Titre: Title:	3D stochastic inversion of gravity data using cokriging and cosimulation
Auteur: Author:	Pejman Shamsipour
Date:	2008
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Shamsipour, P. (2008). 3D stochastic inversion of gravity data using cokriging and cosimulation [Master's thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/8364/

UNIVERSITÉ DE MONTRÉAL

3D STOCHASTIC INVERSION OF GRAVITY DATA USING COKRIGING AND COSIMULATION

PEJMAN SHAMSIPOUR
DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE MINÉRAL)
AOÛT 2008

Your file Votre référence
ISBN: 978-0-494-46080-1
Our file Notre référence
ISBN: 978-0-494-46080-1

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

3D STOCHASTIC INVERSION OF GRAVITY DATA USING COKRIGING AND COSIMULATION

présenté par: SHAMSIPOUR Pejman
en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées
a été dûment accepté par le jury d'examen constitué de:

M. JI Shaocheng, Ph.D, président
M. CHOUTEAU Michel, Ph.D, membre et directeur de recherche
M. MARCOTTE Denis, Ph.D, membre et codirecteur de recherche
M. KEATING Pierre, Ph.D, membre et codirecteur de recherche
M. PILKINGTON Mark, Ph.D, membre

To Farnaz

ACKNOWLEDGMENTS

I would like to express my gratitude to all those who gave me this possibility to complete my thesis. Foremost, I would like to express my thanks to my supervisor, professor Michel Chouteau, for giving me this opportunity to work in his group, for his support and encouragement throughout my research not only as a professor but also as a close friend. Without his guidance and teaching, I would not be able to complete this research.

I wish to express sincere thanks to professor Denis Marcotte, my co-supervisor, who is the best reference for me in Geostatistic forever.

I would like to thank Piere Keating, for his constructive advice that improved my work.

I would like to thank all members of my master defence: professor Mark pilkington and professor Shaocheng Ji for their supportive efforts, valuable comments and all the time they spent to read this thesis.

I wish to thank Research Affiliate Program (RAP) for financial support and Michel Allard from Xstrata zinc company for his guidances and supports. I also want to thank all current and previous members of applied geophysics laboratory of *École polytechnique de Montréal* for their helps and comments. Especially, I would like to thank my friends Abderrezak Bouchedda, Bernard Giroux, Martine Rivest and Jérémie Gaucher. My thanks also go to my beloved mother, Parvaneh, and my father, Hassan. I only hope I have made their trials and support worthwhile. I hope this thesis would fulfill part of my father's wishes for me.

RÉSUMÉ

Le but de cette thèse est de présenter une méthode d'inversion gravimétrique 3D basée sur des approches géostatistiques (le cokrigeage et la simulation conditionnelle).

Dans l'approche proposée, la linéarité entre la densité et la gravité nous permet d'estimer la matrice de covariance des densités à partir des covariances calculées sur les données de gravité.

Le cokrigeage est une méthode d'estimation qui exploite la corrélation spatiale entre une variable secondaire (ici, la gravité) et une variable principale (ici, la densité) afin d'estimer la variable principale en minimisant la variance d'estimation.

Pour connaître la variabilité que peuvent prendre nos densités estimées en fonction du modèle de covariance retenu, nous avons utilisés des simulations géostatistiques. Pour ce faire, l'algorithme de simulation FFT-MA (Fast Fourier Transform - Moving Average) a été utilisé. Celui-ci permet de générer très rapidement, à partir de la matrice de covariance, des simulations non conditionnelles de processus gaussiens non stationnaires, ce qui peut être très avantageux pour des problèmes de grande taille comme l'inversion 3D. Le conditionnement aux données de gravité est ensuite réalisé par cokrigeage.

La méthode proposée a été testée tout d'abord sur deux modèles synthétiques. Le premier consiste en un dyke incliné. Le deuxième modèle est généré par simulation géostatistique. Les résultats obtenus montrent que les approches géostatistiques permettent de réduire les artéfacts dus à la perte de résolution en profondeur, présentent une meilleure robustesse aux bruits et une intégration facile et rapide de contraintes géologiques.

Finalement, notre approche a été appliquée sur des données réelles d'une campagne gravimétrique réalisée par la commission géologique du Canada sur le flanc sud du camp minier de Matagami au nord du Québec. Pour obtenir la carte de l'anomalie

gravimétrique résiduelle, le filtrage par l'opérateur du prolongement vers le haut a été utilisé. Plusieurs approches ont été testées afin de déterminer la profondeur optimale pour cet opérateur. La distribution des densités de la région étudiée est tout d'abord obtenue par cokrigeage. La cosimulation est ensuite utilisée pour générer des cartes de probabilité afin de faciliter l'interprétation. Les résultats ainsi obtenus corroborent bien l'information géologique disponible.

ABSTRACT

The purpose of this thesis is to present an inversion method based on a geostatistical approach (cokriging and conditional simulation) for three dimensional inversion of gravity data including geological constraints.

Cokriging is a method of estimation that minimizes the error variance by applying cross-correlation between several variables. In this study the estimates are derived using gravity data as a secondary variable and the density as the primary variable. In the proposed method, the linearity between gravity and density allows us to obtain a covariance matrix of densities using observed data, i.e, we adjust the density covariance matrix by fitting experimental and theoretical gravity covariance matrices.

To obtain various reasonable solutions in order to see the variability that can be expected from the density covariance model adopted, a geostatistical simulation algorithm is applied. The simulation algorithm used in this thesis is based on the FFT moving average (FFT-MA) generator. Then the simulations are conditioned using cokriging results.

The proposed method is applied to two different synthetic models: 1) the dipping dyke; 2) a stochastic distribution of densities. Then some geological information is added as constraints to the system of cokriging. The results show the ability of the method in fast integration of complex a priori information in the form of covariance functions. The proposed method helps us to modify the lack of resolution at depth and reduce the sensitivity to noise. Increasing the amount of information as constraints also helps to improve the estimation of the density distribution especially at deeper depths.

Finally, the southwest flank of the Matagami mining camp is considered as real data. The best height for upward continuation is studied for generating the residual map. Then our inversion method based on cokriging is applied to these residual

anomalies in order to estimate the density distribution in this region. The co-simulation map is presented and the probability map is plotted in order to have a better interpretation. The results of inversion and simulation methods are in good agreement with the geology of the studied region.

CONDENSÉ EN FRANÇAIS

Introduction

Pour des raisons théoriques et algébriques, l'inversion gravimétrique tridimensionnelle produit malheureusement une infinité de solutions de distribution de densité pour une anomalie donnée.

Beaucoup de travaux ont été effectués ces dernières années pour limiter cette ambiguïté, notamment, en ajoutant des contraintes pour diminuer le nombre de solutions possibles et pour que celles-ci soient les plus réalistes possible. Dans cette étude, nous proposons une approche d'inversion basée sur les méthodes géostatistiques. Il existe de nombreuses applications de la géostatistique en géophysique (Asli et al. (2000); Gloaguen et al. (2005, 2007); Chasseriau et Chouteau (2003) et Giroux et al. (2007)) . Franklin (1970) a été le premier à introduire l'approche stochastique pour la résolution de problémes inverses. Asli et al. (2000) ont utilisé le cokrigeage des densités et de l'anomalie de Bouguer en inversion gravimétrique.

L'approche proposée est basée sur le calcul de la matrice de sensibilité et utilise la linéarité gravité-densité. Cette linéarité permet de modéliser la matrice de covariance des densités à partir des covariances calculées sur les données gravimétriques. Ces matrices de covariance permettent d'effectuer le cokrigeage du champ de densité. Cependant, le cokrigeage ne donne accès qu'à la valeur moyenne des densités. Les simulations conditionnelles permettent de générer des solutions ayant des variations spatiales plus réalistes. De plus, elles permettent d'identifer les caractéristiques stables parmi les champs simulés.

Les zones de densités et/ou de gradients connus, par exemple le long des forages, peuvent être ajoutées comme des contraintes. Ces contraintes sont très faciles à

implanter dans les différents algorithmes utilisés dans cette étude.

Méthodologie proposée

Problème direct en gravimétrie

En inversion gravimétrique tridimensionnelle, il est pratique de subdiviser préalablement la subsurface en blocs auxquels seront assignées des valeurs de densité (paramètres à rechercher). Pour l'instant, on peut seulement mesurer la composante verticale de l'attraction gravitationnelle due aux blocs. Haáz (1953) a publié des solutions pour l'attraction d'un prisme élémentaire de densité ρ à la distance r_{ijk}.

$$g_z = -\Gamma\rho\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\mu_{ijk}\left[x_i ln(y_j + r_{ijk}) + y_i ln(x_j + r_{ijk}) - z_k arctan(\frac{x_i y_j}{z_k r_{ijk}})\right]$$

Pour un prisme dont les dimensions sont: $x_i = x - \xi_i$, $y_j = y - \eta_j$, $z_k = z - \zeta_k$
$i, j, k = 1, 2$
$r_{ijk} = \sqrt{x_i^2 + y_j^2 + z_k^2}$
$\mu_{ijk} = (-1)^i(-1)^j(-1)^k$
Comme l'une des propriétés du champ de pesanteur est d'être additif, l'attraction totale en un point quelconque peut s'écrire comme l'addition des attractions dues à chaque prisme. Ainsi, pour n prismes le champ g_z et donné par:

$$g_z(x, y, z) = \sum_{i=1}^{m} g_i(x, y, z)$$

En considérant qu'il existe n observations de la gravité et m prismes rectangulaires, la relation précédente peut également être écrite sous forme matricielle :

$$g_{n\times 1} = G_{n\times m}\rho_{m\times 1}$$

où: G est la matrice de sensibilité (ou matrice jacobienne).

La méthode proposée

Comme nous l'avons vu, la gravité et la densité sont reliées de façon linéaire. Leurs covariances le sont donc aussi. Nous supposons que l'erreur est de moyenne zéro et est non corrélée avec les données de gravité, i.e. $E[e] = 0$ et $C(g, e) = 0$. Nous supposons également que $E[\rho] = 0$ et $E[g] = 0$, c'est-à-dire que nous traitons les contrastes de gravité et de densité. La matrice de covariance entre les donnés est (Gloaguen et al., 2005):

$$C(g, g) = GC(\rho, \rho)G^T + C_0$$

où $cov(\rho, \rho)$ est la matrice de covariance des densités et C_0 est une matrice diagonale représentant l'erreur sur les données (C_0 est appelée l'effet pépite sur les données). La covariance est maximale a la distance zéro. La façon dont la covariance décroît refléter le degré de ressemblance entre les données en fonction de la distance qui les sépare. La distance a laquelle la covariance devient nulle est appelée la portée. Au-delà. de la portée, les échantillons ne sont plus corrélés. Les modèles admissibles de covariance sont discutés dans (Chilès et Delfiner, 1999). De manière similaire,

$$C(g, \rho) = GC(\rho, \rho)$$

La matrice de covariance des données de gravité est caractérisée par le choix du modéle et de ses paramètres (effet pépite, variance, portée, anisotropie). Une fois que le modéle est sélectionné, les paramètres sont ajustés par une méthode d'optimisation semi-automatique dans l'espace de dimension la taille des paramètres du modèle de covariance. l'algorithme suivant montre les étapes de calcul de la matrice de covariance des paramètres : (Asli et al., 2000):

i. Calculer G (la matrice de sensibilité)

ii. Centrer les données

iii. Calculer la matrice de covariance théorique et expérimentale ($Cgg_{exp} = g \times g^T$ et $Cgg_{th} = GC_{\rho\rho}G^T$)

iv. Transformer les deux matrices en vecteurs v_{th} et v_{exp}

v. Trier v_{th} dans l'ordre décroissant et appliquer le même ordre pour v_{exp}

vi. Classer les vecteurs v_{th} et v_{exp} dans N décalages et calculer la moyenne pour chaque décalage

vii. Minimiser l'erreur pour chaque décalage avec la relation $err[j] = v_{th}[j] - v_{exp}[j]$

Le cokrigeage

Le cokrigeage (Chilès et Delfiner, 1999) est un outil mathématique d'interpolation et d'extrapolation qui utilise la corrélation spatiale entre une variable secondaire (ici, les données de gravité) et une variable principale (ici, les densités) afin d'estimer la variable principale à des endroits non échantillonnés. Le cokrigeage alloue des poids aux données qui minimisent la variance d'estimation. Les poids de cokrigeage simple λ sont donnés par:

$$\lambda = C(g,g)^{-1}C(g,\rho)$$

et:

$$\rho^* = \lambda^T g$$

Les zones de densités et/ou de gradients connus, par exemple le long des forages, peuvent être ajoutées comme des contraintes. Ces contraintes sont très faciles à implanter dans les différents algorithmes utilisés dans cette étude.

Co-Simulation

Par construction, le cokrigeage donne un estimé lissé des densités. Il est parfois souhaitable de connaître les différentes valeurs que peuvent prendre nos estimés en fonction du modèle de covariance retenu. Cela peut être obtenu par simulation géostatistique. Il existe de nombreux algorithmes de simulation (Chilès et Delfiner, 1999). L'algorithme de simulation FFT-MA (Fast Fourier Transform Moving Average) permet de générer très rapidement des simulations non conditionnelles de processus gaussiens stationnaires (Le Ravalec2000). Le conditionnement aux données observées g est ensuite réalisé par cokrigeage. L'algorithme suivant résume les étapes de la méthode de cosimulation(Gloaguen et al., 2005):

i. Calculer G .

ii. Centrer les données

iii. Estimer le modèle de covariance des densités.

iv. Effectuer les simulations non-conditionnelles de densités ρ_s

v. Calculer la gravité simulée aux points mesurés: $g_s = G\rho_s + C_0$

vi. Effectuer les deux cokrigeage des densités avec les donnes mesurées et simulées

vii. Conditionner les simulations par cokrigeage :$\rho_{sc} = \rho^* + (\rho_s - \rho_s^*)$

Modélisation

Nous avons testé les méthodes d'inversion basées sur le cokrigeage et la cosimulation à deux modèles synthétiques:

- un dyke incliné

- un modèle généré par simulation non-conditionnelle étant donné un modèle de variogramme

Les résultats obtenus montrent que les approches géostatistiques permettent de réduire les artéfacts dus à la perte de résolution en profondeur, présentent une meilleure robustesse aux bruits et une intégration facile et rapide de contraintes géologiques. Finalement, les algorithmes développés dans cette thèse ont été appliqués sur des données réelles du camp minier de Matagami(Québec). Le résultat obtenu a été comparé avec l'information géologique connue.

Conclusion

- Les deux algorithmes sont très simples et rapides (non itératif)
- Ils peuvent être appliqués sur des données échantillonnées sur une grille régulière ou non-régulière.
- Le cokrigeage est robuste en présence de bruit.
- L'ajout de contraintes permet d'améliorer l'estimation de la densité, plus particulièrement en profondeur.
- Dans le cas d'un problème de grande taille, la matrice de covariance des paramètres n'a pas besoin d'être calculée explicitement. De plus, les produits du type GKG^T et KG^T peuvent être calculés très rapidement en profitant de la structure particulière (bloc Toeplitz) de la matrice K (Nowaks et al., 2003).
- La méthode proposée est applicable à tout problème inverse linéaire, qu'il soit 2D ou 3D.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

$FFT\text{-}MA$	fast Fourier transform moving average
ST	symmetric Toeplitz matrix
SC	symmetric circulant matrix
CGA	conjugate gradient algorithm
SVD	singular value decomposition
LU	Cholesky decomposition
MAE	mean absolute error
VMS	volcanogenic massive sulfide
$DEXP$	depth from extreme points
SI	structural index
Λ	diagonal matrix of eigenvalues
ρ	density $(kg/m^3 right)$
ρ^*	estimated density
g	gravity $(mGal)$
e	error
γ	Newton's gravitational constant $6.67 \times 10^{11} (N.m/kg^2)$
G	matrix of geometric (sensitivity matrix)
C	covariance function
C_0	nugget effect
$Z(x)$	random function
$\gamma(h)$	variogram
σ^2	variance
σ_k^2	cokriging variance
$E[.]$	expectation function
λ, μ, η	cokriging coefficients

gr	density gradient
F	index matrix of known densities
H	index matrix of known density gradients
n_x, n_y, n_z	number of prisms in x, y, z directions
$F[.]$	3D fast Fourier transform
$F^H[.]$	inverse 3D fast Fourier transform
L	lower triangular matrix
$S(f)$	forward Fourier transform of C
a	variogram range
z_0	depth to source
S_n	structural index

CHAPTER 1

INTRODUCTION

1.1 Introduction

One of the most significant current topics in geophysics is the inverse problem. Inversion of geophysical data is often an ill-posed problem, which means that solutions might not depend continuously on the data. Although a solution that satisfies the observed data can easily be found, there is still the problem of non-uniqueness which is caused by the nature of the physics and the underdetermination of the problem.

The purpose of this thesis is to present an inversion method based on a geostatistical approach (cokriging and conditional simulation) for three dimensional inversion of gravity data including geological constraints. Cokriging is a method of estimation that minimizes the error variance by applying the cross-correlation between several variables. In this case the estimates are derived using the gravity data as secondary variables. In the proposed method, the covariance matrix of densities is determined using observed data. This means that we adjust the density covariance matrix by fitting experimental to theoretical gravity covariance matrices.

The proposed method is applied to synthetic data for two different synthetic models: 1) the dipping dyke; 2) a stochastic distribution of densities. Then some geological information is added as constraints to the system of cokriging. The use of known gradients and densities as constraints is new in this approach and can be useful in the mining industry if a 3D geological model can be built with a GIS.

The geostatistical method allows for fast integration of complex a priori information in the form of covariance functions and training images. The proposed method helps to modify the lack of resolution at depth and reduce the sensitivity to noise. It is non-iterative and hence computationally efficient compared to other inversion methods. It is also simple and flexible.

Stochastic inversion based on cokriging gives a smooth estimate of the density distribution. Using geostatistical simulation algorithms, we can have various reasonable solutions showing the kind of variability that can be expected from the density covariance model. The co-simulation algorithm that is used in this thesis is based on the FFT moving average (FFT-MA) generator. Although, other simulation methods could be used as well.

Finally, the proposed method is applied for the real data. As a case study, we consider the survey data (Bouguer anomaly) from Matagami area in Québec. The studied area is located in the northern part of the Abitibi sub-province, one of the largest Archean greenstone belts in the world.

