F_N$ can be obtained by simulation of a truncated Gaussian field $Y(x)$.

$$F_i = \{x \in \Re^n, a_i < Y(x) \leqslant b_i\} \tag{4.34}$$

The value of an associated standard Gaussian random field in facies F_i is located between thresholds a_i and b_i. If we denote the indicator function of the facies by

$$\mathrm{I}_i(x) = \begin{cases} 1, & x \in F_i \\ 0, & x \notin F_i \end{cases} \tag{4.35}$$

the proportion of each facies will be

$$p_i = \mathrm{E}(I_i(x)) = \mathrm{Prob}\{a_i < Y(x) \leqslant b_i\} \tag{4.36}$$

By inverse function of the cumulative Gaussian distribution, the threshold a_i and b_i can be obtained. Then we can get the non-centered cross covariance of facies F_i and F_j (or non-centered direct covariance if $i = j$) with

$$C_{ij}(h) = \mathrm{Prob}\{a_i < Y(x) \leqslant b_i, a_j < Y(x+h) \leqslant b_j\} \tag{4.37}$$

The indicator cross-variogram of F_i and F_j is derived by

$$\gamma_{ij}(h) = C_{ij}(0) - \frac{C_{ij}(h) + C_{ji}(h)}{2} \tag{4.38}$$

Having the experimental variogram and cross-variogram for indicators, a covariance model for $Y(x)$ can be fitted. In Chapter 7, we fit variogram models for the indicator of two geological domains by minimizing the mean squared error weighted by distance and data pairs. Based on the variogram model, a latent Gaussian field can be simulated by Gibbs sampler under inequality conditions (Freulon and de Fouquet, 1993).

4.2 The Non-stationary covariance functions based on kernel convolution

Based on kernel convolution theory, several researchers developed non-stationary covariance functions in different forms. Only the functions used in this thesis are presented here.

Higdon *et al.* (1999) firstly introduced a non-stationary covariance function by convolving kernel functions

$$C^{NS}(x_i, x_j) = \int_{\Re^n} k_{xi}(u) k_{xj}(u) \mathrm{d}u \tag{4.39}$$

where x_i, x_j are locations in $\Re^n$, and $k_x(\cdot)$ is a kernel function centered at x which has spatially varying covariance matrix $\Sigma(x)$. In order to get a smoothly varying function $k_x(\cdot)$, Higdon *et al.* (1999) parameterized $\Sigma(x)$ in which parameters evolve with location. This non-stationary function is proved to be positive definite in $\Re^n$ by Paciorek and Schervish (2006)

as

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j C^{NS}(x_i, x_j) &= \sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j \int_{\Re^n} k_{xi}(u) k_{xj}(u) \mathrm{d}u \\ &= \int_{\Re^n} \sum_{i=1}^{m} c_i k_{xi}(u) \sum_{j=1}^{m} c_j k_{xj}(u)\, \mathrm{d}u \\ &= \int_{\Re^n} \left(\sum_{i=1}^{m} c_i k_{xi}(u) \right)^2 \mathrm{d}u \end{aligned} \tag{4.40}$$

which is always non-negative.

Paciorek (2003) and Paciorek and Schervish (2006) generalized the kernel convolution form of non-stationary correlation functions. If the isotropic correlation function R^S is positive definite on $\Re^n, n = 1, 2, \cdots$, a non-stationary form can be defined by

$$R^{NS}(x_i, x_j) = \mid \Sigma_i \mid^{\frac{1}{4}} \mid \Sigma_j \mid^{\frac{1}{4}} \mid \frac{\Sigma_i + \Sigma_j}{2} \mid^{-\frac{1}{2}} R^S(\sqrt{Q_{ij}}) \tag{4.41}$$

where $Q_{ij} = (x_i - x_j)^\top (\frac{\Sigma_i + \Sigma_j}{2})^{-1} (x_i - x_j)$ (Σ_i is kernel matrix at x_i), and is also positive definite on $\Re^n, n = 1, 2, \cdots$. By multiplying by a variance parameter σ^2 and replacing R^S by its corresponding form, the non-stationary version of the Matérn covariance function is obtained as

$$C^{NS}(x_i, x_j) = \sigma^2 \frac{1}{\Gamma(\upsilon) 2^{\upsilon - 1}} \mid \Sigma_i \mid^{\frac{1}{4}} \mid \Sigma_j \mid^{\frac{1}{4}} \mid \frac{\Sigma_i + \Sigma_j}{2} \mid^{-\frac{1}{2}} (2\sqrt{\upsilon Q_{ij}})^{\upsilon} K_{\upsilon}(2\sqrt{\upsilon Q_{ij}}) \tag{4.42}$$

where Γ is the gamma function, K_υ is the modified Bessel function of the second kind, and υ is a parameter of the covariance. It includes a non-stationary version of the exponential model when $\upsilon = 0.5$.

4.3 Compactly supported functions

Typical covariance functions with compact support include the spherical family functions (Matheron, 1965), Wu's function (Wu, 1995) and Wendland's function (Wendland, 1995). The construction of these functions share one common path : they are derived by performing certain operators, integral, differential, or/and convolution, on one positive definite radial functions with compact support.

4.3.1 Spherical family model and Euclid's hat

The stationary covariance functions of spherical family model are constructed by auto-convolution of indicator function $w_n(\xi)$ of a sphere in $\Re^n$ (Matheron, 1965),

$$w_n(\xi) = \begin{cases} 1, & \xi \leq \frac{a}{2} \\ 0, & \xi > \frac{a}{2} \end{cases} \tag{4.43}$$

where $\xi =\mid x \mid$ and a is diameter of hypersphere. A general form of covariogram function of $r =\mid h \mid$ in $\Re^n$ is (Chilès and Delfiner, 1999)

$$g_n(r) = \begin{cases} a^n v_{n-1} \int_{r/a}^1 (1-u^2)^{(n-1)/2}\,\mathrm{d}u, & \text{if} \quad r \leq a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{4.44}$$

where v_n stands for the volume of the unit-diameter ball in $\Re^n$ that can be expressed by

$$v_n = \frac{\pi^{n/2}}{2^{n-1} n \Gamma(n/2)} \tag{4.45}$$

where $\Gamma(\cdot)$ is the gamma function. Equation 4.44 is also called "Euclid's hat" (Schaback, 1995) because in Euclid space the function looks like a hat in two dimensions. Particularly when $n = 1, 2, 3$, the covariance function of triangular model valid in $\Re$ is expressed as

$$C_1(r) = \begin{cases} 1 - \frac{r}{a}, & \text{if} \quad r \leq a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{4.46}$$

circular model valid in $\Re^2$

$$C_2(r) = \begin{cases} \frac{\pi}{2}\left[\arccos\left(\frac{r}{a}\right) - \frac{r}{a}\sqrt{1-\frac{r^2}{a^2}}\right], & \text{if} \quad r \leq a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{4.47}$$

and spherical model valid in $\Re^3$

$$C_3(r) = \begin{cases} 1 - \frac{3r}{2a} + \frac{r^3}{2a^3}, & \text{if} \quad r \leq a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{4.48}$$

Radon transform is an integral transform which is also called Montée (Matheron, 1965) in geostatistics. If $f(x_1, x_2, \ldots, x_n)$ is a continuous function with a compact support in $\Re^n$, the Radon transformed function Rf is also a compactly supported function defined in $\Re^{n-1}$ after

integrating out x_n.

$$Rf = \int_R f(x_1, x_2, \ldots, x_n)\,\mathrm{d}x_n \tag{4.49}$$

By repeating the integration on $x_n, x_{n-1}, \ldots, x_{n-m+1}$, one obtains the Radon transform of order m that is defined on space $\Re^{n-m}$.

Let g_n be the covariogram of the indicator function w_n in $\Re^n$. In isotropic case, Radon transformed covariograms of orders 1 and 2 $g_{n,1}(r)$ and $g_{n,2}(r)$ of $g_n(r)$ are calculated in Chilès and Delfiner (1999, p. 73)

$$g_{n,1}(r) = 2\int_0^\infty g_n(\sqrt{r^2+\rho^2})\,\mathrm{d}\rho, \tag{4.50}$$

$$g_{n,2}(r) = 2\pi\int_r^\infty u g_n(u)\,\mathrm{d}u. \tag{4.51}$$

The covariogram $g_{n,m}(r)$ of the m order Radon transform is obtained by repeating Equation 4.50 m times or repeating Equation 4.51 $m/2$ times when m is even. The Radon transform of an isotropic covariogram is isotropic. Moreover, the covariogram function $g_{n,2q}$ is $2q$-times continuously differentiable and the corresponding field is q-times mean-square differentiable (Chilès and Delfiner, 1999).

Matheron (1965) presented two special cases, the covariance function of cubic model and penta-spherical model valid in $\Re^3$ when $q = 1, n = 5$ and $q = 2, n = 7$ respectively.

$$C_{5,2}(r) = \begin{cases} 1 - 7\frac{r^2}{a^2} + \frac{35}{4}\frac{r^3}{a^3} - \frac{7}{2}\frac{r^5}{a^5} + \frac{3}{4}\frac{r^7}{a^7}, & \text{if} \quad r \le a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{4.52}$$

$$C_{7,4}(r) = \begin{cases} 1 - \frac{22}{3}\frac{r^2}{a^2} + 33\frac{r^4}{a^4} - \frac{77}{2}\frac{r^5}{a^5} + \frac{33}{2}\frac{r^7}{a^7} - \frac{11}{2}\frac{r^9}{a^9} + \frac{5}{6}\frac{r^{11}}{a^{11}}, & \text{if} \quad r \le a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{4.53}$$

4.3.2 Wu's function

Starting with a cut-off polynomial function $f_l(r)$

$$f_l(r) = (1 - r^2)_+^l, \quad l \ge 0, \tag{4.54}$$

Wu (1995) constructed a class of functions with compact support by differential operator $\mathcal{D}(f)(r)$ (Equation 2.12) on univariate convolution of the basic function $f_l(r)$ continuous in $\Re^{l-1}$.

$$\varphi_{l,k}^{Wu}(r) = c_{l,k} D^k (f_l * f_l)(r), \quad 0 \le k \le l, \tag{4.55}$$

in which $c_{l,k}$ is a constant ensuring $\varphi_{l,k}^{Wu}(0) = 1$. Wu's function can also be written by a polynomial q_{2l-k} of degree $2l-k$ like

$$\varphi_{l,k}^{Wu}(r) = (1-r)_+^{2l-k+1} q_{2l-k}(r). \tag{4.56}$$

Wu's function $\varphi_{l,k}^{Wu}(r)$ is positive definite in $\Re^{2k+1}$ and possesses $2l-2k$ derivatives.

Moreover, Wu's construction can be connected to Euclid's hat (denoted by X^{2n+1}) defined in odd space dimension $\Re^{2n+1}$ by integral operator $\mathcal{I}(f)(r)$.

$$\varphi_{l,k}^{Wu} \doteq I^{l-k} X^{(2l+1)}. \tag{4.57}$$

in which $\doteq$ means equality up to a constant. Proof of this connection can be found in Wendland (1995). The integral operator $\mathcal{I}(f)(r)$ (Equation 2.11) is a special case of Radon transform with the second order (Equation 4.51). Therefore, as noted by Gneiting (2002), Wu's function corresponds to the spherical family model of Matheron (1965).

4.3.3 Wendland's function

Wendland (1995) developed a group of compactly supported radial basis functions by starting with the truncated power function which is also called Askey's power function (Askey, 1973),

$$\varphi_l(r) = (1-r)_+^l \tag{4.58}$$

defined in $\Re^d$ if $l \geq \lfloor \frac{d}{2} \rfloor + 1$. Here $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. Then applying the integral operator $\mathcal{I}(f)(r)$ (Equation 2.11) repetitively k times on this basic function defines a class of compactly supported functions

$$\varphi_{l,k}^{We}(r) = c_{l,k} \mathcal{I}^k \varphi_l(r), \quad r \geq 0, \quad k = 0, 1, \cdots, \tag{4.59}$$

where $c_{l,k}$ is a constant keeping $\varphi_{l,k}^{We}(0) = 1$ and

$$\mathcal{I}^k \varphi = \mathcal{I}(\mathcal{I}^{k-1}\varphi), \quad k \geq 1. \tag{4.60}$$

It can also be written by a real polynomial $p_{l,k}$ of degree k like

$$\varphi_{l,k}^{We}(r) = (1-r)_+^{l+k} p_{l,k}(r). \tag{4.61}$$

The Wendland's function $\varphi_{l,k}^{We}(r)$ is strictly positive definite on $\Re^d$ when $l = \lfloor \frac{d}{2} \rfloor + k + 1$.

It possesses $2k$ derivatives around zero and $k + l - 1$ derivatives around 1. More details on Wendland's function can be found in Wendland (1995, 1998). Although this family possesses some interesting properties, like a smooth spectral density without zeros (Marcotte, 2015b), it will not be used in the thesis.

CHAPTER 5 ARTICLE 1 : A COMPARISON OF APPROACHES TO INCLUDE OUTCROP INFORMATION IN OVERBURDEN THICKNESS ESTIMATION

Article history : Submitted 14 July 2013, Accepted 4 December 2013, Published online 21 December 2013, Stochastic Environmental Research and Risk Assessment.

Authors : Min Liang, Denis Marcotte and Nicolas Benoit

DOI : 10.1007/s00477-013-0835-6

5.1 Abstract

The estimation of overburden sediment thickness is important in hydrogeology, geotechnics and geophysics. Usually, thickness is known precisely at a few sparse borehole data. To improve precision of estimation, one useful complementary information is the known position of outcrops. One intuitive approach is to code the outcrops as zero thickness data. A problem with this approach is that the outcrops are preferentially observed compared to other thickness information. This introduces a strong bias in the thickness estimation that kriging is not able to remove. We consider a new approach to incorporate point or surface outcrop information based on the use of a non-stationary covariance model in kriging. The non-stationary model is defined so as to restrict the distance of influence of the outcrops. Within this distance of influence, covariance parameters are assumed simple regular functions of the distance to the nearest outcrop. Outside the distance of influence of the outcrops, the thickness covariance is assumed stationary. The distance of influence is obtained through a cross-validation. Compared to kriging based on a stationary model with or without zero thickness at outcrop locations, the non-stationary model provides more precise estimation, especially at points close to an outcrop. Moreover, the thickness map obtained with the non-stationary covariance model is more realistic since it forces the estimates to zero close to outcrops without the bias incurred when outcrops are simply treated as zero thickness in a stationary model.

5.2 Introduction

Many geotechnical, hydrogeological and geophysical applications require a precise estimation of overburden thickness. As an example, in geophysics, overburden thickness must be estimated to guide the interpretation for potential field or electrical methods. In hydrogeology, thickness of the aquifer is an essential part in the development of conceptual models for aquifer recharge estimation and flow and transport models. The thickness measurements are typically obtained from existing databases, report maps and field works. The first step is to carefully validate each data. Then, because the data is sparse, it is important to use efficient interpolation methods and all available information to minimize estimation errors. Kriging is designed to provide the estimation with the lowest estimation error variance (Chilès and Delfiner, 1999). It is based on variogram (or covariance) modeling, under a stationarity hypothesis for the studied variable. Practically, this implies that the spatial correlation between random variables corresponding to two different points only depends on the separation vector of the points.

One readily available information that could be useful to complement the borehole and well data is the known location of outcrops. After all, the overburden thickness at outcrops is obviously zero. One intuitive approach would be to treat outcrops as simply additional zero thickness data. However, such an approach could possibly lead to serious bias, in both the variogram modeling and the kriging, due to the preferential observation of the zero thickness value. It could be useful to simply discard the outcrops for the modeling phase (variogram computation and modeling) and introduce the zeros only for the kriging step. But this does not solve the potential bias problem for the kriging phase.

Dubrule and Kostov (1986) presented kriging under inequality constraints. One idea could be to use the outcrop information as an inequality, i.e. thickness is negative at outcrops and positive elsewhere. However, this approach does not really use explicitly the outcrop information. Moreover, it involves quadratic programming and therefore it is more difficult to apply than kriging as the solution is obtained iteratively.

Another approach is to use constrained simulation using a Gibbs sampler (Freulon and de Fouquet, 1993). In this approach, all points on the map convey the information that it is or not an outcrop. The Gibbs sampler seeks to simulate a Gaussian field of thickness, conditional to the borehole data and the outcrop information, that respects the information at all points. The drawbacks of this approach is CPU time consumption and the necessary Gaussian hypothesis for the Gibbs sampler.

A possibly better approach could be to use a form of non-stationary kriging. Matheron (1973)

developed the IRF-k theory to incorporate a polynomial trend in the mean. Another approach is to link the drift to a secondary variable. The resulting kriging with external drift was applied in hydrogeology by Brochu and Marcotte (2003), and Rivest *et al.* (2008) among others. All these methods would not however avoid the bias introduced by the over-representation of outcrop data.

Studies attempting to adapt kriging specifically to the case of non-stationary covariance model include Haas (1990a) who developed a heterogeneous model by moving window where stationary variogram is allowed locally. He applied this approach to pollutant estimation (Haas, 1990b, 1995). Similar approaches for environmental applications are found in Horta *et al.* (2010) and Soares (2010). Problems with these approaches lies in the estimation of the variogram within each window and the discontinuities it might introduce. Moreover, it is conceptually not satisfying that the covariance between two given points changes as a function of the estimation point.

Sampson and Guttorp (1992) proposed to model non-stationary processes by transforming the geographical space by multidimensional scaling technique to a deformed space then assumed stationary. Environmental applications of this approach are found in (Guttorp and Sampson, 1994; Guttorp *et al.*, 1994). Some authors developed case-specific distance measures that implicitly allow non-stationary covariances (Bailly *et al.*, 2006; Ver Hoef *et al.*, 2006; de Fouquet and Bernard-Michel, 2006; Vera *et al.*, 2008, 2009; Rivest and Marcotte, 2012). Although interesting, it is not clear how to define a distance that would account for the numerous outcrops. Moreoever, a stationary covariance in the deformed space with nugget effect would not remove the bias problem related to the preferential sampling of zeros.

One of the most flexible approaches to model non-stationary fields is the convolution method proposed by Higdon *et al.* (1999). They defined non-stationary covariance structures by convolving spatially varying kernel functions. Some studies focused on extensions and applications of convolution-based models (Fuentes and Smith, 2001; Paciorek, 2003; Paciorek and Schervish, 2004, 2006; Jun and Stein, 2008; Rivest and Marcotte, 2012; Mateu *et al.*, 2013; Shamsipour *et al.*, 2013).

In this paper, a non-stationary model that accounts for the peculiar nature of the outcrop data is proposed. Since the influence of outcrops is expected to be local, the thickness covariance is assumed stationary for the points located outside the distance of influence of outcrops. Within the distance of influence, the nugget is assumed to decrease to zero when reaching the outcrop so as to force the estimation to zero. A jointly coherent and admissible model of covariance for borehole and outcrop points is obtained by the convolution method of Higdon *et al.* (1999).

After a short recall of the main points of the convolution approach, the covariance model parametrization adopted here and the estimation of the meta-parameters are presented. Then, a case study is examined where three different methods are compared : Stationary kriging that ignores the outcrop information (K-S), stationary kriging with outcrop data taken as zero thickness (K-SO) and the non-stationary kriging (K-NS).

5.3 Methodology

5.3.1 Non-Stationary Covariance Model

Non-stationary covariance model C^{NS} can reflect heterogeneity of a variable's spatial structure. The covariance $C^{NS}(x_i, x_j)$ depends separately on the specific locations x_i and x_j, not only on their separation vector $x_i - x_j$ as in the stationary case. Higdon *et al.* (1999) proposed to calculate the non-stationary covariance structures by convolving spatially varying kernel function :

$$C^{NS}(x_i, x_j) = \int_{\Re^2} k_{x_i}(u) k_{x_j}(u) du \tag{5.1}$$

where x_i, x_j and u are locations in $\Re^2$, and $k_x(u)$ is kernel functions centered at x. Note that here the kernel function must be positive definite. An example of a kernel function is the Gaussian kernel,

$$k_x(u) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \ (x-u)^T \Sigma^{-1} (x-u) \tag{5.2}$$

Paciorek and Schervish (2006) proposed the following class of non-stationary correlation functions :

$$R^{NS}(x_i, x_j) = |\Sigma_i|^{1/4} |\Sigma_j|^{1/4} \left| \frac{\Sigma_i + \Sigma_j}{2} \right|^{-1/2} R^S \left(\sqrt{Q_{ij}} \right) \tag{5.3}$$

with quadratic form,

$$Q_{ij} = (x_i - x_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (x_i - x_j) \tag{5.4}$$

where R^S is a stationary correlation function admissible in all dimensions $\Re^p$ for $p = 1, 2, \ldots$, Σ_i is the covariance matrix of the Gaussian kernel centered at x_i. Examples of admissible correlation models are the Gaussian, the exponential, and the Matérn family. Note that the quadratic form Q_{ij} defines a squared distance between points x_i and x_j equivalent to an Euclidean distance after rotation and scaling, analogously to the Mahalanobis distance. A

non-stationary covariance model is obtained by multiplying each entry in R^{NS} by standard deviations $\sigma_i\sigma_j$. Appendix A gives a detailed example of covariance computations for the non-stationary model used in the case study.

For the overburden thickness modeling, the outcrops can be considered to introduce a non-stationarity factor in the model. Indeed, at the outcrops, thickness is identically zero everywhere and the nugget effect is zero as the thickness is perfectly known. It is expected that close to a known outcrop, thickness should be rather small, and the same for the nugget. On the other hand, as soon as one gets far enough from an outcrop, one expects that the outcrop has no more influence on the thickness at that point. These remarks suggest that there could exist an average distance of influence of the outcrops on the thickness, identified as a_{out}. At distances larger than a_{out}, the thickness is assumed to behave like a stationary field with correlation range a where $a > a_{out}$. Within a_{out}, the covariance model is assumed to be of the same type (e.g. exponential), have the same sill C but has a correlation range between a and a_{out} and a nugget effect to sill ratio ($C0/C$) that decreases as the point gets closer to the outcrop. Although many parametrizations exist satisfying these criteria, we adopted the following simple parametrization :

$$a_{NS} = (1 - f(d))a_S + f(d)a_{out} \tag{5.5}$$

$$C0_{NS} = (1 - f(d))C0_S \tag{5.6}$$

$$f(d) = 0.5 + 0.5cos\left(\frac{d}{a_{out}}\pi\right) \ \forall d < a_{out} \tag{5.7}$$

$$f(d) = 0 \ \forall d \geq a_{out} \tag{5.8}$$

where a_S and $C0_S$ are the corresponding parameters of the stationary covariance model, and d is the distance from the closest outcrop to the point considered. Note that when $d > a_{out}$, $f(d) = 0$, hence, the stationary model applies. On the contrary, when $d = 0$, $f(d) = 1$, the NS correlation range equals a_{out} and the nugget effect becomes 0. Therefore, when estimating a point located close to an outcrop, the outcrop will receive a strong kriging weight. Conversely, when kriging a point at a larger distance than a_{out} from an outcrop, the outcrop will usually receive a smaller kriging weight than with the stationary covariance model. Figure 5.1 shows the shape of the function $f(d)$ used in Equations 5.5 and 5.6. Figure 5.2 illustrates the borehole stationary and non-stationary ordinary kriging weights obtained for the simple case of one borehole data (at x=0) and one outcrop (at x=1000). Weights are shown as a function of the position of the estimation point. Clearly, the stationary model will fail to recover zero

thickness close to outcrops, contrary to the non-stationary model.

The stationary covariance function for the overburden thickness with parameters a_S, $C0_S$ and C is obtained by optimizing model parameters in a cross-validation study (see section 5.4.2), after visual determination of the model type and possible anisotropies on the experimental variogram obtained without the outcrop data. Once the stationary model is obtained, the a_{out} parameter is also determined by a cross-validation study that includes the outcrop information.

Impact of the outcrop influence distance parameter a_{out}

To illustrate the impact of parameter a_{out}, consider a 1D profile with two thickness data (at x=100 m and x=1100 m) and two outcrop points (at x=400 m and x=700 m). The estimation by K-S, K-SO and K-NS is computed for different a_{out} values, see Figure 5.3. As already mentioned, the non-stationary covariance model forces the estimation towards zero when the estimated point gets close to an outcrop point. The distance of influence of outcrops (a_{out}) controls the extent of the area directly affected by the outcrop information.

5.3.2 Thickness Estimation

Ordinary kriging (OK) is used to estimate the thickness. With known values $Z(x_i)$ at points x_i, $(i = 1, \cdots, N)$, OK estimates the value at unknown point x_0 by (Chilès and Delfiner, 1999) :

$$\begin{array}{c} Z(x_0)^* = \sum_{i=1}^{N} \lambda Z(x_i) \\ Subject\ to\ \sum_{i=1}^{N} \lambda = 1 \end{array} \tag{5.9}$$

where λ denotes the weights, which are obtained from the OK system :

$$\begin{bmatrix} C(x_1,x_1) & C(x_1,x_2) & \cdots & C(x_1,x_N) & 1 \\ C(x_2,x_1) & C(x_2,x_2) & \cdots & C(x_2,x_N) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C(x_N,x_1) & C(x_N,x_2) & \cdots & C(x_N,x_N) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} C(x_1,x_0) \\ C(x_2,x_0) \\ \vdots \\ C(x_N,x_0) \\ 1 \end{bmatrix} \tag{5.10}$$

where $C(.,.)$ is the covariance function, λ's are weights for observations and μ is the Lagrange multiplier. C can be either stationary or nonstationary, in the latter case obtained from Equation 5.3.

The overburden thickness is interpolated by ordinary kriging based on both stationary and

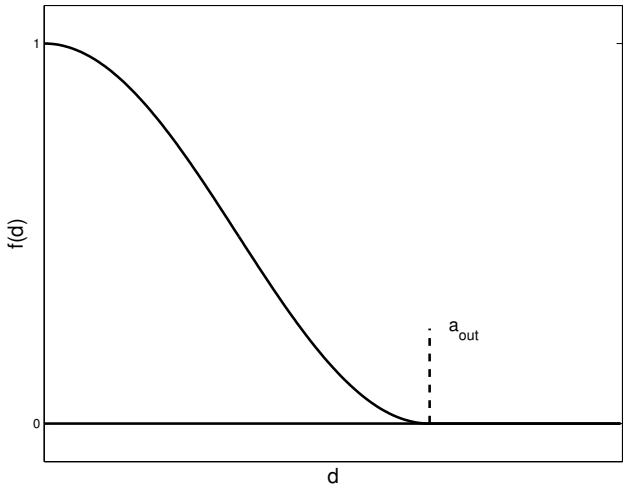

Figure 5.1 $f(d)$ function

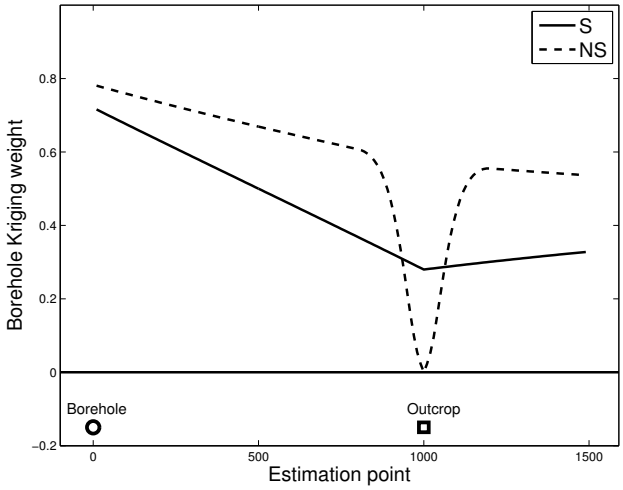

Figure 5.2 Borehole kriging weights for stationary and nonstationary kriging with $a_{out} = 200$ m and exponential stationary covariance with $C_0 = 1, C = 2$ and $a = 2000$ m.

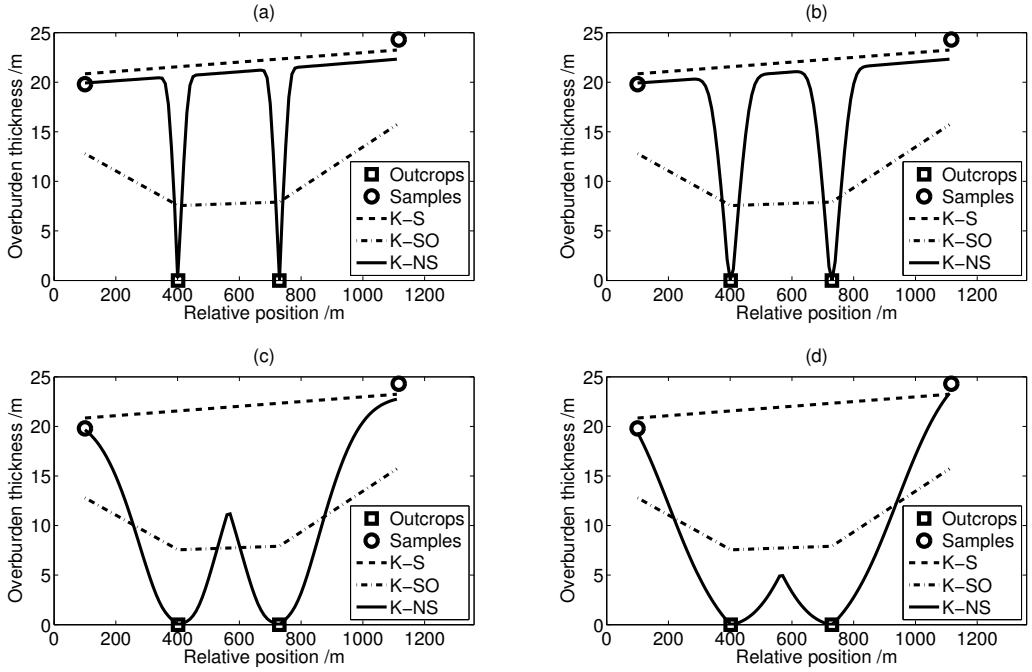

Figure 5.3 Thickness profiles by stationary and nonstationary kriging with a_{out} of (a) 60 m, (b) 120 m, (c) 500 m and (d) 1000 m.

non-stationary (K-NS) covariance models. In the stationary case, two strategies are compared, one ignoring completely the outcrop information (K-S), the other treating outcrops simply as additional data points with zero thickness (K-SO).

5.3.3 Performances Evaluation

To compare performances of K-S, K-SO and K-NS, cross-validation is used (Marcotte, 1995). For this purpose, the whole data set is split in two subsets, a training set and a testing set. Using the training data set and all the outcrops, overburden thickness is estimated at points of the testing set. The estimations are compared with values observed at these points. Two statistics, the mean error (ME) and the mean absolute error (MAE), are used to evaluate the estimation accuracy of the three approaches. The ME and MAE are calculated using the following formulas :

$$ME = \frac{1}{n}\sum_{i=1}^{n}(Z_i^* - Z_i) \tag{5.11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Z_i^* - Z_i| \tag{5.12}$$

where Z_i^* is the estimated value ; Z_i is the real data ; n is the total number of data in the testing set. These statistics are computed over the whole testing set or only with the points within distance a_{out} of an outcrop. Recall that the testing set is not involved at all in the initial determination of the covariance parameters and the a_{out} parameter, therefore the above statistics represent a fair assessment of the model performances.

5.4 Case Study

5.4.1 Study area

The overburden thickness data, includes 1051 boreholes located in Montérégie Est, in the south of Québec (Carrier *et al.*, 2013). The study area covers approximately 1500 km^2. A set of 526 boreholes is selected to form the training set. The remaining 525 boreholes form the testing set. In addition, there are 2065 quasi-point outcrops collected from geological maps. The data and outcrops locations are shown in Figure 5.4.

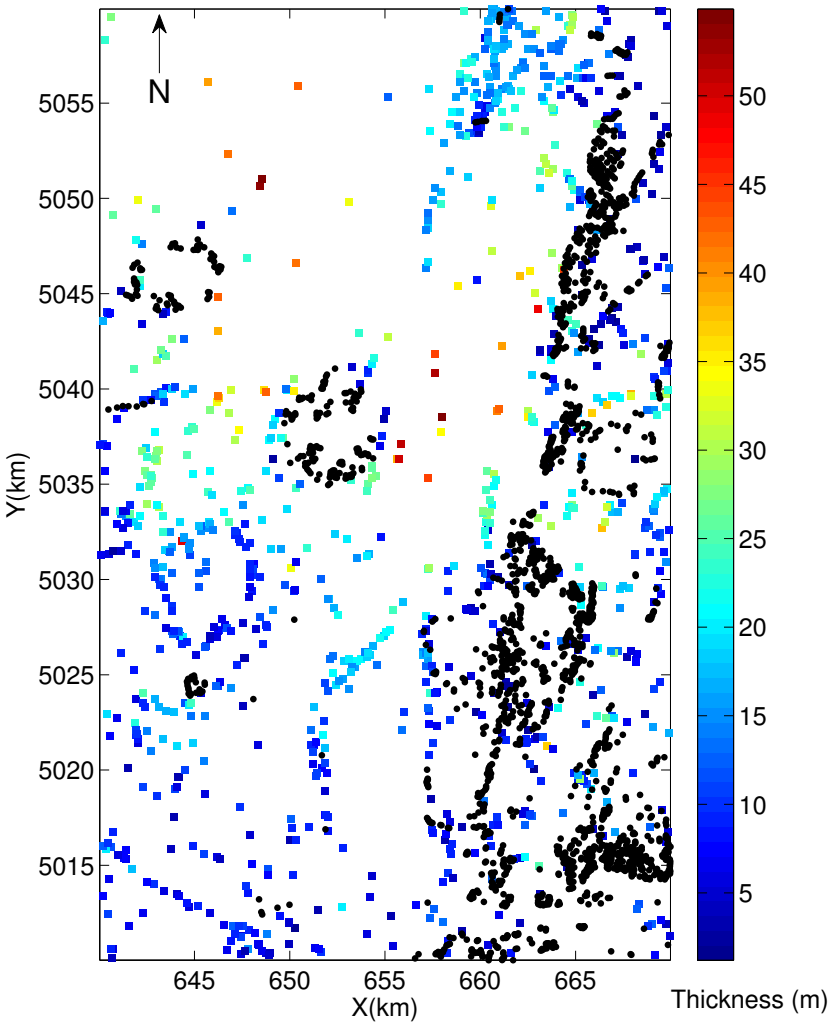

Figure 5.4 Map of study area and sample data. Borehole data as squares and outcrops as black dots

5.4.2 Model Parameters and results

Using only the data in the training set, the experimental variogram on overburden thickness is computed. An exponential isotropic function provides the optimal MAE for cross-validation on the training set. This stationary covariance model has parameters nugget $C0_S = 9.5$ m^2, sill $C = 78$ m^2 and correlation range $a_S = 6686$ m. These parameters are used in all the rounds of cross-validation.

Estimation of a_{out}

For points close to outcrop data, the type of covariance model is assumed remain an exponential isotropic model. The model parameter a_{out} is varied over a series of values between 20 m and 6000 m. The training set MAE cross-validation results are shown in Figure 5.5. Clearly the MAE reaches a minimum around $a_{out} = 120$ m.

Comparison between the estimations with non-stationary and stationary covariances

With the optimized covariance parameters, performance of K-NS, K-S and K-SO are compared. To increase statistics robustness, we considered 100 random drawings to form as many different training sets of size 526 used to estimate the 525 points of the testing set. Note however that the covariance models is not re-estimated for each training set and the same exponential isotropic model with nugget $C0_S = 9.5$ m^2, sill $C = 78$ m^2 and correlation range $a_S = 6686$ m is used. Table 5.1 lists the mean statistics for the 100 rounds cross-validation obtained for the three methods.

The method of stationary kriging with outcrops data treated as zero values (K-SO) shows the largest MAE. Moreover, the estimates appear biased towards underestimation (ME < 0) due to the preferential sampling of zero values corresponding to outcrops. This indicates that, when not used correctly, more information (the 2065 quasi-point outcrops) does not necessarily brings higher precision. The non-stationary method (K-NS) shows the lowest MAE among the three methods. The method appears also unbiased (ME≈ 0). The relative improvements in MAE are respectively 0.2% and 16% compared to K-S and K-SO. When computing statistics over the testing points within distance a_{out} from the nearest outcrop, the MAE improves by 13.8% and 35.5% compared to K-S and K-SO. Note that in average 16 boreholes per run (min 8, max 24) were found within the distance a_{out} from an outcrop leading to a total of 1631 points over the 100 runs used to compute the statistics ME_{120m} and MAE_{120m}. As expected, the improvement brought by K-NS compared to K-S is larger close to

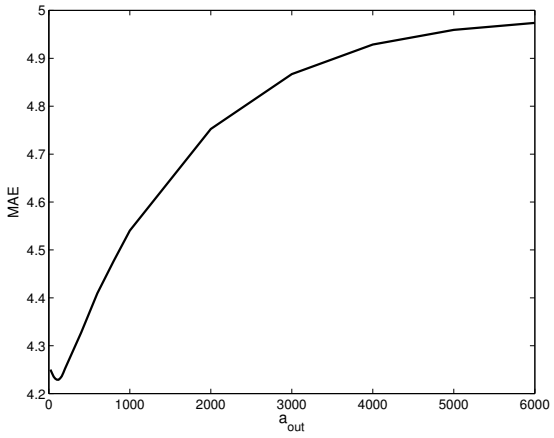

Figure 5.5 Training set cross-validation MAE for the non-stationary covariance model as a function of a_{out}.

the outcrops. Moreover, the ME of K-NS for these points remains close to zero whereas K-S overestimates the thickness (ME=2.8 m) due to not incorporating the outcrop information. Hence, the K-NS is as precise as K-S away from the outcrops and more precise close to the outcrops. K-NS is more precise than K-SO in all circumstances, the latter being seriously biased.

Figure 5.6 shows the maps of overburden thickness produced by the three methods. The map of K-SO exhibits a strong influence of outcrops. The effects of zeros is spreading at a large distance of the outcrops due to the large range of the stationary model. On the contrary, the maps of K-NS and K-S appear visually quite similar, except for estimation points close to outcrops where the outcrop effect is only visible on K-NS. Figure 5.7 shows the correlations between the 3 maps as a function of the distance to the nearest outcrop. Only the points within that distance are kept when computing correlation. Clearly, differences between K-NS and K-S are apparent only for points close to outcrops.

Table 5.1 Statistics of estimates by stationary (K-S and K-SO) and non-stationary kriging (K-NS)

	K-S	K-SO	K-NS
ME	-0.0031	-3.0073	-0.0636
MAE	4.2376	5.0349	4.2289
ME_{120m}	2.8212	-7.7798	0.2255
MAE_{120m}	5.8271	7.7915	5.0235

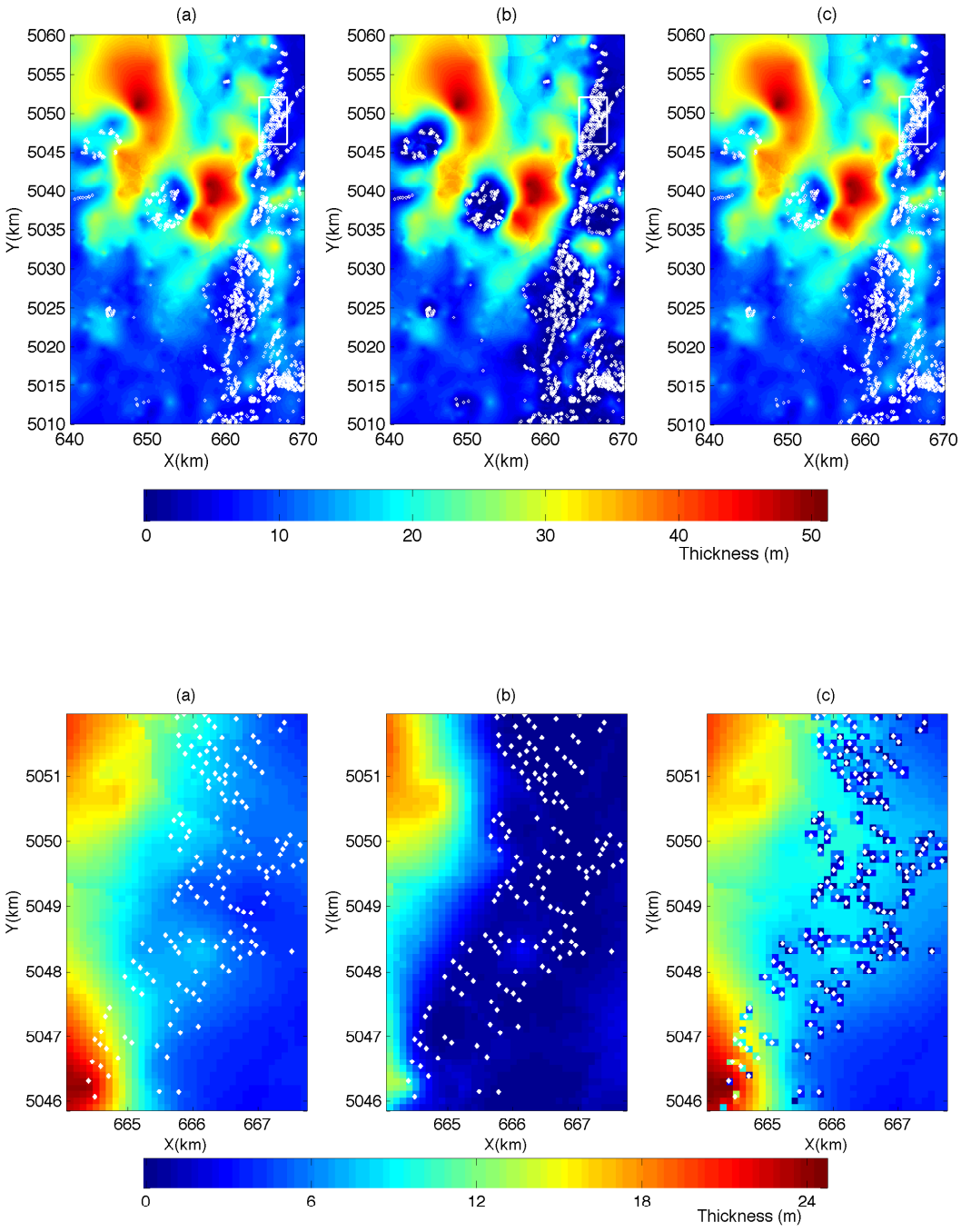

Figure 5.6 Estimation maps by (a) K-S, (b) K-SO and (c) K-NS. White dots represent outcrops. Top row, the entire study area. Bottom row, zoom in the outlined rectangle.

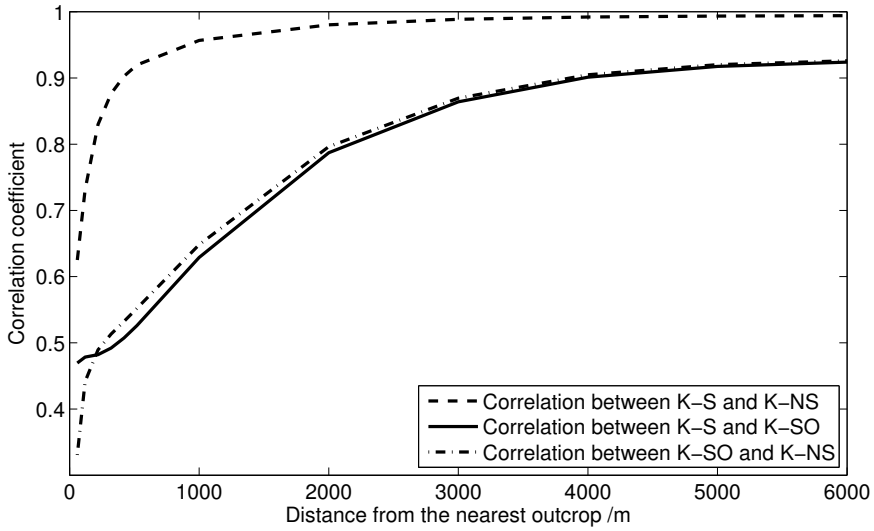

Figure 5.7 Correlations between K-S, K-SO and K-NS as a function of the distance to the nearest outcrop. Only the points within that distance are kept when computing the correlations.

5.5 Discussion

The non-stationary kriging approach proposed succeeds to account for the information about presence of outcrops without introducing a bias toward underestimation of the overburden thickness. This bias is obvious when the outcrops are intuitively treated as zero overburden thickness and used in ordinary kriging. Note that the bias would remain with alternative interpolation approaches such as simple kriging with a spatially varying mean or universal kriging. In the NS model, the smaller correlation range associated with the outcrop points limits their distance of influence, but keeps the outcrop signature when the estimation point gets close to it. The estimation map appears very similar to the map that ignores completely the outcrop information (K-S). However, as the estimation point gets close to the outcrop, the outcrop influence increases gradually so as to retrieve exactly a zero overburden thickness when reaching the outcrop. This approach proved more precise than the two contenders considered. It is definitively more precise than K-SO but only slightly more precise than K-S (see Table 5.1). Compared to K-S, the main merit of the approach is to provide more realistic estimates close to the outcrops.

A simple method based on a training set MAE cross-validation was used to estimate the overburden stationary covariance and the non-stationary covariance associated with the outcrops. Parameters for the stationary model were obtained by cross-validation and visual assessment

of the experimental variograms computed with borehole data only. For the non-stationary model, the covariance parameters are parametrized as functions of the distance to the closest outcrop. The parameter a_{out} controls the distance of influence of the outcrops. This parameter is determined by the cross-validation performance of the corresponding non-stationary kriging for the thickness estimation in a training set. The optimal value found is small (a_{out} = 120 m), indicating a small distance of influence of the outcrops. Moreover, because the number of boreholes found within 120 m from an outcrop is limited, it precludes parameterizing the NS covariance more finely.

Only non-stationarity of the covariance was considered. It could be argued that non-stationarity of the mean could be present in the vicinity of outcrops. Attempts made with kriging with external drift where the drift is assumed to vary linearly with the distance to an outcrop failed to improve the estimation. Moreover, this approach was less efficient than the proposed NS model to force naturally the estimation to zero at outcrops.

5.6 Conclusion

The non-stationary model presented is helpful to include the outcrop information without introducing the kind of bias that is observed when outcrops are simply treated as additional zero data. It involves a single additional parameter a_{out} that can easily be estimated by cross-validation. It helps produce more realistic maps and improve the thickness estimation precision, especially within the distance of influence of the outcrops.

5.7 Acknowledgements

This research was made possible by research grants provided by Chinese Scholarship Council and NSERC.

5.8 Appendix A

The computation of the non-stationary covariance is illustrated along a profile in 2D. Three borehole data, one outcrop point and one estimation point are respectively located at x_1(100,0), x_2(600,0), x_3(1000,0), x_{out}(500,0) and x_0(450,0). The points x_2, x_{out} and x_0 are within the distance of influence of the outcrop here selected as a_{out}=120 m. Following Equations 5.5 to 5.8, the covariance parameters on these points can be determined. Table 5.2 gives the resulting covariance parameters associated to each point.

The same stationary and isotropic covariance model as for the case study is used (exponential

Table 5.2 Non-Stationary covariance model parameters for each point

Point	d (m)	f(d)	a(d) (m)	C(d) (m^2)	C0(d) (m^2)
x_1	400	0	6686	78	9.5
x_2	100	0.067	6246.2	78	8.864
x_3	500	0	6686	78	9.5
x_{out}	0	1	120	78	0
x_0	50	0.629	2553.3	78	3.521 (not used)

model with $C0_S = 9.5$ m^2, $C = 78$ m^2 and $a_S = 6686$ m). Then data-to-data and data-to-estimation point covariance are calculated following Equation 5.3 and 5.4. In the equation, x_i and x_j represent the 2×1 vector of coordinates at points i and j. In the model, the NS covariance at point x_i is considered isotropic. Hence, the kernel matrix Σ_i is a diagonal matrix with a_i^2 repeated along the diagonal, and similarly for Σ_j (with a_i obtained from Table 5.2). As an example, for the pair of points $x_2 - x_{out}$, one computes :

$$\begin{aligned} Q_{2,out} &= [100, 0] \left(\frac{\begin{bmatrix} 6246.2^2 & 0 \\ 0 & 6246.2^2 \end{bmatrix} + \begin{bmatrix} 120^2 & 0 \\ 0 & 120^2 \end{bmatrix}}{2} \right)^{-1} \begin{bmatrix} 100 \\ 0 \end{bmatrix} \\ &= 5.1243 \times 10^{-4} \end{aligned} \tag{5.13}$$

Hence, one gets using Equation 5.3 :

$$\begin{aligned} R_{2,out}^{NS} &= 6246.2 \times 120 \times (5.1243 \times 10^{-8}) \times \exp(-\sqrt{5.1243 \times 10^{-4}}) \\ &= 0.0375 \end{aligned} \tag{5.14}$$

Finally, as the point variance was assumed stationary at 78 m^2

$$C_{2,out}^{NS} = 0.0375 * \sqrt{(78)(78)} = 2.9289 \text{ m}^2$$

The resulting data-to-data and data-to-estimation point covariances are listed in Table 5.3 for K-NS and in Table 5.4 for K-SO. For that particular example, the outcrop kriging weight for K-NS is less than with K-SO. However, as the estimation point gets closer to the outcrop, the weight assigned to the outcrop point increases (see Figure 5.8).

The model is more general than the above example suggests as the kernel matrix Σ_i can be

Table 5.3 Non-Stationary covariances for K-NS and kriging weights

	$C(x_i, x_j)$			$C(x_i, x_0)$		
	x_1	x_2	x_3	x_{out}	x_0	λ
x_1	87.5	72.0322	68.1764	2.5719	48.5711	0.2130
x_2	72.0322	86.8636	73.1542	2.9289	52.9483	0.4861
x_3	68.1764	73.1542	87.5	2.5181	46.6377	0.0880
x_{out}	2.5719	2.9289	2.5181	78	7.1159	0.2129

Table 5.4 Stationary covariances for K-SO and kriging weights

	$C(x_i, x_j)$			$C(x_i, x_0)$		
	x_1	x_2	x_3	x_{out}	x_0	λ
x_1	87.5	72.3797	68.1764	73.4704	74.0219	0.2420
x_2	72.3797	87.5	73.4704	76.8421	76.2696	0.2633
x_3	68.1764	73.4704	87.5	72.3797	71.8404	0.1346
x_{out}	73.4704	76.8421	72.3797	87.5	77.4189	0.3601

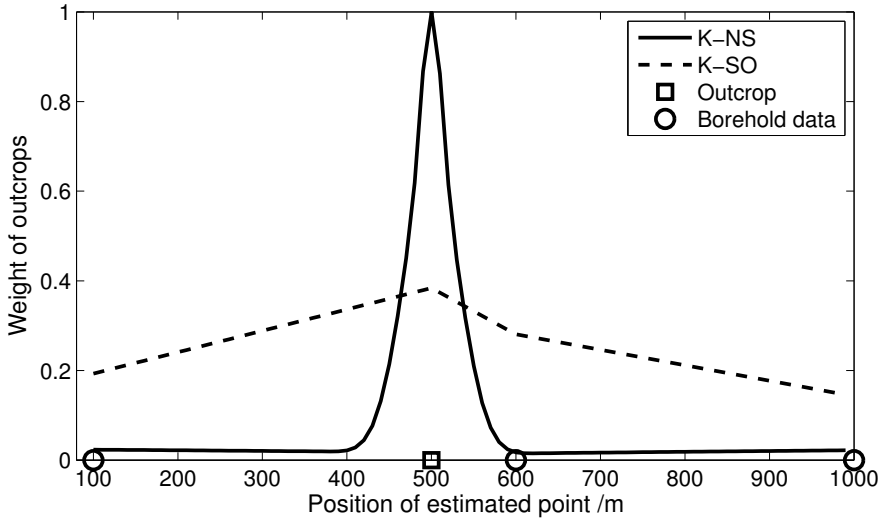

Figure 5.8 Kriging weight assigned to outcrop as a function of the estimation point for K-SO and K-NS

any symmetric positive-definite matrix instead of a diagonal matrix. This would enable to model spatially varying orientation and anisotropy ratio of the ellipses of ranges (i.e. spatially varying geometric anisotropy). As an example, at point x_i with correlation ranges of 1000 m and 200 m along azimuths 30^o and 120^o respectively, the associated kernel matrix is :

$$\begin{aligned} \Sigma_i &= U\Lambda U^T \\ &= \begin{bmatrix} \cos(30) & \sin(30) \\ -\sin(30) & \cos(30) \end{bmatrix} \begin{bmatrix} 200^2 & 0 \\ 0 & 1000^2 \end{bmatrix} \begin{bmatrix} \cos(30) & -\sin(30) \\ \sin(30) & \cos(30) \end{bmatrix} \\ &= \begin{bmatrix} 280000 & 415692 \\ 415692 & 760000 \end{bmatrix} \end{aligned} \tag{5.15}$$

where U is a rotation matrix.

Similarly, the variance can be modeled as a spatially varying function like here for the range and the nugget effect. Rivest and Marcotte (2012) shows an example where the variance of a contaminant is described by a decreasing function of the distance to the source of the contaminant.

References

BAILLY, J. S., MONESTIEZ, P. and LAGACHERIE, P. (2006). Modelling spatial variability along drainage networks with geostatistics. *Mathematical Geosciences*, 38, 515–539.

BROCHU, Y. and MARCOTTE, D. (2003). A simple approach to account for radial flow and boundary conditions when kriging hydraulic head fields for confined aquifers. *Mathematical Geology*, 35, 111–139.

CARRIER, M.-A., LEFEBVRE, R., RIVARD, C. and PARENT, M. (2013). Atlas du projet Montérégie Est, sud du Québec, Canada. Technical report Rapport final INRS R-1412.

CHILÈS, J. P. and DELFINER, P. (1999). *Geostatistics : Modeling Spatial Uncertainty*. John Wiley & Sons.

DE FOUQUET, C. and BERNARD-MICHEL, C. (2006). Modèles géostatistiques de concentrations ou de débits le long des cours d'eau. *Comptes Rendus Geoscience*, 338, 307–318.

DUBRULE, O. and KOSTOV, C. (1986). An interpolation method taking into account inequality constraints : I. methodology. *Mathematical geology*, 18, 33–51.

FREULON, X. and DE FOUQUET, C. (1993). Conditioning a gaussian model with inequalities. *Geostatistics Troia'92*, Springer. 201–212.

FUENTES, M. and SMITH, R. L. (2001). A new class of nonstationary spatial models. Technical report, North Carolina State University, Raleigh, NC.

GUTTORP, P., MEIRING, W. and SAMPSON, P. D. (1994). A space-time analysis of ground-level ozone data. *Environmetrics*, 5, 241–254.

GUTTORP, P. and SAMPSON, P. D. (1994). Methods for estimating heterogeneous spatial covariance functions with environmental applications. *Handbook of Statistics*, 12, 661–689.

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

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

HAAS, T. C. (1995). Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *Journal of the American Statistical Association*, 90, 1189–1199.

HIGDON, D., SWALL, J. and KERN, J. (1999). Non-stationary spatial modeling. *Bayesian statistics*, 6, 761–768.

HORTA, A., CAEIRO, M. H., NUNES, R. and SOARES, A. (2010). Simulation of continuous variables at meander structures : Application to contaminated sediments of a lagoon. P. M. Atkinson and C. D. Lloyd, editors, *geoENV VII - Geostatistics for Environmental Applications*, Springer Netherlands, no. 16 Quantitative Geology and Geostatistics. 161–172.

JUN, M. and STEIN, M. L. (2008). Nonstationary covariance models for global data. *The Annals of Applied Statistics*, 2, 1271–1289.

MARCOTTE, D. (1995). Generalized cross-validation for covariance model selection. *Mathematical geology*, 27, 659–672.

MATEU, J., FERNÁNDEZ-AVILÉS, G. and MONTERO, J. M. (2013). On a class of nonstationary, compactly supported spatial covariance functions. *Stochastic Environmental Research and Risk Assessment*, 27, 297–309.

MATHERON, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5, 439.

PACIOREK, C. J. (2003). *Nonstationary Gaussian processes for regression and spatial modelling*. PhD Dissertation, Department of Statistics, Carnegie Mellon University.

PACIOREK, C. J. and SCHERVISH, M. J. (2004). Nonstationary covariance functions for gaussian process regression. *In Proc. of the Conf. on Neural Information Processing Systems (NIPS)*. MIT Press, Cambridge, 273–280.

PACIOREK, C. J. and SCHERVISH, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17, 483–506.

RIVEST, M. and MARCOTTE, D. (2012). Kriging groundwater solute concentrations using flow coordinates and nonstationary covariance functions. *Journal of Hydrology*, 472–473, 238–253.

RIVEST, M., MARCOTTE, D. and PASQUIER, P. (2008). Hydraulic head field estimation using kriging with an external drift : A way to consider conceptual model information. *Journal of Hydrology*, 361, 349–361.

SAMPSON, P. D. and GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 87, 108–119.

SHAMSIPOUR, P., MARCOTTE, D., CHOUTEAU, M., RIVEST, M. and BOUCHEDDA, A. (2013). 3D stochastic gravity inversion using nonstationary covariances. *Geophysics*, 78, G15–G24.

SOARES, A. (2010). Geostatistical methods for polluted sites characterization. P. M. Atkinson and C. D. Lloyd, editors, *geoENV VII - Geostatistics for Environmental Applications*, Springer Netherlands, no. 16 Quantitative Geology and Geostatistics. 187–198.

VER HOEF, J. M., PETERSON, E. and THEOBALD, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological statistics*, 13, 449–464.

VERA, J. F., MACÍAS, R. and ANGULO, J. M. (2008). Non-stationary spatial covariance structure estimation in oversampled domains by cluster differences scaling with spatial constraints. *Stochastic Environmental Research and Risk Assessment*, 22, 95–106.

VERA, J. F., MACÍAS, R. and ANGULO, J. M. (2009). A latent class MDS model with spatial constraints for non-stationary spatial covariance estimation. *Stochastic Environmental Research and Risk Assessment*, 23, 769–779.

CHAPTER 6 ARTICLE 2 : A CLASS OF NON-STATIONARY COVARIANCE FUNCTIONS WITH COMPACT SUPPORT

Article history : Submitted 15 January 2015, Accepted 6 June 2015, Published online 23 June 2015, Stochastic Environmental Research and Risk Assessment.

Authors : Min Liang and Denis Marcotte
DOI : 10.1007/s00477-015-1100-y

6.1 Abstract

This article describes the use of non-stationary covariance functions with compact support to estimate and simulate a random function. Based on the kernel convolution theory, the functions are derived by convolving hyperspheres in $\Re^n$ followed by a Radon transform. The order of the Radon transform controls the differentiability of the covariance functions. By varying spatially the hyperspheres radius one defines non-stationary isotropic versions of the spherical, the cubic and the penta-spherical models. Closed-form expressions for the non-stationary covariances are derived for the isotropic spherical, cubic, and penta-spherical models. Simulation of the different non-stationary models is easily obtained by weighted average of independent standard Gaussian variates in both the isotropic and the anisotropic cases. The non-stationary spherical covariance model is applied to estimate the overburden thickness over an area composed of two different geological domains. The results are compared to the estimation with a single stationary model and the estimation with two stationary models, one for each geological domain. It is shown that the non-stationary model enables a reduction of the mean square error and a more realistic transition between the two geological domains.