1.2 Literature Review

Generally, in inverse problems, we want to describe measured data d by finding model parameters m characterizing some physical process which reduces to solving the system:

$$Gm + e = d \tag{1.1}$$

Where G is a forward mapping operator (linear or nonlinear) that maps the model parameters m into observations d and e is the error in the measured data.

In 3D gravity inversion, there are different ways that the model can be defined.

One flexible way is to describe the model by a grid of prismatic cells. In this thesis, we use the approach of a grid of prismatic cells (Haáz, 1953) because of its great flexibility. In this technique, the subsurface is divided into prisms of known sizes and positions and the density contrasts are the parameters to estimate.

When the number of parameters is larger than the number of observations at surface, the system does not provide enough information to determine uniquely all model parameters. In this situation, the problem is said to be underdetermined. Besides, by defining the gravity field by Gauss's theorem (Blakely, 1995), there are no assumptions about the shape of the sources or distribution of density, hence theoretically the approach causes the ambiguity as well. As a result, although a solution that satisfies the observed data can easily be found, there is still the problem of non-uniqueness which is caused by the nature of the physics and the underdetermination of the problem.

Many strategies can be used to deal with the non-uniqueness problem in gravity inversion. They all involve some kind of constraints to limit the resulting solution. Green (1975) uses an appropriate weighting matrix to fix some of the parameters when geological or density information are available. Last and Kubik (1983) applied a compact solution with a minimum volume constraint. Smoothness or roughness of density distribution which control gradients of parameters in spatial directions has been used in magnetic inversion by Pilkington (1997). Li and Oldenburg (1998) counteract the decreasing sensitivities of cells with depth by weighting them with an inverse function of depth. Another 3D inversion technique allowing definition of depth resolution is proposed by Fedi and Rapolla (1999). Prior information in the form of parameter covariances can be included (Tarantola and Valette, 1982) to orient the search for a solution. Montagner and Jobert (1988) used exponential covariance functions in which the rate of exponential decay determines the correlation length of the parameters.

Geostatistical applications in geophysical inversion were applied by Asli et al. (2000), Gloaguen et al. (2005, 2007) and Giroux et al. (2007). Linear stochastic inversion was first described by Franklin (1970). Asli et al. (2000) cokriged gravity anomalies to obtain cell densities. Chasseriau and Chouteau (2003) proposed the method for of 3D inversion of gravity data using a model of covariance.

The covariance model also allows stochastic imaging of inverted fields by geostatistical simulations. There exist many efficient simulation algorithms (Chilès and Delfiner, 1999). The Fast Fourier Transform Moving Average simulation (FFT-MA) is a fast simulation algorithm for generating regular grid nonconditional Gaussian stationary processes (Le Ravalec et al., 2000).

The proposed method in this thesis is stochastic 3D inversion using cokriging and co-simulation.

1.3 Thesis Plan

This thesis has been organized and divided into six chapters. The first chapter gives a brief overview of the thesis (introduction) and then literature review of the subject. In the second chapter, at first we describe the forward modeling in the gravity method and then we explain some issues about the geometry of the domain which includes padding, optimal cell size and kernel calculation with less memory.

In chapter 3, we explain the main theoretical concepts of the whole thesis. At first, we explain the concept of cokriging and then we introduce a stochastic inversion approach based on cokriging. We introduce the v-v plot (Asli et al., 2000) method that is used to estimate the covariance matrix of parameters. We also explain how to add some constraints to our cokriging system to improve the performance of inversion and consequently parameter estimation. Then, we present

a method for efficient calculation of the gravity-density covariance matrix and the gravity-gravity covariance matrix which is based on circulant embedding and the fast Fourier transform (FFT)(Nowak et al., 2003). In this chapter, we also describe co-simulation based on the FFT moving average (FFT-MA) generator. We use this method as a geostatistical simulation algorithm to obtain various reasonable solutions in order to see the variability that can be expected from the density covariance model adopted.The conditioning is done by cokriging(Chilès and Delfiner, 1999).

In chapter 4, we apply the proposed inversion method based on cokriging to two different synthetic models. We use a dipping dyke model and then a synthetic model generated by non-conditional LU simulation with a specific variogram model. In both cases, co-simulation based on FFT-MA is applied as well.

In chapter 5, we consider the Matagami region in north-west of Quebec. Based on the gravity measurements surveyed from this region, we apply upward continuation to these anomalies in order to extract the residual anomalies. Then our inversion method based on cokriging is applied to these residual anomalies in order to estimate the density distribution in this region. Then, co-simulation based on FFT-MA is applied as well. Based on the results of inversion and also co-simulation, we try to interpret the geology in this region using the estimated and simulated densities.

Finally the last chapter includes conclusions of this thesis.

CHAPTER 2

THREE-DIMENSIONAL GRAVITY MODELING

The main purpose in applied geophysics is to detect changes in the physical properties of rocks lying beneath the surface. One of the physical properties which changes from one rock type to another is density. The aim of gravity surveying is to measure variations in the Earth's gravitational field caused by changes in the density of subsurface rocks. Gravity surveys have been used in the investigation of wide range of scale such as tectonic studies, mineral exploration, engineering and environmental problems, oil and gas exploration.

The gravity method is one of the oldest geophysical methods. Its theory is well established and the acquisition of data can be relatively simple and low cost. However, many earth scientists do not feel confident with this method because of the non-uniqueness of possible interpretations.

2.1 Forward Modeling

The purpose of forward modeling is to compute the gravimetric response g at the surface due to a density distribution in the sub-surface ρ.

$$g = G\rho \tag{2.1}$$

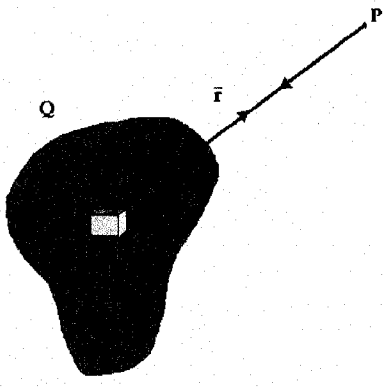

Figure 2.1 The magnitude of gravitational attraction of a mass (Q) on a unit mass (P) at distance r

where G is the matrix of the geometric terms (kernel).

Using Newton's law, the magnitude of attraction of a mass (Q) on a unit mass (P) at distance r is given by:

$$g(P) = \gamma\rho(Q)\int_v \frac{\hat{r}}{r^2}dv \tag{2.2}$$

This process has been shown in Figure 2.1 where γ is Newton's gravitational constant ($6.67 \times 10^{11}(N.m^2.Kg^{-2})$), ρ is density and dv is a volume element.

Considering the two coordinate systems (ξ, η, ζ) for the densities and (x, y, z) for the field (z is oriented downward), the vertical component g_z at point (x, y, z) due to a 3D body of density $\rho(\xi, \eta, \zeta)$ and volume v is given by:

$$g_z(x, y, z) = -\gamma\int\int\int \rho(\xi, \eta, \zeta)\frac{z-\zeta}{r^3}d\xi d\eta d\zeta \tag{2.3}$$

The most common method of evaluating the above equation is to break the 3D domain down into geometrically simple bodies having constant density. In our case, and for the sake of simplicity, the domain studied is divided into a finite number of rectangular prisms of uniform density. It should be noted this classic method can consume memory space for large set of observations. However, using

the symmetry of the kernel helps to minimize the amount of memory. In this particular example, several equivalent analytical forms were proposed for equation 2.3 (see (Li and Chouteau, 1998) for a critical review of different formulas). Following the formula of Haáz (1953), we have:

$$g_z = -\gamma\rho \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \mu_{ijk}\left[x_i ln(y_j + r_{ijk}) + y_i ln(x_j + r_{ijk}) - z_k arctan(\frac{x_i y_j}{z_k r_{ijk}})\right] \tag{2.4}$$

With,

$x_i = x - \xi_i$, $y_j = y - \eta_j$, $z_k = z - \zeta_k$ $i, j, k = 1, 2$

$r_{ijk} = \sqrt{x_i^2 + y_j^2 + z_k^2}$

$\mu_{ijk} = (-1)^i(-1)^j(-1)^k$

The response at point (x, y, z), of all the prisms making up the body, is the sum of the contribution of each prism:

$$g(x, y, z) = \sum_{i=1}^{m} g_i(x, y, z) \tag{2.5}$$

Considering that there are n gravity observations and m rectangular prisms, the preceding relationship can also be written in the matrix form:

$$g_{n\times 1} = G_{n\times m}\rho_{m\times 1} \tag{2.6}$$

With G, the matrix of the geometric terms.

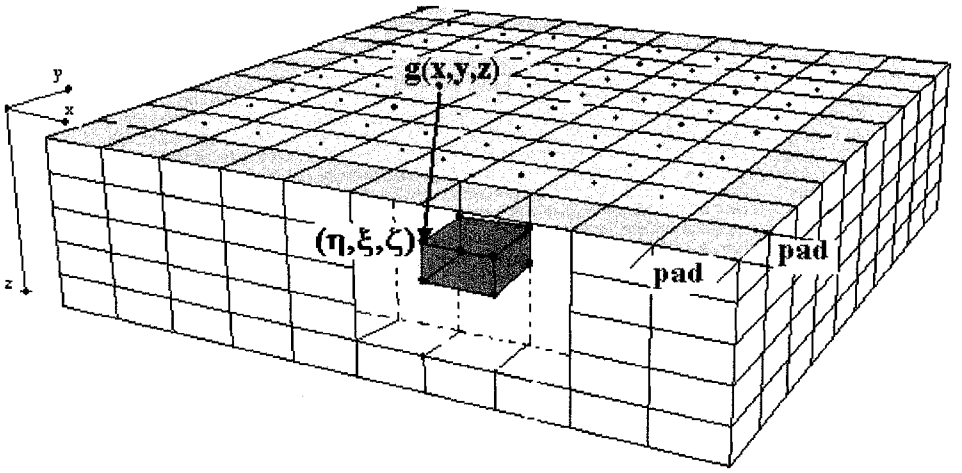

Figure 2.2 Discretization of the subsurface with prisms. Gravity stations $g(x, y, z)$ are located at the center of the upper face of prisms in the top layer. The stations lie on a horizontal grid at a constant elevation $z = 0$. n_x and n_y are the number of observations in x and y directions respectively, and $n_{cx} = n_x + 2pad$, $n_{cy} = n_y + 2pad$, n_{cz} the number of cells in x, y and z directions. pad indicates the cells added around the domain of observation and they are shown with darker color.

2.2 Geometry of domain

Even though this chapter deals with forward modeling, we prefer to introduce two important issues with regards to the geometry of domain here that are crucial in inversion problems. These issues are padding and optimal cell size and they have been investigated by Boulanger and Chouteau (2001) in detail.

2.2.1 Padding

There is a method to obtain automatically the dimensions of the inversion domain in the x and y directions extending the gravity data grid with cells around (padding). This method helps us to avoid possible distortion along the boundary. Based on this method, we can calculate the number of cells to add around the domain which we refer to it as *pad*.

Considering the extra cells, the number of cells in x and y directions is equal $nx + pad$ and $ny + pad$ respectively, where nx and ny are the number of cells in x and y directions, and nz is the number of cells in z direction. The details can be found in chapter 5.

2.2.2 Optimal cell size

One of the important problems in inversion is the determination of the best cell size for a given problem. Two competing criterion have to be considered:

- Short wavelengths present in the observed data must be modeled by considering sufficiently small cell sizes.

- Considering sufficiently large cell sizes such that the computation time be reasonable.

Boulanger and Chouteau (2001)proposed that the maximum cell size which is able to fit the shortest wavelength in the data should be less than $1.2dx$ where dx is the distance between gravity observations.

2.3 Symmetry of kernel and less memory

Boulanger and Chouteau (2001) showed for a given station on the surface, the kernel is symmetric with respect to the vertical z axis. Based on his approach, we only need to calculate the first row of the matrix G. The rest of the elements of G_{ij} can be determined by choosing the appropriate value of j. G_{1j} is stored in a three dimensional array $G(ig, jg, kg)$ where:
$ig = \{1, 2, ..., nx + pad\}$, $jg = \{1, 2, ..., ny + pad\}$ and $kg = \{1, 2, ..., nz\}$.
The observation indices are:
$io = \{1, 2, ..., nx + pad\}$, $jo = \{1, 2, ..., ny + pad\}$
and the parameter indices are:
$ip = \{1, 2, ..., nx + pad\}$, $jp = \{1, 2, ..., ny + pad\}$, $kp = \{1, 2, ..., nz\}$.

Five loops operating on jo, io, kp, jp, ip indices respectively are used to determine the indices
$ig = |(ip - pad) - io| + 1$, $jg = |(jp - pad) - jo| + 1$ and $kg = kp$
which call the appropriate value of sensitivity. To store $G(ig, jg, kg)$, the required memory is a $(1 \times M)$ vector with $M = (nx + pad) \times (ny + pad) \times nz$ instead of a $(N \times M)$ matrix.

CHAPTER 3

METHODOLOGY OF STOCHASTIC INVERSION BY COKRIGING AND CO-SIMULATION

The purpose of this chapter is to describe the methodologies that are used in this thesis. In this chapter we shall explain the main theoretical concepts of the whole thesis. At first, the concept of cokriging is explained and then a stochastic inversion approach based on cokriging is introduced. In this method we introduce the v-v plot method that is used to estimate the covariance matrix of the parameters. We also explain how to add some constraints to our cokriging system to improve the performance of inversion and consequently parameter estimation. Then, we present a method for efficient calculation of the gravity-density covariance matrix and the gravity-gravity covariance matrix which is based on circulant embedding and the fast Fourier transform (FFT). Finally, co-simulation based on the FFT moving average (FFT-MA) generator is described. Then, we use this method as a geostatistical simulation algorithm to obtain various reasonable solutions in order to see the variability that can be expected from the density covariance model adopted.

3.1 Geostatistics concept

The aim of geostatistics is to provide quantitative descriptions of natural variables distributed in space or in time and space. Matheron (1965) created the term regionalized variable to define numerical function $z(x)$ depending on a continuous space index x. In fact, geostatistics is the application of probabilistic methods to

regionalized variables. The variogram is the structure function which can describe how the values at two points behave as the separation between these points increases. Under the condition of second order stationarity (intrinsic assumption), the variogram is given by:

$$\gamma(h) = \frac{1}{2} E\left[Z(x+h) - Z(x)\right]^2 \tag{3.1}$$

For modeling typically isotropic variograms, we need to define three parameters:

- Range: The range is the distance after which no more correlation exists between two observations.
- Sill: The sill is the limit value of the variogram at the distant equal to the range. If it exists, the sill represents the variance of the random variable.
- Nugget effect: the nugget effect represents the very small scale variation, positioning errors, errors of analysis and analytical precision.

Often, the variogram differs according to the direction considered. The 3D variogram model allowing for directional dependency can be defined with 6 parameters (3 ranges and 3 rotations) instead of a single range. Considering x, y and z axes as ellipsoid axes, we can calculate the range in any spatial directions from a_x, a_y and a_z (see Figure 3.1). We can then estimate the parameter covariance at all distances and in all directions of the parameter space. Each element c_{ij} (covariance between the ith and jth cells) of the covariance matrix of parameters corresponding to a distance and a direction, can then be determined. In the case that the principal axes of the ellipsoid are not parallel to the reference axes, a transformation by rotation is applied in order to align the reference axes to the principal ellipsoid axes.

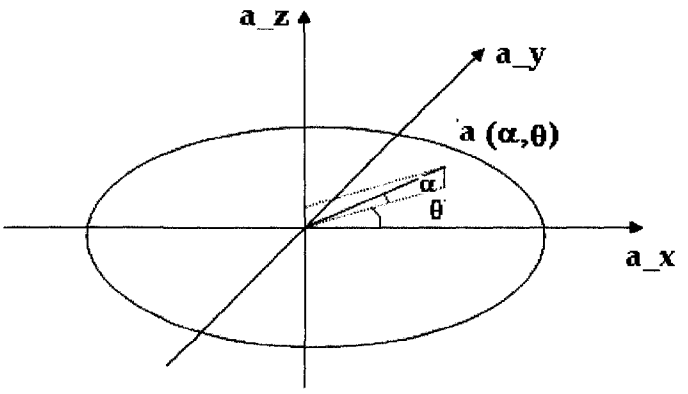

Figure 3.1 Ellipsoid with ranges equal to a_x, a_y and a_z respectively along the three main axes.

Three rotations around each reference axis are needed. The new coordinates are calculated in this new reference system.

When the variogram has a sill, the covariance model $C(h)$ is linked with the model variogram by : $\gamma(h) = \sigma^2 - C(h)$. The covariance function, $C(h)$, measures the covariance of the variable of interest at two points as a function of the distance between these points. This function is usually decreasing with increasing distance and for a critical distance, the range, becomes null. At the distance zero, the covariogram value is simply the variance of the variable of interest (Chilès and Delfiner, 1999). The covariance function is defined as:

$$\begin{cases} E[Z(x)] = m \\ E[(Z(x) - m)(Z(x) - m)] = E[Z(x)Z(x+h)] - m^2 = C(h) \end{cases} \tag{3.2}$$

where $E[Z(x)]$ is the mathematical expectation of the random function $Z(x)$. The mean of each variable $Z(x)$ within the domain is equal to a constant m, $Z(x)$ is the value of the function at the position x, $Z(x+h)$ is the value of the function at

the position $x + h$.

The cross-covariance (or cross-correlation) function measures the covariances between two different variables as a function of the distance vector between the points where these variables are defined. The cross-covariogram, $C_{12}(h)$ of a set of two random functions $Z_1(x)$ and $Z_2(x)$ is defined as (Chilès and Delfiner, 1999):

$$\begin{cases} E[Z_i(x)] = m_i, \quad i = 1, 2 \\ E[(Z_1(x) - m_1)(Z_2(x+h) - m_2] = C_{12}(h) \end{cases} \tag{3.3}$$

where the mean of each variable $Z_i(x)$ at any point of the domain is equal to a constant m_i.