6.2 Introduction

The spatial structure of a Gaussian random field $Z(x) : x \in \Re^n$ is fully characterized by its covariance function. In most popular geostatistical methods, such as kriging approach, the covariance function $C(\cdot, \cdot)$ is assumed to be stationary which means the covariance $C(x_i, x_j)$ is a function of the vector distance $h_{ij} = x_j - x_i$ or of its norm in the isotropic case. Most of the time, the Euclidean norm is used although Christakos and Papanicolaou (2000) indicate

other norms can be used. The hypothesis of stationarity is practical but often too stringent to fully account for the heterogeneous characteristics of a field.

Haas (1990a) developed a heterogeneous model by moving a window where stationarity is assumed locally. Developments and applications of this method can be found in Haas (1990b, 1995), Horta *et al.* (2010) and Soares (2010). Another important method to include non-stationarity is the spatial transformation technique based on multidimensional scaling (Sampson and Guttorp, 1992). Further applications and extensions of this approach are found in Guttorp *et al.* (1994); Smith (1996); Meiring *et al.* (1997); Damian *et al.* (2001); Schmidt and O'Hagan (2003); Boisvert and Deutsch (2008).

Higdon *et al.* (1999) proposed a non-stationary Gaussian covariance model by convolving spatially-varying kernels. Paciorek (2003) extended this kernel convolution method and produced explicit non-stationary correlation functions with locally varying geometric anisotropies. By evolving the stationary spectrum over space, Pintore and Holmes (2004) constructed closed form non-stationary covariance functions that allowed the degree of differentiability to vary in space. Stein (2005) proposed a very flexible non-stationary spatial covariance functions for Matérn model that enables the differentiability and the local anisotropy to vary spatially. Paciorek and Schervish (2006) built a class of non-stationary covariance functions including a non-stationary version of Matérn covariance model and rational quadratic covariance model. Following the general non-stationary covariance function of Stein (2005), Mateu *et al.* (2013) constructed a class of non-stationary covariance functions by combining completely monotonic functions and compactly supported covariance functions. However, the resulting non-stationary covariance functions appear to be not compactly supported. Applications of non-stationary models using the above models can be found in Paciorek and Schervish (2006); Jun and Stein (2008); Rivest and Marcotte (2012); Shamsipour *et al.* (2013); Liang *et al.* (2014).

A covariance function with compact support means that the covariance between two points becomes zero when their distance exceeds a certain threshold. This important property enables to obtain a sparse covariance matrix which can reduce significantly the memory requirements when dealing with large datasets. Several methods were proposed to create stationary covariance functions with compact support (Moreaux, 2008). One trivial approach is to set to zero the covariances below a certain threshold (Rygaard-Hjalsted *et al.*, 1997). However, this naive approach does not preserve positive definiteness of the covariance function (Furrer *et al.*, 2006). Better approaches use tapering of covariance functions that ensures positive definiteness and compactness. Furrer *et al.* (2006) used tapering functions such as the spherical and Wendland covariance functions, to obtain stationary compact versions of the Matérn fa-

mily covariance functions. They gave conditions on the tapering functions to ensure positive definiteness of the resulting compact covariance. The estimates obtained are asymptotically optimal, but the kriging variances from the tapered model are biased and need rescaling by a factor depending on the data locations. Conditions on tapering functions for other covariance functions than the Matérn class could not be derived by Furrer *et al.* (2006).

Direct construction of stationary compactly supported covariance functions are described by Matheron (1965), Sansò and Schuh (1987), Wu (1995), Wendland (1995), and Gneiting (2002) among others. Matheron (1965) proposed to create compactly supported covariance functions by 'La montée' (an integral operator) and 'La descente' (a differential operator) and gave many examples in isotropic and stationary case. Wu (1995) constructed a compactly supported function by differential operator $D(f)(r) = -\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}f(r)$ on univariate convolution of a cut-off polynomial function $f_l(r)$ continuous in $\Re^{l-1}$,

$$f_l(r) = (1 - r^2)_+^l, \quad l \geq 0, \tag{6.1}$$

$$\varphi_{l,k}^{Wu}(r) = c_{l,k} D^k (f_l * f_l)(r), \quad 0 \leq k \leq l, \tag{6.2}$$

in which $c_{l,k}$ is a constant ensuring $\varphi_{l,k}^{Wu}(0) = 1$. Wu's function $\varphi_{l,k}^{Wu}(r)$ is positive definite in $\Re^{2k+1}$ and possesses $2l - 2k$ derivatives. Moreover, Wu's construction can be connected to Euclid's hat $X^{(2l+1)}$ by integral operator $I(f)(r) = \int_r^\infty uf(u)\mathrm{d}u$,

$$\varphi_{l,k}^{Wu} = I^{l-k} X^{(2l+1)}. \tag{6.3}$$

Euclid's hat $X^{(2n+1)}$ is generated by d-variate auto-convolution of the characteristic function of the unit ball. The I operator is a special case of Radon transform with even order. Therefore, as noted by Gneiting (2002), Wu's function corresponds to the spherical family of Matheron (1965). This family is reviewed in Section 6.3.

Wendland (1995) constructed a group of functions with compact supports by starting with the truncated power function which is also called Askey's power function (Askey, 1973) ,

$$\varphi_l(r) = (1 - r)_+^l, \quad l > 0. \tag{6.4}$$

Then applying the integral operator $I(f)(r)$ repeatedly on this basic function defines a class of compactly supported functions

$$\varphi_{l,k}^{We}(r) = \mathcal{I}^k \varphi_l(r). \tag{6.5}$$

Wendland function $\varphi_{l,k^{We}}$ are strictly positive definite on $\Re^d$ if and only if $l \geq \frac{d+1}{2} + k$ and possesses $2k$ derivatives around zero and $k + l - 1$ derivatives around 1.

Gneiting (2002) presented turning bands operator to model correlation functions with some negative values (also called hole effect model). Starting with a compactly supported function φ_s defined in $\Re^s$, the transformed function still has compact support and is positive definite in $\Re^{s-2}$. Bellier and Monestiez (2010) proposed to model this hole effect by auto-convolution of two concentric spheres. Their covariance model has a single wave effect and a compact support. More classes of Radial basis functions with compact support can be found in Buhmann (2003). Porcu *et al.* (2013) proposed multivariate stationary covariance models with compact support based on scale mixtures of Askey functions (Askey, 1973). Based on functions of Wendland (1995) and Gneiting (2002), Kleiber and Porcu (2015) developed locally-stationary admissible multivariate covariance models with compact support.

Chilès and Delfiner (1999, p. 491) describe how to simulate non-stationary compact models of the spherical family by dilution with fixed functions defined over hyperspheres of radius varying in space. However, they did not provide closed-form expressions for the resulting non-stationary covariances. Moreover, as the dilution function varies in space, the mean as well as the covariance vary in space with this approach. In this article we adopt a slightly different and simpler approach. Instead of using dilution functions, equivalent weighting functions are defined. The weights are normalized so as to provide a unit variance for the simulated field at all locations, hence defining a non-stationary correlation function. In addition, a closed-form for non-stationary compactly supported covariance model is obtained for the spherical model in $\Re^n$ by convolving local indicator functions. For the Radon transformed models (cubic and penta-spherical models) closed-form expressions are derived in $\Re^3$. Moreover, the numerical approximation by discrete convolution of the dilution functions is straightforward, as well as the extension to the anisotropic case.

After a short review of stationary covariance functions of the spherical family, the theory of the non-stationary form and its computation are presented. The closed-form expressions for the non-stationary isotropic spherical, cubic and penta-spherical models are derived. Simple examples of non-stationary covariances are presented. A method for non-stationary simulation based on moving average with normalized non-stationary dilution function is described. A case study for the estimation of overburden thickness in an area composed of two distinct geological domains is presented.

6.3 Stationary covariance functions of the spherical family

The compactly supported covariance functions of the spherical family were first proposed by Matheron (1965). These functions are piecewise polynomial functions. Radon transform is used to define covariance functions with higher differentiability.

6.3.1 The spherical family

Let $w(x)$ be a function in $\Re^n$. The n-dimensional auto-convolution of the function $w(x)$ can be denoted as :

$$g(h) = \int w(x)w(x+h)dx. \tag{6.6}$$

The function $g(h)$ is known as the covariogram in geostatistics. This function is positive definite by construction, therefore it defines a valid stationary covariance function. As an example, $w(x)$ can be an indicator function on a bounded domain D,

$$w(x) = \begin{cases} 1, & x \in D \\ 0, & x \notin D \end{cases} \tag{6.7}$$

Then, the covariogram $g(h)$ is the Lebesgue measure of $D \cap D_{-h}$. For example when D is defined as a line segment in 1D, $g(h)$ represents the length of the intersection between two identical segments, one being translated by h. When D is a circle in 2D, $g(h)$ is the surface of intersection of two identical circles with one circle translated by h. Similarly, when D is a sphere in 3D, $g(h)$ is the intersected volume of two balls with offset h. These cases generate respectively the well known triangular model, circular model, and spherical model. In $\Re^n$, the corresponding covariogram $g(h)$ is the volume of intersection of the two hyperspheres with offset h.

The widely used spherical family model was proposed by Matheron(1965). A general form of covariogram function of $r = |h|$ in $\Re^n$ is (Chilès and Delfiner, 1999, p. 81)

$$g_n(r) = \begin{cases} a^n v_{n-1} \int_{r/a}^1 (1-u^2)^{(n-1)/2}\,\mathrm{d}u, & \text{if} \quad r \le a, \\ 0, & \text{if} \quad r > a, \end{cases} \tag{6.8}$$

where v_n stands for the volume of the unit-diameter ball in $\Re^n$ that can be expressed by

$$v_n = \frac{\pi^{n/2}}{2^{n-1} n \Gamma(n/2)}. \tag{6.9}$$

where $\Gamma(\cdot)$ is the gamma function. The family of spherical covariance functions have compact

support, the ball of radius a in $\Re^n$.

6.3.2 Covariance functions obtained by Radon transform

Radon transform is an integral transform which is also called Montée in geostatistics (Matheron, 1965). If $f(x_1, x_2, \ldots, x_n)$ is a continuous function with a compact support in $\Re^n$, the Radon transformed function Rf is also a compactly supported function defined in $\Re^{n-1}$ after integrating out x_n.

$$Rf = \int_R f(x_1, x_2, \ldots, x_n)\, \mathrm{d}x_n \tag{6.10}$$

By repeating the integration on $x_n, x_{n-1}, \ldots, x_{n-m+1}$, one obtains the Radon transform of order m that is defined on space $\Re^{n-m}$.

Let g_n be the covariogram of the indicator function w_n in $\Re^n$. In isotropic case, Radon transformed covariograms of orders 1 and 2 $g_{n,1}(r)$ and $g_{n,2}(r)$ of $g_n(r)$ are calculated by Chilès and Delfiner (1999, p. 73)

$$g_{n,1}(r) = 2\int_0^\infty g_n(\sqrt{r^2+\rho^2})\, \mathrm{d}\rho, \tag{6.11}$$

$$g_{n,2}(r) = 2\pi \int_r^\infty u g_n(u)\, \mathrm{d}u. \tag{6.12}$$

The covariogram $g_{n,m}(r)$ of the m order Radon transform is obtained by repeating Equation 6.11 m times or repeating Equation 6.12 $m/2$ times when m is even. The Radon transform of an isotropic covariogram is isotropic. Moreover, the covariogram function $g_{n,2q}$ is $2q$-times continuously differentiable and the corresponding field is q-times mean-square differentiable (Chilès and Delfiner, 1999). The Radon transform of the geometric covariogram is equal to the geometric covariogram of the Radon transformed variable. Therefore, applying the Radon transform of order $m = 2q$ to the indicator function gives a dilution function $\tilde{w}_{n,2q}(x)$ defined in $\Re^{n-2q}$. Closed-form expression for the dilution function is given by Chilès and Delfiner (1999, p. 82).

$$\tilde{w}_{n,2q}(x) = \begin{cases} v_{2q}(a^2 - 4|x|^2)^q, & \text{if} \quad |x| \le \frac{a}{2}, \\ 0, & \text{if} \quad |x| > \frac{a}{2}, \end{cases} \tag{6.13}$$

where $v_{2q}$is the volume of the unit-diameter ball of $\Re^{2q}$.

In section 6.4.1, the dilution functions are used as weighting functions in the moving average method of simulation. Hence, a more convenient normalizing constant is defined so as to ensure unit variance at each point. This is obtained by ensuring the sum of squares of the

weights equals one.

$$w_{n,2q}(x) = \begin{cases} c_{n,q}(a^2 - 4|x|^2)^q, & \text{if} \quad |x| \leq \frac{a}{2}, \\ 0, & \text{if} \quad |x| > \frac{a}{2}, \end{cases} \tag{6.14}$$

Table 6.1 gives the normalizing constant $c_{n,q}$ for the spherical, cubic and penta-spherical models with correlation range a corresponding to the diameter of the corresponding hyper-sphere.

6.4 Non-stationary compactly supported covariance functions

In the stationary case, the isotropic spherical covariance model is proportional to the volume of intersection of two hyperspheres of equal diameter a translated by h. Anisotropy is introduced by replacing the hyperspheres by hyper-ellipsoids as the support for the indicator function. As suggested by Chilès and Delfiner (1999), an immediate generalization in the non-stationary case is to allow the hyperspheres or hyper-ellipsoids to have different sizes and orientations (for the hyper-ellipsoids) defined as a function of the location x.

6.4.1 Non-stationary isotropic covariance model by convolution

The non-stationary random variable $Z(x)$ can be obtained by weighted average as $Z(x_i) = \int w_i(x_i - u)Y(x_i - u)\mathrm{d}u, i = 1, \cdots, n$ where $Y(x_i - u), i = 1, \cdots, n$ are zero mean and unit variance independent Gaussian variables. Considering two specific points x_i and x_j, one has :

$$C^{NS}(x_i, x_j) = \int_{\Re^{n-2q}} w_i(x_i - u)w_j(x_j - u)\mathrm{d}u. \tag{6.15}$$

Table 6.1 Normalizing constants in Eq. 6.14

Case	$c_{n,q}$
Spherical (n=3, q=0)	$\left(\frac{6}{\pi a^3}\right)^{0.5}$
Cubic (n=5, q=1)	$\left(\frac{105}{4\pi a^7}\right)^{0.5}$
Penta-spherical (n=7, q=2)	$\left(\frac{3465}{64\pi a^{11}}\right)^{0.5}$

Following Paciorek and Schervish (2006), for a set of m points, one has :

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j C^{NS}(x_i, x_j) &= \sum_{i=1}^{m}\sum_{j=1}^{m} c_i c_j \int w_i(x_i - u) w_j(x_j - u) \mathrm{d}u \\ &= \int \sum_{i=1}^{m} c_i w_i(x_i - u) \sum_{j=1}^{m} c_j w_j(x_j - u)\, \mathrm{d}u \\ &= \int \left(\sum_{i=1}^{m} c_i w_i(x_i - u) \right)^2 \mathrm{d}u \end{aligned} \tag{6.16}$$

which is always non-negative. Hence the non-stationary covariance function is admissible. Globally the covariance function has a compact support less or equal to the maximum range found over the field.

6.4.2 Closed-form expression in the non-stationary isotropic case for the spherical model

Specifically the covariance function of spherical model is the n-dimensional convolution of indicator functions of two spheres defined in $\Re^n$. It corresponds to the volume of intersection of the spheres. The volume of intersection of the two hyperspheres can be calculated by the sum of the volumes of two hyper caps of hyperspheres (Figure 6.1). A simple form of the hyper cap volume was proposed by Li (2011). In the following, one assumes $r_i \leq r_j$ without loss of generality. Let P be an intersection point of the two hyperspheres in (b) and (c) of Figure 6.1. The co-altitude angle $\angle P x_i x_j$ is denoted as ϕ. For the smaller cap ($0 < \phi \leq \pi/2$, marked by *s-cap*$_i$ in (c) of Figure 6.1) trimmed from sphere i, the volume can be calculated by :

$$V_n^{s-cap_i}(r_i, r_j, h) = \frac{1}{2} \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2}) \tag{6.17}$$

where h is the separation vector between the centers of the two spheres and $sin^2\phi$ can be expressed by $1 - (\frac{h^2 + r_i^2 - r_j^2}{2hr_i})^2$. Γ is the gamma function, and $I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2})$ denotes the incomplete beta function for $sin^2\phi$ in bound$(\frac{n+1}{2}, \frac{1}{2})$. The two functions can be easily and rapidly computed.

The volume of the larger cap (denoted by *l-cap*$_i$ in (b) of Figure 6.1) from sphere i is simply the volume of the small hypersphere i minus the volume of the small cap :

$$\begin{aligned} V_n^{l-cap_i}(r_i, r_j, h) &= V_n^{sph_i}(r_i) - V_n^{s-cap_i}(r_i, r_j, h) \\ &= \tfrac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n - \tfrac{1}{2} \tfrac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n I_{sin^2\phi}(\tfrac{n+1}{2}, \tfrac{1}{2}) \\ &= \tfrac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n \left(1 - \tfrac{1}{2} I_{sin^2\phi}(\tfrac{n+1}{2}, \tfrac{1}{2})\right) \end{aligned} \tag{6.18}$$

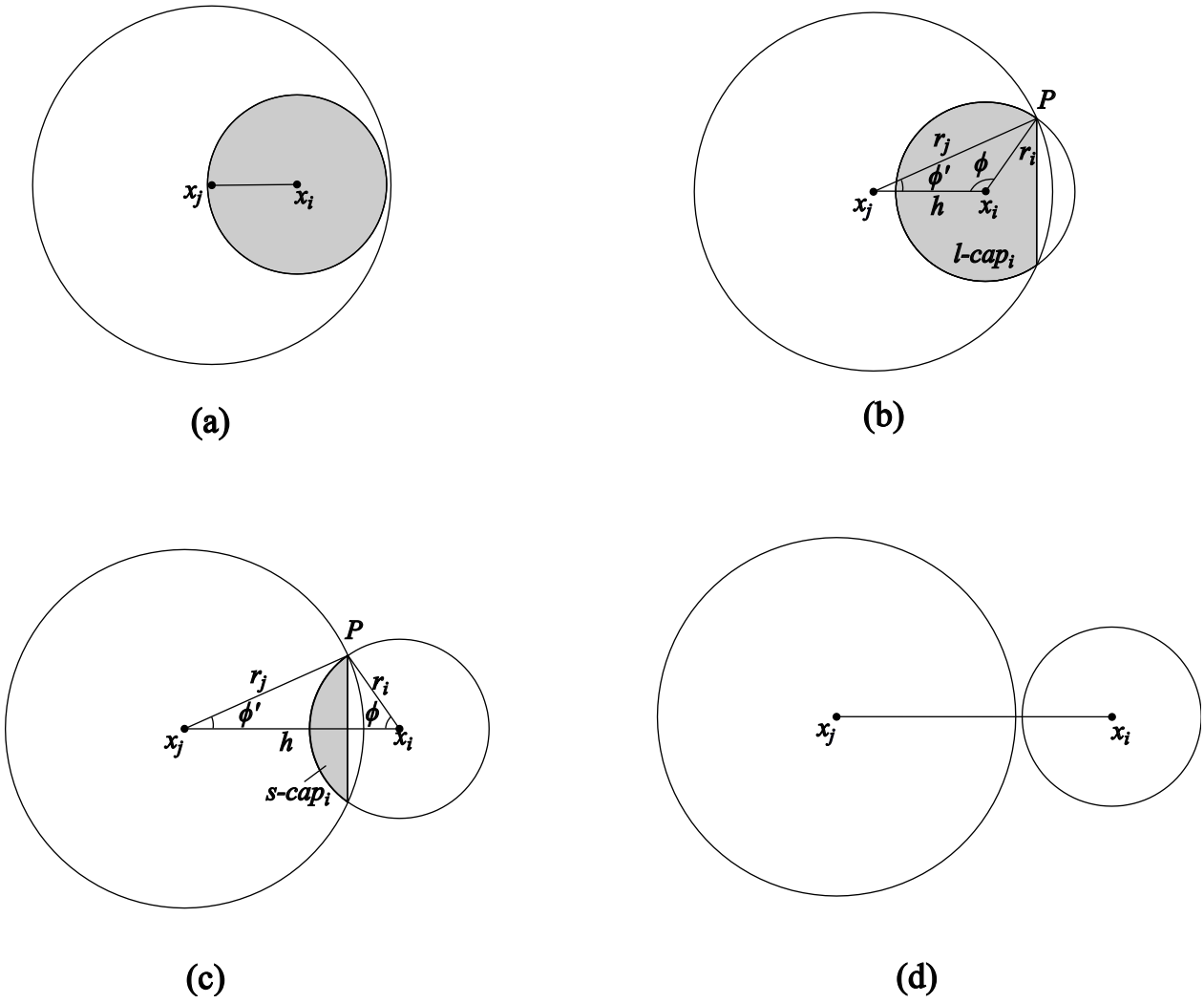

Figure 6.1 General view of circle-circle intersection

Therefore, considering that the volumes of both caps are calculated with different formulas, the intersection of the two hyperspheres show four different domains of definition (as illustrated with the circle-circle intersection in Figure 6.1) : (a) the larger sphere encloses entirely the smaller sphere ; (b) the larger sphere encloses the larger cap of the smaller sphere ; (c) the larger sphere encloses the smaller cap of the smaller sphere ; and (d) no intersection. The volume of intersection of spheres $V_n(r_i, r_j, h)$, is described by a piecewise function.

$$V_n(r_i, r_j, h) = \begin{cases} \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n, & 0 \leq h < r_j - r_i \\ \frac{1}{2}\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_j^n I_{sin^2\phi'}(\frac{n+1}{2}, \frac{1}{2}) + & \\ \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n \left(1 - \frac{1}{2} I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2})\right), & r_j - r_i \leq h < \sqrt{r_j^2 - r_i^2} \\ \frac{1}{2}\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_j^n I_{sin^2\phi'}(\frac{n+1}{2}, \frac{1}{2}) + & \\ \frac{1}{2}\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r_i^n I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2}), & \sqrt{r_j^2 - r_i^2} \leq h < r_j + r_i \\ 0, & h \geq r_j + r_i \end{cases} \tag{6.19}$$

where r_i and r_j are respectively the radius of the smaller and larger hyperspheres, $sin^2\phi$ and $sin^2\phi'$can be replaced by $1 - (\frac{h^2+r_i^2-r_j^2}{2hr_i})^2$ and $1 - (\frac{h^2+r_j^2-r_i^2}{2hr_j})^2$ respectively. In 3D case, the volume of intersection of two spheres simplifies to :

$$V_3(r_i, r_j, h) = \begin{cases} \frac{4\pi}{3} r_i^3, & 0 \leq h < r_j - r_i \\ \frac{\pi}{12h}(r_i + r_j - h)^2 & \\ (h^2 + 2hr_i - 3r_i^2 + 2hr_j + 6r_i r_j - 3r_j^2), & r_j - r_i \leq h < r_j + r_i \\ 0, & h \geq r_j + r_i \end{cases} \tag{6.20}$$

The correlation function of x_i and x_j in $\Re^n$ is obtained by normalizing the covariance function by the standard deviation functions at x_i and x_j. The standard deviation function $\sigma(x_i)$ actually is the square root of the volume of the sphere centered at x_i. Thus after normalizing by $\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}(r_i r_j)^{\frac{n}{2}}$, a non-stationary correlation function is obtained :

$$R^{NS}(x_i, x_j) = \begin{cases} (\frac{r_i}{r_j})^{\frac{n}{2}}, & 0 \leq h < r_j - r_i \\ \frac{1}{2}(\frac{r_j}{r_i})^{\frac{n}{2}} I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2}) + & \\ (\frac{r_i}{r_j})^{\frac{n}{2}} \left(1 - \frac{1}{2} I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2})\right), & r_j - r_i \leq h < \sqrt{r_j^2 - r_i^2} \\ \frac{1}{2}(\frac{r_j}{r_i})^{\frac{n}{2}} I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2}) + & \\ \frac{1}{2}(\frac{r_i}{r_j})^{\frac{n}{2}} I_{sin^2\phi}(\frac{n+1}{2}, \frac{1}{2}), & \sqrt{r_j^2 - r_i^2} \leq h < r_j + r_i \\ 0, & h \geq r_j + r_i \end{cases} \tag{6.21}$$

Note that the above function is not equal to one when $h = 0$. This is because the smaller

sphere represents only a part of the larger sphere, hence is different and the correlation is not 1. This could correspond for example to two points spatially close but on either sides of a fault. On the contrary, when the transition between the ranges is gradual the non-stationary correlation function goes to 1 as h goes to 0.

The variance can also vary in space. This defines the following non-stationary covariance model with compact support :

$$C^{NS}(x_i, x_j) = \sigma(x_i)\sigma(x_j)R^{NS}(x_i, x_j) \tag{6.22}$$

6.4.3 Other non-stationary isotropic models of the spherical family model

In a similar way as for the isotropic spherical model, the non-stationary covariance function C^{NS} can be computed by evaluating $C^{NS}(x_i, x_j) = \int w_i(x_i - u)w_j(x_j - u)\mathrm{d}u$ where the weight functions are given by Eq. 6.14. These functions were evaluated by applying the convolution theorem that states that the convolution of two functions in the spatial domain is a simple product in the spectral domain. Hence, the Fourier transform of each weight function were multiplied and the inverse Fourier transform was obtained. The resulting rather lengthy, but simple, expressions are presented in the appendix.

6.4.4 Computation of non-stationary anisotropic spherical family covariance models

In the anisotropic case, the dilution (or weight) function is defined over an hyper-ellipsoid. As this ellipsoid change orientations in space, there is no simple formulas to compute the volume of intersection, even for the spherical model (indicator dilution function). However, it can be computed numerically. First, a fine 3D grid is selected in the original coordinate system. Then, for each point x_i in turn, the local grid including the ellipsoid found at x_i is rotated according to the direction of anisotropy and scaled so that the ellipsoid becomes a unit sphere. The weight function $w(\cdot)$ is then obtained from the isotropic closed-form expression given in Eq. 6.14 evaluated at the points of the local transformed grid but reported at the points of the original grid, giving $w(x_i - x)$. Finally, the covariance is obtained numerically as $\widetilde{C}^{NS}(x_i, x_j) = \sum_{x \in grid} w_i(x_i - x)w_j(x_j - x)\Delta v$ where Δv is the volume of one discretizing grid cell.

6.5 Examples

6.5.1 Correlation

The correlation is a function of the distance between the points and their correlation ranges. The correlation between two points $R^{NS}(x_i, x_j)$ with ranges $a_i = 1$ and $a_j \leq 1$ is shown for the spherical case in Figure 6.2. When $a_i \neq a_j$, the non-stationary correlation function has compact support $(a_i + a_j)/2$. As the distance between two points decreases, the correlation increases. When the distance $h_{ij} < (a_i - a_j)/2$, the volume of intersection of the two spheres is the volume of the small sphere. This volume of intersection is proportional to the covariance between the two variables associated with points x_i and x_j. As the variance of each variable is proportional to the volume of each sphere, the covariance can be normalized at $\sqrt{a_j^3/a_i^3}$ to obtain a NS correlation function.

Figure 6.3 gives the correlation between x_i and x_j where the correlation range at x_j is obtained by linear interpolation between range 1 at x_i and range 0.5 for $h_{ij} = 1$. Spherical, cubic and penta-spherical models are represented. The correlation curve changes smoothly as a function of the location considered. For comparison, the stationary correlation functions with intermediate range 0.75 are also illustrated.

6.5.2 Unconditional simulation

As indicated in section 6.4.1, the non-stationary random function $Z(x)$ can be obtained by weighted average as $Z(x_i) = \int w_i(x_i - u)Y(x_i - u)\mathrm{d}u$ where $Y(x_i - u), i = 1, \cdots, n$ are independent standard Gaussian variables. The weight functions are given in Eq. 6.14 after suitable rotation and scaling as described in section 6.4.4.

Figure 6.4 shows 2D realizations from non-stationary isotropic models of the spherical family by this approach. The simulated fields of cubic model and penta-spherical model are smoother than spherical model since they have 1 time and 2 times mean-square differentiability. Figure 6.5 compares the theoretical variogram to the experimental ones for 100 realizations on grids 1000*200 from a spherical model where range changes from 30 on top to 5 on bottom. The average variogram fits well the theoretical curve.

Non-stationary anisotropic realizations of the spherical, cubic and penta-spherical models are shown in Figure 6.6. In this figure, the range is increasing from 5 to 50 going from the circle center to the border and the direction of anisotropy is radial.