3.2 Theory and formulation of cokriging

Cokriging (Chilès and Delfiner, 1999) is a mathematical interpolation and extrapolation tool that uses the spatial correlation between the secondary variables and a primary variable to improve estimation of the primary variable at unsampled locations. The cokriging method weights data so as to minimize the estimation variance (the cokriging variance). We denote the principal variable by $Z(x)$ with nz observations and the secondary variable by $Y(x)$, with ny observations (for simplicity we just suppose one secondary variable here, although it is possible to apply cokriging with many secondary variables). In simple cokriging, with known mean m_z and m_y, the cokriging estimate is:

$$Z_0^* = m_z + \sum_{i=1}^{nz} \beta_i (Z_i - m_z) + \sum_{i=1}^{ny} \alpha_i (Y_i - m_y) \tag{3.4}$$

The estimation variance is given as follows:

$$var(Z_0 - Z_0^*) = var(Z_0) + \sum_{i=1}^{nz}\sum_{j=1}^{nz}\beta_i\beta_j cov(Z_i, Z_j) + \sum_{i=1}^{ny}\sum_{j=1}^{ny}\alpha_i\alpha_j cov(Y_i, Y_j) \quad (3.5)$$
$$+2\sum_{i=1}^{nz}\sum_{j=1}^{ny}\beta_i\alpha_j cov(Z_i, Y_j) - 2\sum_{i=1}^{nz}\beta_i cov(Z_0, Z_i) - 2\sum_{i=1}^{ny}\alpha_i cov(Z_0, Y_i)$$

The minimum of this function is obtaind by canceling its partial derivatives with respect to the weights:

$$\sum_{j=1}^{nz}\beta_j cov(Z_i, Z_j) + \sum_{j=1}^{ny}\alpha_j cov(Z_i, Y_j) = cov(Z_0, Z_i)\ \forall i = 1...nz \quad (3.6)$$

$$\sum_{j=1}^{nz}\beta_j cov(Y_i, Z_i) + \sum_{j=1}^{ny}\alpha_j cov(Y_i, Y_j) = cov(Z_0, Z_i)\ \forall i = 1...ny \quad (3.7)$$

In general, we can have the matrix form for the cokriging system described above. For p variables under second-order of stationarity and if the global means are known, the simple cokriging estimate can be written (Marcotte, 1991):

$$Z_0^* = \Gamma^T z_\mu + \mu \quad (3.8)$$

where Z_0^* is the estimated vector of the p variables at point x_0 of coordinates $[x_1, x_2, ..., x_d]$, and Γ is the solution matrix $(np \times np)$ of the simple kriging system where n is the number of samples of each variable. z_μ is the $np \times 1$ data vector centered around the mean vector μ. Γ is obtained from the following equation:

$$\mathbf{K}\Gamma = \mathbf{K_0} \quad (3.9)$$

where $\mathbf{K}$ is the $(np * np)$ matrix of point-point covariances for the p variables, $\mathbf{K}$ is formed of $n * n$ submatrix $\mathbf{K_{ij}}$ of size $p * p$ giving the covariances between point

i and point j for the p variables. $\mathbf{K}_0$ is the $np \times p$ matrix of point-(point or block to estimate) covariances, $\mathbf{K}_0$ is formed of n matrices $\mathbf{K}_{0i}$ of size $p \times p$ giving the covariances between point i and point or block to estimate for the p variables.

The simple cokriging variances of individual variable are given by:

$$\sigma_k^2 = \sigma^2 - diag(\mathbf{K}_0^T \Gamma) \tag{3.10}$$

Where $diag(\mathbf{K}_0^T \Gamma)$ is the $p \times 1$ vector formed of diagonal elements of $K_0^T \Gamma$.

A useful tool to verify the performance of the cokriging is cross-validation. It consists of cokriging the measured values after removal of data one at a time for each variable. If the difference between measured and calculated data is small, then, the cokriging performs well. Also, the standardized errors of prediction should give a variance close to one. If this is not the case, then, the coregionalization model should be revised.

3.2.1 Gravity cokriging system

In our cokriging system, the primary variable is density(ρ) and the secondary variable is gravity(g). The cokriging system used here has the particularity that the estimate is obtained only from the secondary variables.
In order to minimize estimation variance, the optimal weights are obtained upon differentiation of $var(\rho_0 - \rho_0^*)$ with respect to α that from now on we refer to as λ. Assuming we work with gravity and density contrasts, i.e., $E[\rho] = E[g] = 0$, the matrix form of the estimation variance can be written:

$$E((\rho_0 - \rho_0^*)(\rho_0 - \rho_0^*)^T) = C_{\rho\rho} - C_{\rho g}\lambda - \lambda^T C_{g\rho} + \lambda^T C_{gg}\lambda. \tag{3.11}$$

Minimization of the above variance will give us the simple cokriging solution and based on that we can obtain the optimal weights:

$$C_{gg}\lambda = C_{g\rho} \tag{3.12}$$

Finally, the estimate of densities is obtained from the gravity data using the optimal weights:

$$\rho^* = \lambda^T g \tag{3.13}$$

3.3 Inversion by Cokriging

Here, we suppose that we have N gravity data and we want to estimate the densities of M cells. If we assume a stationary model, $C_{\rho\rho}$, for the block density covariances, then the gravity-gravity covariance matrix and the gravity-density covariance matrix can be written as follows:

- $C_{g\rho} = GC_{\rho\rho}$. Where, G is the N $\times$ M kernel.
- $C_{gg} = GC_{\rho\rho}G^T + C_{e_g e_g}$. Where, $C_{e_g e_g}$ is the diagonal variance matrix of the error on the measured gravity data g.

The covariances C_{gg} and $C_{g\rho}$ are not stationary. It should be noted that for the same horizontal spacing, the covariance varies as a function of the depth. In addition, defining a domain with finite extension for the densities causes a border effect which consequently makes the covariances become a function of the exact positions of the points relative to the limits of the domain studied. Because of the non-stationary nature of the gravity-gravity and gravity-density covariances, estimators such as the traditional variogram or the covariance function, cannot

be used directly. Therefore, the model parameters for density covariance must, themselves, be estimated by inversion. The proposed approach for estimating the density model is based on the concept of the V-V plot (Asli et al., 2000) and is explained in the next section.

After estimating the density model, its covariance matrix is calculated. Knowing $C_{\rho\rho}$, the anomaly covariances and density-anomaly cross-covariances can be calculated, which in turn, allows a cokriging system to be constructed. From this cokriging system, we can calculate the coefficients λ and the density distribution based on the gravity observations:

$$C_{gg}\lambda = C_{g\rho} \tag{3.14}$$

$$\rho^* = \lambda^T g \tag{3.15}$$

As it can be seen, the estimation in this cokriging system is only based on the secondary variables (gravity data). In this inversion method, the density contrast is estimated, not the density itself. Here, it is supposed that the gravity anomaly represents a residual with mean of 0 and therefore a simple cokriging system can be used. Alternatively, ordinary cokriging with the non-bias condition $\sum \lambda_i = 0$ can be used. Because of lack of density measurements, a trend cannot be directly imposed on the density according to the depth. The trend must, therefore, be added as a later step to the values estimated by the simple or ordinary cokriging (Asli et al., 2000).

3.3.1 Similarities between the formulation of inversion by cokriging and Bayesian formulation

It should be mentioned that there exists relation between the inversion by cokriging and Bayesian formulation proposed by Tarantola and Valette (1982). As we mentioned:

$$\rho^* = \lambda^T g \tag{3.16}$$

Where:

$$\lambda = C_{gg}^{-1} C_{g\rho} \tag{3.17}$$

So, we can write the density estimation as:

$$\rho^* = (C_{gg}^{-1} C_{g\rho})^T g = \left[(C_{g\rho})^T (C_{gg}^{-1})^T\right] g \tag{3.18}$$

Having equations :

$$C_{gg} = G C_{\rho\rho} G^T + C_{e_g e_g} \tag{3.19}$$

$$C_{g\rho} = G C_{\rho\rho} \tag{3.20}$$

Finally, the density distribution can be expressed as:

$$\rho^* = \left[(G C_{\rho\rho})^T ((G C_{\rho\rho} G^T + C_{e_g e_g})^{-1})^T)\right] g = \left[C_{\rho\rho} G^T (G C_{\rho\rho} G^T + C_{e_g e_g})^{-1}\right] g \tag{3.21}$$

Final equation (3.21) is the Bayesian formulation which is proposed by Tarantola and Valette (1982).

It should be mentioned that when data is measured without error ($C_{e_g e_g} = 0$), then we have $g = G\rho^*$, i.e., the gravity is perfectly recovered by the inverted cokriged model.

3.3.2 V-V plot method

The V-V plot method (Asli et al., 2000) is used to estimate the density variogram model. In this method, at first a density variogram model is assumed. Based on this model, the theoretical gravity-gravity covariance values are calculated as explained before. Assuming we work with density contrasts, i.e. $E[g] = 0$ and $E[\rho] = 0$, where $E[.]$ is mathematical expectation of variables of gravity and density, the observed data covariance matrix is simply:

$$C_{gg-exp} = g.g^T \tag{3.22}$$

Then, the theoretical covariance values are arranged in decreasing order and grouped in classes. For each of these classes, the mean is calculated. Likewise, the individual pairs forming the gravity data are arranged and grouped using the same order and the same classes. Therefore, for a single theoretical structural distance, a value similar to an experimental variogram is obtained.

If the initial density-density covariance model is appropriate, the experimental gravity variogram versus theoretical variogram graph (V-V plot) should show a dispersion around a straight line at 45^o. It might be more useful to draw experimental and theoretical covariance values together and test that they have a good correlation. If the correlation is not good enough, the proposed density-density covariance model is modified and the procedure is repeated.

The algorithm for V-V plot is as follows:

i. Calculate the matrices Cgg_{exp} and Cgg_{th}.

ii. Transform both matrices into vectors v_{th} and v_{exp}.

iii. Sort v_{th} in decreasing order and apply the same ordering to v_{exp}.

iv. Bin the vectors v_{th} and v_{exp} in N_{lag} lags and compute the mean for each lag.

v. Minimize the error for each lag with the relation $err[j] = v_{th}[j] - v_{exp}[j]$.

The adjustment of density-density covariance model may be made manually or automatically. In either case, it should be noted that during the fitting of experimental and theoretical gravity covariance values, it is more important to have a really good correlation at smaller structural distances then at larger structural distances. In the following section, we explain briefly the Simplex method for automatic V-V plot.

3.3.2.1 Simplex method for V-V plot

As we mentioned before, the covariances C_{gg} and $C_{g\rho}$ are not stationary. So, we use the V-V plot method to find the appropriate density covariance matrix. In this method we adjust the density covariance matrix such that experimental and theoretical gravity covariance matrices have good correlation. In order to do this, we can use the simplex method to minimize the difference between experimental and theoretical gravity covariance values.

The simplex search method was first proposed by Nelder and Mead (1965). It is an enormously popular direct search method for multidimensional unconstrained minimization. The Nelder-Mead method attempts to minimize a scalar-valued nonlinear function of n real variables using only function values, without any derivative information (explicit or implicit). The Nelder-Mead method thus falls in the general class of direct search methods. If n is the length of x, a simplex in n-dimensional space is characterized by the $n + 1$ distinct vectors that are its vertices. In two-space, a simplex is a triangle; in three-space, it is a pyramid. At each step of the search, a new point in or near the current simplex is generated. The function

value at the new point is compared with the function's values at the vertices of the simplex and, usually, one of the vertices is replaced by the new point, giving a new simplex. This step is repeated until the diameter of the simplex is less than the specified tolerance.

In MATLAB software, Optimization Toolbox provides simplex search method using the `fminsearch` command. `fminsearch` finds the minimum of an unconstrained multi-variable function derivative-free, starting at an initial estimate. This is generally referred to as unconstrained nonlinear optimization.

3.4 Inversion by Cokriging using Constraints

There are some methods to remedy the problem of non-uniqueness and improve the performance of gravity inversion by adding constraints. One of the big advantages of inversion by cokriging is that we can easily add constraints as primary or secondary variables to our cokriging system. In fact, including constraints is straightforward as they are nothing else than new data used in cokriging. Here we suppose that we have two kind of constraints:

- N_F fixed densities ρ_F of certain cells.
- $N_{Gr} = N_{Gr_x} + N_{Gr_y} + N_{Gr_z}$ known density gradients $gr^T = [gr_x^T gr_y^T gr_z^T]$ between two adjacent cells in the x, y, z directions.

Using the above constraints, the estimate for densities is obtained using the similar cokriging system as before but here we will have two more parameters as secondary variables:

$$\rho = \lambda^T g + \mu^T \rho_F + \eta^T gr \tag{3.23}$$

Here, the estimation variance is minimized with respect to the weights λ, μ, η. The simple cokriging system is given by:

$$\begin{bmatrix} C_{g,g} & C_{g,\rho_F} & C_{g,gr} \\ C_{\rho_F,g} & C_{\rho_F,\rho_F} & C_{\rho_F,gr} \\ C_{gr,g} & C_{gr,\rho_F} & C_{gr,gr} \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \\ \eta \end{bmatrix} = \begin{bmatrix} C_{g,\rho} \\ C_{\rho_F,\rho} \\ C_{gr,\rho} \end{bmatrix} \tag{3.24}$$

The covariance matrices for the above equation are given by:

1. $C_{g,\rho} = GC_{\rho,\rho}$. Where, G is the N×M kernel.
2. $C_{g,g} = GC_{\rho,\rho}G^T + C_{e_g e_g}$. Where, $C_{e_g e_g}$ is the positive semi-definite diagonal variance matrix of the error on the measured gravity data g.
3. $C_{\rho_F,\rho} = FC_{\rho,\rho}$. Where, F is an $N_F \times M$ matrix and all of its coefficients have value 1 when we know the density of the cell and 0 elsewhere.
4. $C_{g,\rho_F} = C_{g,\rho}F^T$. We have $C_{\rho_F,g} = C^T_{g,\rho_F}$.
5. $C_{\rho_F\rho_F} = FC_{\rho\rho}F^T + C_{e_F e_F}$. Where, $C_{e_F e_F}$ is the positive semi-definite diagonal variance matrix of the error on the fixed densities ρ_F.
6. $C_{gr,\rho} = HC_{\rho,\rho}$. Where, H is an $N_{Gr} \times M$ matrix and all of its coefficients have value -1 and 1 where the gradient between two cell is known and 0 elsewhere.
7. $C_{gr,\rho_F} = HC_{\rho,\rho}F^T$. We have $C_{\rho_F,gr} = C^T_{gr,\rho_F}$.
8. $C_{gr,g} = HC_{\rho,\rho}G^T$. We have $C_{gr,g} = C^T_{g,gr}$.
9. $C_{gr,gr} = HC_{\rho,\rho}H^T + C_{e_{gr}e_{gr}}$. Where, $C_{e_{gr}e_{gr}}$ is the positive semi-definite diagonal variance matrix of the error on the fixed gradients gr.

The solution of the cokriging system, the estimated densities, can be written:

$$\rho^* = \begin{bmatrix} C_{g,\rho}^T & C_{\rho_F,\rho}^T & C_{gr,\rho}^T \end{bmatrix} \begin{bmatrix} C_{g,g} & C_{g,\rho_F} & C_{g,gr} \\ C_{\rho_F,g} & C_{\rho_F,\rho_F} & C_{\rho_F,gr} \\ C_{gr,g} & C_{gr,\rho_F} & C_{gr,gr} \end{bmatrix}^{-1} \begin{bmatrix} g \\ \rho_F \\ gr \end{bmatrix} \tag{3.25}$$

The inverse matrix calculation is done by singular value decomposition (SVD). This method has been preferred to the conjugate gradient algorithm (CGA) used by Li and Oldenburg (1998) and Boulanger and Chouteau (2001), because it is numerically stable and it is a standard tool for small inverse problems. However, for large problems, preconditioned CGA is needed.

3.5 Efficient calculation of the gravity-density covariance matrix and the gravity-gravity covariance matrix

During inversion using cokriging, it takes several matrix-matrix multiplications to compute the gravity-density covariance matrix and the gravity-gravity covariance matrix. To compute these matrices, we need to calculate and save the density-density covariance matrix. For large numbers (n) of density values, the required storage and computational costs to compute these matrices, proportional to n^2, becomes prohibitive.

Nowak et al. (2003) proposed a collection of highly efficient spectral methods to compute these matrices, based on circulant embedding and the fast Fourier transform (FFT). These methods can be used whenever the densities are a stationary random variable on a regular equi-spaced grid. Using his proposed method, computational costs are reduced from $O(n^2)$ to $O(nlog_2 n)$ and storage requirements are reduced from $O(n^2)$ to $O(n)$.

In his paper, Nowak et al. (2003) demonstrated the application to two dimensional (2D) domains. Here, we extend his approach to three-dimensional (3D) domains. 2D problems lead to symmetric block circulant matrices with circulant blocks (SCC), while 3D problems lead to level 3 blocking structures (SCCC).

The general procedure for 3D applications can be summarized as follows:

- Define the density as a discretized random space variable such that its auto-covariance matrix K has symmetric level 3 blocking Toeplitz structure (STTT).
- Embed the STTT matrix K into a larger matrix with symmetric level 3 blocking circulant structure (SCCC).
- Use spectral methods to calculate all matrix-matrix products.
- Extract the results from the embedded matrices.

An $\hat{n_x} \times \hat{n_x}$ symmetric Toeplitz (ST) matrix has the following structure:

$$\begin{bmatrix} t_0 & t_1 & \dots & t_{\hat{n_x}-1} \\ t_1 & t_0 & \dots & t_{\hat{n_x}-2} \\ \dots & & \dots & \\ t_{\hat{n_x}-1} & t_{\hat{n_x}-2} & \dots & t_0 \end{bmatrix} \tag{3.26}$$

For circulant embedding, the rows have to be periodic. To embed ST to SC matrices, we should extend the first row by appending the elements $t_1, t_2, ..., t_{\hat{n_x}-2}$ in reverse order to obtain a series $c_0, c_1, ..., c_{n_x}, ..., c_1$ where $n_x = \hat{n_x} - 1$, corresponding to mirroring the covariance function to make it periodic. Symmetric Toeplitz matrices can be completely determined by their first row (column). So, we only need to calculate the first column of the density-density covariance matrix K which

is an $N \times N$ matrix where $N = n_x n_y n_z \times n_x n_y n_z$ matrix. n_x, n_y and n_z are the number of prisms in x, y and z directions respectively. To embed this first column, we first embed every n_x element to SC blocks. So, we have $n_y n_z$ SC blocks of size $2n_x \times 1$. Then, we extend every n_y blocks of size $2n_x \times 1$ to obtain a periodic series of blocks including n_z SCC blocks of size $4n_x n_y \times 1$. After that, we extend these n_z SCC blocks to obtain a periodic series of SCCC blocks and the overall size will be $8N \times 1$. We refer to the embedded version of K as $K2$.

To maintain the compatibility of matrix dimensions for multiplication, for example for the calculation of KG^T, the matrix G^T has to undergo the same embedding by zero-padding all entries corresponding to the new entries in the SCCC covariance matrix $K2$. We recall that G^T is an $N \times M$ matrix where M is the number of observations and N is the number of prisms. We reshape every column sized $N \times 1$ into an $n_x \times n_z n_y$ matrix. Every row of the new matrix is reshaped into an $n_y \times n_z$ matrix, padded with zeros to obtain an $2n_y \times 2n_z$ matrix, and reshaped back to $1 \times 4n_y n_z$ row. Now, we have an $n_x \times 4n_y n_z$ matrix. Every column of the new matrix is padded with zeros to obtain a $2n_x \times 4n_y n_z$ matrix. We reshape this matrix back to an $8N \times 1$ row. We repeat this process with every column of G^T and finally we will have an $8N \times M$ matrix which we call $G2^T$. The zero padding is to suppress the influence of the new elements of $K2$ during matrix multiplication.

To calculate the matrix-matrix multiplication of KG^T, we can split it up into single vector-matrix multiplications:

$$Ku_p, p = 1, ..., M \tag{3.27}$$

where u_k is the kth column of G^T.

Using the diagonalization theorem (Nowak et al., 2003), for an $N \times N$ SCCC matrix

K we have:

$$K = F^H \Lambda F \tag{3.28}$$

where Λ is the diagonal matrix of eigenvalues and F is the 3D Fourier-matrix (v=Fu where u and v are the vectors obtained by rearranging the matrices G^T and $FFT(G^T)$ column-wise) and $F^H = F^{-1}$ is the Hermitian transpose of F.

$$FK = \Lambda F \tag{3.29}$$

Because one column of K contains all information, $FK_1 = \Lambda F_1$ and all entries of F equal $n^{\frac{-1}{2}}$.

$$\lambda = n^{\frac{1}{2}} F K_1 \tag{3.30}$$

where λ is an $n \times 1$ vector of the eigenvalues. Now, we can calculate matrix-vector multiplication:

$$Ku_p = (F^H \Lambda F)u_p = F^H \Lambda (Fu_p) = F^H (\Lambda v) = F^H [\lambda_1 v_1, ..., \lambda_n v_n]^T \tag{3.31}$$

Fu_p is calculated using 3D FFT and $F^H[.]$ is calculated using inverse 3D FFT.

Because of embedding, the result of KG^T is also embedded. We have to reverse the embedding process to extract the actual result.

3.6 LU simulation

In this thesis, in order to generate the synthetic densities, we used the LU simulation method. Method LU (Cholesky) is the easiest method for programming, and probably most effective for small fields (Chilès and Delfiner, 1999). This method can be applied for conditional as well as non-conditional simulation.

The non-conditional simulation algorithm is written as follow:
Suppose the density random variable, $\rho(x)$, is Gaussian with mean 0 and has covariance $C(h)$. For a simulation of n points:

i. Knowing the density variogram model and the size of the grid, construct density covariance matrix $C_{\rho\rho}$.

ii. Apply the decomposition $C_{\rho\rho} = LL'$ where L is a lower triangular matrix.

iii. Generate a vector Y with n elements that are identically independent distributed $(i.i.d)$ with $N(0,1)$.

iv. Calculate $\rho = LY$.

v. It can be verified that $E[\rho\rho'] = E[LYY'L'] = LE[YY']L' = LIL' = LL' = C_{\rho\rho}$.

3.7 Co-simulation based on the FFT moving average (FFT-MA) generator

Inversion using cokriging gives a smooth estimate of the density model. Sometimes it is desirable and useful to obtain various reasonable solutions in order to see the variability that can be expected from the density covariance model adopted. This can be achieved using geostatistical simulation algorithms rather than cokriging. Among different efficient simulation algorithms (Chilès and Delfiner, 1999), the Fast Fourier Transform Moving Average simulation (FFT-MA) is a fast simulation algorithm for generating regular grid non-conditional Gaussian stationary processes (Le Ravalec et al., 2000).

3.7.1 FFT moving average (FFT-MA) generator

The moving average method uses the n-dimensional covariance function C instead of the covariance matrix. This covariance operator can be written as follow:

$$C = g * \overline{g} \tag{3.32}$$

Where $\overline{g}(x)$=g(-x). If we can calculate the n-dimensional function g, we can generate a Gaussian random field y with mean m and covariance matrix C:

$$y = m + g * z \tag{3.33}$$

Where z is a n-dimensional Gaussian white noise.

The basic idea of FFT-MA algorithm is to determine the convolution product $g * z$ in the frequency domain. Based on the Fourier analysis theorem, any stationary process has a covariance function C of the form:

$$C(x) = \int_{-\infty}^{+\infty} S(f) exp(2i\pi f.x) df \tag{3.34}$$

Where S is the spectral density function depending on frequency f and the integral is a n-fold integral. S and C have the same information in different spaces. C is the inverse Fourier transform of S, while S is the forward Fourier transform of C:

$$S(f) = \int_{-\infty}^{+\infty} C(x) exp(-2i\pi f.x) dx \tag{3.35}$$

If we put $Y = FFT(y - m) = FFT(g * z) = GZ$, $Z = FFT(z)$ and $G = FFT(g)$, then the expectation of the variance of Y is (Le Ravalec et al., 2000):

$$E\left[Y(f)\overline{Y(f')}\right] = \begin{cases} 0 & \text{if } f \neq f' \\ G(f)\overline{G(f)} = S(f) & \text{if } f = f' \end{cases} \tag{3.36}$$

We can write the results within a discretized framework. For simplicity, here we assume one dimension, but the generalization to two and three dimensions is easy. Because of the periodicity implied by using Fourier transforms, we should oversize the desired field. For example, for a stochastic process of size Ndx where N is the number of points and dx the sampling rate, we generate a process of size $Ndx + lc$, where lc is the correlation length. After simulation, we discard the additional points, leaving the first and last points of the sequence uncorrelated. If we consider a sequence of N_1 equidistant points and sampling rate of dx_1, the frequency rate is $df1 = 1/(N_1 dx_1)$ and the frequency range $[-1/(2dx_1); 1/(2dx_1)]$.

$$S(j_1) = dx_1 \sum_{k_1=1}^{N_1} C(k_1) exp(-2i\pi \frac{k_1 j_1}{N_1}) \tag{3.37}$$

$$C(k_1) = \frac{1}{N_1 dx_1} \sum_{j_1=1}^{N_1} S(j_1) exp(2i\pi \frac{k_1 j_1}{N_1}) \tag{3.38}$$

From the above framework, equation (3.32) can be rewritten as:

$$S(j_1) = \frac{1}{dx_1} G(j_1)\overline{G(j_1)} \tag{3.39}$$

The only requirement on function g is that equation (3.32) is satisfied. Because there is no unique solution to this equation, additional constraints are considered to get functions with particular forms.

$$G(j_1) = \sqrt{dx_1 S(j_1)} \tag{3.40}$$

We can generate unconditional Gaussian random fields according to the following

steps:

i. Construction of the sampled covariance C.

ii. Generation of the normal deviates z on the grid.

iii. Calculation of the Fourier transforms of z and C, giving Z and the power spectrum S, respectively.

iv. Derivation of G from S using equation 3.40.

v. Multiplication of G by Z.

vi. Inverse Fourier transform of $(G.Z)$ giving $g * z$.

vii. Derivation of y from equation 3.33.

Because the computations are made using FFT, the simulation is fast and stable. In this method, structural parameters (the covariance) are treated separately from the random parameters. In order to have different realizations with the same covariance, we only have to update the random vector. Moreover, it is possible to keep the same random vector and change only the covariance for example to eliminate random fluctuations in a sensitivity analysis of the covariance model.

3.7.2 Co-simulation using FFT-MA generator

The FFT-MA algorithm generates non-conditional simulations ρ_s. The computed gravity data of each simulation are not linked to the measured gravity data. It is necessary to post-condition the simulation by cokriging. First, the cokriging of densities ρ^* with measured gravity data g is performed. Then, for each realization ρ_s, the gravity data are computed using $g_s = G\rho_s + e$, where e is drawn from

a $N(0, C0)$. Keeping the same cokriging weights, cokriging of the densities ρ_s^* with computed simulated gravity data g_s is performed. Finally, the conditional simulated densities ρ_{sc} are:

$$\rho_{sc} = \rho^* + (\rho_s - \rho_s^*) \tag{3.41}$$

where ρ_{sc} has the desired covariance. When the gravity covariances have zero nugget effect, the measured gravity data and the computed gravities of each conditional simulation are exactly the same. When a nugget effect is present in the observed gravity data, the computed gravity data are different from the measured gravity data by an amount compatible with the level of error described by the nugget effect. Using the conditional simulation approach, we can produce many statistically equivalent density models.

CHAPTER 4

APPLICATION TO SYNTHETIC MODELED DATA

In this chapter, we apply the proposed method based on cokriging and cosimulation to two kinds of synthetic data:

1. Dipping dyke
2. Synthetic data generated by LU simulation

At first, the model of dipping dyke is selected. The model is studied without and with constraints. Gradients in between some blocks are considered as constraints in this model and their effects are described. Finally, the co-simulation method based on FFT-MA is applied to investigate the behavior of simulation in this model. In the second attempt, we generate the synthetic model using LU simulation with a specific variogram model. Then, we apply the proposed inversion method without and with constraints. The constraints considered in this case are the known densities in boreholes. The errors obtained for the densities as function of depth are examined. All histograms including cokriging and simulation histograms are examined as well.

4.1 Application to Dipping Dyke

We use a simple 45^o dipping short dyke with uniform density contrast 1500 kg/m^3 with respect to an homogeneous background (see Figure 4.1). We select this model (Chasseriau and Chouteau (2003); Li and Oldenburg (1998)) in order to test the

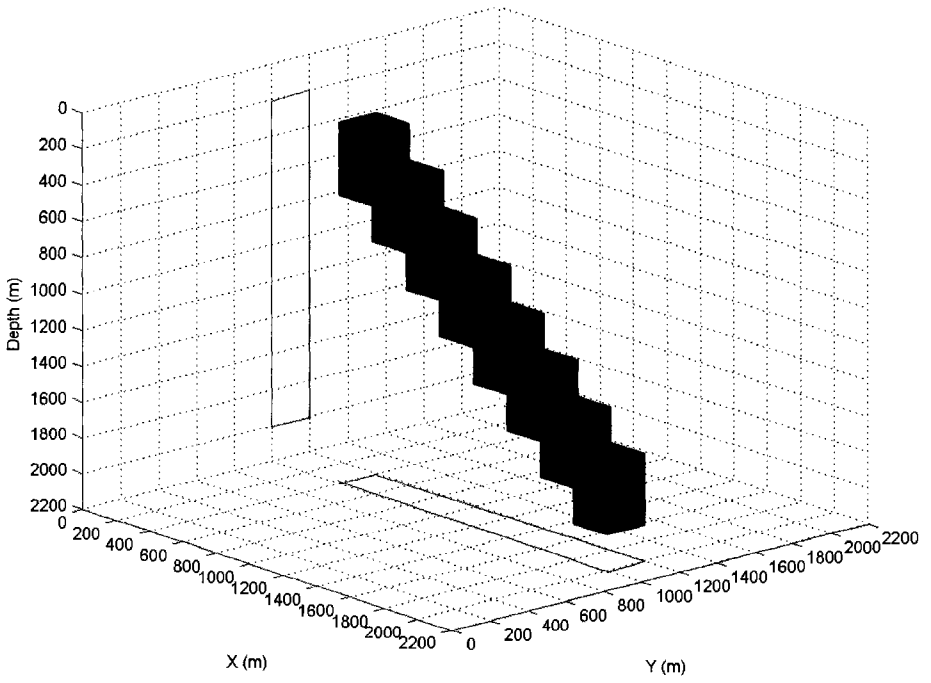

Figure 4.1 Model of dipping dyke ($\rho = 1500kg/m^3$) in homogeneous background

performance of the proposed algorithm (inversion by cokriging) to resolve depth, anisotropy and dip. The 3D domain is divided into $22 \times 22 \times 22 = 10648$ cubic prisms. The dimensions of each elementary cubic prism are $100 \times 100 \times 100$ meters. Therefore, dimensions of the modeled domain are $2.2 \times 2.2 \times 2.2$ km. Figure 4.2 shows the anomaly calculated from this synthetic model. The surface gravity response was computed using equation $g = G\rho$ at 484 sites. The data is located at the center of the top face of cells in the first layer just above the ground surface in order to avoid singularities in the computation of the geometric matrix (G). The range of gravity values is from 0.2028 $mGal$ to 2.0306 $mGal$. We assume that data are uncorrelated and use a diagonal data noise covariance matrix with $\sigma_d^2 = 0.01$ $(mGal)^2$.

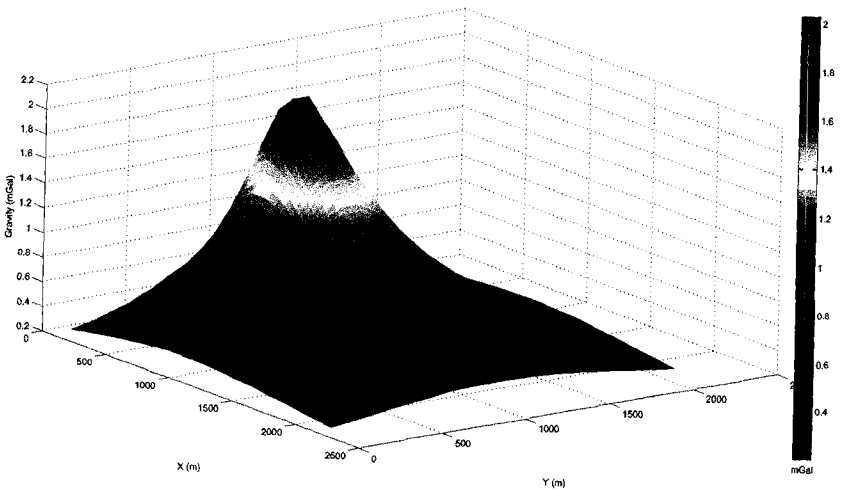

Figure 4.2 Surface anomaly caused by the dyke

Here, we try to estimate the density variogram of the dyke. The experimental variogram is calculated in three directions: dyke direction, y direction and direction perpendicular to the plane formed by the two first axes. Figure 4.3 shows the experimental variograms in these three directions. In dyke direction, we have an exponential covariance model with a range of 1500 meters and a sill of $28000(kg/m^3)^2$. In the y-direction and in the direction perpendicular to the plane formed by the two first axes, we obtain an exponential model respectively with ranges of 200 and 230 meters. The nugget effect is null in the three directions. From these modelled variograms, we conclude that the dyke has an exponential anisotropic covariance model with an average variance of $20000(kg/m^3)^2$. We can apply this variogram to calculate the covariance matrix of densities. Then using cokriging method, we can reconstruct the density distribution. However, in reality we don't have all information of subsurface to calculate the variogram. In this situation we have to carefully borrow (use) all spatial geological information to model the variogram. In the worst case, when we don't have any geological information, we attempt to

find the covariance matrix directly from observed data.

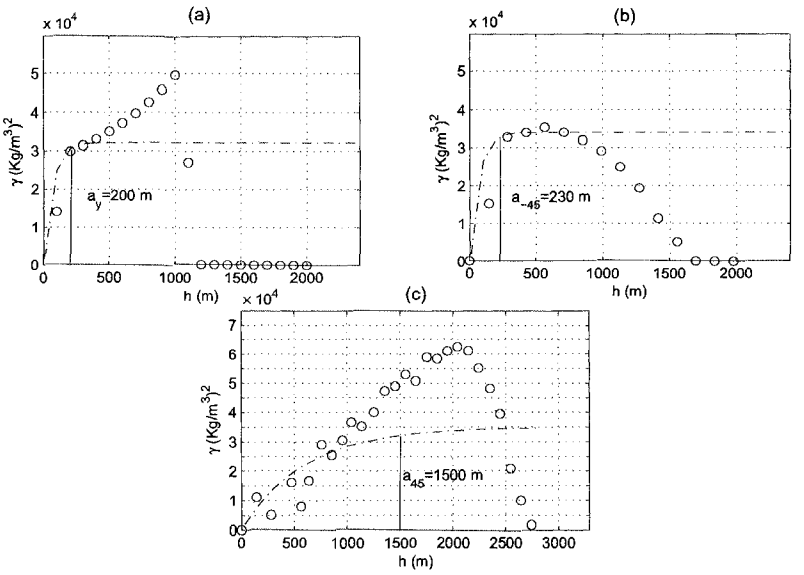

Figure 4.3 Variograms for the dyke model. The experimental variograms fit an exponential model variogram. In the y axis direction (a), the range is $200m$ and in direction perpendicular to the dyke (b), the range is $230m$. The range is $1500m$ along the 45^o dip (c).

The inversion of the calculated anomalies to estimate the density distribution will be performed using the following steps:

1. First we assume a variogram model for the density distribution and we calculate the density covariance matrix $C_{\rho\rho}$.

2. We calculate the experimental gravity covariance matrix C_{gg-exp} using the gravity data. We also calculate the theoretical gravity covariance matrix C_{gg-th} using the density covariance matrix $C_{\rho\rho}$ by the following formula: $C_{gg-th} = GC_{\rho\rho}G^T$.