Simulation times obtained for the non-stationary simulation are reported in Table 6.2. The computation times increase linearly with the number of points to simulate and are propor-

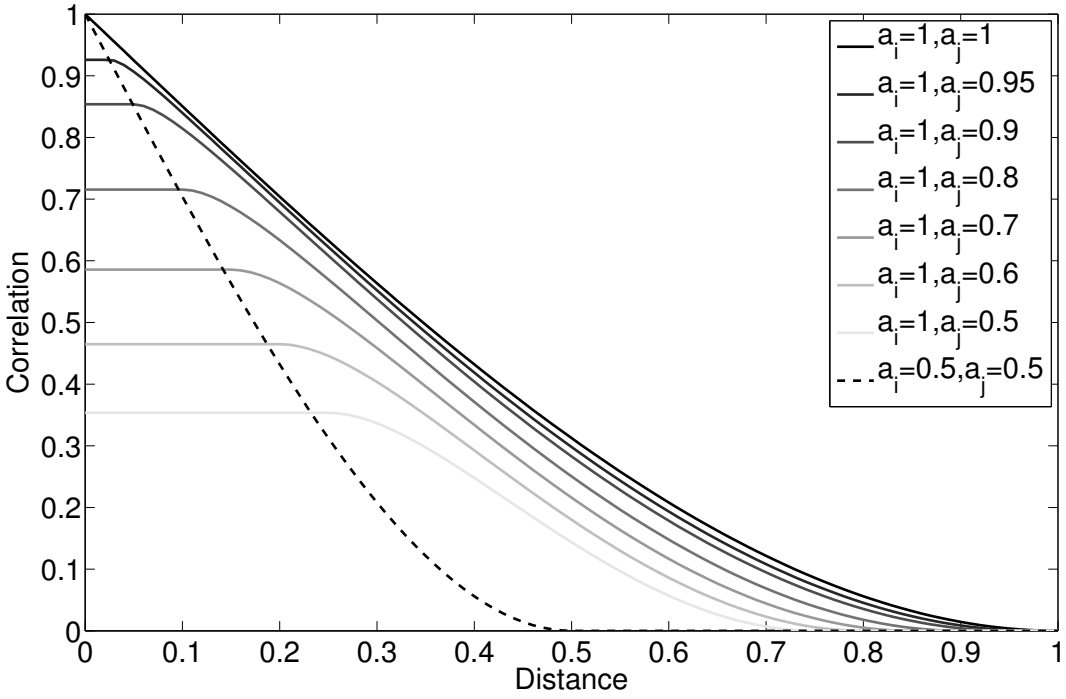

Figure 6.2 Spherical NS correlation between points x_i (with range a_i) and point x_j (with range a_j) as a function of distance between the points ($||x_i - x_j||$). For $||x_i - x_j|| < (a_i - a_j)/2$ (with $a_i \geq a_j$), the correlation is constant at $\sqrt{a_j^3/a_i^3}$.

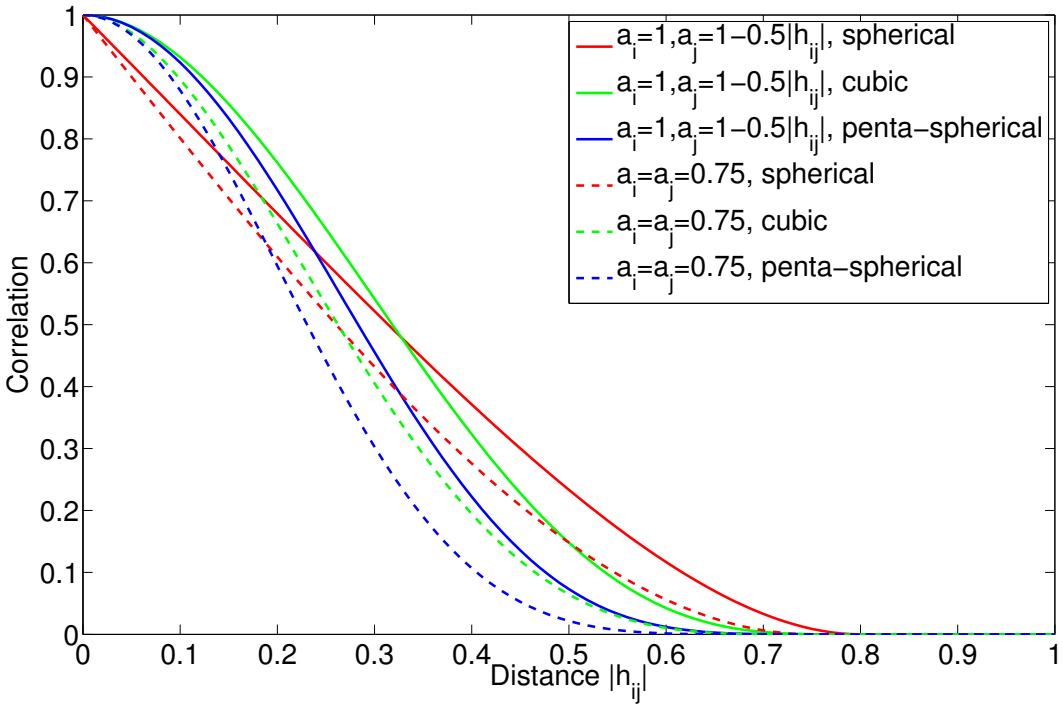

Figure 6.3 NS correlation ((solid lines) and stationary correlation (dashed lines) between points x_i (where range $a_i = 1$) and point x_j (where range $a_j = 1 - 0.5|h_{ij}|$) as a function of distance between the points ($h_{ij} = ||x_i - x_j||$) for the spherical, the cubic and the penta-spherical models. For comparison, the stationary correlation functions with an intermediate range of 0.75 are also illustrated.

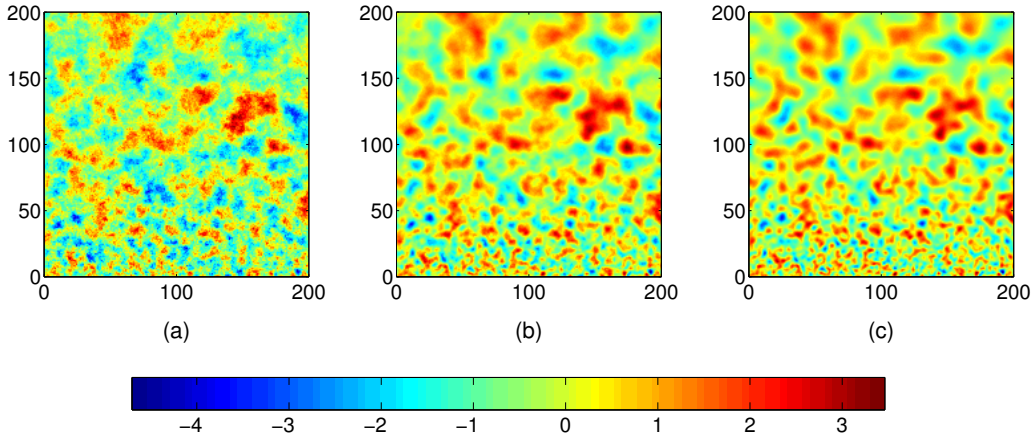

Figure 6.4 Non-stationary isotropic realizations by (a) spherical model, (b) cubic model and (c) penta-spherical model. In all simulations, range of model is changing from 30 on top to 5 on bottom.

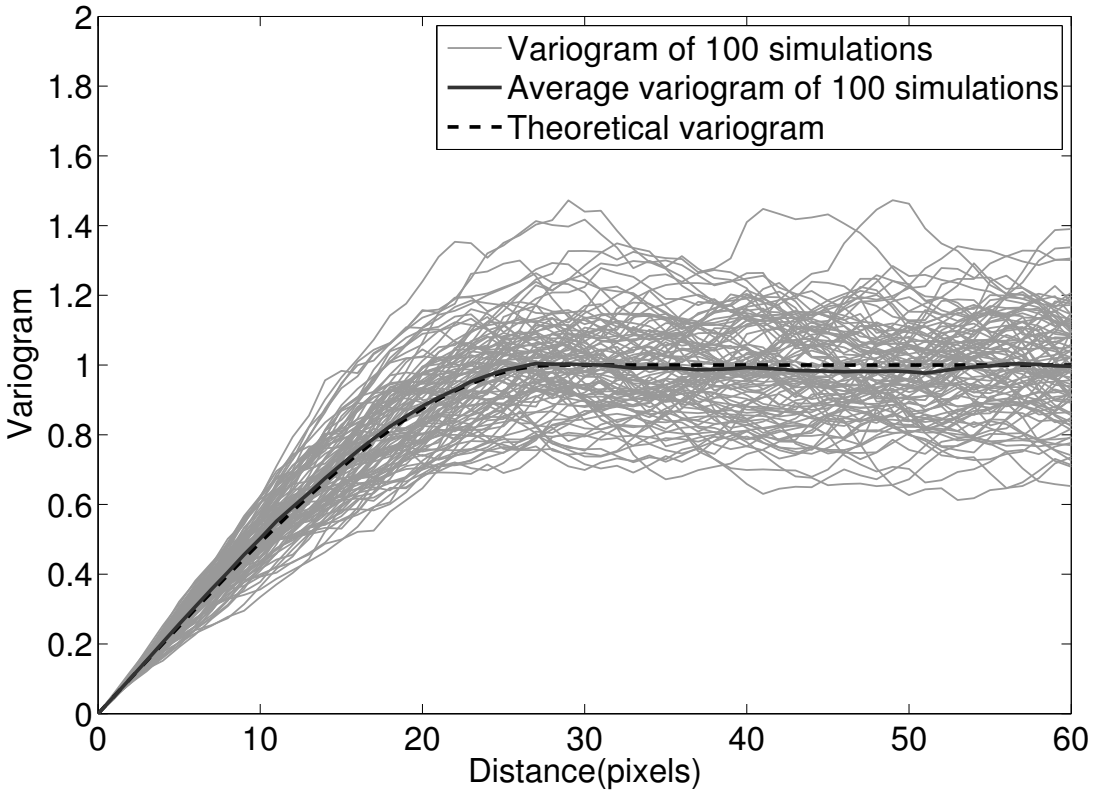

Figure 6.5 Comparison of simulated variogram and the theoretical. For each realization, the variogram is computed along the vertical between the pixels on row $y = 200$ (where $a = 30$) and the pixels on row $y = 200 - distance$ (where $a = 30 - \frac{distance}{200} \times (30 - 5)$).

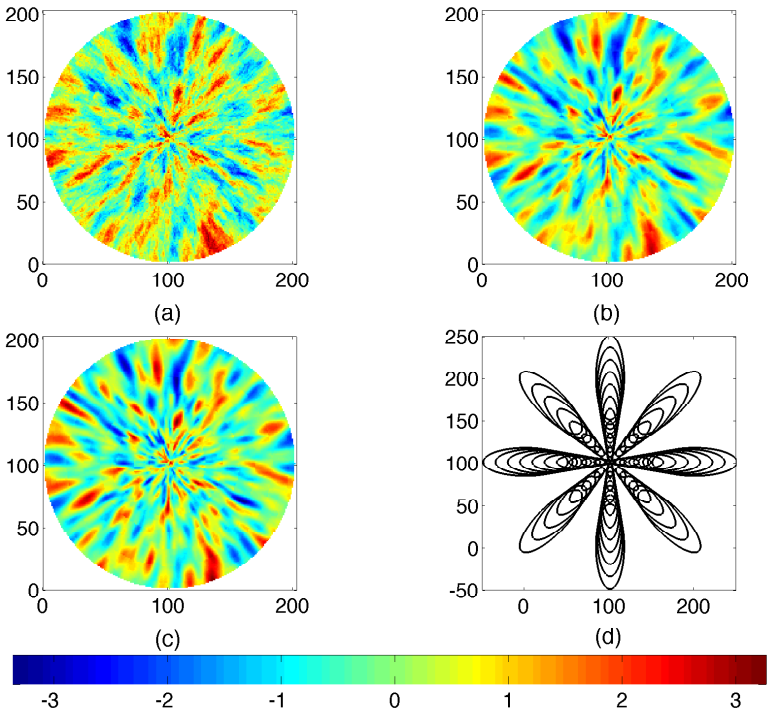

Figure 6.6 Non-stationary anisotropic realizations by (a) spherical model, (b) cubic model and (c) penta-spherical model. In all simulations, the azimuth and range of a point are functions of location. The direction of the main continuity follows the azimuth and the range increases linearly from 5 pixels at the center to 50 pixels at the circumference. The range in the tangential direction is set to 1/3 the range in the radial direction. (d) shows the local structure by illustrating the local support of the weight function.

tional to $(a/dx)^3$ where a is the range and dx is the discretization mesh used. Therefore, the proposed method of simulation would not be efficient for fields with large range and fine mesh. In that case, an alternative could be to envisage a non-stationary sequential Gaussian (NS-SGS) method using covariances evaluated numerically. However, the numerical evaluation of all NS-covariances could be computationally as challenging as the approach described here unless closed-form expressions are available. Therefore, it is advantageous, in presence of large ranges, to resort to combination of isotropic NS spherical, cubic, and penta-spherical together with isotropic or anisotropic Matérn or Gaussian NS covariance models for which closed-form expressions exist. Note finally that the moving average simulation method leads to easy parallelization as the weight functions and the moving average can be computed simultaneously and independently at each point of the simulation.

6.5.3 Conditioning the realizations

The conditioning of the realizations to the observed data is done by NS-kriging in a similar way as for the stationary case (Chilès and Delfiner, 1999). The only modification to bring is to replace the stationary covariance by the NS-covariance in the kriging system. Figure 6.7 shows an example of three conditional realizations obtained in 1D for a cubic model with range abruptly changing at $x = 50$.

6.5.4 Sparse covariance matrix

Compactly supported covariance function creates sparse covariance matrix. The sparsity depends of the ratio of the average range to the field size. Taking an example of 10 thousands points on a mesh $100 * 100$, the covariance matrix has size $10^4 * 10^4$. Table 6.3 gives the covariance matrix sparsity, and memory consumption ratio for three stationary ranges. Clearly, substantial gains in memory can be achieved when sparsity is important, i.e. when range is small compared to the field size.

Table 6.2 Computation time (seconds) of isotropic simulations by spherical model

Range \ Size of simulation		100 by 100	200 by 200	400 by 400	800 by 800
NS	a changes from 5 to 20	6	23	89	350
isotropic	a changes from 5 to 40	19	72	287	1126
simulation	a changes from 5 to 80	160	646	2535	10240

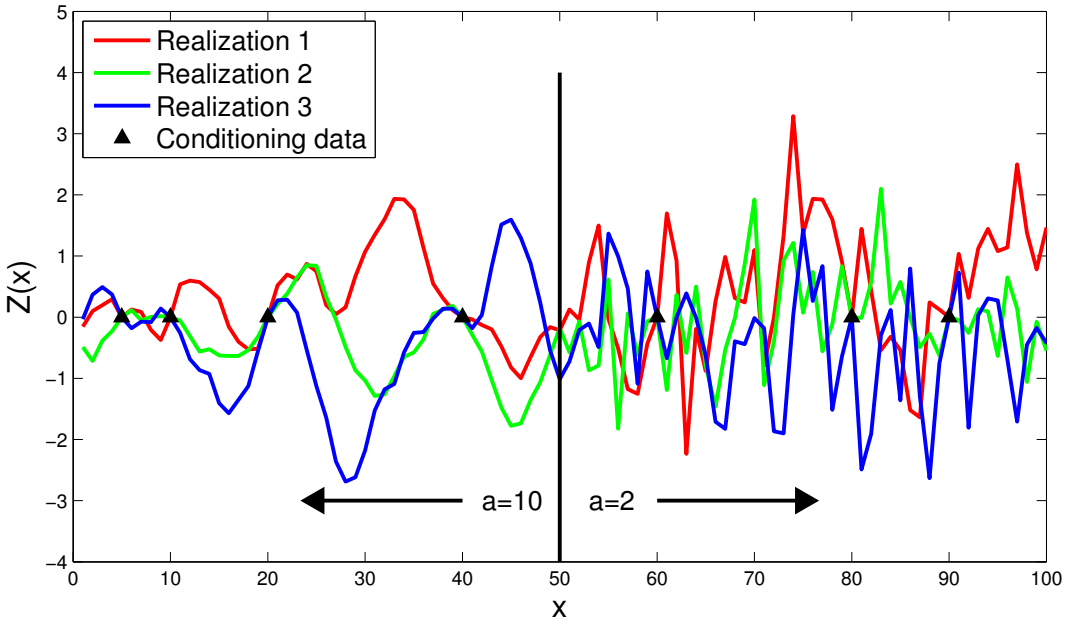

Figure 6.7 Three 1D conditional realizations for cubic model with range 10 for $x < 50$ and range 2 for $x > 50$.

Table 6.3 Sparsity and memory consumption for a simulated field of size 100×100

Range	Sparsity (%)	Memory consumption ratio full/sparse
10	97.2	17.8
20	89.6	4.8
30	78.6	2.3

6.6 Case study

The data on overburden thickness in Montérégie Est in the south of Québec is used as a simple application of non-stationary spherical covariance functions. There are 831 observations in the study area which possesses two distinct geological domains, as illustrated in Fig. 6.8. The spatial distributions of overburden in these two domains differ, a larger range being observed in the western domain. The results of kriging overburden thickness with the stationary and the non-stationary models are compared.

In the stationary case, the global covariance model is obtained from all the thickness data in the studied area by leave-one-out cross validation (Marcotte, 1995). An isotropic spherical function provides the optimal mean square error with parameters nugget $10m^2$, total sill $33m^2$ and correlation range $6km$. Then, the procedure is repeated to estimate separately the spatial structure in both geological domains. In the western domain, the isotropic spherical covariance function has parameter nugget $3.5m^2$, total sill $34.5m^2$ and isotropic correlation range $20km$. In the eastern domain, the covariance function is still isotropic spherical and has nugget $3m^2$, total sill $54m^2$ but correlation range of only $1km$.

To allow for a gradual transition between the two domains, a transition area is defined around their contact. The three parameters sill, range and ratio nugget/total sill are assumed to vary smoothly and continuously within the transition zone, from the values on the western zone to the values in the eastern zone. The variation is parametrized with the following function for a parameter p :

$$\begin{aligned} &p(x) = p_w, \ \forall x \in \text{Western part} \\ &p(x) = p_e, \ \forall x \in \text{Eastern part} \\ &p(x) = 1/2 \left(p_w + p_e + \cos(\tfrac{d_e \pi}{width})(p_e - p_w)\right) \end{aligned} \tag{6.23}$$

where p_w and p_e are parameter in the western and eastern domain, d_e is the distance from the eastern edge of the transition zone to the point and $width$ is the width of the transition zone. Fig. 6.9 shows an example of range evolution. The width of transition zone centered at boundary is considered as a parameter of the non-stationary model. Taking an example where the transition zone has width $3km$, Figure 6.10 shows the covariance contours at a few selected points.

Cross-validation results are compared for different widths of the transition zone. The results obtained are also compared to a global stationary kriging ignoring the geological domains and one separate stationary kriging for each domain. Cross-validation statistics are computed for the whole area and within a subarea around the geological contact, see Table 6.4.

The non-stationary kriging does not show significant improvement for the mean absolute

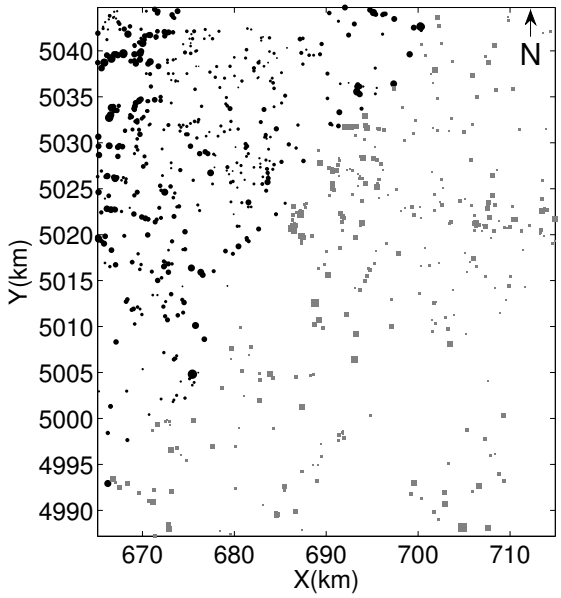

Figure 6.8 Map of study area and sample data. Size of symbols is proportional to thickness value. Black and gray symbols represent sample data in different geological domain.

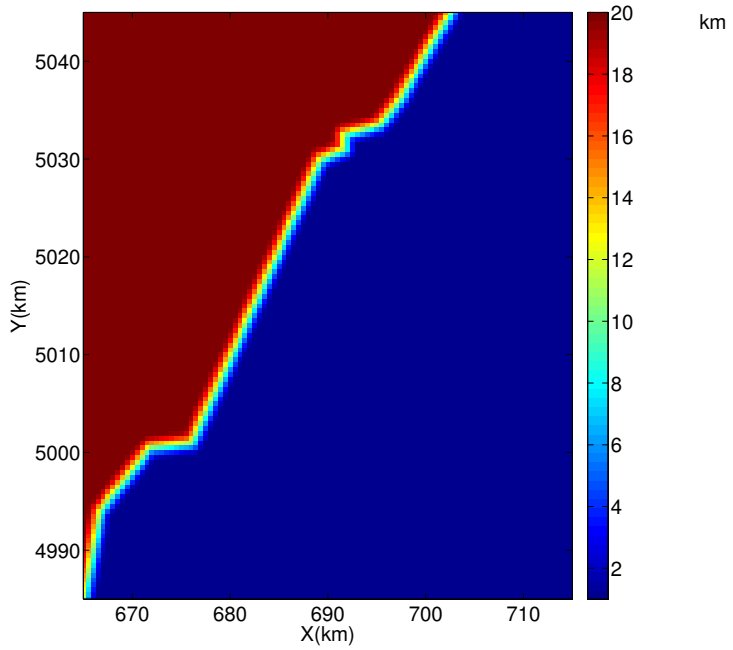

Figure 6.9 Map of correlation range evolution.

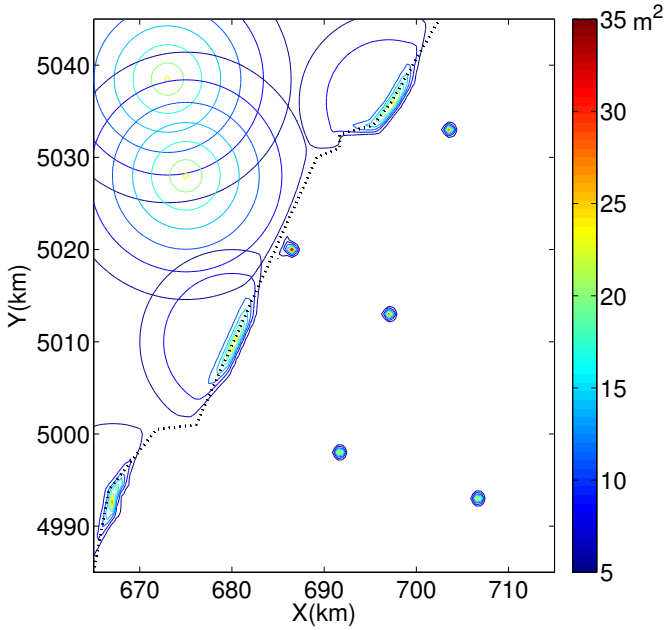

Figure 6.10 Covariance contours on 10 points in the study area. The black dashed line represents the boundary of the two geological domains.

error (MAE) over the whole area. The best reduction in MAE over the whole area of the non-stationary kriging with respect to global stationary kriging and local stationary-kriging are about 1.3% and 0.4% respectively for a width of approximately 2 km. However, on the subarea centered at the geological contact, the stationary global model shows the largest MAE. The non-stationary kriging has worse estimation than stationary kriging with local models when the transition area is taken wider than $4km$ but is better when a narrower transition area is assumed. The best reduction in MAE reaches 13 % and 5.9 % for the subarea around the contact compared to global and local stationary krigings. Similar observations can be made for the root mean square error (RMSE).

When the width of transition area reduces to 0, statistics of non-stationary kriging are close to the stationary case with local models. They are not identical however as in one case (local stationary models) only the data in the same domain as the estimation point are taken, whereas in the non-stationary case, data from both domains are used.

Figure 6.11 shows the maps obtained using a global stationary model (a), local stationary models (b) and a non-stationary model with transition zones of total width 3 km (c). Parts surround by white rectangles are enlarged in Figure 6.12. The map of the global stationary model is smooth, but differences of spatial structures between the two geological domains

Table 6.4 Statistics of estimates by stationary and non-stationary kriging

		On whole region		On an area centered at boundary with width $3km$ including 70 data	
		MAE	RMSE	MAE	RMSE
Stationary kriging with global model		4.71	6.58	6.39	7.96
Stationary kriging with local models		4.67	6.64	5.93	7.97
Non-stationary kriging with transition area width d	$d = 10km$	4.71	6.67	6.22	7.98
	$d = 8km$	4.72	6.66	6.14	7.87
	$d = 6km$	4.70	6.64	6.06	7.73
	$d = 4km$	4.67	6.61	5.85	7.48
	$d = 3km$	4.65	6.58	5.69	7.26
	$d = 2km$	4.66	6.57	5.70	7.17
	$d = 1km$	4.65	6.58	5.58	7.32
	$d = 0km$	4.68	6.64	5.87	7.89

are not visible. The map from stationary kriging with a local model per geological domain shows local feature in each domain. However, the boundary shows a sharp change that is not expected for a variable like the overburden thickness. The map from the non-stationary model keeps the geological domain characteristics and improves the smoothness at the boundary, as seen in Figure 6.12. Figure 6.13 shows the kriging standard deviation of global, local stationary and non-stationary models. The estimated variance in the eastern part is higher than in the western part due to the shorter range and high sill of the local model in this part. The standard deviation in (c) shows a transitional zone at the boundary of the two geological domains.

6.7 Conclusion and discussion

In this article, the family of non-stationary covariance functions with compact support suggested by Chilès and Delfiner (1999) was studied. Closed-form expressions for the non-stationary isotropic spherical, cubic and penta-spherical models were derived. For the anisotropic case, numerical approximation based on the known weighting functions are used. The models so defined (and their numerical approximation) were shown to be admissible. The non-stationary compact support models constitute a valuable addition to the set of existing unbounded non-stationary models (e.g. exponential, Gaussian and Matérn class) for which closed-form expressions are available for both the isotropic and anisotropic cases. The new NS covariance functions with compact support provide a wide class of functions that can be used to taper the unbounded NS covariance models for which anisotropic closed-form expressions exist.

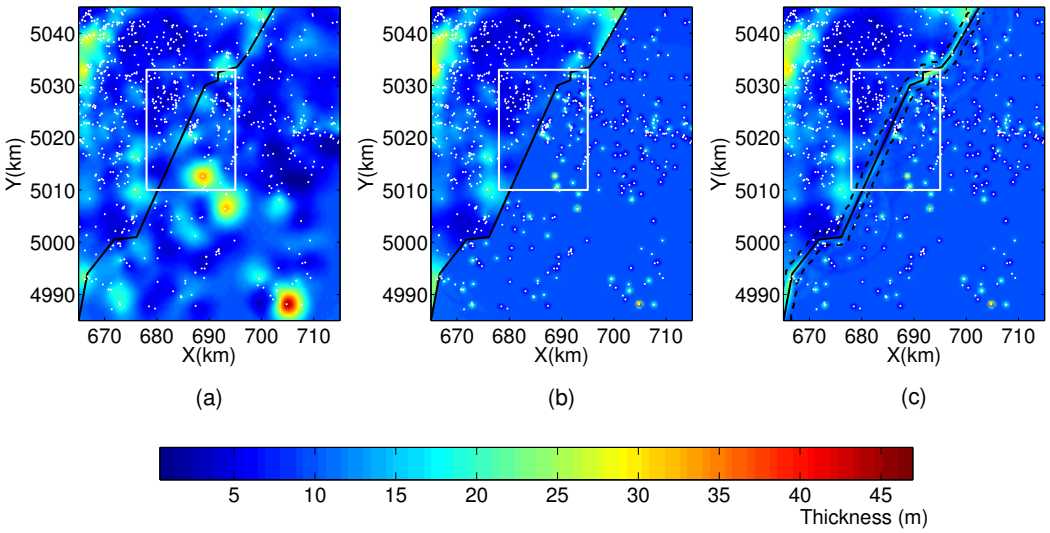

Figure 6.11 Estimation maps at the whole area by (a) stationary kriging with global model, (b) stationary kriging with local model in each domain and (c) non-stationary kriging with a total 3 km wide transition area (between dashed lines), centered at the contact between the geological domains, where the model parameters change continuously. The white points represent data locations. The black line indicates the contact of two geological domains.

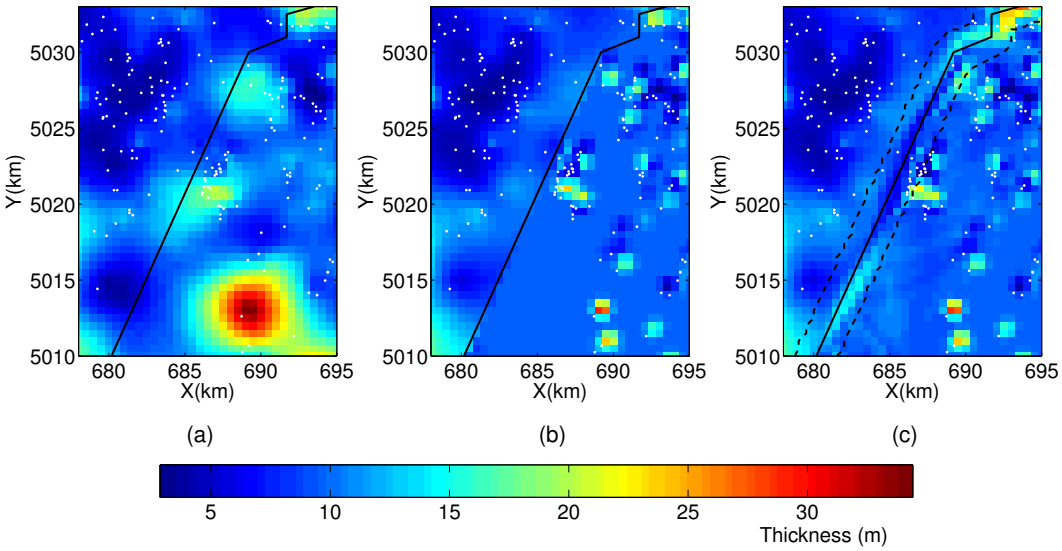

Figure 6.12 Zoomed maps (area outlined in Fig. 6.11) by (a) stationary kriging with global model, (b) stationary kriging with local model in each domain and (c) non-stationary kriging with a transition area in which the model parameters change continuously.

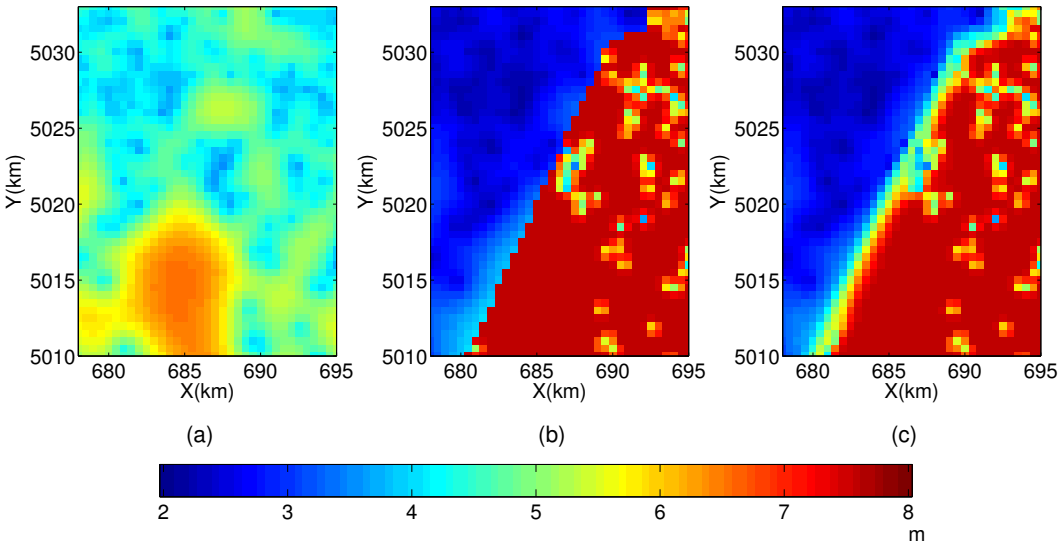

Figure 6.13 Kriging standard deviation by (a) stationary kriging with global model, (b) stationary kriging with local model in each domain and (c) non-stationary kriging with a transition area in which the model parameters change continuously.

This would allow to generate anisotropic NS tapered covariance.

The spherical family of covariance functions is formed of polynomial functions. The differentiability of the function is controlled by the dimension n of hyperspheres and the order of the Radon transform. Allowing the diameter of the hyperspheres to vary in space enables to define a non-stationary covariance function in R^n. After applying the Radon transform of order $2q$, the non-stationary covariance function is admissible in $\Re^{n-2q}$ and is $2q$ times differentiable. One advantage of the models with compact support is the reduction in memory requirement due to the sparsity of the resulting covariance matrix.

It was not possible to obtain the closed-form expressions of the non-stationary covariances in the anisotropic case. However, all the covariances of the spherical family (including cubic and penta-spherical models) can be evaluated numerically easily as the associated weighting functions are known. The same functions also allow easy simulation of the non-stationary fields by moving average over an independent zero-mean, unit-variance Gaussian field. Alternatively, the numerical values of the covariance can be computed and used within SGS or Cholesky simulation methods by simply replacing the stationary covariances by non-stationary ones.

Admittedly, the choice of parametrization of the spatial variation of the NS covariance parameters (the supra-model) and the estimation of its parameters remain key elements controlling the quality and success of any NS modeling approach. This remains an open area of research. In the application example, the choice of imposing a transition zone was dictated by geological considerations. The form of the supra-model was chosen so as to ensure continuity with the geological domains at the fringe of the transition zone and the parameters were chosen by cross-validation. These choices led to improvements in RMSE over the transition zone of 8.8% compared to either a single zone stationary kriging or to separate stationary kriging within each geological domain.

6.8 Acknowledgment

This research was financed by FQRNT and NSERC research grants of the second author. The constructive comments of two anonymous reviewers helped improve significantly the manuscript.

6.9 Appendix - Closed-form expressions for the cubic and the penta-spherical models

The closed-form expressions for the 3D non-stationary isotropic cubic covariance functions with supports a_1 and a_2 and with $a_1 > a_2$ are given, after normalization to a unit sill, by :

$$C(a_1,a_2,r) = \begin{cases} \frac{28r^2+3a_2^2-7a_1^2}{3a_2^2-7a_1^2} & 0 \le r \le \frac{(a_1-a_2)}{2} \\ -\frac{(a_1+a_2-2r)^4}{256a_2^5\left(-7a_1^2+3a_2^2\right)r}[35a_1^4 - 4a_1^3(35a_2+26r) & \\ +2a_1^2\left(105a_2^2+68a_2r+4r^2\right) & \\ +(a_2-2r)^2\left(35a_2^2+36a_2r+12r^2\right) & \\ +4a_1\left(-35a_2^3+34a_2^2r+60a_2r^2+24r^3\right)] & \frac{(a_1-a_2)}{2} \le r \le \frac{(a_1+a_2)}{2} \\ 0 & \frac{(a_1+a_2)}{2} \le r \end{cases} \tag{6.24}$$

Similarly, the closed-form expressions for the penta-spherical model are :

$$C(a_1,a_2,r) = \begin{cases} \frac{99a_1^4+15a_2^4+440a_2^2r^2+1584r^4-66a_1^2\left(a_2^2+12r^2\right)}{99a_1^4-66a_1^2a_2^2+15a_2^4} & 0 \le r \le \frac{(a_1-a_2)}{2} \\ -\frac{(a_1+a_2-2r)^6}{1536a_2^7\left(33a_1^4-22a_1^2a_2^2+5a_2^4\right)r}[231(a_1-a_2)^6 - 12(a_1+a_2) & \\ (89a_1^4-392a_1^3a_2+622a_1^2a_2^2-392a_1a_2^3+89a_2^4)r & \\ +36\left(29a_1^4+28a_1^3a_2-178a_1^2a_2^2+28a_1a_2^3+29a_2^4\right)r^2 & \\ +32(a_1+a_2)\left(29a_1^2-140a_1a_2+29a_2^2\right)r^3 & \\ -48\left(13a_1^2+70a_1a_2+13a_2^2\right)r^4 & \\ -960(a_1+a_2)r^5 - 320r^6] & \frac{(a_1-a_2)}{2} \le r \le \frac{(a_1+a_2)}{2} \\ 0 & \frac{(a_1+a_2)}{2} \le r \end{cases} \tag{6.25}$$

Figure 6.14 illustrates the resulting covariance functions for the case a_1=1 and $a_2 = 0.5$, 0.7 and 0.9 for the cubic and the penta-spherical model.

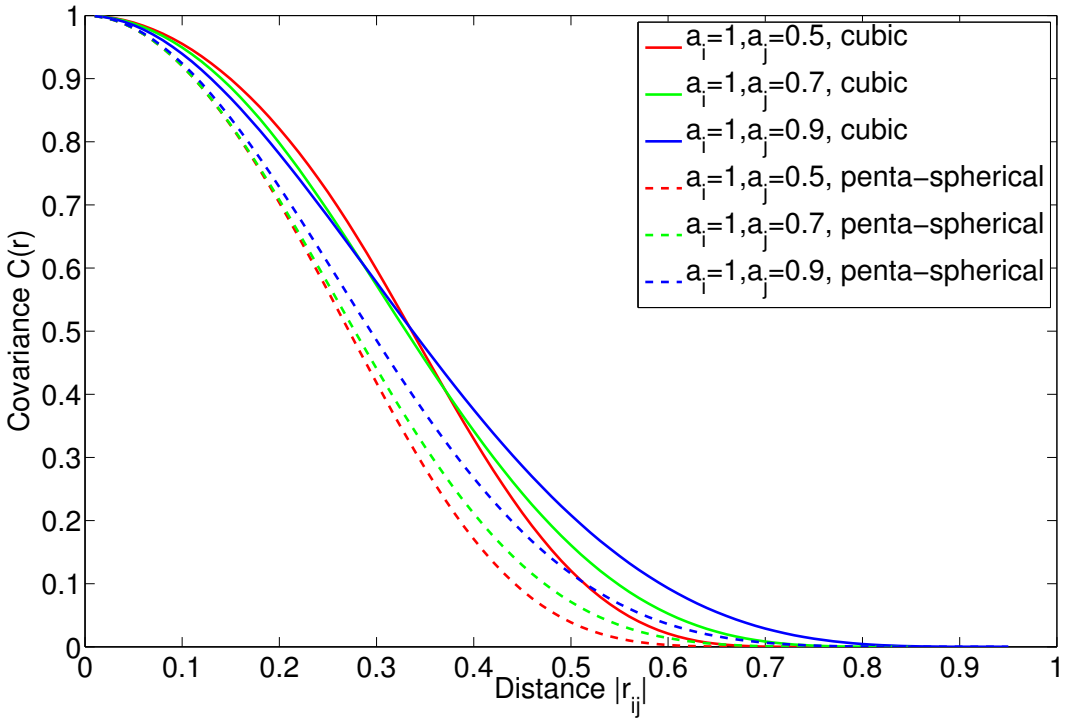

Figure 6.14 NS isotropic covariance for the cubic and the penta-spherical model for various supports.

References

ASKEY, R. (1973). Radial characteristic functions. Technical report, University of Wisconsin.

BELLIER, E. and MONESTIEZ, P. (2010). A spatial covariance model with a single wave effect and a finite range. *Statistics and Probability Letters*, 80, 1343–1347.

BOISVERT, J. and DEUTSCH, C. V. (2008). Shortest anisotropic path to reproduce complex geological features. *8th International Geostatistics Congress, Santiago, Chile*. vol. 27, 1041–1046.

BUHMANN, M. D. (2003). *Radial basis functions : theory and implementations*, vol. 12. Cambridge university press.

CHILÈS, J. P. and DELFINER, P. (1999). *Geostatistics : Modeling Spatial Uncertainty*. John Wiley & Sons.

CHRISTAKOS, G. and PAPANICOLAOU, V. (2000). Norm-dependent covariance permissibility of weakly homogeneous spatial random fields. *Stoch Environ Res Risk Assess*, 14, 1–8.

DAMIAN, D., SAMPSON, P. D. and GUTTORP, P. (2001). Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics*, 12, 161–178.

FURRER, R., GENTON, M. G. and NYCHKA, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15, 502–523.

GNEITING, T. (2002). Compactly supported correlation functions. *Journal of Multivariate Analysis*, 83, 493–508.

GUTTORP, P., MEIRING, W. and SAMPSON, P. D. (1994). A space-time analysis of ground-level ozone data. *Environmetrics*, 5, 241–254.

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

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

HAAS, T. C. (1995). Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *Journal of the American Statistical Association*, 90, 1189–1199.

HIGDON, D., SWALL, J. and KERN, J. (1999). Non-stationary spatial modeling. *Bayesian statistics*, 6, 761–768.

HORTA, A., CAEIRO, M. H., NUNES, R. and SOARES, A. (2010). Simulation of continuous variables at meander structures : Application to contaminated sediments of a lagoon. P. M. Atkinson and C. D. Lloyd, editors, *geoENV VII - Geostatistics for Environmental Applications*, Springer Netherlands, no. 16 Quantitative Geology and Geostatistics. 161–172.

JUN, M. and STEIN, M. L. (2008). Nonstationary covariance models for global data. *The Annals of Applied Statistics*, 2, 1271–1289.

KLEIBER, W. and PORCU, E. (2015). Nonstationary matrix covariances : compact support, long range dependence and quasi-arithmetic constructions. *Stochastic Environmental Research and Risk Assessment*, 29, 193–204.

LI, S. (2011). Concise formulas for the area and volume of a hyperspherical cap. *Asian Journal of Mathematics and Probability Letters Statistics*, 4, 66–70.

LIANG, M., MARCOTTE, D. and BENOIT, N. (2014). A comparison of approaches to include outcrop information in overburden thickness estimation. *Stochastic Environmental Research and Risk Assessment*, 28, 1733–1741.

MARCOTTE, D. (1995). Generalized cross-validation for covariance model selection. *Mathematical geology*, 27, 659–672.

MATEU, J., FERNÁNDEZ-AVILÉS, G. and MONTERO, J. M. (2013). On a class of nonstationary, compactly supported spatial covariance functions. *Stochastic Environmental Research and Risk Assessment*, 27, 297–309.

MATHERON, G. (1965). *Les variables régionalisées et leur estimation*. PhD Dissertation, Faculté des Sciences, Université de Paris.

MEIRING, W., MONESTIEZ, P., SAMPSON, P. D. and GUTTORP, P. (1997). Developments in the modelling of non stationary spatial covariance structure from space-time monitoring data. Ernest Y. Baafi and Neil A. Schofield, editors, *Geostatistics Wollongong 96*, Springer. 162–173.

MOREAUX, G. (2008). Compactly supported radial covariance functions. *Journal of Geodesy*, 82, 431–443.

PACIOREK, C. J. (2003). *Nonstationary Gaussian processes for regression and spatial modelling*. PhD Dissertation, Department of Statistics, Carnegie Mellon University.

PACIOREK, C. J. and SCHERVISH, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17, 483–506.

PINTORE, A. and HOLMES, C. (2004). Non-stationary covariance functions via spatially adaptive spectra. Technical report, Department of Statistics, University of Oxford.

PORCU, E., DALEY, D. J., BUHMANN, M. and BEVILACQUA, M. (2013). Radial basis functions with compact support for multivariate geostatistics. *Stochastic Environmental Research and Risk Assessment*, 27, 909–922.

RIVEST, M. and MARCOTTE, D. (2012). Kriging groundwater solute concentrations using flow coordinates and nonstationary covariance functions. *Journal of Hydrology*, 472–473, 238–253.

RYGAARD-HJALSTED, C., CONSTABLE, C. G. and PARKER, R. L. (1997). The influence of correlated crustal signals in modelling the main geomagnetic field. *Geophysical Journal International*, 130, 717–726.

SAMPSON, P. D. and GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 87, 108–119.

SANSÒ, F. and SCHUH, W.-D. (1987). Finite covariance functions. *Bulletin Géodésique*, 61, 331–347.

SCHMIDT, A. M. and O'HAGAN, A. (2003). Bayesian inference for non-stationary spatial covariance structure via spatial deformations. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 65, 743–758.

SHAMSIPOUR, P., MARCOTTE, D., CHOUTEAU, M., RIVEST, M. and BOUCHEDDA, A. (2013). 3D stochastic gravity inversion using nonstationary covariances. *Geophysics*, 78, G15–G24.

SMITH, R. (1996). Estimating nonstationary spatial correlations. *Preprint, University of North Carolina*, 76.

SOARES, A. (2010). Geostatistical methods for polluted sites characterization. P. M. Atkinson and C. D. Lloyd, editors, *geoENV VII - Geostatistics for Environmental Applications*, Springer Netherlands, no. 16 Quantitative Geology and Geostatistics. 187–198.

STEIN, M. L. (2005). Nonstationary spatial covariance functions. *Technical report, University of Chicago, Center for IntegratingStatistical and Environmental Science.*

WENDLAND, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in computational Mathematics*, 4, 389–396.

WU, Z. (1995). Compactly supported positive definite radial functions. *Advances in Computational Mathematics*, 4, 283–292.

CHAPTER 7 ARTICLE 3 : SIMULATION OF NON-LINEAR COREGIONALIZATION MODELS BY FFTMA

Article history : Submitted to Computers & Geosciences on 6 June 2015.

Authors : Min Liang, Denis Marcotte and Pejman Shamsipour

7.1 Abstract

A fast and efficient method to simulate multivariate fields with non-linear models of coregionalization (N-LMC) is described. The method generalizes FFTMA to the multivariate simulation of the N-LMC with symmetric cross-covariances, hence the name GFFTMA. It allows for example to use an exponential model as the direct covariance for the main variable, a Cauchy model for the secondary variable and a K-Bessel model for the cross-covariance. Each covariance and cross-covariance are Fast Fourier Transformed (FFT) to get the discrete spectral densities. Then the spectral matrix is eigen-decomposed at each frequency separately to provide the square root matrix and to enforce positive-definiteness in cases where small negative eigenvalues are found. Finally the simulated spectrum is obtained as multiplication of the root matrix and the white noise coefficients. The method is particularly fast for covariances having derivatives at the origin and/or for covariances with long range. Hence, two-variable 2D fields of 100 million pixels with all-Gaussian or all-cubic covariances and cross-covariance are both simulated in less than 200 s. The CPU-time increases only as $Nlog(N)$ (N, the number of points to simulate). Additional realizations are obtained at a low marginal cost as the eigen-decomposition step needs to be done only once for the first realization. The main limitation of the approach is its rather stringent memory requirement. Synthetic examples illustrate the simulations of N-LMC with two and three variables for different combinations of seven available models. It shows that the theoretical models are all well reproduced. An illustrative case-study on overburden thickness simulation is provided where the secondary information consists of a latent Gaussian variable identifying the geological domain.

7.2 Introduction

In geology, mining, petroleum, hydrogeology and other applications, it is common to observe several secondary variables that are spatially correlated to the main variable of interest. Often, the secondary variables are more exhaustively sampled than the main variable. To improve precision, it is important to include this information in the simulation of the main variable. Up to now, essentially two cases of figure are encountered in applications. The first one is when a physical relationship exists between the main variable and the secondary ones. This situation is common in geophysics (Asli *et al.*, 2000; Shamsipour *et al.*, 2010, 2011) and in hydrogeology (Ahmed and de Marsily, 1989; Dong, 1990). The second case of figure is the modeling of the statistical link by the linear model of coregionalization (LMC) (Myers, 1983; Marcotte, 1991; Journel and Huijbregts, 1978; Wackernagel, 2003) where it is assumed that all variables being studied share the same spatial structures. Advantages of LMC are important : unequalled ease of verification of admissibility (Goulard and Voltz, 1992) and possibility to use a large variety of simulation algorithms, including the efficient turning bands method (Matheron, 1973; Chilès and Delfiner, 2012) and FFTMA (Le Ravalec-Dupin *et al.*, 2000) as all what is required is to combine linearly independent univariate simulations. The main disadvantage of LMC is the rather severe restriction it imposes that the variables share common spatial structures. It is indeed frequent that one observes a smoother behavior on the secondary variables than on the main variable. Using the LMC in this case incurs an important loss of information.

Recently Marcotte (2015) proposed a new tool facilitating the verification of admissibility for non-LMC models (N-LMC) with symmetrical cross-covariances. A direct follow-up is to find an efficient method of simulation of the N-LMC. Shinozuka (1971) developed a continuous spectral method to simulate a multivariate homogeneous process by a series of cosine functions. He proposed to use Cholesky decomposition of the spectral density matrix, at selected frequencies. Then, the study was extended to non-homogeneous oscillatory processes characterized by an evolutionary power spectrum (Shinozuka and Jan, 1972). Mejía and Rodríguez-Iturbe (1974) focused on discussing connection of correlation and spectrum of a random field and provided a simulation method by sampling from the spectral density functions. Zagayevskiy (2015) use the spectral method with turning bands to simulate LMC. Mantoglou (1987) and Emery *et al.* (2015) implement variants of the continuous spectral method (Shinozuka and Jan, 1972) with turning bands that can both simulate N-LMC.

Alternatively, Pardo-Igúzquiza and Chica-Olmo (1993) developed a Fourier integral method for unconditional simulation of random fields. The FFT was used in both numerical calculations of density spectral function and generation of realizations. Chilès (1995) used discrete

spectral method based on the FFT for simulation of intrinsic random functions. Chilès and Delfiner (1997) propose various tools to simulate by FFT while minimizing aliasing effects due to the asymptotic ranges of many covariance functions. Le Ravalec-Dupin *et al.* (2000) proposed to integrate FFT and the moving average method in a fast and flexible method they named FFTMA. Applications of FFTMA can be found in Le Ravalec-Dupin *et al.* (2001), Gloaguen *et al.* (2005) and Shamsipour *et al.* (2011). Le Ravalec-Dupin and Da Veiga (2011) extended FFTMA for cosimulation of two variables that are linearly correlated. In addition to being limited to two variables, their method was based on the Markov-Bayes approximation, so it did not allow full control and generality of the simulated cross-covariances.

In this contribution, based on the discrete spectral method, the FFTMA is adapted to the unconditional simulation of multivariate fields with the N-LMC. One requirement on the coregionalization is the cross-covariances must be symmetric. The square root of the spectral matrix is obtained at each frequency sampled by the FFT. Then spectrum of white noise fields for n variables are computed by FFT and combined with the square root matrices (one per frequency) so as to generate the desired direct and cross-structures. To increase the applicability of the FFTMA, we also present a simple idea to extend the simulation to points that do not fall on a regular grid. This allows the post-conditioning by cokriging to be performed to obtain conditional realizations at any desired point.

After reviewing the FFTMA algorithm, we describe the necessary modifications to simulate N-LMC. We then discuss implementation details related to the type of simulated covariance, the admissibility issue, the computing time and the solution proposed for the case of samples not on a regular grid. We show that the program, named GFFTMA, reproduces the desired N-LMC. CPU time and memory requirements are examined. Finally, a case study is presented.

7.3 Methodology

The moving average method (MA) was presented for simulation of one dimensional Gaussian random fields in Journel (1974). Then, it was extended by Oliver (1995) to two and three dimensions. A zero mean Gaussian random field $z(x)$ with covariance $C(h)$ (h is a distance vector between x_i and x_j) is generated by

$$z(x) = g(h) * y(x). \tag{7.1}$$

where $*$ is the convolution operator, $y(x)$ is a Gaussian white noise and $g(h)$ is a convolution root of the covariance, i.e. $C(h) = g(h) * \hat{g}(h)$ in which $\hat{g}(h) = g(-h)$. For symmetric covariances and cross-covariances, it is possible to choose $g(h) = g(-h)$, hence a symme-

tric convolution root g exists (although non-symmetric roots also exist). Oliver (1995) listed covariance functions and their auto-convolution root in two and three dimensions. Le Ravalec-Dupin *et al.* (2000) recognized the spatial convolution of Equ. 7.1 is most efficiently computed in the spectral domain by using FFT, hence the method FFTMA was developed. Oliver (2003) described the use of Cholesky decomposition to simulate 2D N-LMC models combining different model types for the direct covariances or different ranges.

The convolution theorem, either in its continuous or discrete versions, states that the Fourier transform of the convolution of two functions is the product of the Fourier transform of the two functions (Cooley *et al.*, 1969; Priestley, 1982; Le Ravalec-Dupin *et al.*, 2000; Rao *et al.*, 2011) :

$$\mathcal{F}(g_1 * g_2) = \mathcal{F}(g_1) \cdot \mathcal{F}(g_2). \tag{7.2}$$

In the univariate case with symmetrical covariance $C = g * g$ where g is the symmetric square root. As the spectral density function $S = \mathcal{F}(C) = \mathcal{F}(g * g) = [\mathcal{F}(g)]^2$, one has $\mathcal{F}(g) = \sqrt{S}$. Thus to obtain the root spectrum, it suffices, in the univariate case, to take the square root of the spectral density function of the covariance. Having the root spectrum, one simply has to multiply it with the spectrum of a Gaussian white noise and then take the inverse Fourier transform to get back in the spatial domain. The final result is the convolution of the covariance root with a Gaussian white noise (a moving average), ensuring correct reproduction of the desired covariances.

7.3.1 The FFTMA in the multivariate case

Le Ravalec-Dupin and Da Veiga (2011) present an approximate method to simulate sequentially two variables with different structures and a given correlation. Their approach is to simulate a first variable and then to simulate the second one conditional to the first one using a Markov-Bayes hypothesis for the cross-covariance. It is only an approximate method due to the Markov-Bayes hypothesis that does not allow full control on the simulated cross-covariances. For example, when the first variable has a spherical covariance, the second variable could not be simulated exactly with a Gaussian covariance. Here, we seek to generalize the simulation to any number of variables and to allow any covariance and cross-covariance, with the only restriction that the coregionalization must be an admissible model. For this, FFTMA is used to simulate directly the joint distribution of the n variables with imposed covariances and cross-covariances. The idea is based on the fact that knowledge of direct and cross-covariances is equivalent to knowledge of the direct and cross-spectral densities. Hence, the goal is to simulate n variables having completely specified spectral matrices at all frequencies.

The covariance matrix function $C(h)$ is :

$$\begin{pmatrix} C_{1,1}(h) & C_{1,2}(h) & \cdots & C_{1,n}(h) \\ C_{2,1}(h) & C_{2,2}(h) & \cdots & C_{2,n}(h) \\ \vdots & \vdots & \ddots & \vdots \\ C_{n,1}(h) & C_{n,2}(h) & \cdots & C_{n,n}(h) \end{pmatrix} \tag{7.3}$$

Assuming symmetry of cross-covariances, the corresponding spectral matrix $S(f)$ (f is frequency) is obtained by separate Fourier transform of each of the $n \times (n+1)/2$ covariance functions :

$$\begin{pmatrix} S_{1,1}(f) & S_{1,2}(f) & \cdots & S_{1,n}(f) \\ S_{2,1}(f) & S_{2,2}(f) & \cdots & S_{2,n}(f) \\ \vdots & \vdots & \ddots & \vdots \\ S_{n,1}(f) & S_{n,2}(f) & \cdots & S_{n,n}(f) \end{pmatrix} \tag{7.4}$$

where $S_{ij}(f) = S_{ji}(f)$ as the Fourier transform of a real even function is also real and even.

At any given frequency f, by Bochner's theorem, the spectral matrix must be positive semi-definite. Then one can use the square root matrix decomposition of $S(f)$ as : $S(f) = U(f)U(f)^T$ where ^{T} denotes the transpose operator, $U = VD^{1/2}V^T$ and D and V are respectively the diagonal eigenvalue matrix and the orthonormal eigenvector matrix, i.e. one has $S = VDV^T$ and $VV^T = V^TV = I$. Now, consider a zero-mean unit-variance Gaussian white noise $y(x)$ in the spatial domain. It is well known that its Fourier transform $\widetilde{y}(f)$ is also a zero-mean unit-variance Gaussian white noise. Hence, one has :

$$E[U(f)\widetilde{y}(f)\widetilde{y}(f)^TU(f)^T] = U(f)U(f)^T = S(f) \tag{7.5}$$

where $\widetilde{y}(f)$ is a column vector of size n containing the Fourier transform coefficients at frequency f for the corresponding n independent Gaussian white noises. This equation leads to the following algorithm in the discrete case :

- Compute the FFT of each of the $n \times (n+1)/2$ periodized covariances and cross-covariances evaluated on the regular grid (N points) ; this provides $S(f)$.
- Generate n independent white noise vectors on N grids and compute their FFT $\widetilde{y}(f)$.
- For all frequencies, respectively compute the eigenvalue-eigenvector decomposition of the spectral matrix $S(f)$ and get the $U(f) = V(f)D(f)^{1/2}V(f)^T$. The eigen-decomposition is done on the small matrix of size $n \times n$ at each frequency. For all frequencies, $U(f)$ consists of $n \times n$ matrices.
- Calculate the spectrum of simulation at all frequencies f as $\widetilde{z}(f) = U(f)\widetilde{y}(f)$.

— Take the separate inverse FFT on each component of vectors $\widetilde{z}(f)$ to obtain the simulation $z(x)$ in spatial domain that have the right covariances and cross-covariances.

The reader will note that the above procedure is the direct analog of what is done by FFTMA in the univariate case where $S(f)$ is a scalar, and one computes $\sqrt{S(f)}\widetilde{y}(f)$ and then takes the inverse Fourier transform. In the multivariate case $S(f)$ is a matrix, and the square root is replaced by the square root matrix decomposition. The Cholesky decomposition of $S(f)$ could also have been used, and in fact it was used in the initial versions of GFFTMA. However, it appeared that the Cholesky decomposition was numerically unstable in some models at high frequencies where all the spectral values are close to zero.

7.3.2 Post-conditioning by cokriging

The classical method of post-conditioning by cokriging (Chilès and Delfiner, 2012) is used, either by simple cokriging or by ordinary kriging or kriging with a trend (Emery, 2007). Therefore,

$$Z_{CS}(x) = Z^*(x) + (Z_S(x) - Z_S^*(x)) \tag{7.6}$$

where $Z^*(x)$ is the cokriging estimate at simulated point x using the data, $Z_S^*(x)$ is the cokriging estimate at simulated point x using the simulated values at sample points and $Z_S(x)$ are unconditional simulated values obtained with GFFTMA.

An important limitation of the FFTMA is to simulate only on a regular grid. This is a problem for the post-conditioning step as one also needs simulated values at sampled points. Moreover, they must have the right structure and be compatible with all the points already simulated on the regular grid. Relying on the screening effect approximation, one idea is to locally propagate the unconditional simulated vectors of the regular grid to sample points. By Cholesky or SGS method, values on the sample points can be simulated conditioned by the neighboring regular grid vectors. This is repeated for each sample point in turn. Sample points that are simulated by SGS or Cholesky are added to the already simulated vectors for the next sample points. At the end of the procedure, correlated random vectors with approximately the right covariance structure and compatible with the regular grid are available at each sample point. One can then proceed with the post-conditioning as usual. Note that this step is relatively fast compared to a standard SGS as only the sample points need to be simulated, not the grid points that are already simulated by GFFTMA.