3. We use the V-V plot method that was discussed in chapter 3 with $Nlag = 50$ and we try to adjust the density model such that experimental and theoretical variograms become almost identical. After applying this method, we conclude that the density variogram model is anisotropic and exponential with $C_0 = 62(kg/m^3)^2$, $C = 19000(kg/m^3)^2$ and $a_{45} = 1200m$, $a_{-45} = 200m$ and $a_y = 200m$. The covariance vectors v_{exp} and v_{th} are shown in Figure 4.4. It can be seen from this figure that the correlation between them is 98.74% and the Mean Absolute Error (MAE) is $0.04(mGal)^2$. This means that there is a good fit between experimental and theoretical covariance matrices.

   After selecting the density variogram model, we will calculate the density covariance matrix and using that we will calculate $C_{g\rho} = GC_{\rho\rho}$ and $C_{gg} = GC_{\rho\rho}G^T$. Now, we can use the cokriging system equation $C_{gg}\lambda = C_{g\rho}$ to find coefficients λ and then we can find an estimate for the density distribution: $\rho^* = \lambda^T g$. The estimated density distribution at section $y = 1100m$ is shown in Figure 4.5.

   By comparing Figure 4.1 and Figure 4.5, we can see that the inversion method using cokriging is a very effective method for density estimation. The shape

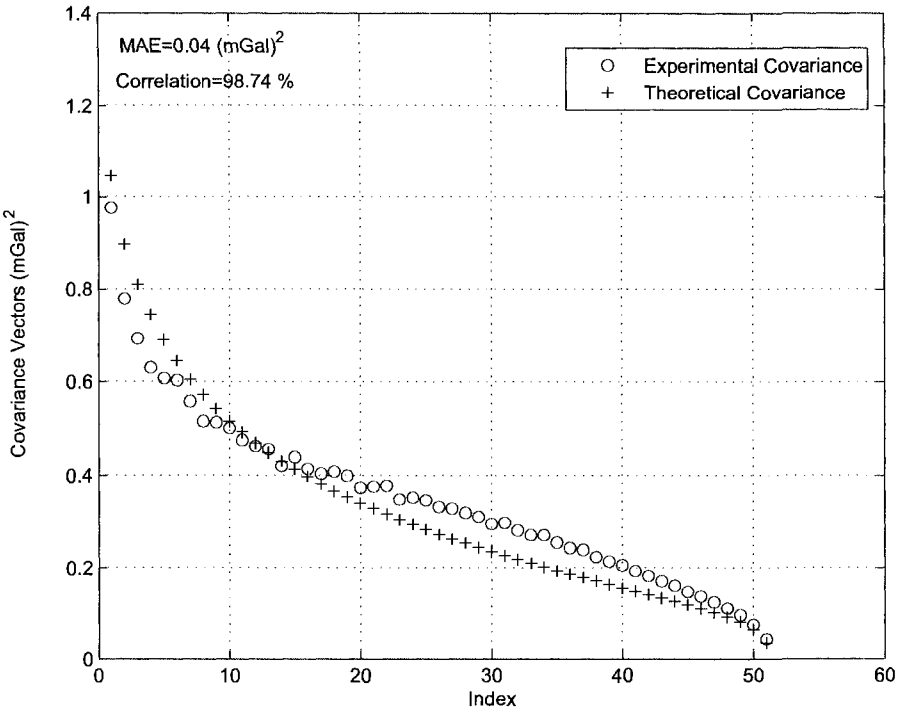

Figure 4.4 Fit of the experimental and theoretical gravity covariance matrices. Density variogram model: anisotropic and exponential with $C0 = 62(kg/m^3)^2$, $C = 1900(kg/m^3)^2$ and $a_{45} = 1200m$, $a_{-45} = 200m$ and $a_y = 200m$.

and the density contrast of the dyke are not exactly recovered but we can retrieve the dip and the position of the dyke using the stochastic approach.

After finding the density distribution, we can calculate the gravity data. This gravity is obtained by multiplying the geometric matrix (G) by the inverted density vector (ρ^*). In Figure 4.6, we compare this calculated gravity with our input (observed) gravity. It can be seen from this figure that these two gravity data are identical and the MAE is equal $2.43 \times 10^{-7} mGal$.

We have also applied direct inversion method in (Tarantola and Valette (1982);Menke

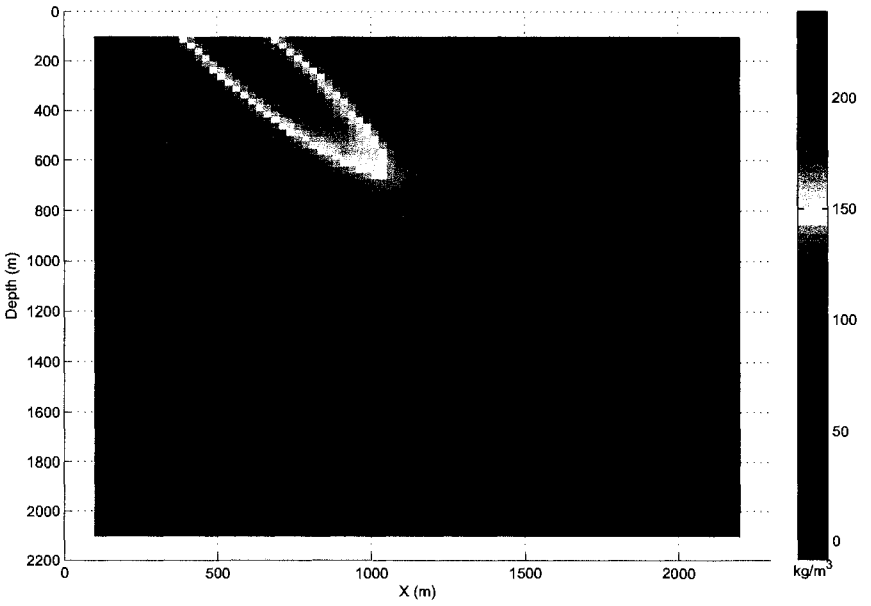

Figure 4.5 The estimated density distribution at section $y = 1100m$ using inversion by cokriging.

(1989)) assuming $C_{\rho\rho}$ =I to the surface gravities to estimate the density distribution. After 60 iterations, the estimated density at section $y = 1100m$ has been shown in Figure 4.7. Comparing this figure with the result of inversion by cokriging, we can see that the stochastic approach leads to a major improvement in the image reconstruction of the structure causing the gravity anomaly. The density contrast is also better estimated, basically due to anisotropy identified in the recovered model.

4.1.1 Co-simulation of dipping dyke model

As we mentioned in chapter 3, cokriging gives a smooth estimate of the density distribution. Using geostatistical simulation algorithms, we can have various rea-

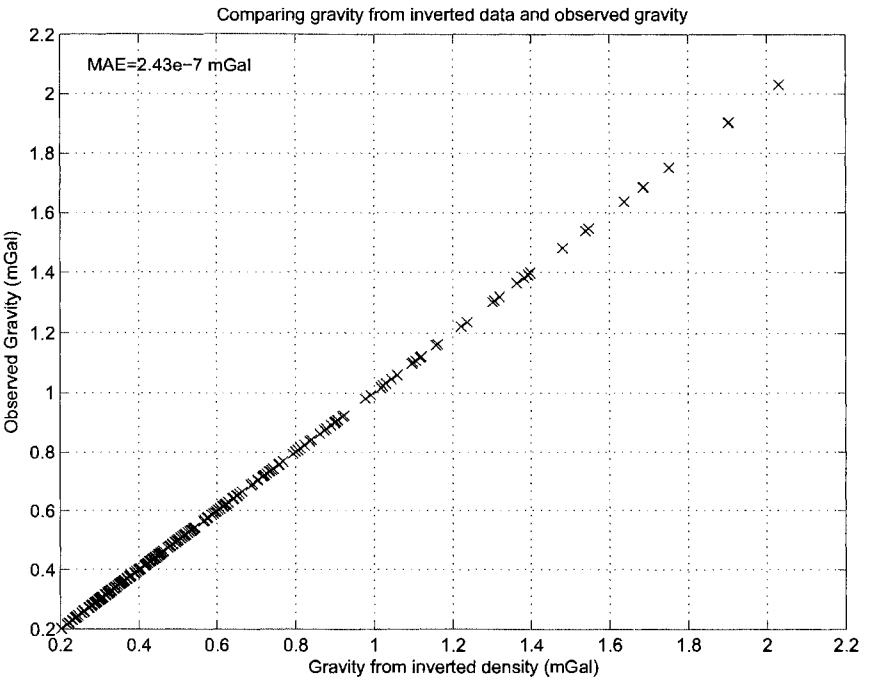

Figure 4.6 Comparing the input gravity with gravity calculated from inverted densities.

sonable solutions showing the kind of variability that can be expected from the density covariance model. We use the same density variogram model obtained during inversion by cokriging. Then using the FFT-MA method, we provide various non-conditional density simulations. We start post-conditioning with the cokriging of densities obtained from the observed gravity from the previous section. One of the conditional simulated densities with good correlation with the density distribution at section $y = 1100m$ has been shown in Figure 4.8. We have also compared the observed gravity with the gravity from conditional simulated densities in Figure 4.9. As we can see from this figure and in general the gravity from conditional simulated densities is identical with the observed gravity in absence of he noise.

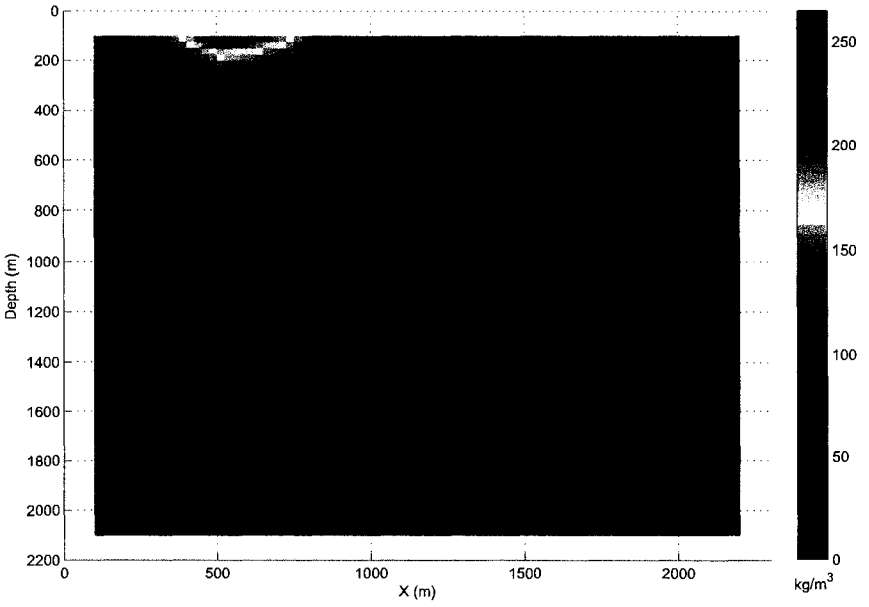

Figure 4.7 Density distribution at section $y = 1100m$ estimated using direct inversion ($C_{\rho\rho}$ =I).

4.1.2 Application to dipping dyke by adding gradient constraints

In this section, we will add some constraints (known gradients) to our inversion system and we will show that knowing extra information about the system will help us to find the density distribution of the dyke more accurately.

Here, we use the same domain and the same gravity observations. However, we assume that there exists some known gradients between blocks as constraints. For example, these gradients could be known from borehole logs or core examination. Figure 4.10 shows the domain and the location of the known gradients. These known gradients are equal to 1500 kg/m^3 and they are between the following pairs of blocks: $(1250, 1050, 1050)$ and $(1250, 1050, 950)$,
$(1350, 1050, 1050)$ and $(1350, 1050, 950)$,

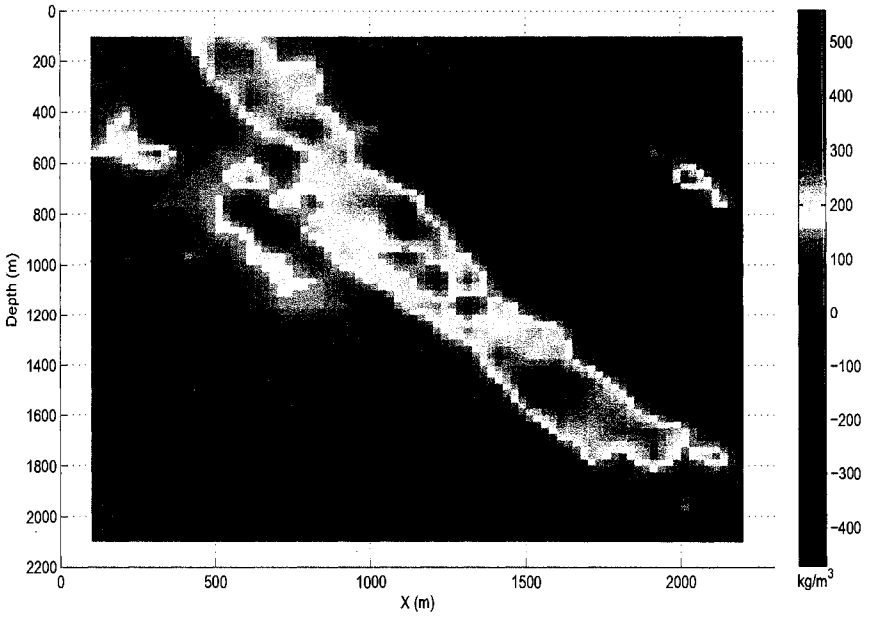

Figure 4.8 The conditional simulated densities at section $y = 1100m$.

$(1250, 1050, 1350)$ and $(1250, 1050, 1450)$,
$(1350, 1050, 1350)$ and $(1350, 1050, 1450)$.

We apply inversion by cokriging using constraints that were discussed before in chapter 3 in order to find a density distribution using the following steps:

1. We use the same variogram model for the density distribution, the same density covariance matrix $C_{\rho\rho}$ and consequently the same theoretical gravity covariance matrix C_{gg-th}.

2. We form a matrix H for the known gradients. H is an $N_G \times M$ matrix where N_G is the number of known density gradients and M is the total number of prisms. Each row of H corresponds to one of the known gradients and all of

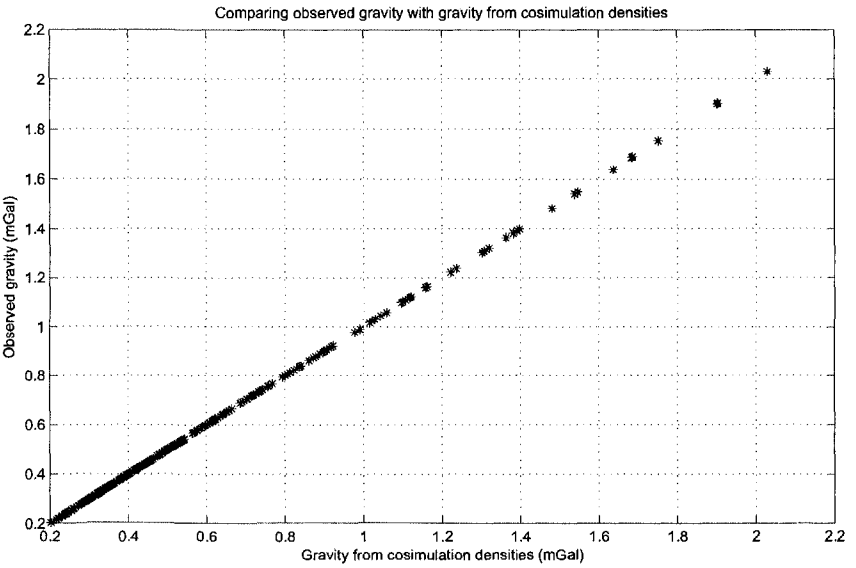

Figure 4.9 Comparing the observed gravity with the gravity from conditional simulated densities.

its coefficients have value -1 and 1 where the gradient between two cell is known and 0 elsewhere.

3. We calculate $C_{g,\rho} = GC_{\rho\rho}$, $C_{g,g} = GC_{\rho\rho}G^T$, $C_{gr,\rho} = HC_{\rho,\rho}$, $C_{gr,g} = HC_{\rho,\rho}G^T$, $C_{gr,g} = C^T_{g,gr}$, $C_{gr,gr} = HC_{\rho,\rho}H^T + C_{e_{gr}e_{gr}}$. Where, $C_{e_{gr}e_{gr}}$ is the diagonal variance matrix of the error on the fixed gradients gr. Here, we suppose that there is no error on gradients.

4. Now, we can use the cokriging system using constraints:

$$\begin{bmatrix} C_{g,g} & C_{g,gr} \\ C_{gr,g} & C_{gr,gr} \end{bmatrix} \begin{bmatrix} \lambda \\ \eta \end{bmatrix} = \begin{bmatrix} C_{g,\rho} \\ C_{gr,\rho} \end{bmatrix} \tag{4.1}$$

After finding coefficients λ and μ, we can find an estimate for density distribution:

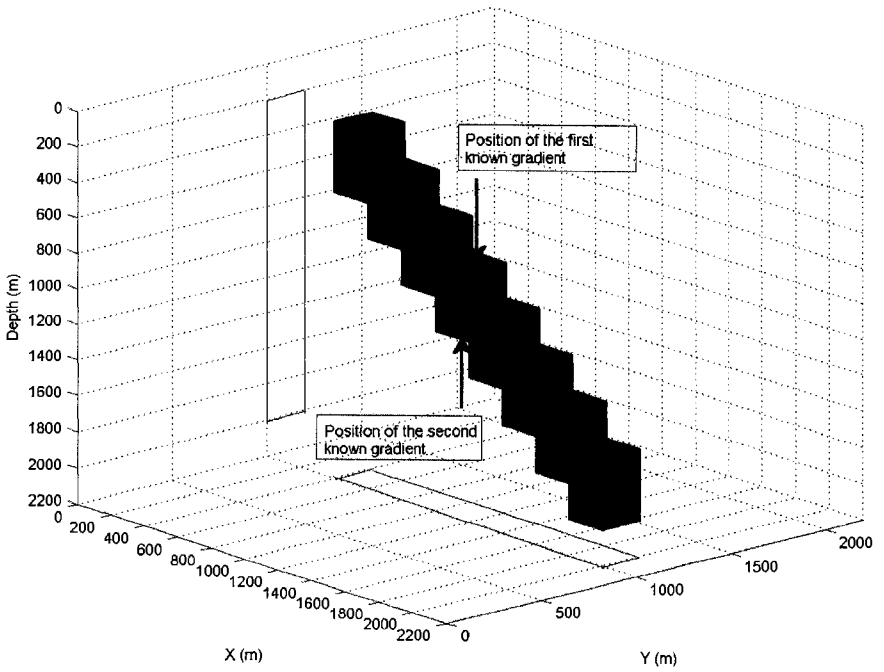

Figure 4.10 The domain selected for dipping dyke and the location of the known gradients

$$\rho^* = \begin{bmatrix} C_{g,\rho}^T & C_{gr,\rho}^T \end{bmatrix} \begin{bmatrix} C_{g,g} & C_{g,gr} \\ C_{gr,g} & C_{gr,gr} \end{bmatrix}^{-1} \begin{bmatrix} g \\ gr \end{bmatrix} \tag{4.2}$$

$$\rho = \lambda^T g + \eta^T gr \tag{4.3}$$

The estimated density distribution at section $y = 1100m$ is shown in Figure 4.11. By comparing Figure 4.11, 4.5 and Figure 4.1, we can see that the inversion using cokriging, when we know some constraints provides a more accurate estimate of the density distribution. The dyke is more accurately located and the density contrasts are very clear.