7.3.3 Models with asymptotic range

Following Le Ravalec-Dupin *et al.* (2000) and Chilès and Delfiner (2012), the size of the field to simulate internally must be at least $L = b + a$ where b is the size of the desired simulation in a given direction and a is the correlation range in this direction. For covariances that possess only an asymptotic range, the practical range is used (i.e. the distance where the covariance represents only 5 % of the sill). One exception is for the Gaussian covariance model where it was found that the practical range was not sufficient to ensure good covariance reproduction at small distances. For this model, the practical range was extended from $\sqrt{3}a$ to $3a$. Note that for all models with asymptotic ranges, the periodicity of the FFT introduces a discontinuity of the slope of the periodized covariances at mid-distance of the simulated field (see Fig. 7.1). This discontinuity causes the simulated models to become numerically not admissible as some small negative spectral values appear in the spectrum. The solution adopted here is to replace the direct spectral negative values by very small positive values and the corresponding cross-spectral values by zeros. In addition, all negative eigenvalues found during the eigen-decomposition are simply replaced by zeros. This ensures numerical admissibility of each simulated N-LMC. Examples in the next section show these simple corrections to ensure positive-definiteness at all frequencies of the FFT do not introduce a measurable bias on the covariance. All the simulated covariances and cross-covariances do not depart significantly from the intended theoretical N-LMC.

To account for possible geometric anisotropies in different directions and of covariances formed of multiple nested components, the maximum range or practical range among all directions and all components of the N-LMC was retained as the unique parameter a used to define the extension of the simulated grid.

7.3.4 Test of GFFTMA

The examples presented in Figs 7.2 to 7.5 seek to illustrate the versatility of GFFTMA and the quality of reproduction of the multivariate models. The models are described in details in Table 7.1. In each case, 200 realizations of a square field of size 500 x 500 are obtained for isotropic cases with two variables (Figs 7.2 and 7.3) and three variables (Fig. 7.4) and for an anisotropic case with two variables (Fig. 7.5). The direct and cross-variograms of the first 25 realizations are illustrated together with the mean variogram computed over the 200 realizations. Models used include mixtures of spherical, exponential, Gaussian, cubic, penta, generalized Cauchy and K-Bessel. The coregionalizations are all N-LMC as different structures appear on the direct and cross-variograms. The models are checked to be admissible with program TASC3D (Marcotte, 2015). This program uses the theoretical 3D spectral

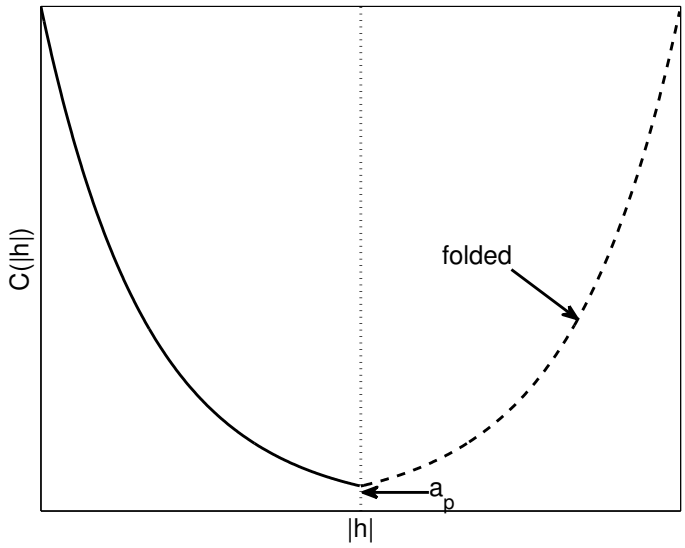

Figure 7.1 Example of slope discontinuity at practical range a_p after periodization of the exponential model.

densities of the same models available in GFFTMA to verify the positive-definiteness of the spectral matrices at any set of user specified frequencies. It allows any number of components in the direct and cross-covariance models of the N-LMC, each component being isotropic or anisotropic.

7.3.5 Computing time

Because of the use of FFT, GFFTMA is probably unequaled for the computing time required to get the unconditional simulations. The post-conditioning is generally the slower step, although this step can be also quite fast when dual kriging is used in a global neighborhood (Royer and Vierra, 1984; Davis and Grivet, 1984). Fig. 7.6-a shows the well known $Nlog(N)$ CPU time relationship of the FFT transform (N the number of simulated pixels) as the correlation between CPU time and $Nlog(N)$ reaches 0.999. Note that one realization for two variables on 100 million pixels is obtained in approximately 200 seconds for the cubic and Gaussian models with range 100, 500 seconds for the spherical model with range 200 and 2800 seconds for the spherical model with range 100 pixels. All computations are done on an Intel Xeon (2.13 GHz). The most consuming computation is the eigen-decomposition. We stress that this computation step is done only once to get the square root matrices. When the model has a linear behavior at the origin and a small range, the eigen-decomposition must be computed at almost all frequencies (which is the total number of pixels in the simulation). On

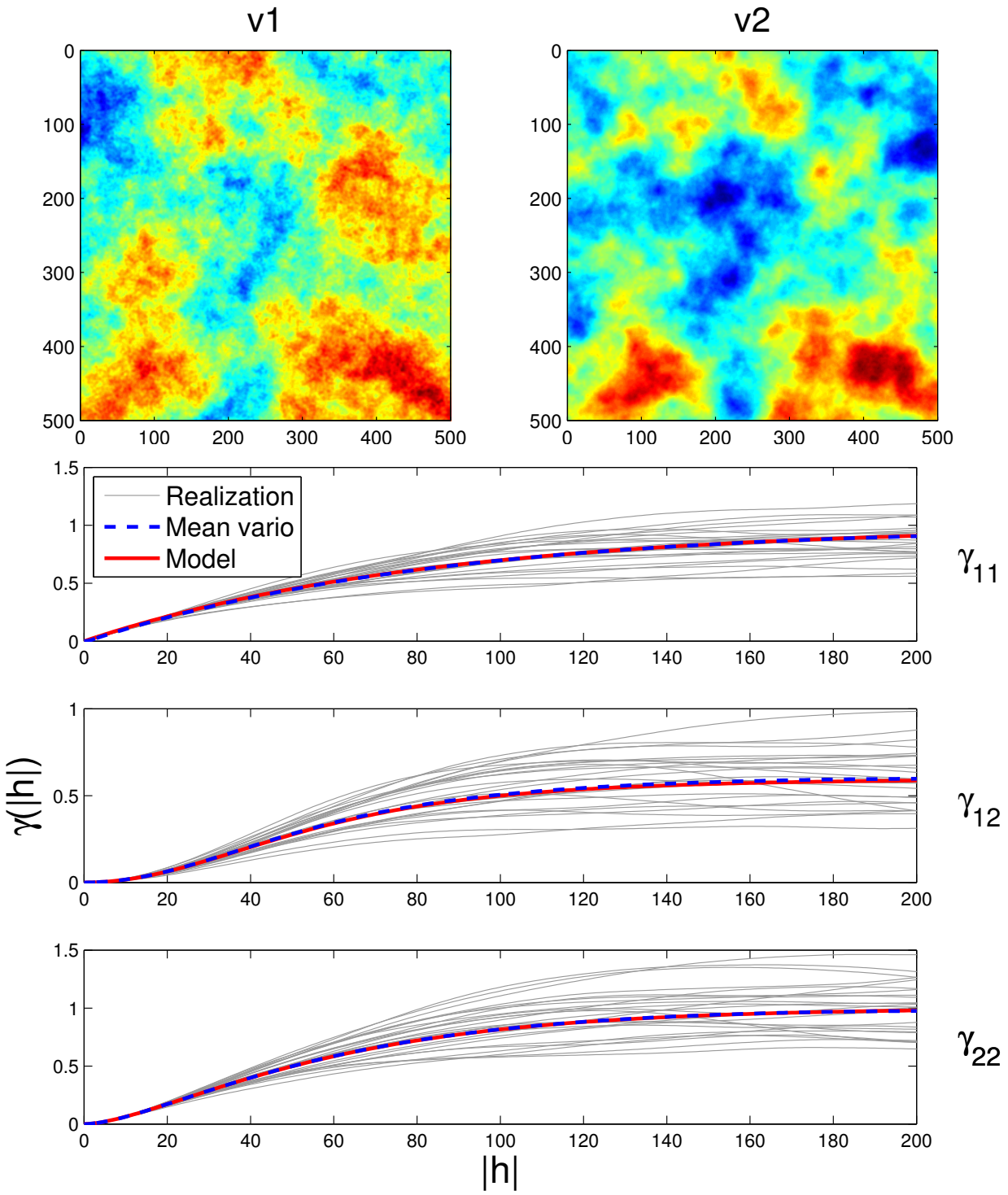

Figure 7.2 One realizations of v1 and v2 (top row) and the direct and-cross variograms. Case 1 of Table 7.1 mixing exponential, Generalized Cauchy with $\nu = 2$ and K-Bessel with $\nu = 1$. Mean variogram is computed by combining E-W and N-S directions over 200 realizations. Only the first 25 individual realization variograms (light gray) are shown.

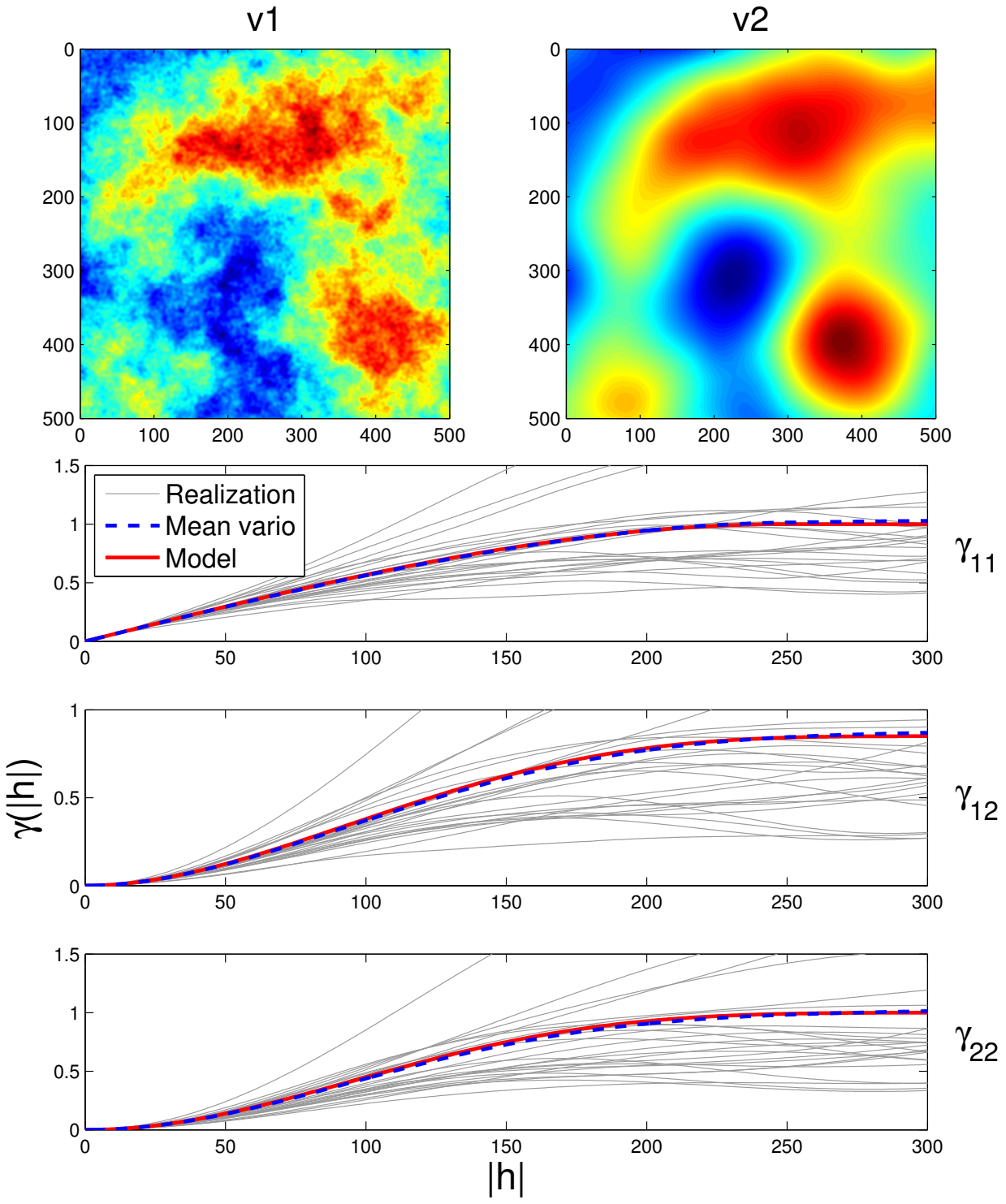

Figure 7.3 One realization of v1 and v2 (top row) and the direct and cross-variograms. Case 2 of Table 7.1. Mean variogram is computed by combining E-W and N-S directions over 200 realizations. Only the first 25 individual realization variograms (light gray) are shown.

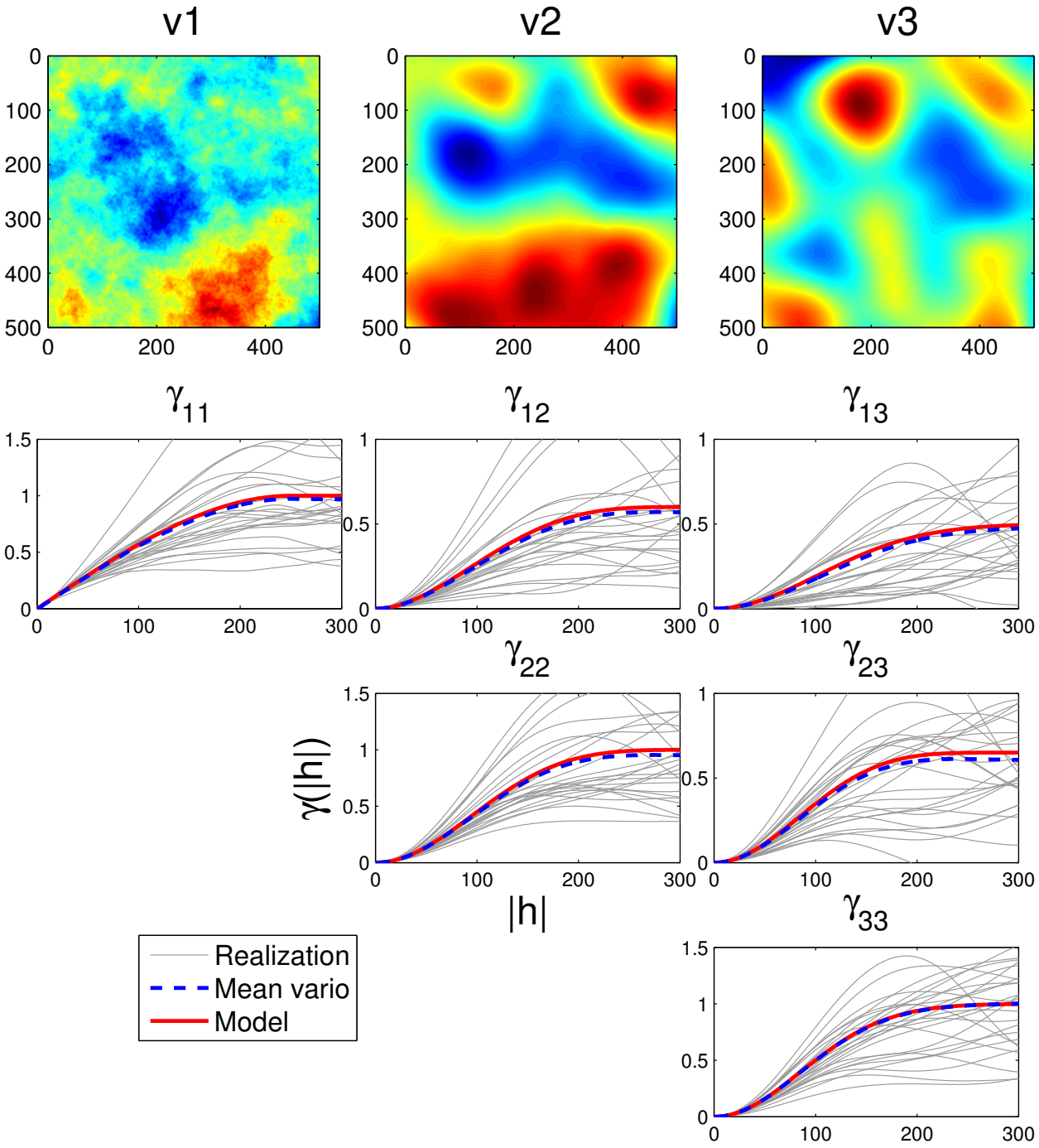

Figure 7.4 One realization of v1, v2 and v3 (top row) and the direct and cross-variograms. Case 3 of Table 7.1. Mean variogram is computed by combining E-W and N-S directions over 200 realizations. Only the first 25 individual realization variograms (light gray) are shown.

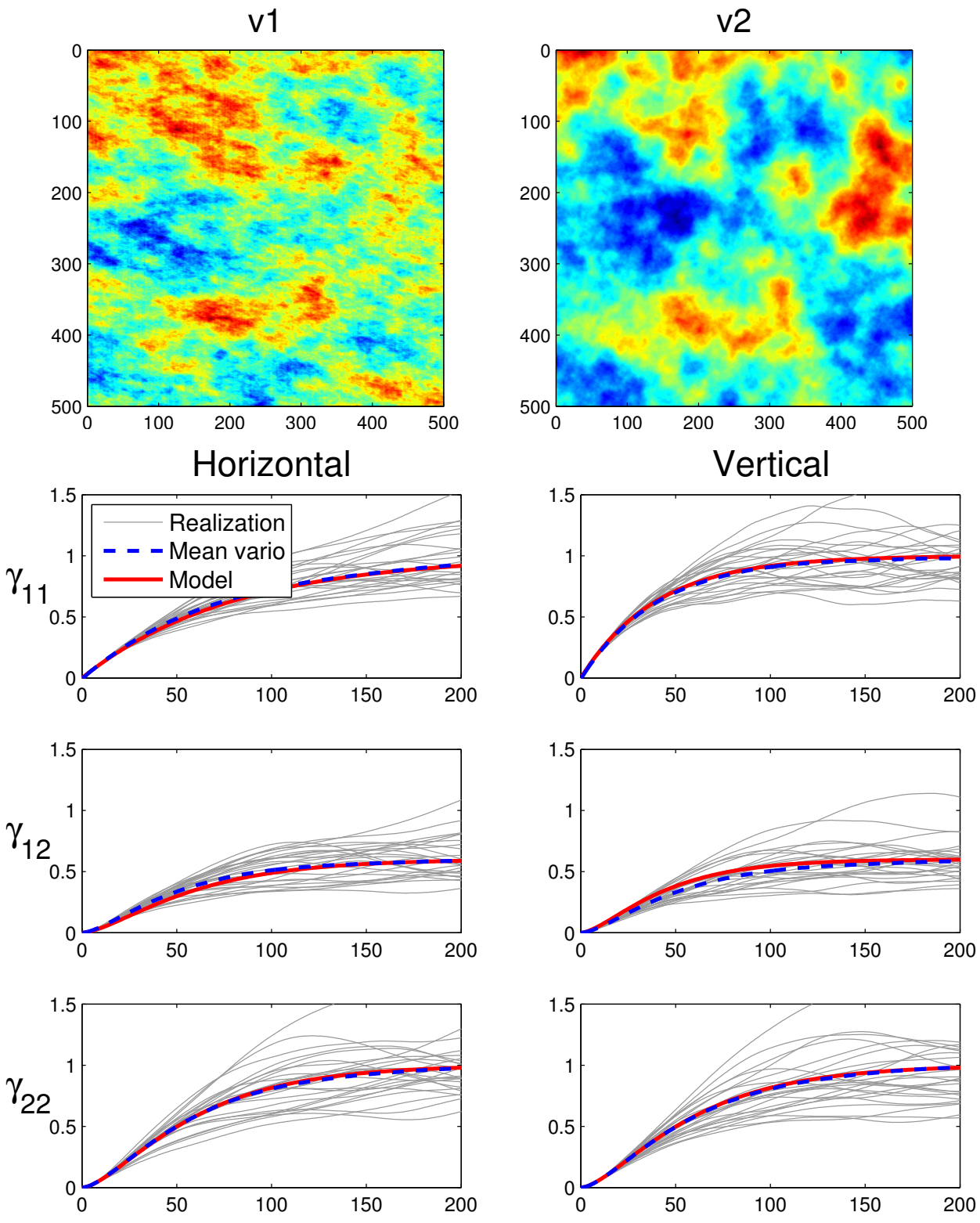

Figure 7.5 One realization of v1 and v2 (top row) and the direct and cross-variograms. Case 4 (with geometric anisotropy) in Table 7.1. Horizontal direction (left column) and vertical direction (right column). Mean variogram is computed over 200 realizations. Light gray : variograms for the first 25 realizations

Table 7.1 Models used in Figs 7.2-7.5 and Fig. 7.9

Case	Figure	Model description
1	Fig. 7.2	v1-v1 : Exponential ($a_p = 250, C = 1$) v1-v2 : Cauchy ($a_p = 155, C = 0.6, \nu = 2$) v2-v2 : K-Bessel ($a_p = 158, C = 1; \nu = 1$)
2	Fig. 7.3	v1-v1 : Spherical ($a = 250, C = 1$) v1-v2 : Cubic ($a = 310, C = 0.85$) v2-v2 : Penta ($a = 350, C = 1$)
3	Fig. 7.4	v1-v1 : Spherical ($a = 250, C = 1$) v1-v2 : Cubic ($a = 310, C = 0.5$) v1-v3 : Gaussian ($a_p = 250, C = 0.5$) v2-v2 : Penta ($a = 350, C = 1$) v2-v3 : Penta ($a = 310, C = 0.8$ v3-v3 : Gaussian ($a = 208, C = 1$)
4	Fig. 7.5	v1-v1 : Exponential ($a_{p,hor} = 240, a_{p,vert} = 120, C = 1$) v1-v2 : K-Bessel ($a_{p,hor} = 158, a_{p,vert} = 118.5, C = 0.6, \nu = 1$) v2-v2 : K-Bessel ($a_{p,iso} = 158, C = 1; \nu = 1$)
5	Fig. 7.9	v1-v1 : Nugget ($C = 0.15$)+Exponential ($a_p = 27km, C = 0.45$) v1-v2 : Nugget(C=0.001) + Gaussian ($a_p = 32.6km, C = 0.3$) v2-v2 : Nugget(C=0.001)+ Gaussian ($a_p = 24.6km, C = 1$)

the contrary, with differentiable variograms at the origin and/or for large range covariances, most frequencies have a quite small contribution that can be disregarded without discernable effects. In all cases, every additional realization is obtained in a comparatively negligible time that is independent of the covariance model used, as each realization calls for only two FFTs and the product of the square root matrices with the Gaussian white noise FFT coefficients. As expected, the simulation time increases linearly with the number of realizations (Fig. 7.6-b).

7.3.6 Memory usage

No doubt, the main limitation of the approach is the memory usage, especially when simulating large grids for models with long practical range. The maximum memory mobilized by the current implementation of GFFTMA is approximately $32p^2(n_s + n_a)^d n_{sim}$ where p is the number of variables, n_s is the number of pixels simulated, n_a is the number of pixels required to reach the maximal practical range, d is the dimension of the studied field and n_{sim} is the number of realizations to simulate. Table 7.2 gives the approximate maximum size that can be simulated using only the available RAM memory in 2D and 3D as a function of the memory available above the memory required to run the operating system and Matlab's overhead.

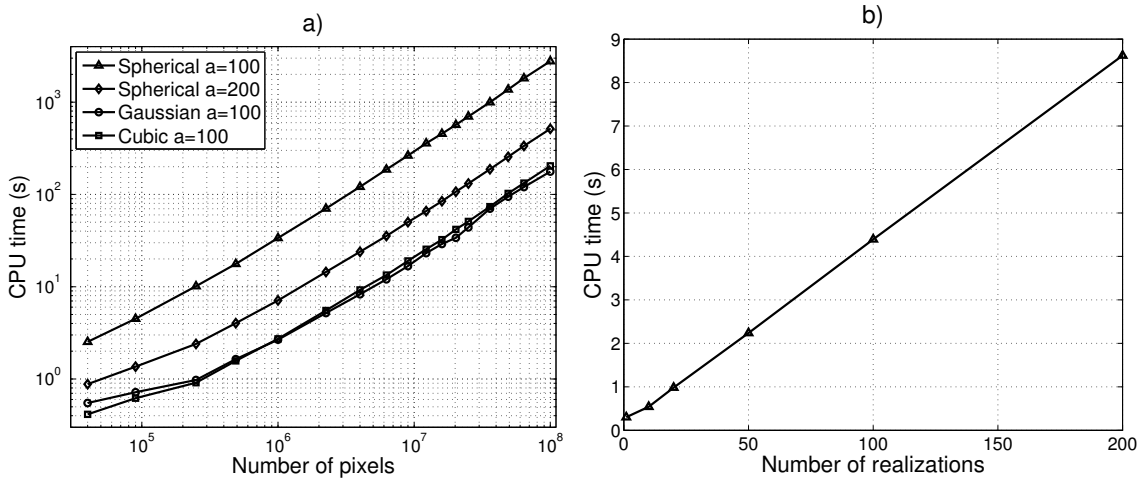

Figure 7.6 Evolution of computing time as a function of a) the number of simulated pixels and b) the number of realizations (for a field of 200 x 200 pixels). Simulation of two variables for four different models in a) and a spherical model with range 100 in b).

For larger fields, the program has to use a swap file, which slows considerably the execution time. Clearly, results of Table 7.2 indicate GFFTMA is more suitable for simulation in 2D than in 3D, although mid-size problems in 3D can still be treated.

Table 7.2 Maximum size of simulated field as a function of available RAM above overhead memory required by operating system and Matlab (for $n_{sim} = 1, n_a = 100, p = 2$)

Available RAM	2D	3D
4G	5400 × 5400	210 × 210 × 210
16G	11000 × 11000	400 × 400 × 400
48G	19000 × 19000	620 × 620 × 620

7.4 Case Study - Overburden thickness simulation

The case study bears on the cosimulation of overburden thickness and surface lithological information represented by a Gaussian latent variable. The overburden thickness data includes 4730 boreholes in Montérégie Est, in the south of Québec. The study area covers approximately $14000 km^2$ split into two main geological domains. The geological domain located at northwest (domain A) has generally higher overburden thickness than the rest (domain B). The sample locations are shown in Fig. 7.7.

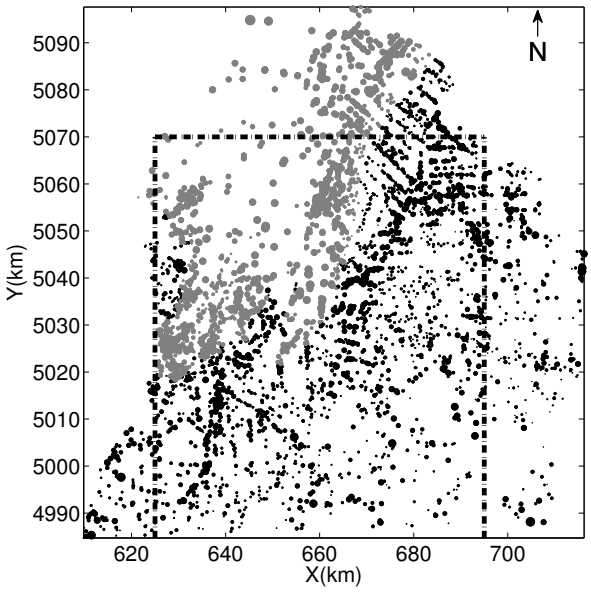

Figure 7.7 Map of sample data. Gray dots : boreholes in geological domain A, black dots : boreholes in geological domain B. The area covered by simulation in Fig. 7.10 is outlined (dashed line)

A bivariate field is to be simulated by the GFFTMA. The logarithm of overburden thickness is treated as the first variable. Another known information is the indicator of geological domain of the studied area. A number of approaches can be used in the simulation of the geological types or facies, such as sequential indicator simulation (Journel and Isaaks, 1984; Deutsch and Journel, 1998), multiple point statistics (Guardiano and Srivastava, 1993; Ortiz and Deutsch, 2004), truncated Gaussian simulation (Matheron *et al.*, 1987), and truncated pluriGaussian simulation (Galli *et al.*, 1994). In this article, we used the truncated Gaussian simulation, i.e. the geological domain indicator is obtained by thresholding a continuous latent Gaussian variable at the sampling points. First, the fixed threshold of the Gaussian field has been determined using declusterized proportions of each geological domain. Then the latent variable variogram is chosen so as to match the indicator variogram and log(thickness)-indicator cross-variogram. Having the variogram model of the latent variable, the Gibbs sampler provides latent Gaussian at sampling points in agreement with the geological domain indicator. For each conditional realization, a different Gibbs sampler realization of the latent Gaussian variable at sample points is used.

Figure 7.8 shows the direct and cross variograms of the log(thickness) and the latent Gaussian variable. The experimental variogram of overburden log(thickness) is well adjusted by an

exponential model. The latent variable was chosen to have a Gaussian covariance model because it provided better fit to the geological domain indicator variogram and log(thickness)-indicator cross-variogram. The fit obtained after the Gibbs sampling reproduces the assumed model as expected. Finally, the cross-variogram is also well fitted by a Gaussian model. Because the two variables have different direct variogram structures, they define a N-LMC model with zero-lag correlation of $R = 0.38$.

Two series of unconditional realizations of logarithm of overburden thickness were obtained on a regular grid : an univariate simulation by FFTMA and the bivariate N-LMC simulation involving geological domain by GFFTMA. Then, the unconditional simulated values were extended to the sample point locations by local LU simulation.

Figure 7.9 shows direct and cross variograms of the two variables for 30 unconditional realizations obtained in cosimulation of N-LMC. The average variograms of the 30 realizations coincide almost perfectly with the models. Each realization is then post-conditioned by co-kriging using the same N-LMC model.

7.4.1 Comparison of statistics of conditional realizations by N-LMC model and univariate simulation

To assess the gain obtained by considering the geological information, the full dataset is split into two parts : a small part (10%) is used as conditioning data and the rest (90%) is used as test data. Conditional realizations are obtained at the test data locations for two different cases : one using the univariate simulation of log(thickness) and the second using N-LMC simulation of log(thickness) with the Gaussian latent variable as secondary information at both conditioning points and test data locations. Then the final simulations of log(thickness) for overburden are transformed to the thickness. Table 7.3 shows mean statistics of 30 realizations obtained with the two instances. The mean absolute error (MAE) of simulation with N-LMC shows an improvement of 3.9% to 4.3% compared to the univariate simulation, indicating that despite the small zeros-lag correlation ($R = 0.38$), the geological information remains useful.

Table 7.3 Mean MAE (in m) for 30 realizations

	On 981 testing points within 5km distance from boundary	On all 4257 testing points
Case 1 (Univariate simulation)	6.49	6.95
Case 2 (N-LMC simulation)	6.21	6.68
% improvement with N-LMC	4.3 %	3.9 %

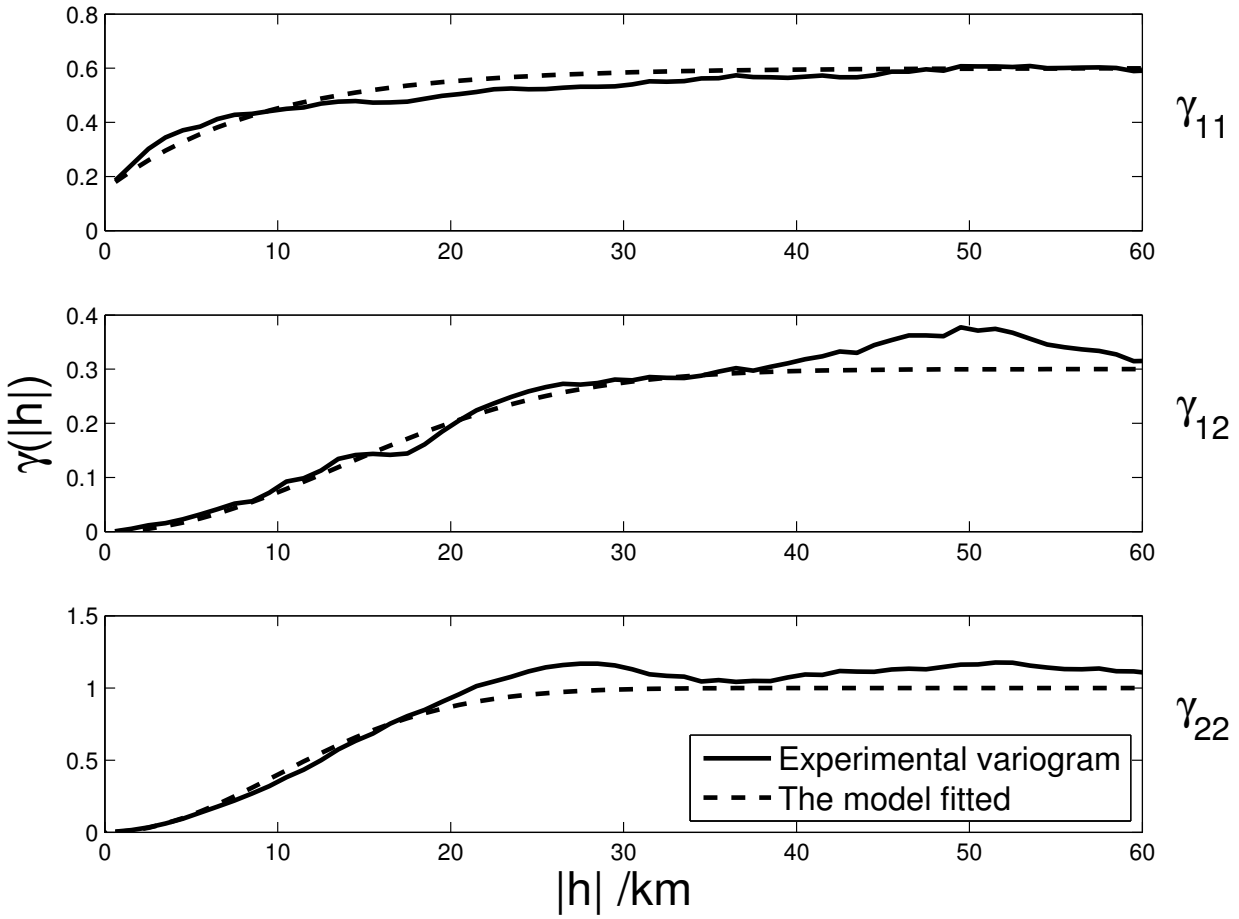

Figure 7.8 Model fitting for direct and cross variograms of log(thickness) and one latent field representing the geological domain information.

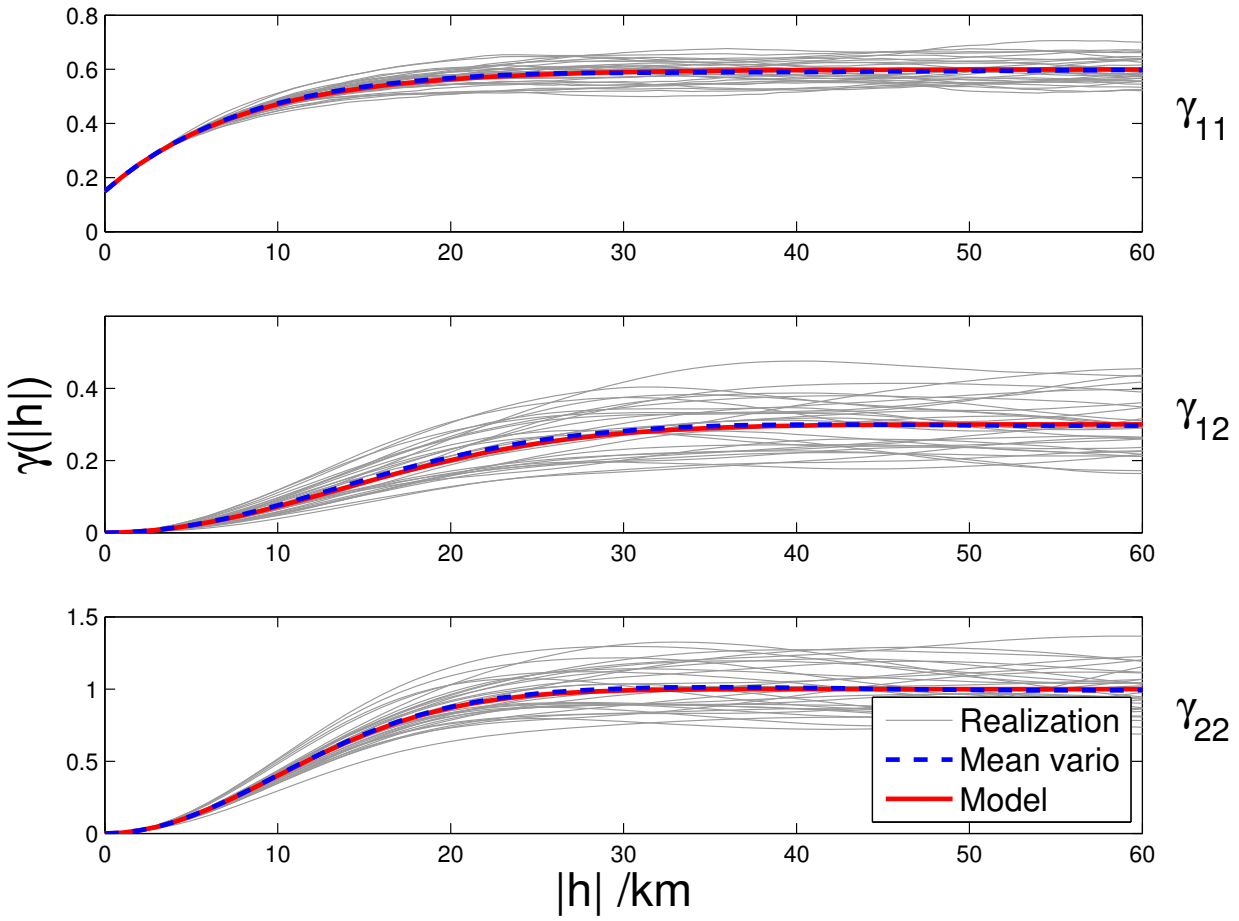

Figure 7.9 Comparison of the direct and cross variograms of log(thickness) and the latent variable between 30 realizations and the theoretical model.

Figure 7.10 shows 3 realizations of the univariate simulation (left) and the N-LMC (right), still using 10% of the thickness data and all the latent data. Clearly, the realizations of the N-LMC show larger thickness toward the North than do the univariate realizations. Hence, one concludes that the geological information conveyed by the latent variable has a noticeable impact on the simulations despite the low correlation between the log(thickness) and the latent variable.

7.5 Discussion

Examples were provided illustrating simulation of the models available in GFFTMA. The average experimental direct and cross-variograms reproduced almost perfectly the theoretical model in all cases presented, (see Figs 7.2 to 7.5) even for N-LMC known to be not strictly admissible like the N-LMC models having a compactly supported direct covariance function (e.g. models in Figs 7.3-7.4). These models are known to be non admissible (Marcotte, 2015), because at high frequencies the spectral densities of the compactly supported covariance function possess multiple zeros, hence the spectral density matrix cannot be positive semidefinite at these frequencies. However, this fact has little consequence in practice as seen by the quality of the reproduction of all direct and cross-covariances. Moreover, the adopted correction for the small negative eigenvalues renders these models admissible at the sampled scale. In the post-conditioning by cokriging, there were no instance where the cokriging matrix was found not positive definite.

The eigen-decomposition is done at each frequency where the amplitude exceeds the threshold $A_{max,i}/10^6$ where $A_{max,i}$ is the maximum amplitude found in the direct spectrum of variable i. As soon as one threshold is exceeded, the spectral matrix is eigen-decomposed. As a consequence, the CPU times of GFFTMA are largely related to the behavior at the origin of the models present in the N-LMC. Linear behavior variograms have larger high frequencies, hence eigen-decomposition must be computed at more frequencies. For a same behavior at the origin, the smaller the range with respect to simulated field size, the larger the CPU time. Nevertheless, the CPU times remain tractable and always keep the $Nlog(N)$ dependency where N is the number of pixels to simulate. Moreover, the eigen-decomposition is done only once, irrespective of the number of realizations. Previous trials with Cholesky decomposition, instead of eigen-decomposition, to compute the square root matrices were significantly faster but proved numerically unstable in many circumstances. For this reason, it is preferable to stick to the slower eigen-decomposition.

The memory requirements are rather important as the field size to simulate internally in GFFTMA to reduce aliasing is equal to the size of the field to simulate plus the maximal

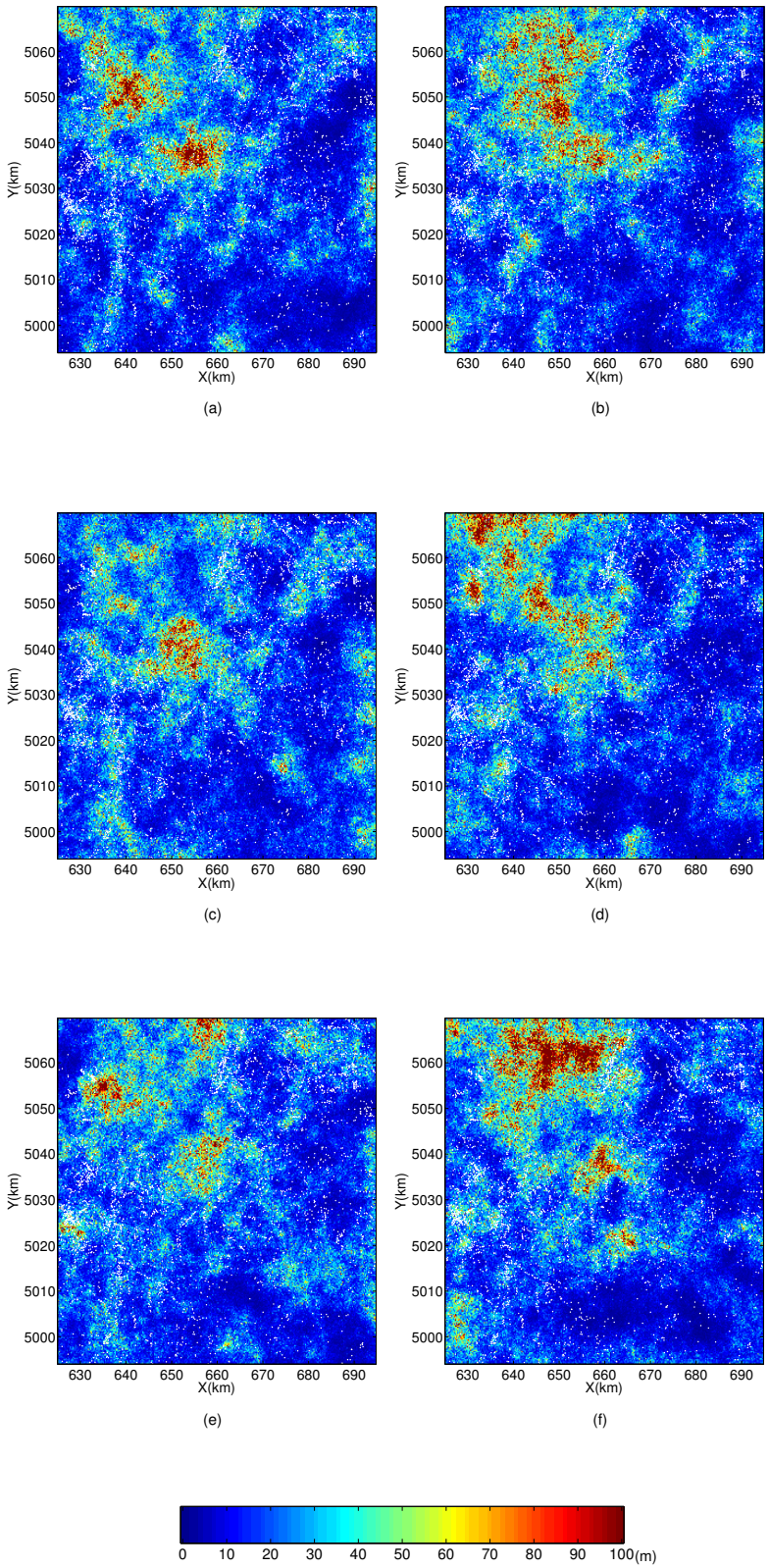

Figure 7.10 Realizations of univariate simulation (left) and N-LMC simulation (right). White dots represent data locations.

practical range found among all model components. As an example, with a maximal practical range of 100, 48G of RAM enables simulation of a field comprising approximately 238 million pixels in 3D, corresponding to 620 pixels along each dimension. One avenue to limit the size of the simulation may be to try to approximate models displaying asymptotic sills (exponential, Gaussian, Cauchy and K-Bessel) by combination of models with compact support (spherical, cubic and penta). This could reduce significantly the size of the field to simulate internally, although it is not clear if the number of frequencies to eigen-decompose does also diminish.

7.6 Conclusion

The GFFTMA program can generate unconditional simulations in both isotropic and anisotropic situations for any number of variables and any number of covariance components for each direct and cross-covariances. It enables to simulate N-LMC with symmetric cross-covariances. Basic covariance models included in GFFTMA are nugget, exponential, Gaussian, spherical, cubic, penta, generalized Cauchy and K-Bessel (Matérn). If needed, other models can be added easily with a single line of code. Simulated examples show that GFFTMA succeed to produce realizations having on average the desired N-LMC.

7.7 Acknowledgments

This research was financed by NSERC (RGPIN-2015-06653) and by Polytechnique Montréal. The constructive comments of three anonymous reviewers were helpful to improve the manuscript.

7.8 Appendix - usage of GFFTMA

The program GFFTMA is a Matlab function used to produce unconditional simulations of N-LMC models. For conditional simulation, it must be completed with other functions performing 1- extension of the unconditional grid to sample points not lying on the grid and 2-the post-conditioning by cokriging with the same N-LMC model. The check of the admissibility of the N-LMC model can be done with program TASC3D (Marcotte, 2015).

GFFTMA is called with :

datasim=gfftma(model,seed,nsim,vsiz,grid,thres)

The input *model* is a cell variable of size $n \times n$, n being the number of variables. Each model $\{i, j\}$ possesses p_{ij} elementary components. As GFFTMA assumes symmetric cross-covariances, model $\{i, j\}$ must be the same as model $\{j, i\}$. Each component is specified

with 4 parameters (when all components of all models are isotropic) or 6 parameters in the 2D anisotropic case and 9 parameters in the 3D anisotropic case. In the isotropic case, the four parameters are : 1-model type, 2-range, 3-shape, 4-sill. In the 2D anisotropic case, the 6 parameters are 1-model type, 2- first range, 3- second range, 4-anticlockwise rotation, 5- shape parameter and 6-sill. In the 3D anisotropic case, the 9 parameters are : 1-model type, 2-first range, 3-second range, 4-third range, 5-rotation around x, 6-rotation around y, 7 rotation around z, 8-shape, and 9-sill. The shape parameter is used only for model types 7 (Cauchy) or 8 (K-Bessel). The rotations follow the same convention and order as described in Marcotte (1991), i.e. all anticlockwise rotations, z first, then y, then x. The model types are 1-nugget, 2-exponential, 3-Gaussian, 4-spherical, 5-cubic, 6-penta, 7-Cauchy 8-K-Bessel. As soon as one model component is anisotropic, all models (even isotropic ones) must be specified with the anisotropic convention.

The input *seed* is the seed of the random number generator.

The input *nsim* is the number of realizations required.

The input $1 \times d$ vector *vsiz* allows the user to specify the number of pixels to use internally for the simulation. This enables to fix the random numbers to the same values for two different calls to GFFTMA by specifying the same seed and the same *vsiz* . The internal number of pixels must be chosen so as to prevent aliasing, i.e. it must be larger than practical range (in pixels) plus the desired size (in pixels) to simulate. Normal usage is to leave this vector empty. The program then computes the required number of pixels.

The input *grid* is a vector. In 1D it contains [nx,dx], in 2D it contains [nx,dx,ny,dy] and in 3D [nx,dx,ny,dy,nz,dz], where nx is the number of pixels desired in the x direction and dx is the distance between consecutive pixels along x, and similarly for directions y and z.

The input *thres* is the multiplying factor applied to the maximal spectral density value. The spectral matrices where $S_i(f) > thres * \underbrace{max}_{f} S_i(f)$ for at least one i are eigen-decomposed. When *thres* is left empty, the default value $thres = 10^{-6}$ is used. Using larger values reduces the number of required eigen-decompositions and thus speeds up the simulation but it may affect the correct reproduction of the linear behavior at short distances of the exponential and spherical models.

The output *datasim* is a cell array of size $nsim \times nvar$ with the results of the simulation.

References

AHMED, S. and DE MARSILY, G. (1989). Co-kriged estimates of transmissivities using jointly water level data. M. Armstrong, editor, *Geostatistics*, Springer Netherlands, vol. 4 of *Quantitative Geology and Geostatistics*. 615–628.

ASLI, M., MARCOTTE, D. and CHOUTEAU, M. (2000). Direct inversion of gravity data by cokriging. *Geostatistics 2000, Cape Town, W.J. Kleingeld and D.G. Krige (eds), South Africa*. 64–73.

CHILÈS, J. P. (1995). Quelques méthodes de simulation de fonctions aléatoires intrinsèques. *Cahiers de Géostatistique, Fasc*, 5, 97–112.

CHILÈS, J. P. and DELFINER, P. (1997). Discrete exact simulation by the fourier method. E. Baafi and N. Schofield, editors, *Geostatistics, Wollongong96*. 258–269.

CHILÈS, J. P. and DELFINER, P. (2012). *Geostatistics : Modeling Spatial Uncertainty*. John Wiley & Sons.

COOLEY, J. W., LEWIS, P. A. and WELCH, P. D. (1969). The Fast Fourier Transform and its applications. *Education, IEEE Transactions on*, 12, 27–34.

DAVIS, M. and GRIVET, C. (1984). Kriging in a global neighborhood. *Mathematical Geology*, 16, 249–265.

DEUTSCH, C. V. and JOURNEL, A. G. (1998). *GSLIB-Geostatistical Software Library and User's Guide*. Oxford University Press.

DONG, A. (1990). *Estimation géostatistique des phénomènes régis par des équations aux dérivées partielles*. PhD Dissertation, Centre de géostatistique, Fontainebleau.

EMERY, X. (2007). Conditioning simulations of Gaussian random fields by ordinary kriging. *Mathematical Geology*, 39, 607–623.

EMERY, X., ARROYO, D. and PORCU, E. (2015). An improved spectral turning-bands algorithm for simulating stationary vector gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, in press.

GALLI, A., BEUCHER, H., LE LOCH, G. and DOLIGEZ, B. (1994). The pros and cons of the truncated Gaussian method. *Geostatistical simulations*, Springer. 217–233.

GLOAGUEN, E., MARCOTTE, D., CHOUTEAU, M. and PERROUD, H. (2005). Borehole radar velocity inversion using cokriging and cosimulation. *Journal of Applied Geophysics*, 57, 242–259.

GOULARD, M. and VOLTZ, M. (1992). Linear coregionalization model : Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology*, 24, 269–286.

GUARDIANO, F. B. and SRIVASTAVA, R. M. (1993). Multivariate geostatistics : beyond bivariate moments. *Geostatistics Troia92*, Springer. 133–144.

JOURNEL, A. and ISAAKS, E. (1984). Conditional indicator simulation : application to a Saskatchewan uranium deposit. *Mathematical Geology*, 16, 685–718.

JOURNEL, A. G. (1974). Geostatistics for conditional simulation of ore bodies. *Economic Geology*, 69, 673–687.

JOURNEL, A. G. and HUIJBREGTS, C. J. (1978). *Mining Geostatistics*. Academic Press, London.

LE RAVALEC-DUPIN, M. and DA VEIGA, S. (2011). Cosimulation as a perturbation method for calibrating porosity and permeability fields to dynamic data. *Computers & Geosciences*, 37, 1400–1412.

LE RAVALEC-DUPIN, M., HU, L. and NOTINGER, B. (2001). Stochastic reservoir modeling constrained to dynamic data : Local Calibration and Inference of Structural Parameters. *SPE Journal*, 6, 25–31.

LE RAVALEC-DUPIN, M., NOETINGER, B. and HU, L. (2000). The FFT moving average generator : An efficient numerical method for generating and conditioning Gaussian simulations. *Mathematical Geology*, 32, 701–722.

MANTOGLOU, A. (1987). Digital simulation of multivariate two- and three-dimensional stochastic processes with a spectral turning bands method. *Mathematical Geology*, 19, 129–149.

MARCOTTE, D. (1991). Cokriging with Matlab. *Computers & Geosciences*, 17, 1265–1280.

MARCOTTE, D. (2015). Tasc3d : A program to test the admissibility in 3d of non-linear models of coregionalization. *Computers & Geosciences*, 83, 168–175.

MATHERON, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5, 439.

MATHERON, G., BEUCHER, H., DE FOUQUET, C., GALLI, A., GUERILLOT, D. and RAVENNE, C. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs. *paper SPE*, 16753, 22–25.

MEJÍA, J. M. and RODRÍGUEZ-ITURBE, I. (1974). On the synthesis of random field sampling from the spectrum : An application to the generation of hydrologic spatial processes. *Water Resources Research*, 10, 705–711.

MYERS, D. E. (1983). Estimation of linear combinations and co-kriging. *Mathematical Geology*, 15, 633–637.

OLIVER, D. S. (1995). Moving averages for gaussian simulation in two and three dimensions. *Mathematical Geology*, 27, 939–960.

OLIVER, D. S. (2003). Gaussian cosimulation : Modelling of the cross-covariance. *Mathematical Geology*, 35, 681–698.

ORTIZ, J. M. and DEUTSCH, C. V. (2004). Indicator simulation accounting for multiple-point statistics. *Mathematical Geology*, 36, 545–565.

PARDO-IGÚZQUIZA, E. and CHICA-OLMO, M. (1993). The Fourier Integral Method : An efficient spectral method for simulation of random fields. *Mathematical Geology*, 25, 177–217.

PRIESTLEY, M. (1982). *Spectral analysis and time series*. No. v. 1-2 Probability and mathematical statistics. Academic Press.

RAO, K. R., KIM, D. N. and HWANG, J. J. (2011). *Fast Fourier Transform-Algorithms and Applications*. Springer Science & Business Media.

ROYER, J. J. and VIERRA, P. C. (1984). Dual formalism of kriging. V. G., D. M., J. A. G. and M. A., editors, *Geostatistics for natural resources characterization Part 2*. NATO-ASI, 691–702.

SHAMSIPOUR, P., CHOUTEAU, M. and MARCOTTE, D. (2011). 3D stochastic inversion of magnetic data. *Journal of Applied Geophysics*, 73, 336–347.

SHAMSIPOUR, P., MARCOTTE, D., CHOUTEAU, M. and KEATING, P. (2010). 3D stochastic inversion of gravity data using cokriging and cosimulation. *Geophysics*, 75, I1–I10.

SHINOZUKA, M. (1971). Simulation of multivariate and multidimensional random processes. *J. Acoustical Society of America*, 49, 357–368.

SHINOZUKA, M. and JAN, C. M. (1972). Digital simulation of random processes and its applications. *Journal of Sound and Vibration*, 25, 111–128.

WACKERNAGEL, H. (2003). *Multivariate Geostatistics, 2nd edition*. Springer Science & Business Media.

ZAGAYEVSKIY, Y. (2015). *Multivariate Grid-Free Geostatistical Simulation of Natural Phenomena*. PhD Dissertation, University of Alberta, Edmonton.

CHAPTER 8 GENERAL DISCUSSION

This dissertation provides non-stationary covariance functions and a simulation approach for N-LMC in modeling natural phenomena. In the univariate case, a class of non-stationary covariance functions with compact support is proposed. In the multivariate case, the GFFTMA is developed for simulation of multivariate fields that do not require the assumption of linear model of coregionalization.

In applications of the non-stationary model in this thesis, parameters of the covariance model were estimated by cross-validation. In the first application of overburden thickness estimation incorporating outcrops information (Section 5.3.1), the influence radius of an outcrop is defined and estimated by cross-validation. Beyond the influence circle, the overburden thickness distribution is assumed stationary. In the influence circle, the covariance model is non-stationary. The correlation range and nugget of the non-stationary model are treated as functions of distance to the nearest outcrop. That is enough to obtain a smooth estimation around outcrops and zero thickness at outcrops. In the second application of non-stationary covariance model (Section 6.6), the overburden thickness is estimated over an area composed of two geological domains. Between these two domains, a transition zone was dictated by geological considerations. The width of the transition zone is estimated by cross-validation. The parameters of covariance model, the correlation range, ratio of nugget to sill and sill, are functions of the distance to the edge of the transition zone. This model ensures a smooth transition between the two geological domains. The estimation of parameters of the non-stationary covariance model is important as it controls the quality and success of any non-stationary modeling approach. Higdon *et al.* (1999) characterized the non-stationary model by 8 parameters and estimated them in a hierarchical model. Paciorek and Schervish (2006) integrated the non-stationarity of the model by 7 parameters and estimated them with a Bayesian approach. In the applications presented in this dissertation, the parameters were estimated by cross-validation. In transition zones, based on the geological knowledge the parameters are considered as functions of distance to a critical point or line, the points of outcrops in the first case and the boundary of two geological domains in the second case. Maximum likelihood can also be used in a Gaussian context. More research is needed about the best method to estimate the non-stationary covariance models parameters. However probably more important is the expert knowledge and the statistical tools that lead to the choice of the supra-model governing the spatial variations of the covariance parameters.

The closed-form of non-stationary covariance functions with compact support presented is

only for the isotropic case. In the anisotropic case, the kernel function is the hyper-ellipsoid defined in $\Re^n$. Till now there is not explicit form of intersection of two hyper-ellipsoids. One way to calculate the non-stationary covariance of spherical family model in anisotropic case is the numerical approach provided in Section 6.4.4. Another way is tapering the isotropic spherical family functions by non-stationary covariance models for which anisotropic closed-form expressions exist, such as exponential, Gaussian or Matérn class. Then the combined model can describe the anisotropic spatial structure. In addition, the tapered functions still have the compact support and the non-negative definiteness. However, it is not clear whether or not all the properties of the anisotropic spherical family model can be recovered with the tapering approach.