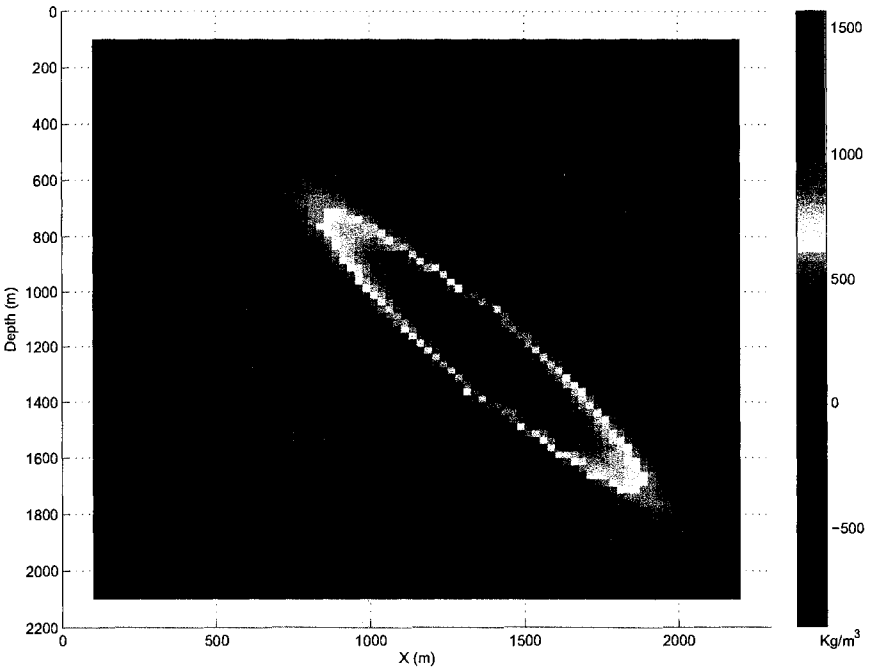

Figure 4.11 The estimated density distribution at section $y = 1100m$ using inversion by cokriging with constraints (known gradients).

4.2 Application to Synthetic Data generated by stochastic method

Densities for $1000m \times 1000m \times 200m$ prisms were generated by non-conditional LU simulation using a spherical model with $C = 60000(kg/m^3)^2$, $a_{h,45} = 5000m$, $a_{h,135} = 9000m$, $a_{vert} = 4000m$ where C is the variogram sill, and $a_{h,45}$, $a_{h,135}$, a_{vert} are the variogram ranges along horizontal directions 45^o and 135^o and vertical direction respectively. The 3D domain is divided into $11 \times 11 \times 21 = 2541$ cubic prisms. Therefore, dimensions of the modeled domain are $11 \times 11 \times 4.2Km$. The LU simulation method has been explained in chapter 3.

Generated density values ρ by LU simulation for our case are shown in Figure 4.12 at four different sections, $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$. Using density values, we can calculate the synthetic gravity data using equation $g = G\rho$, where, G is the matrix of geometric terms that are calculated using equation 2.6. Here, we suppose that gravity data are in the center top face of the surface prisms. From now on, we assume that the generated gravity are real and we would like to invert them in order to estimate the density distribution.

4.2.1 Inversion of synthetic data

We use inversion by cokriging that was discussed before in order to find the density distribution. We first calculate the experimental gravity covariance matrix C_{gg-exp} using the gravity data assuming that they have zero mean. Using the V-V plot method with $Nlag = 100$, we adjust the density covariance matrix $C_{\rho\rho}$ such that experimental and theoretical variograms become almost identical. The adjusted density variogram model is spherical with $C = 90000(kg/m^3)^2$ and $a_{h,45} = 6900m$, $a_{h,135} = 8900m$, $a_{vert} = 4000m$. The covariance vectors v_{exp} and v_{th} are shown in Figure 4.13. It can be seen from this figure that the correlation between them is

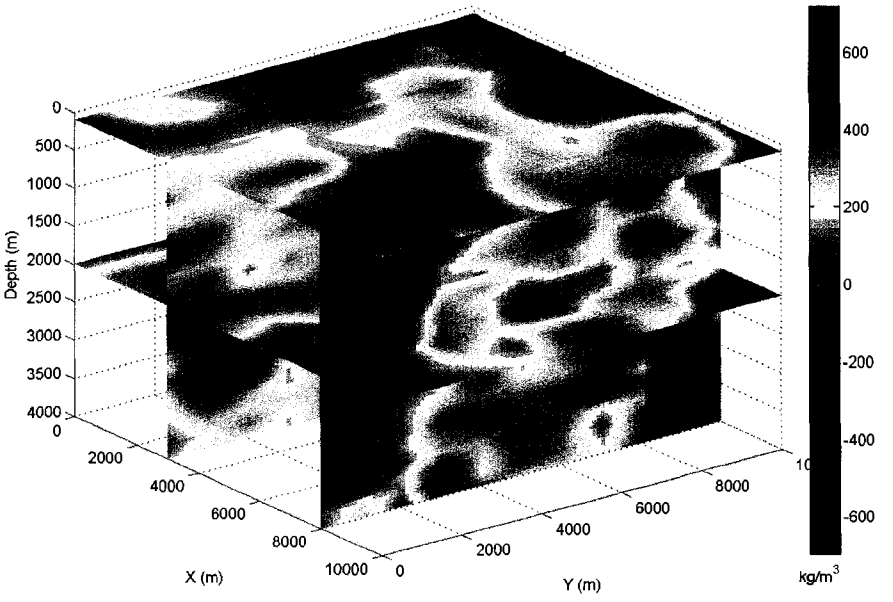

Figure 4.12 Generated density by LU simulation at four different sections, $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$

0.9466.

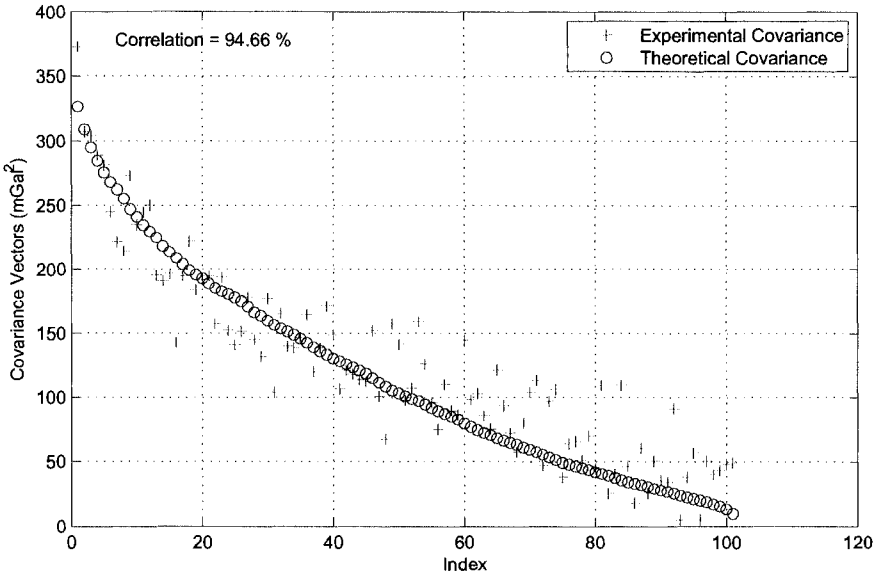

Figure 4.13 Fit of the experimental and theoretical gravity covariance matrices. Density variogram model: anisotropic and spherical with $C0 = 100(kg/m^3)^2$, $C = 90000(kg/m^3)^2$ and $a_{h,45} = 6900m$, $a_{h,135} = 8900m$ and $a_{vert} = 4000m$.

After selecting the density variogram model, we will calculate the density covariance matrix and using that we will calculate $C_{g\rho} = GC_{\rho\rho}$ and $C_{gg} = GC_{\rho\rho}G^T$. Now, we can use the cokriging system equation $C_{gg}\lambda = C_{g\rho}$ to find coefficients λ and then we can find an estimate for density distribution: $\rho^* = \lambda^T g$. The estimated density distribution at 4 sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$ is shown in Figure 4.14.

By comparing Figure 4.12 and Figure 4.14, we can see that inversion method using cokriging is a very effective method to estimate the shape and the density contrast. The obtained density distribution is smoother than the original model; however, the main features are recovered.

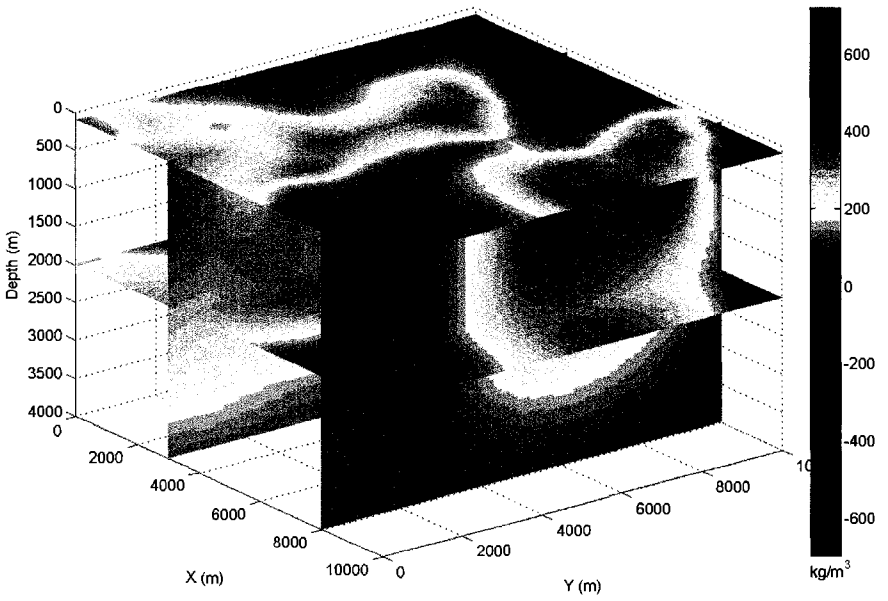

Figure 4.14 The estimated density distribution at 4 sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$ using inversion by cokriging.

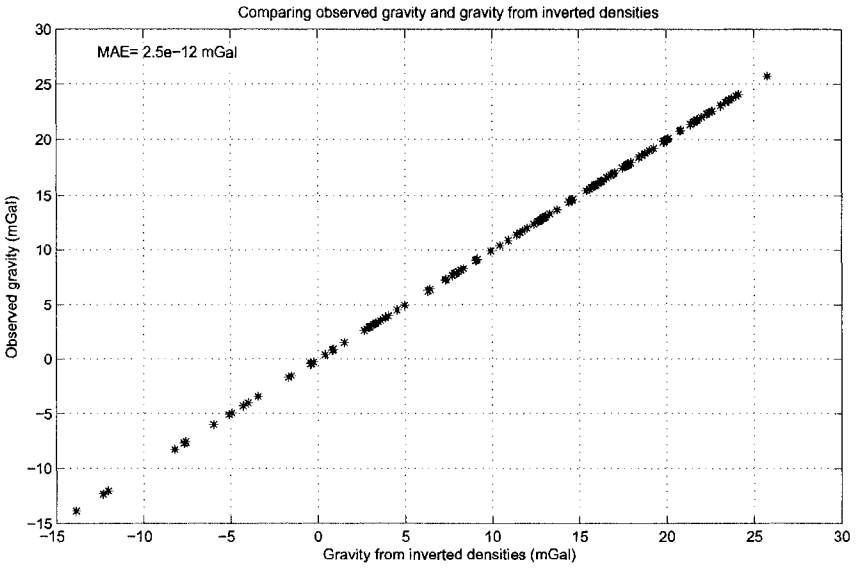

Figure 4.15 Comparing the input gravity with gravity calculated from inverted densities.

In Figure 4.15, we compare the input gravity with gravity of inverted densities. It can be seen from this figure that these two gravity data are almost identical with MAE equal $2.5 \times 10^{-12} mGal$.

4.2.2 Co-simulation of the synthetic model

As for the dyke model, we use the geostatistical simulation algorithm (FFT-MA) and generate 200 conditional simulated density distributions. One of the conditional simulated densities with good correlation with the density distribution at 4 sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$ is shown in Figure 4.16.

We also show the histogram of the synthetic densities, the inverted densities using cokriging and also the conditional simulated densities in Figure 4.17. From this

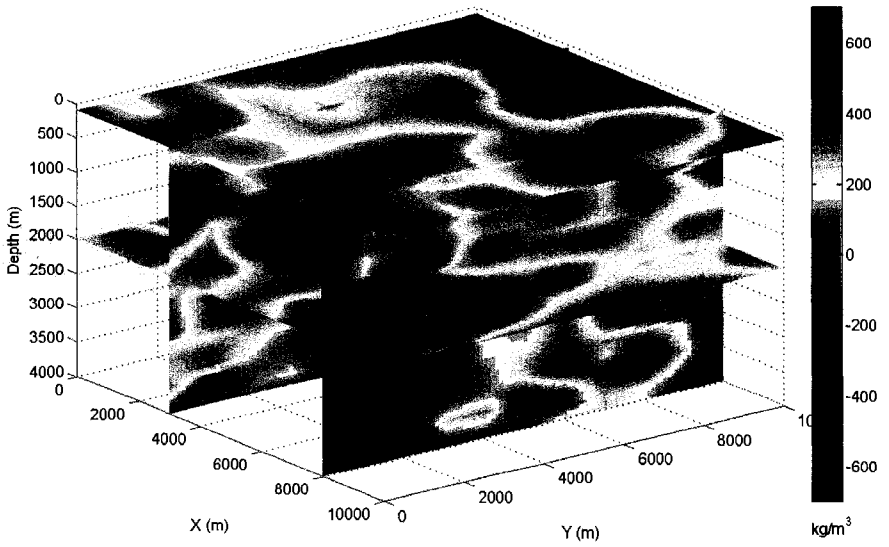

Figure 4.16 The conditional simulated densities at sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$.

figure, we can see that the histogram of the simulated densities shows a better match with the synthetic densities, in term of the mean, variance and distribution compared with the inverted data using cokriging.

4.2.3 Inversion on synthetic data using constraints

In this section, we will add some constraints to our system and we will show that knowing extra information about the system will help us to find the density distribution more accurately.

Here, we use the same domain and the same density model as before. However, we assume that there are 5 boreholes in our domain and we know the densities along these boreholes. We assume this information allows us to estimate properly the density of all the prisms along these boreholes. The 5 boreholes are at points

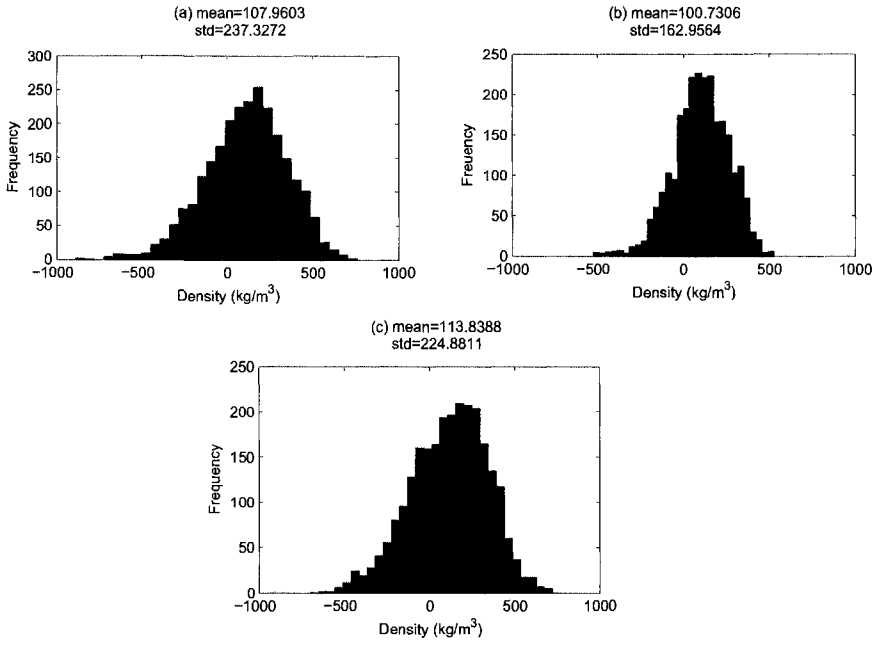

Figure 4.17 The histogram of (a) the conditional simulated densities, (b) the inverted densities using cokriging and (c) the synthetic densities

$(x = 2000m, y = 2000m)$, $(x = 2000m, y = 8000m)$, $(x = 5000m, y = 5000m)$, $(x = 8000m, y = 2000m)$, $(x = 8000m, y = 8000m)$. Figure 4.18 shows a plan view of the domain and the location of the boreholes.