For simulation of a continuous variable in univariate case, the dilution approach provides a general method to simulate the non-stationary model (Chilès and Delfiner, 1999). It is used in the moving average model that focuses on the non-stationarity of the parent random function associated with the variable of interest. In the simulation, the non-stationarity is considered in the dilution functions. For usual covariance models, the Matérn, Gaussian and spherical family model, the dilution functions are already derived and listed in Chilès and Delfiner (1999). For unusual model, Ehm *et al.* (2004) discussed what kind of covariance function can be written as auto-convolution of kernel functions therefore has a dilution function. Besides simulation of non-stationarity, another advantage of the dilution method is that it is adapted in the anisotropic case. The ratio and direction of the anisotropy are treated in the dilution function by scaling and rotation (seen in Section 6.4.4). Based on the dilution method, in Chapter 6 a weighted average method was used in simulation of non-stationary spherical family model in both isotropic and anisotropic cases. The computation time relates to the range of model to simulate and the discrete grids. Therefore, for a covariance function with small correlation range, this method is recommended for its efficiency and adaptation of non-stationarity and anisotropy of the model.

The GFFTMA method proposed in Chapter 7 is efficient in simulation of fields with N-LMC. The computation time of the GFFTMA depends on the models of coregionalization and the number of points to simulate. With the same number of grids to simulate, the GFFTMA is faster for simulating the model having derivatives at the origin than the model with linear behavior. When the models of N-LMC are determined, the computation time has the relation $Nlog(N)$ with the number of points to simulate N. Replacing eigen-decomposition by Cholesky decomposition in computation of the square root matrices would significantly save the CPU time, but it is proved numerically unstable in many circumstances. The main limitation of the GFFTMA is the rather stringent memory requirement. As seen in Section 7.3.6, the GFFTMA requires large memory for simulation of models with long practical range

in 3D. In these cases, the turning band methods (Emery *et al.*, 2015; Marcotte, 2015a) are recommended.

The GFFTMA method cannot be extended to a non-stationary case. As mentioned by Oliver (1995), in the moving average method the covariance function C is considered to be decomposed instead of the covariance matrix. The GFFTMA is developed based on moving average method. The first step of GFFTMA is to calculate the discrete covariance on the grid to simulate. In the non-stationary case, the covariance C varies with distance vector and locations. Thus it is not possible to determine the covariances on the target grids without considering locations. Available methods for simulation of non-stationary N-LMC will be discussed in the next part.

For the multivariate model, the main problems are testing the admissibility of the model and simulation. With the stationary assumption of the model, Marcotte (2015b) proposed an algorithm to test the positive definiteness of N-LMC in $\Re^3$. With respect to the simulation of N-LMC, the GFFTMA proposed in this thesis is a fast and efficient method. Based on the turning bands method, Emery *et al.* (2015) and Marcotte (2015a) recently developed approaches for simulation of N-LMC. In the non-stationary circumstance, first an open question is how to ensure the admissibility of the model. On simulation of a non-stationary N-LMC, the Cholesky decomposition can be used, as it focuses on decomposition of the covariance matrix which can include the non-stationarity. Yet the covariance matrix for Cholesky decomposition has order of $N \times N$ (N is the number of points to simulate). Therefore it cannot accommodate a field with more than 10^4 points. Another method can be used in non-stationary simulation of N-LMC is the sequential Gaussian simulation (SGS). However it is sensitive to the number and location of neighbors used in kriging. Also it cannot reproduce the exact structure for some models (Chilès and Delfiner, 2012; Emery, 2004b, 2010).

In conclusion, the methods proposed in this thesis help the study of natural phenomena. The main contributions of this thesis, the main limitations and the future work requires are described in the next chapter.

CHAPTER 9 CONCLUSION AND FUTURE WORK

9.1 Conclusion

This thesis has provided new geostatistical methods for the study of natural phenomena. In the univariate case, first a non-stationary model is developed by incorporating outcrops information in estimation of overburden thickness. Also, I proposed a class of non-stationary covariance functions with compact support and a simulation method based on moving averages with the dilution function. In the multivariate case, a fast and efficient simulation method of the N-LMC is developed. The method generalized the FFTMA to the multivariate case. All the specific objectives described in Chapter 1 were fulfilled. I highlight the following contributions :

In the first paper, I provided a new method to incorporate outcrop information in the estimation of overburden thickness. The new method avoids the bias introduced by common practice which either ignores the outcrops or treats the outcrops as zero thickness data.

In the second article, I presented and tested a new class of non-stationary covariance functions with compact support. Moreover, the closed-form expressions were derived in the isotropic case and a practical method of simulation was developped in the anisotropic case as well.

In the third article, the general FFT-MA (GFFTMA), an efficient simulation method for multiple variables was proposed. This method can simulate with non-linear model of coregionalization (N-LMC), which allows different types of covariance models for different variables.

9.2 Limitations and future work

The proposed methods have several drawbacks which may be considered as topics of future work.

1. The main limitation of the GFFTMA (seen in Section 7.3.6) is its rather stringent memory requirement. The memory requirement is related to the size of the field in practical simulation, the sum of size of the target field and the maximal practical range found among all model components. To reduce the memory requirement, one possible extension of GFFTMA in future is to limit the size of the simulation by replacing models displaying asymptotic sills (exponential, Gaussian, Cauchy and K-Bessel) with combination of models with compact support (spherical, cubic and penta-spherical). However the CPU time may not reduce, as it is not clear if the number of frequencies

to eigen-decompose does also diminish.

2. In N-LMC, the coregionalization model cannot strictly be admissible if the direct covariance models are pure spherical, cubic or penta model except the LMC. In spectral domain, the spectral densities of the spherical family models possess multiple zeros at high frequencies (Marcotte, 2015b). At these frequencies, the spectral density matrix may be negative. Taking an example of two variables, $S_{11} \cdot S_{22} - S_{12}^2 < 0$ at the frequency of $S_{11} = 0$ meanwhile $S_{12} \neq 0$. The GFFTMA treats this problem by adopting correction for the small negative eigenvalues in the eigen-decomposition of spectral matrices. Then these models are forced admissible at the sampled scale. Although this problem does not have large influence on simulation by GFFTMA and post-conditioning by cokriging, to obtain a positive semi-definite N-LMC, other compactly supported function is suggested to study in future, for example the Wendland functions, which do not show zeros in the spectral densities (Marcotte, 2015b).

3. The GFFTMA method is used in simulation of N-LMC with stationary covariances. In simulation of a non-stationary multivariate field with N-LMC, the Cholesky decomposition and the sequential Gaussian simulation (SGS) can be used. However, since the limitation of the Cholesky decomposition in small simulated field (less than 10^4 points) (Chilès and Delfiner, 2012) and the sensitivity of SGS on the number and location of neighbors used in kriging (Chilès and Delfiner, 2012; Emery, 2004b, 2010), a new efficient simulation method is required to be investigated.

4. On estimation of overburden thickness, two secondary variables were incorporated, one is outcrop information and the other is geological domain. Besides these variables, some other information may also help to understand the overburden distribution. Geophysical data can be such useful information. For example, a case study of overburden thickness estimation is done by Chouteau *et al.* (2013) incorporating airborne time-domain electrical-magmetic data in North-West Abitibi of Québec in Canada.

REFERENCES

AHMED, S. and DE MARSILY, G. (1989). Co-kriged estimates of transmissivities using jointly water level data. M. Armstrong, editor, *Geostatistics*, Springer Netherlands, vol. 4 of *Quantitative Geology and Geostatistics*. 615–628.

ARMSTRONG, M., GALLI, A., BEUCHER, H., LOC'H, G.AND RENARD, D., DOLIGEZ, B., ESCHARD, R. and GEFFROY, F. (2011). *Plurigaussian Simulations in Geosciences*. Springer Berlin Heidelberg, Berlin, Heidelberg.

ASKEY, R. (1973). Radial characteristic functions. Technical report, University of Wisconsin.

ASLI, M., MARCOTTE, D. and CHOUTEAU, M. (2000). Direct inversion of gravity data by cokriging. *Geostatistics 2000, Cape Town, W.J. Kleingeld and D.G. Krige (eds), South Africa*. 64–73.

BAILLY, J. S., MONESTIEZ, P. and LAGACHERIE, P. (2006). Modelling spatial variability along drainage networks with geostatistics. *Mathematical Geosciences*, 38, 515–539.

BELLIER, E. and MONESTIEZ, P. (2010). A spatial covariance model with a single wave effect and a finite range. *Statistics and Probability Letters*, 80, 1343–1347.

BLOUIN, M., GLOAGUEN, E., BELLEFLEUR, G. and PUGIN, A. (2013). Characterizing seismic anisotropy in unconsolidated marine clay using 9c vsp data. *2013 SEG Annual Meeting*. Society of Exploration Geophysicists, 22-27 September, Houston, Texas.

BOCHNER, S. (1937). *Lectures by S. Bochner on Fourier Analysis, 1936-1937*. Edwards Brothers, Printers.

BOISVERT, J. and DEUTSCH, C. V. (2008). Shortest anisotropic path to reproduce complex geological features. *8th International Geostatistics Congress, Santiago, Chile*. vol. 27, 1041–1046.

BROCHU, Y. and MARCOTTE, D. (2003). A simple approach to account for radial flow and boundary conditions when kriging hydraulic head fields for confined aquifers. *Mathematical Geology*, 35, 111–139.

BUHMANN, M. D. (2003). *Radial basis functions : theory and implementations*, vol. 12. Cambridge university press.

CARON, R., SAMSON, C., CHOUTEAU, M. and BATES, M. (2013). Designing a htem system for mapping glaciolacustrine overburden thickness. *2013 SEG Annual Meeting*. Society of Exploration Geophysicists, 22-27 September, Houston, Texas.

CARRIER, M.-A., LEFEBVRE, R., RIVARD, C. and PARENT, M. (2013). Atlas du projet Montérégie Est, sud du Québec, Canada. Technical report Rapport final INRS R-1412.

CHILÈS, J. P. (1995). Quelques méthodes de simulation de fonctions aléatoires intrinsèques. *Cahiers de Géostatistique, Fasc*, 5, 97–112.

CHILÈS, J. P. and DELFINER, P. (1997). Discrete exact simulation by the fourier method. E. Baafi and N. Schofield, editors, *Geostatistics, Wollongong96*. 258–269.

CHILÈS, J. P. and DELFINER, P. (1999). *Geostatistics : Modeling Spatial Uncertainty*. John Wiley & Sons.

CHILÈS, J. P. and DELFINER, P. (2012). *Geostatistics : Modeling Spatial Uncertainty*. John Wiley & Sons.

CHOPIN, N. (2011). Fast simulation of truncated gaussian distributions. *Statistics and Computing*, 21, 275–288.

CHOUTEAU, M., BOUDOUR, Z., PARENT, M. and MARCOTTE, D. (2013). Estimation of overburden thickness using airborne time-domain em data and a few drill hole data. *13th SAGA Biennial Conference and Exhibition*. South African Geophysical Association, Mpumalanga, South Africa.

CHRISTAKOS, G. and PAPANICOLAOU, V. (2000). Norm-dependent covariance permissibility of weakly homogeneous spatial random fields. *Stoch Environ Res Risk Assess*, 14, 1–8.

COOLEY, J. W., LEWIS, P. A. and WELCH, P. D. (1969). The Fast Fourier Transform and its applications. *Education, IEEE Transactions on*, 12, 27–34.

CRESSIE, N. (1993). *Statistics for spatial data*. J. Wiley.

DAMIAN, D., SAMPSON, P. D. and GUTTORP, P. (2001). Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics*, 12, 161–178.

DAVIS, M. and GRIVET, C. (1984). Kriging in a global neighborhood. *Mathematical Geology*, 16, 249–265.

DE FOUQUET, C. and BERNARD-MICHEL, C. (2006). Modèles géostatistiques de concentrations ou de débits le long des cours d'eau. *Comptes Rendus Geoscience*, 338, 307–318.

DEUTSCH, C. V. and JOURNEL, A. G. (1998). *GSLIB-Geostatistical Software Library and User's Guide*. Oxford University Press.

DEUTSCH, J. L. and DEUTSCH, C. V. (2014). A multidimensional scaling approach to enforce reproduction of transition probabilities in truncated plurigaussian simulation. *Stochastic Environmental Research and Risk Assessment*, 28, 707–716.

DONG, A. (1990). *Estimation géostatistique des phénomènes régis par des équations aux dérivées partielles*. PhD Dissertation, Centre de géostatistique, Fontainebleau.

DUBRULE, O. and KOSTOV, C. (1986). An interpolation method taking into account inequality constraints : I. methodology. *Mathematical geology*, 18, 33–51.

EHM, W., GNEITING, T. and RICHARDS, D. (2004). Convolution roots of radial positive definite functions with compact support. *Transactions of the American Mathematical Society*, 356, 4655–4685.

EMERY, X. (2004a). Properties and limitations of sequential indicator simulation. *Stochastic Environmental Research and Risk Assessment*, 18, 414–424.

EMERY, X. (2004b). Testing the correctness of the sequential algorithm for simulating gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 18, 401–413.

EMERY, X. (2007a). Conditioning simulations of Gaussian random fields by ordinary kriging. *Mathematical Geology*, 39, 607–623.

EMERY, X. (2007b). Simulation of geological domains using the plurigaussian model : new developments and computer programs. *Computers & Geosciences*, 33, 1189–1201.

EMERY, X. (2007c). Using the gibbs sampler for conditional simulation of gaussian-based random fields. *Computers & Geosciences*, 33, 522–537.

EMERY, X. (2008a). Statistical tests for validating geostatistical simulation algorithms. *Computers & Geosciences*, 34, 1610–1620.

EMERY, X. (2008b). A turning bands program for conditional co-simulation of cross-correlated gaussian random fields. *Computers & Geosciences*, 34, 1850–1862.

EMERY, X. (2010). Multi-gaussian kriging and simulation in the presence of an uncertain mean value. *Stochastic Environmental Research and Risk Assessment*, 24, 211–219.

EMERY, X., ARROYO, D. and PELÁEZ, M. (2014). Simulating large gaussian random vectors subject to inequality constraints by gibbs sampling. *Mathematical Geosciences*, 46, 265–283.

EMERY, X., ARROYO, D. and PORCU, E. (2015). An improved spectral turning-bands algorithm for simulating stationary vector gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, in press.

EMERY, X. and LANTUÉJOUL, C. (2006). TBSIM : A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers & Geosciences*, 32, 1615–1628.

EMERY, X. and SILVA, D. A. (2009). Conditional co-simulation of continuous and categorical variables for geostatistical applications. *Computers & Geosciences*, 35, 1234–1246.

FREULON, X. and DE FOUQUET, C. (1993). Conditioning a gaussian model with inequalities. *Geostatistics Troia'92*, Springer. 201–212.

FUENTES, M. (2001). A high frequency kriging approach for non-stationary environmental processes. *Environmetrics*, 12, 469–483.

FUENTES, M. and SMITH, R. L. (2001). A new class of nonstationary spatial models. Technical report, North Carolina State University, Raleigh, NC.

FURRER, R., GENTON, M. G. and NYCHKA, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15, 502–523.

GALLI, A., BEUCHER, H., LE LOCH, G. and DOLIGEZ, B. (1994). The pros and cons of the truncated Gaussian method. *Geostatistical simulations*, Springer. 217–233.

GALLI, A. and GAO, H. (2001). Rate of convergence of the gibbs sampler in the gaussian case. *Mathematical Geology*, 33, 653–677.

GEMAN, S. and GEMAN, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-6, 721–741.

GLOAGUEN, E., MARCOTTE, D. and CHOUTEAU, M. (2005a). A non-linear gpr tomographic inversion algorithm based on iterated cokriging and conditional simulations. O. Leuangthong and C. Deutsch, editors, *Geostatistics Banff 2004*, Springer Netherlands, vol. 14 of *Quantitative Geology and Geostatistics*. 409–418.

GLOAGUEN, E., MARCOTTE, D., CHOUTEAU, M. and PERROUD, H. (2005b). Borehole radar velocity inversion using cokriging and cosimulation. *Journal of Applied Geophysics*, 57, 242–259.

GNEITING, T. (1998). Closed form solutions of the two-dimensional turning bands equation. *Mathematical Geology*, 30, 379–390.

GNEITING, T. (2002). Compactly supported correlation functions. *Journal of Multivariate Analysis*, 83, 493–508.

GOULARD, M. (1989). Inference in a coregionalization model. *Geostatistics*, Springer. 397–408.

GOULARD, M. and VOLTZ, M. (1992). Linear coregionalization model : Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology*, 24, 269–286.

GUARDIANO, F. B. and SRIVASTAVA, R. M. (1993). Multivariate geostatistics : beyond bivariate moments. *Geostatistics Troia92*, Springer. 133–144.

GUTTORP, P., MEIRING, W. and SAMPSON, P. D. (1994). A space-time analysis of ground-level ozone data. *Environmetrics*, 5, 241–254.

GUTTORP, P. and SAMPSON, P. D. (1994). Methods for estimating heterogeneous spatial covariance functions with environmental applications. *Handbook of Statistics*, 12, 661–689.

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

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

HAAS, T. C. (1995). Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *Journal of the American Statistical Association*, 90, 1189–1199.

HIGDON, D., SWALL, J. and KERN, J. (1999). Non-stationary spatial modeling. *Bayesian statistics*, 6, 761–768.

HORTA, A., CAEIRO, M. H., NUNES, R. and SOARES, A. (2010). Simulation of continuous variables at meander structures : Application to contaminated sediments of a lagoon. P. M. Atkinson and C. D. Lloyd, editors, *geoENV VII - Geostatistics for Environmental Applications*, Springer Netherlands, no. 16 Quantitative Geology and Geostatistics. 161–172.

HUNTER, J. A., PULLAN, S. E., BURNS, R. A., GAGNE, R. M. and GOOD, R. L. (1984). Shallow seismic reflection mapping of the overburden-bedrock interface with the engineering seismograph-some simple techniques. *Geophysics*, 49, 1381–1385.

JOURNEL, A. and ISAAKS, E. (1984). Conditional indicator simulation : application to a Saskatchewan uranium deposit. *Mathematical Geology*, 16, 685–718.

JOURNEL, A. G. (1974). Geostatistics for conditional simulation of ore bodies. *Economic Geology*, 69, 673–687.

JOURNEL, A. G. and HUIJBREGTS, C. J. (1978). *Mining Geostatistics*. Academic Press, London.

JUN, M. and STEIN, M. L. (2008). Nonstationary covariance models for global data. *The Annals of Applied Statistics*, 2, 1271–1289.

KITANIDIS, P. K. (1993). Generalized covariance functions in estimation. *Mathematical Geology*, 25, 525–540.

KLEIBER, W. and PORCU, E. (2015). Nonstationary matrix covariances : compact support, long range dependence and quasi-arithmetic constructions. *Stochastic Environmental Research and Risk Assessment*, 29, 193–204.

LANTUÉJOUL, C. and DESASSIS, N. (2012). Simulation of a gaussian random vector : A propagative version of the gibbs sampler. *The 9th International Geostatistics Congress*. Oslo, Norway, 174–181.

LANTUÉJOUL, C. (2002). *Geostatistical Simulation : Models and Algorithms*. Springer.

LE LOC'H, G. and GALLI, A. (1997). Truncated plurigaussian method : theoretical and practical points of view. *Geostatistics wollongong*, 96, 211–222.

LE RAVALEC-DUPIN, M. and DA VEIGA, S. (2011). Cosimulation as a perturbation method for calibrating porosity and permeability fields to dynamic data. *Computers & Geosciences*, 37, 1400–1412.

LE RAVALEC-DUPIN, M., HU, L. and NOTINGER, B. (2001). Stochastic reservoir modeling constrained to dynamic data : Local Calibration and Inference of Structural Parameters. *SPE Journal*, 6, 25–31.

LE RAVALEC-DUPIN, M., NOETINGER, B. and HU, L. (2000). The FFT moving average generator : An efficient numerical method for generating and conditioning Gaussian simulations. *Mathematical Geology*, 32, 701–722.

LI, S. (2011). Concise formulas for the area and volume of a hyperspherical cap. *Asian Journal of Mathematics and Probability Letters Statistics*, 4, 66–70.

LIANG, M., MARCOTTE, D. and BENOIT, N. (2014). A comparison of approaches to include outcrop information in overburden thickness estimation. *Stochastic Environmental Research and Risk Assessment*, 28, 1733–1741.

MANNSETH, T. (2014). Relation between level set and truncated pluri-gaussian methodologies for facies representation. *Mathematical Geosciences*, 46, 711–731.

MANTOGLOU, A. (1987). Digital simulation of multivariate two- and three-dimensional stochastic processes with a spectral turning bands method. *Mathematical Geology*, 19, 129–149.

MANTOGLOU, A. and WILSON, J. L. (1982). The Turning Bands Method for simulation of random fields using line generation by a spectral method. *Water Resources Research*, 18, 1379–1394.

MARCOTTE, D. (1991). Cokriging with Matlab. *Computers & Geosciences*, 17, 1265–1280.

MARCOTTE, D. (1995). Generalized cross-validation for covariance model selection. *Mathematical geology*, 27, 659–672.

MARCOTTE, D. (2015a). Spatial turning band simulation of anisotropic non linear models of coregionalization with symmetric cross-covariances. *Computers & Geosciences*. Manuscript submitted for publication.

MARCOTTE, D. (2015b). Tasc3d : A program to test the admissibility in 3d of non-linear models of coregionalization. *Computers & Geosciences*, 83, 168–175.

MARIETHOZ, G., RENARD, P., CORNATON, F. and JAQUET, O. (2009). Truncated plurigaussian simulations to characterize aquifer heterogeneity. *Groundwater*, 47, 13–24.

MATEU, J., FERNÁNDEZ-AVILÉS, G. and MONTERO, J. M. (2013). On a class of non-stationary, compactly supported spatial covariance functions. *Stochastic Environmental Research and Risk Assessment*, 27, 297–309.

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

MATHERON, G. (1965). *Les variables régionalisées et leur estimation*. PhD Dissertation, Faculté des Sciences, Université de Paris.

MATHERON, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5, 439.

MATHERON, G., BEUCHER, H., DE FOUQUET, C., GALLI, A., GUERILLOT, D. and RAVENNE, C. (1987). Conditional simulation of the geometry of fluvio-deltaic reservoirs. *paper SPE*, 16753, 22–25.

MAUS, S. (1999). Variogram analysis of magnetic and gravity data. *Geophysics*, 64, 776–784.

MEIRING, W., MONESTIEZ, P., SAMPSON, P. D. and GUTTORP, P. (1997). Developments in the modelling of non stationary spatial covariance structure from space-time monitoring data. Ernest Y. Baafi and Neil A. Schofield, editors, *Geostatistics Wollongong 96*, Springer. 162–173.

MEJÍA, J. M. and RODRÍGUEZ-ITURBE, I. (1974). On the synthesis of random field sampling from the spectrum : An application to the generation of hydrologic spatial processes. *Water Resources Research*, 10, 705–711.

MOREAUX, G. (2008). Compactly supported radial covariance functions. *Journal of Geodesy*, 82, 431–443.

MYERS, D. E. (1983). Estimation of linear combinations and co-kriging. *Mathematical Geology*, 15, 633–637.

OLIVER, D. S. (1995). Moving averages for gaussian simulation in two and three dimensions. *Mathematical Geology*, 27, 939–960.

OLIVER, D. S. (2003). Gaussian cosimulation : Modelling of the cross-covariance. *Mathematical Geology*, 35, 681–698.

ORTIZ, J. M. and DEUTSCH, C. V. (2004). Indicator simulation accounting for multiple-point statistics. *Mathematical Geology*, 36, 545–565.

PACIOREK, C. J. (2003). *Nonstationary Gaussian processes for regression and spatial modelling*. PhD Dissertation, Department of Statistics, Carnegie Mellon University.

PACIOREK, C. J. and SCHERVISH, M. J. (2004). Nonstationary covariance functions for gaussian process regression. *In Proc. of the Conf. on Neural Information Processing Systems (NIPS)*. MIT Press, Cambridge, 273–280.

PACIOREK, C. J. and SCHERVISH, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17, 483–506.

PARDO-IGÚZQUIZA, E. and CHICA-OLMO, M. (1993). The Fourier Integral Method : An efficient spectral method for simulation of random fields. *Mathematical Geology*, 25, 177–217.

PINTORE, A. and HOLMES, C. (2004). Non-stationary covariance functions via spatially adaptive spectra. Technical report, Department of Statistics, University of Oxford.

PORCU, E., DALEY, D. J., BUHMANN, M. and BEVILACQUA, M. (2013). Radial basis functions with compact support for multivariate geostatistics. *Stochastic Environmental Research and Risk Assessment*, 27, 909–922.

PRIESTLEY, M. (1982). *Spectral analysis and time series*. No. v. 1-2 Probability and mathematical statistics. Academic Press.

PRIESTLEY, M. B. (1965). Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 27, 204–237.

RAO, K. R., KIM, D. N. and HWANG, J. J. (2011). *Fast Fourier Transform-Algorithms and Applications*. Springer Science & Business Media.

RIMSTAD, K. and OMRE, H. (2014). Generalized gaussian random fields using hidden selections. *ArXiv e-prints*.

RIVEST, M. and MARCOTTE, D. (2012). Kriging groundwater solute concentrations using flow coordinates and nonstationary covariance functions. *Journal of Hydrology*, 472–473, 238–253.

RIVEST, M., MARCOTTE, D. and PASQUIER, P. (2008). Hydraulic head field estimation using kriging with an external drift : A way to consider conceptual model information. *Journal of Hydrology*, 361, 349–361.

ROBERTS, G. O. and SAHU, S. K. (1997). Updating schemes, correlation structure, blocking and parameterization for the gibbs sampler. *Journal of the Royal Statistical Society. Series B (Methodological)*, 291–317.

ROYER, J. J. and VIERRA, P. C. (1984). Dual formalism of kriging. V. G., D. M., J. A. G. and M. A., editors, *Geostatistics for natural resources characterization Part 2*. NATO-ASI, 691–702.

RYGAARD-HJALSTED, C., CONSTABLE, C. G. and PARKER, R. L. (1997). The influence of correlated crustal signals in modelling the main geomagnetic field. *Geophysical Journal International*, 130, 717–726.

SAMPSON, P. D. and GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association*, 87, 108–119.

SANSÒ, F. and SCHUH, W.-D. (1987). Finite covariance functions. *Bulletin Géodésique*, 61, 331–347.

SCHABACK, R. (1995). Creating surfaces from scattered data using radial basis functions. *Mathematical methods for curves and surfaces*, 477–496.

SCHMIDT, A. M. and O'HAGAN, A. (2003). Bayesian inference for non-stationary spatial covariance structure via spatial deformations. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 65, 743–758.

SHAMSIPOUR, P., CHOUTEAU, M. and MARCOTTE, D. (2011). 3D stochastic inversion of magnetic data. *Journal of Applied Geophysics*, 73, 336–347.

SHAMSIPOUR, P., MARCOTTE, D., CHOUTEAU, M. and KEATING, P. (2010). 3D stochastic inversion of gravity data using cokriging and cosimulation. *Geophysics*, 75, I1–I10.

SHAMSIPOUR, P., MARCOTTE, D., CHOUTEAU, M., RIVEST, M. and BOUCHEDDA, A. (2013). 3D stochastic gravity inversion using nonstationary covariances. *Geophysics*, 78, G15–G24.

SHINOZUKA, M. (1971). Simulation of multivariate and multidimensional random processes. *J. Acoustical Society of America*, 49, 357–368.

SHINOZUKA, M. and JAN, C. M. (1972). Digital simulation of random processes and its applications. *Journal of Sound and Vibration*, 25, 111–128.

SMITH, R. (1996). Estimating nonstationary spatial correlations. *Preprint, University of North Carolina*, 76.

SOARES, A. (2010). Geostatistical methods for polluted sites characterization. P. M. Atkinson and C. D. Lloyd, editors, *geoENV VII - Geostatistics for Environmental Applications*, Springer Netherlands, no. 16 Quantitative Geology and Geostatistics. 187–198.

STEIN, M. L. (2005). Nonstationary spatial covariance functions. *Technical report, University of Chicago, Center for IntegratingStatistical and Environmental Science*.

STREBELLE, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34, 1–21.

THRUSH, P. W. (1968). A dictionary of mining, mineral and related terms.

VER HOEF, J. M., PETERSON, E. and THEOBALD, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological statistics*, 13, 449–464.

VERA, J. F., MACÍAS, R. and ANGULO, J. M. (2008). Non-stationary spatial covariance structure estimation in oversampled domains by cluster differences scaling with spatial constraints. *Stochastic Environmental Research and Risk Assessment*, 22, 95–106.

VERA, J. F., MACÍAS, R. and ANGULO, J. M. (2009). A latent class MDS model with spatial constraints for non-stationary spatial covariance estimation. *Stochastic Environmental Research and Risk Assessment*, 23, 769–779.

WACKERNAGEL, H. (2003). *Multivariate Geostatistics, 2nd edition*. Springer Science & Business Media.

WENDLAND, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in computational Mathematics*, 4, 389–396.

WENDLAND, H. (1998). Error estimates for interpolation by compactly supported radial basis functions of minimal degree. *Journal of Approximation Theory*, 93, 258 – 272.

WU, Z. (1995). Compactly supported positive definite radial functions. *Advances in Computational Mathematics*, 4, 283–292.

XU, C., DOWD, P. A., MARDIA, K. V. and FOWELL, R. J. (2006). A flexible true plurigaussian code for spatial facies simulations. *Computers & geosciences*, 32, 1629–1645.

YAO, T. and JOURNEL, A. G. (1998). Automatic modeling of (cross) covariance tables using fast fourier transform. *Mathematical Geology*, 30, 589–615.

ZAGAYEVSKIY, Y. (2015). *Multivariate Grid-Free Geostatistical Simulation of Natural Phenomena*. PhD Dissertation, University of Alberta, Edmonton.