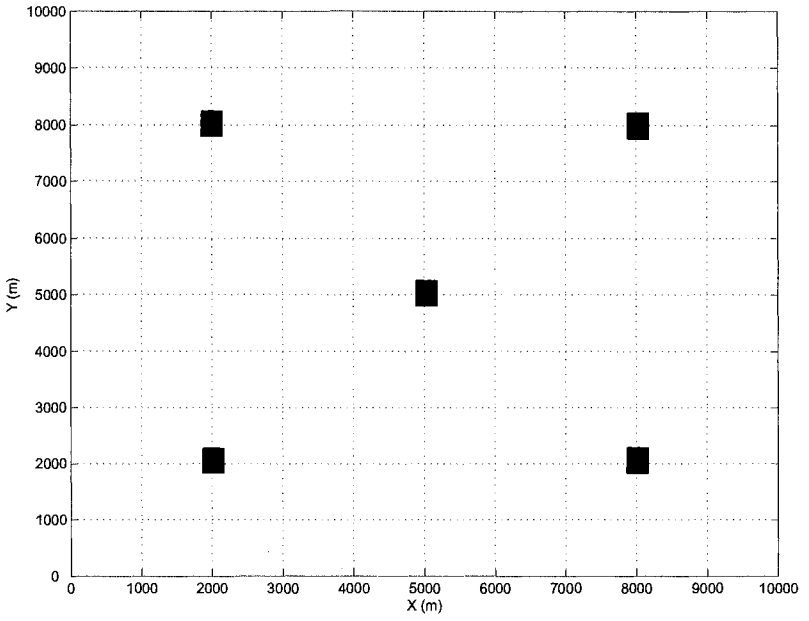

Figure 4.18 Plan view of the domain and the location of the boreholes.

So, we know the densities of $5 \times 21 = 105$ prisms and we can use them as constraints to solve our cokriging equation during inversion. We apply inversion by cokriging using constraints that was discussed before in chapter 3 in order to find the density distribution. We use the same variogram model for the density distribution, the same density covariance matrix $C_{\rho\rho}$ and consequently the same theoretical gravity covariance matrix C_{gg-th}. We form a matrix F for the known densities. F is an $N_F \times M$ matrix where N_F is the number of known densities and M is the total number of prisms. Each row of F corresponds to one of the known densities and all of its elements have value 0 except the one corresponds to the position of the

known density that is 1. Now, we can use the cokriging system using constraints to find the density distribution:

$$\rho^* = \begin{bmatrix} C_{g,\rho}^T & C_{\rho_F,\rho}^T \end{bmatrix} \begin{bmatrix} C_{g,g} & C_{g,\rho_F} \\ C_{\rho_F,g} & C_{\rho_F,\rho_F} \end{bmatrix}^{-1} \begin{bmatrix} g \\ \rho_F \end{bmatrix} \tag{4.4}$$

$$\rho = \lambda^T g + \mu^T \rho_F \tag{4.5}$$

The estimated density distribution is shown in Figure 4.19. By comparing Figures 4.19, 4.14 and Figure 4.12, we can see that using cokriging, when we know some constraints provides more accurate estimate of the density distribution. In order to see the effect of constraints in inversion by cokriging, we consider two different situations. At first we suppose that we only have the information from 4 boreholes, excepting the one at point $(x = 5000m, y = 5000m)$. Then we suppose that we have the information of all 5 boreholes. Figure 4.20 shows the synthetic density and the inverted density distribution with and without constraints at section $y = 5000m$. As we can see from this figure, increasing the amount of information as constraints will improve the estimation of density distribution especially at larger depths.

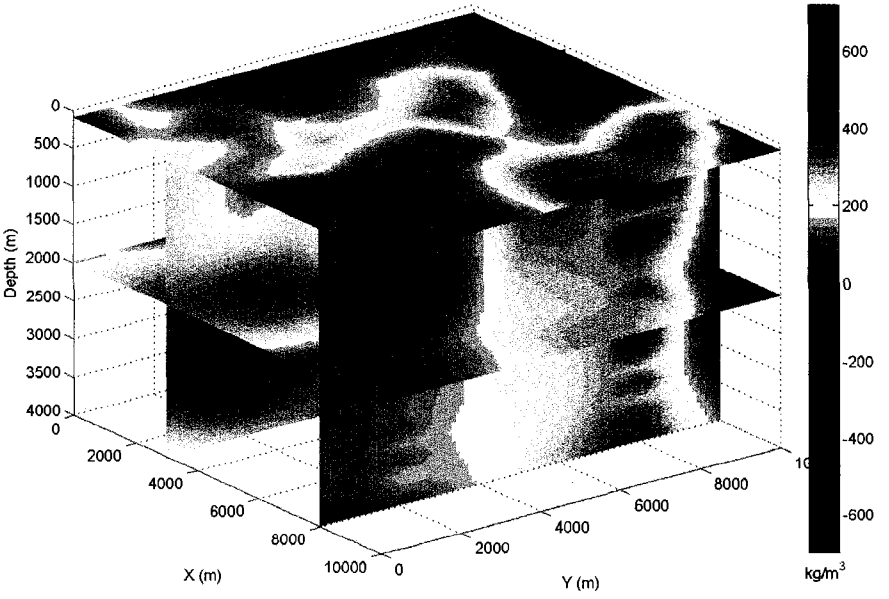

Figure 4.19 The estimated density distribution at 4 sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$ using inversion by cokriging with constraints (5 boreholes).

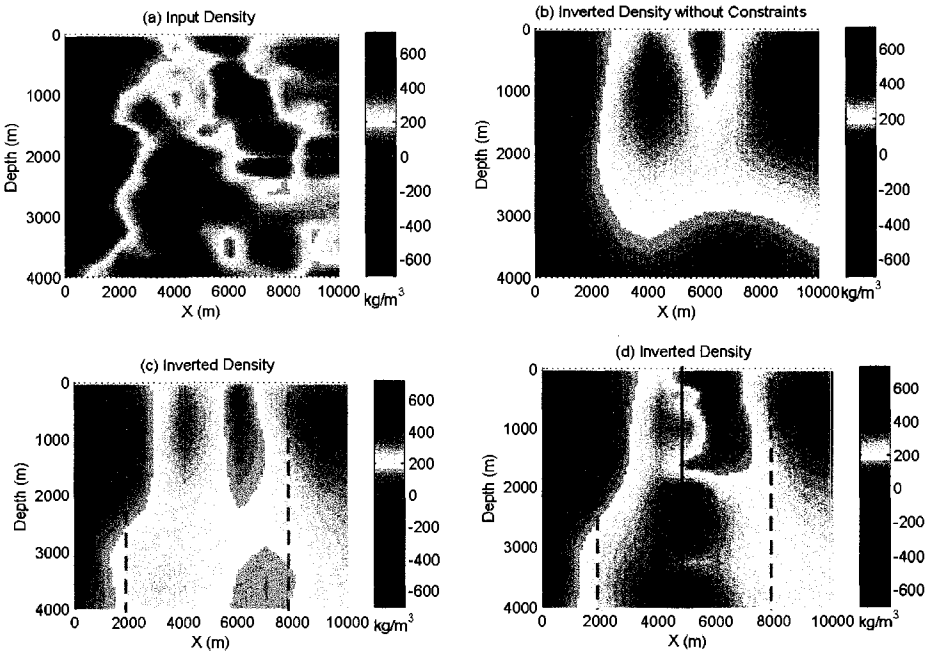

Figure 4.20 (a)The synthetic density, (b)the inverted density without constraints, (c)the inverted density with constraints (4 boreholes) and (d)the inverted density with constraints (5 boreholes) at section $y = 5000m$.

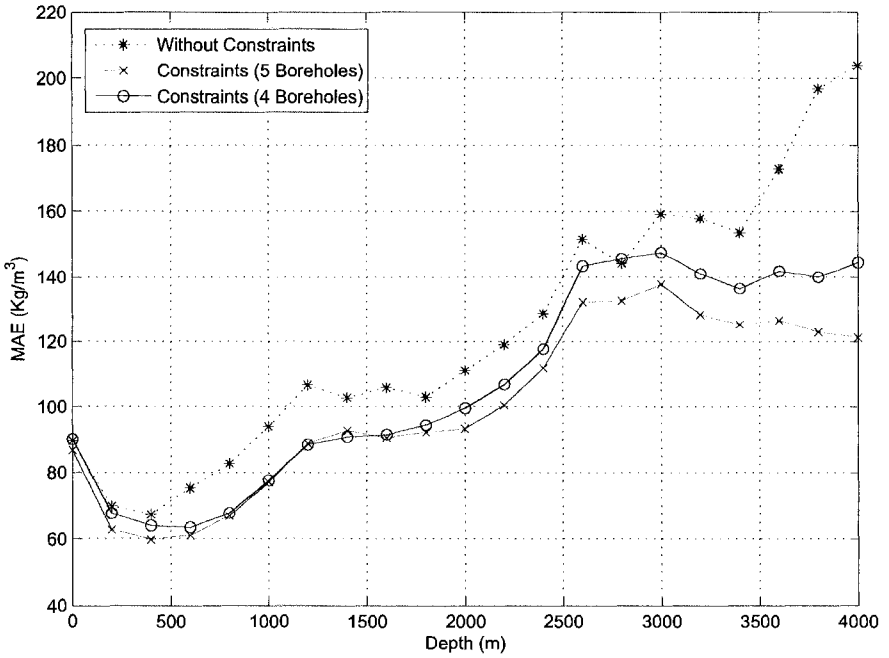

Figure 4.21 The mean absolute errors (MAE) obtained for the densities as a function of depth. Three situations: without constraints, with constraints (4 boreholes) and with constraints (5 boreholes).

Figure 4.21 shows the mean absolute errors (MAE) obtained for the densities as a function of depth. We can see from this figure that in general MAE's tend to increase with depth. In addition, increasing the amount of information as constraints helps reduce the MAE and generates, as would be expected, solutions more closely approaching the simulated reality at depth.

Figure 4.22 shows the synthetic densities versus estimated densities by cokriging with and without constraints. As we can see from this figure, by adding constraints estimated densities become closer to the synthetic densities.

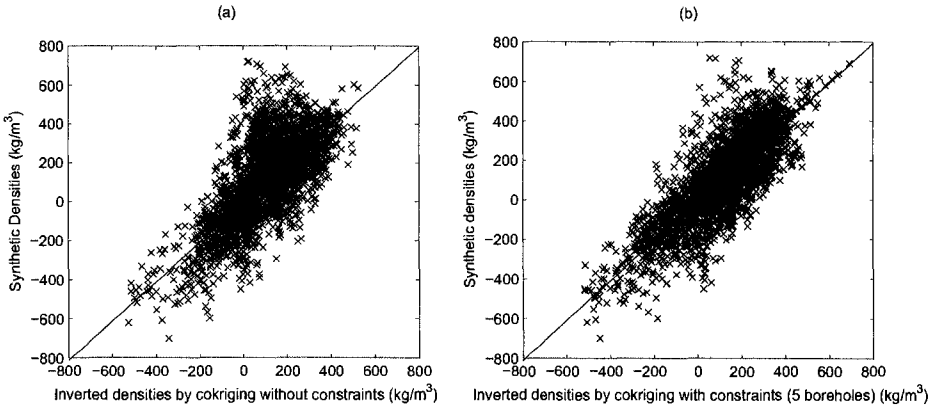

Figure 4.22 The synthetic densities versus estimated densities by cokriging (a) without constraints, (b) with constraints (5 boreholes).

4.2.3.1 Co-simulation of the synthetic model using constraints

Again, we assume information from 5 boreholes as our constraints and we use FFT-MA method and generate 200 conditional simulated density distributions. One of the conditional simulated densities showing good correlation with the density distribution in 4 sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$ is shown in Figure 4.23.

The synthetic densities, the inverted densities using cokriging with constraints and also the conditional simulated densities with constraints are shown in Figure 4.24. From this figure, we can see that the histogram of the simulated densities shows a better match with the synthetic density distribution, in term of the mean, variance and distribution compared with the inverted data using cokriging. Also by comparing this figure with 4.17, we can see that by adding constraints, the mean,

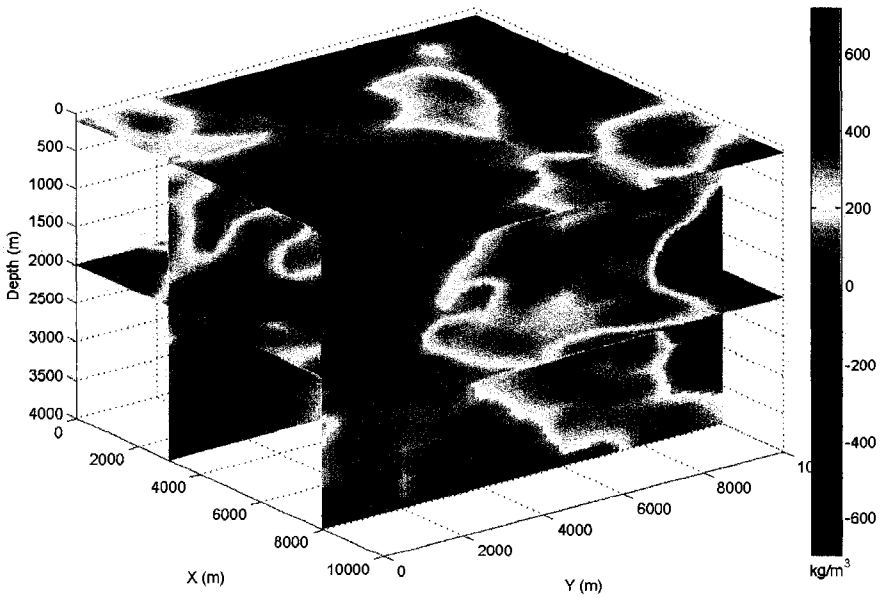

Figure 4.23 The conditional simulated densities with constraints (5 boreholes) at sections $z = 200m$, $z = 2000m$, $x = 3000m$ and $x = 8000m$.

variance and distribution of estimated densities become closer to the synthetic densities.

Figure 4.25 shows the simulated density distribution with and without constraints at one section ($y = 5000m$) for comparison. Only known densities are used as constraints. Adding other constraints such as known gradients can be done the same way.

Here, we should mention that geostatistical simulation of the density distribution (ρ) allows computing the probability that the density falls under or above a given range (threshold). As an example, Figure 4.26 shows the maps of the probability that the density falls below 200, 300 and 400 kg/m^3 at the section $y = 5000m$.

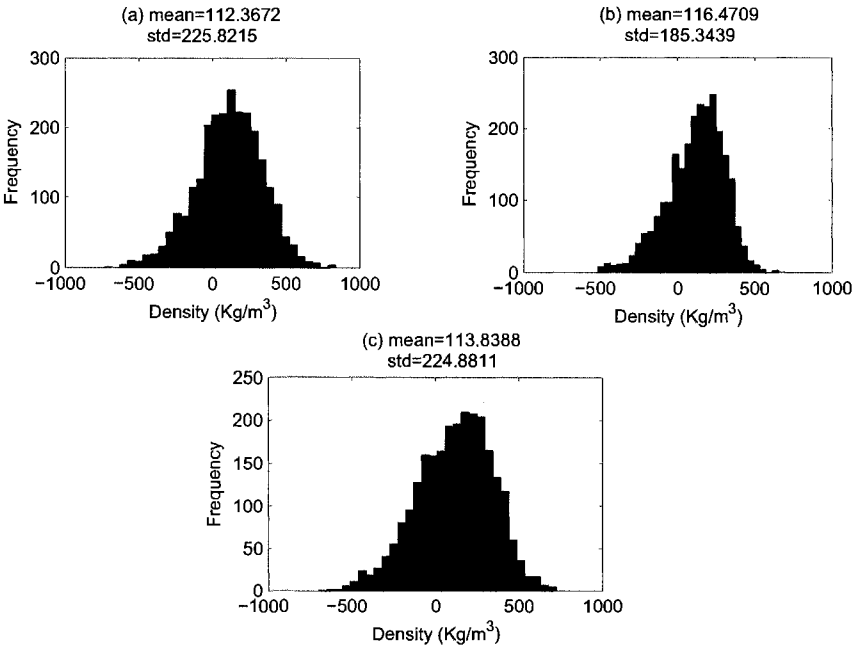

Figure 4.24 The histogram of (a)the conditional simulated densities with constraints (5 boreholes), (b)the inverted densities using cokriging with constraints (5 boreholes) and (c)the initial densities

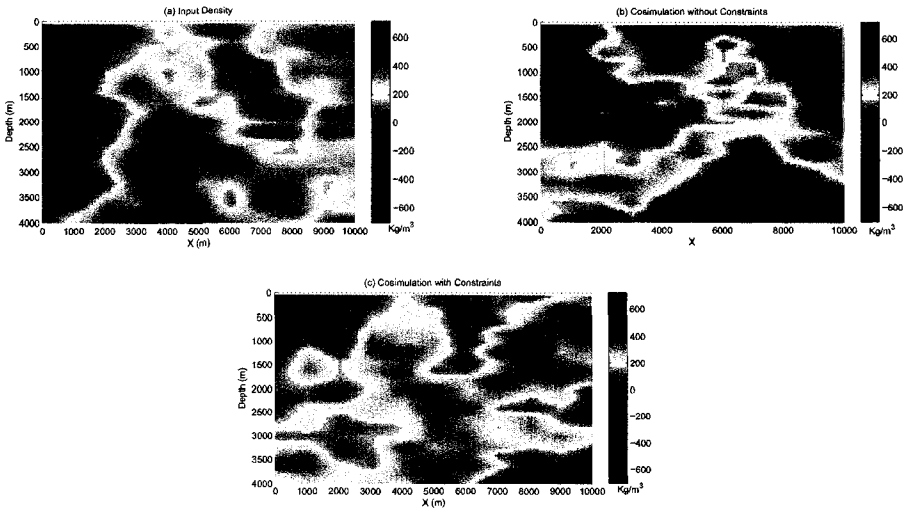

Figure 4.25 (a)Synthetic density distribution, (b) the conditional simulated densities without constraints and (c)the conditional simulated densities with constraints (5 boreholes) at section $y = 5000m$

4.3 Summary and discussion

We separate this section in two different parts:

Results of testing the inversion algorithm by cokriging:

- The non-iterativeness of the method allows us to easily repeat the program.
- The algorithm is very flexible for both regular and non-regular grids.
- The algorithm is tested in presence of noise and its performance is stable.
- For obtaining a good fit between theoretical and experimental covariance matrix, it is necessary to carefully consider the geological information available from the domain under study (dyke model).

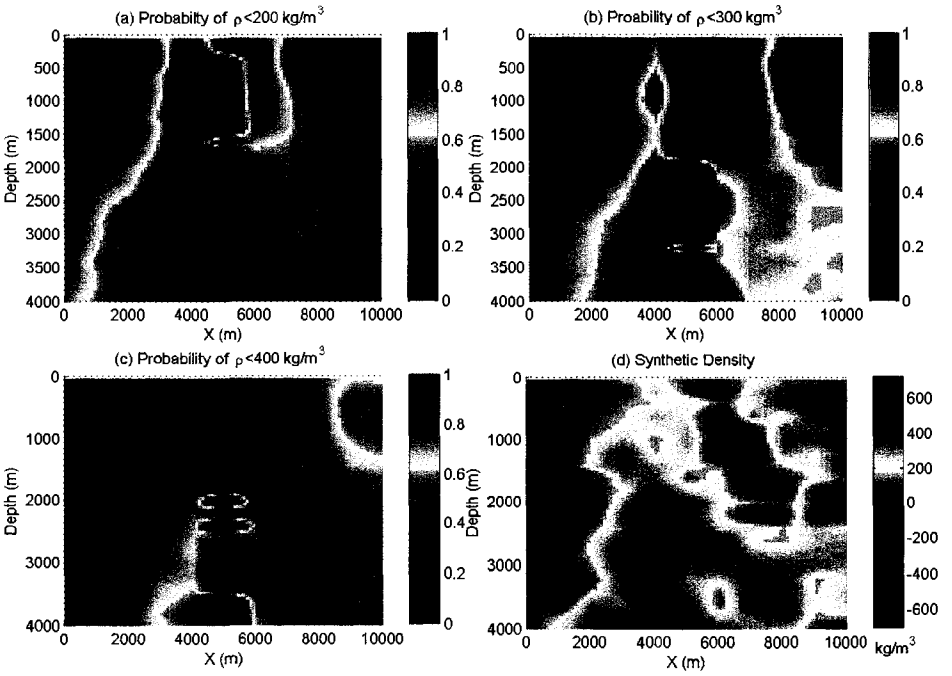

Figure 4.26 (a)Probability of $\rho < 200kg/m^3$, (b) Probability of $\rho < 300kg/m^3$ (c)Probability of $\rho < 400kg/m^3$ and (d) Synthetic Density. All the figures are at section $y = 5000m$

- For fitting the theoretical and experimental covariance matrices with an automatic algorithm (Simplex method), we have to be careful about the initial model that we input to the algorithm. We should be aware about giving more weight to the first part of variogram (the parts corresponding to small distances).

Comparing the result of different methods applied on synthetic data:

- For compact bodies without anisotropy, the least squares inversion method gives good results with a good spatial resolution only in the x and y directions. However, they can not provide accurate estimate for the orientation of the body and also do not show good resolution in depth.
- We saw that the orientation of dyke can be recognized better with stochastic methods.
- We found out that the histogram of the simulated densities shows a better match with the histogram of the synthetic densities, in term of the mean, variance and distribution compared with the inverted data using cokriging.
- Increasing the amount of information as constraints will improve the estimation of density distribution especially at larger depths. For the dyke model, we used some known gradients as the constraints and we saw that they had a very positive effect on the inversion by cokriging and on the reconstruction of density distribution contrasts. For the synthetic model, we used some borehole densities as the constraints and we saw their effectiveness in estimation of the density distribution.
- Geostatistical simulation provides probability maps with different thresholds that can be used as a tool in the interpretation of the inversion results.

CHAPTER 5

APPLICATION TO REAL DATA

The purpose of this chapter is to apply the proposed method based on cokriging and co-simulation to gravity data measured in the Matagami area in order to determine the subsurface structure. We first introduce the location of the study area and its geological setting. Then, we prepare the Bouguer gravity anomaly data in order to make them suitable for use by the proposed algorithm. This preparation comprises the re-sampling with kriging and upward continuation for separating the residual and the regional gravity data. The methods of cokriging inversion and co-simulation are applied to the prepared data and all the results are shown in 3D. Finally, using the results of these methods, we interpret the density distribution in the studied area.

5.1 Study area: Matagami region

The study area is located in the Matagami region(west of southern flank of Matagami mining camp), which is, due to intense mining and exploration activity, one of the most studied areas of the Abitibi area. The area of the gravity survey extends from -77° 11′ to -77° 36′ latitude and from 49° 38′ to 49° 49′ longitude. 1570 gravimetric measurements with about a $600m$ spacing are available for the entire studied region. The data distribution is irregular. The data were corrected for drift, free air and latitude to yield the Bouguer anomaly. The distribution of the Bouguer anomaly values can be seen in the Figure 5.1, where the range is -63.5 to $-44mGal$.

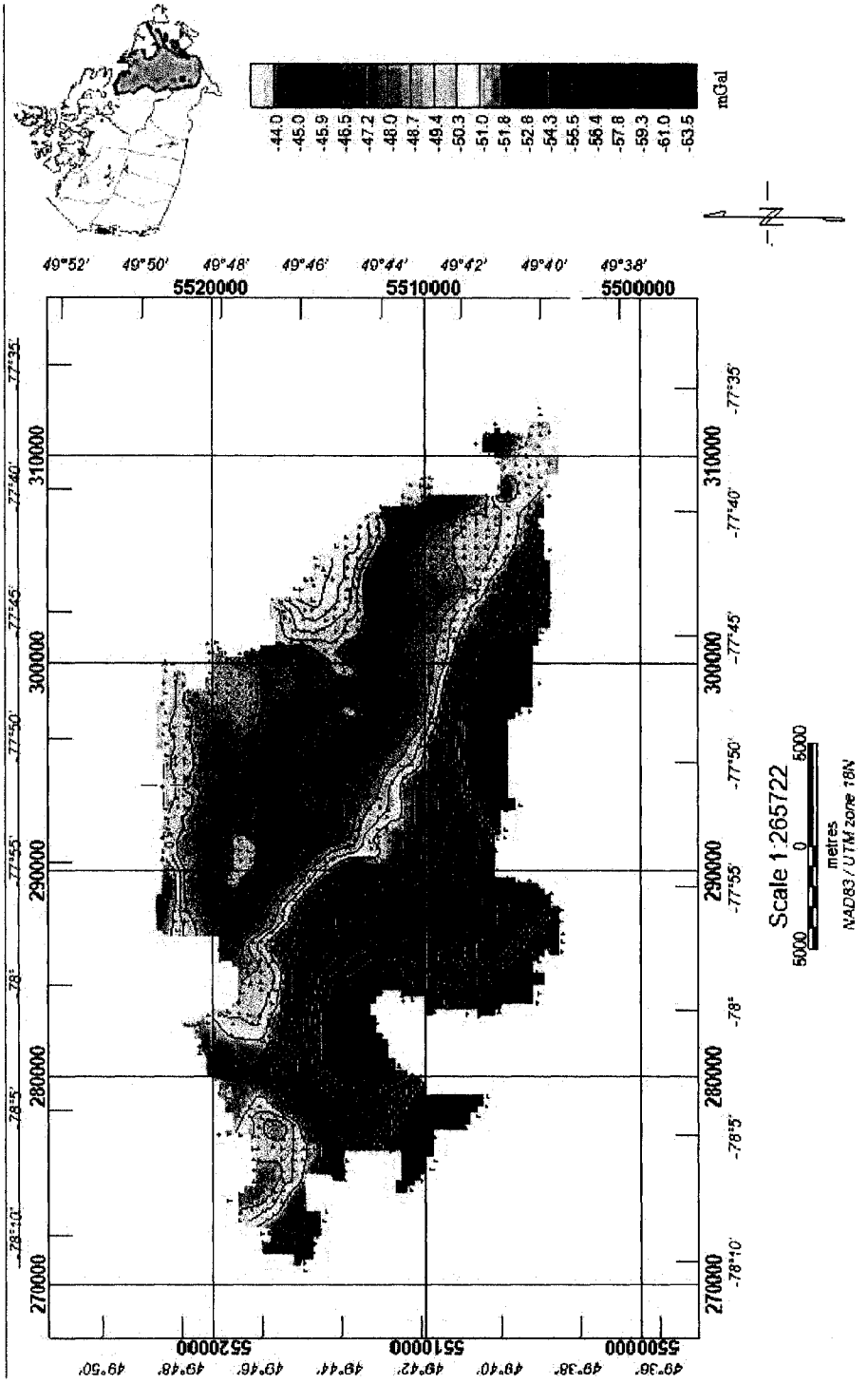

Figure 5.1 Distribution of Bouguer anomaly from the region of Matagami

5.1.1 Geological setting

The study area is located in the northern part of the Abitibi Sub-province, one of the largest Archean greenstone belts in the world. Many volcanogenic massive sulfide (VMS) deposits have been identified at the contacts of bimodal volcanic sequences in the Archean-age Abitibi greenstone belt, which is located on the border of the province of Ontario and Québec. The Matagami volcanic complex of northern Abitibi belt is formed by two major phases of volcanism (Piché et al., 1993) :

- The initial phase produced basaltic volcanism and formed the Wabassee Group.
- The late phase is dominated by rhyolites of the Watson Lake Group.

A cherty, sulphidic chemical sediment known as the Key Tuffite marks the contact and discontinuity between the two groups. This thin horizon $(0.6 - 6m)$ is the primary exploration target because it hosts most of the ore bodies discovered in the area (Calvert and Li, 1999). Exploration of the Key Tuffite by systematic drilling is very expensive especially with increasing depth (up to 1600 m), thus use of geophysical methods can be a very efficient solution. The Key Tuffite is too thin to be resolved by this gravity survey, but the important issue is that the Key Tuffite is located at the contact between the mafic Wabassee Group and the felsic Watson Lake Group. The mafic Wabasee Group has a high density, so it produces gravity highs. On other hand, the felsic Watson Lake Group (rhyolite) has a low density, so it produces gravity lows. The geological setting of the studied area is shown in Figure 5.2.

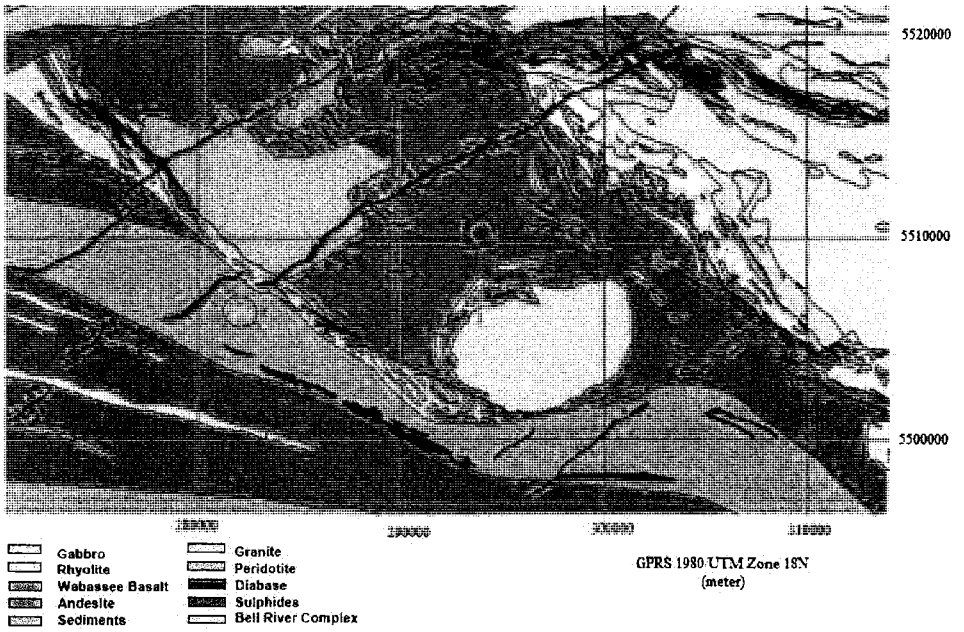

Figure 5.2 The geological setting of the studied area

5.2 Preparing the gravity data for inversion

Even though the method proposed can be applied to irregular observation data, we re-sample data using kriging so that they have a regular spacing so that we can use fast Fourier transforms for upward continuation. Kriging methods provide minimum errors for re-sampling. Marcotte and Chouteau (1993) proposed a method for gravity data transformation by kriging that can be applied even on non-regular grids.

The second preparation method is separating the residual and the regional anomalies. This task is subjective. Thus methods like smoothing and upward continuation are examined. Then, some methods were applied to calculate the maximum height for upward continuation.

Finally, we should mention that the study area is covered by overburden of glacial origin consisting of clay, sand, gravel and till of variable thickness from 0 to $60m$. Most of this information is deduced from EM methods surveyed by Xtrata Zinc. To correct for the problem of overburden, we remove the effect of gravity caused by overburden from the measured gravity data. We calculated the gravity caused by overburden and the range was between 0 to 3 mGal at the worst case with the average density of the soil. It should be mentioned that in some locations, the gravity caused by overburden was small (below 1 mGal) and negligible.

5.2.1 Re-sampling by kriging

Here, the Bouguer anomaly is interpolated, by kriging, on a $600m$ grid, giving 2061 kriged points. The experimental variogram of the gravity data is estimated by a gravimetric model. The omni-directional gravity variogram with $a = 4900m$ and $C = 68(mGal)^2$ is shown in Figure 5.3.

5.2.2 Separating the residual and the regional Fields

The subject of regional and residual field separation has been a vital subject in gravity and magnetics for decades. The procedures of separating the regional and the residual fields may be divided into two groups: those which favor a mathematical approach (polynomial fitting, spectral and space-domain filtering, upward continuation), and those which involve a more and less visual smoothing on profiles or maps. None of these procedures is devoid of assumptions and each method may yield non-unique results. For example, in the spectral factorization technique, the filter design is dictated by the clarity of the slope change between the short and long-wavelength features and, in the case of the upward continuation technique,

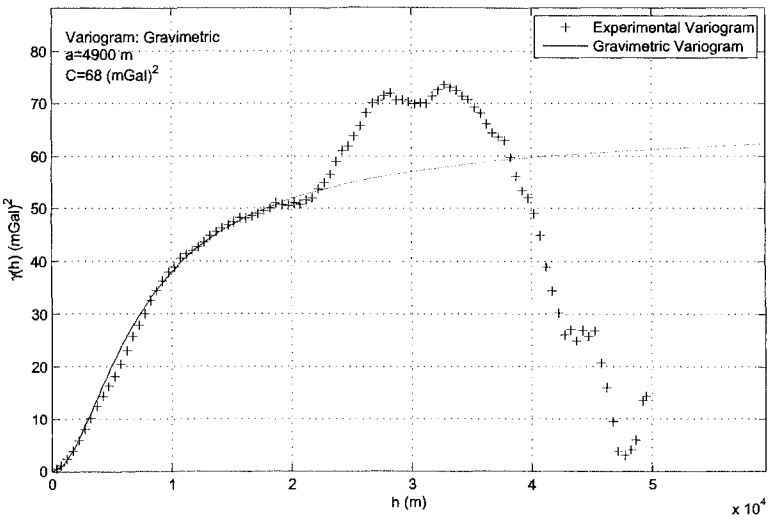

Figure 5.3 The experimental variogram of the gravity data along with the modeled gravimetric variogram with $C = 68mGal^2$, $a = 4900m$.

by the choice of the continuation height. The graphical method is empirical and clearly non-unique (Gupta and Ramani, 1980).

Without entering into a general discussion of the various possible definitions of regional and residual fields, the term residual is used for the field modeled by a shallow to intermediate source distribution, and the term regional for the field with a deeper origin. Here, we try to find the residual field of our Bouguer anomalies that is used by inversion to estimate the density distribution of our region. We applied two methods to separate the regional and residual fields: smoothing and upward continuation that will be explained in the next sections.

5.2.2.1 Smoothing

Under conditions such as our case, and indeed in any situation for which the survey in a broad area is available, the residual field may be found by subtracting this data (as regional field) from the Bouguer anomaly. The gravity anomaly grid of Canada is available at a $2km$ interval for all provinces. This grid can be applied as a regional anomaly to be subtracted from our local anomaly in Matagami area. The result will be the residual field.

The $2km$ gravity anomaly grid of Canada for our region and its subtraction from our Bouguer anomaly, are shown respectively in Figure 5.4 and 5.5.

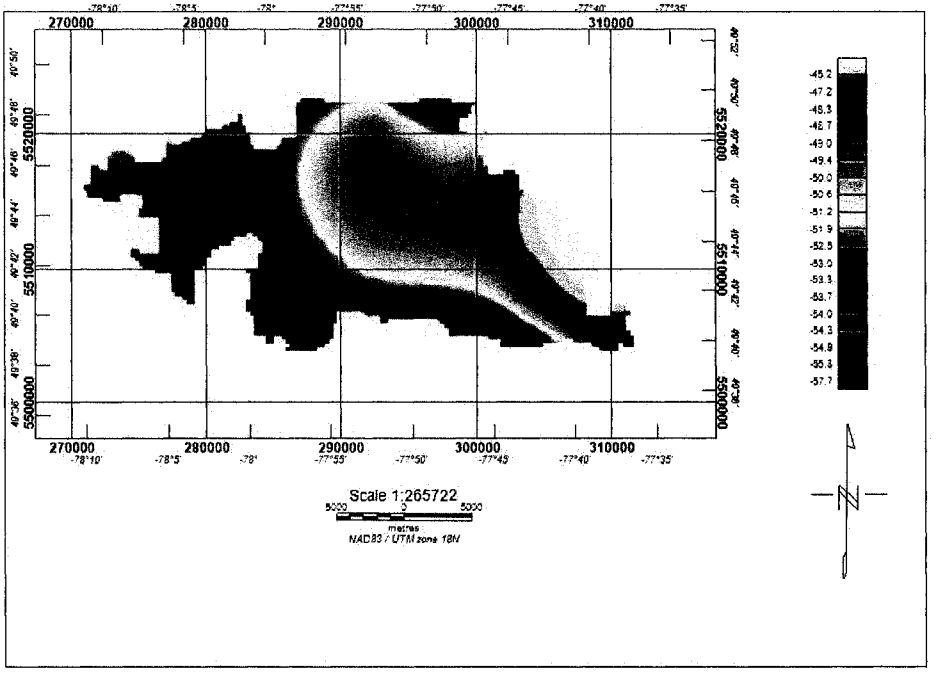

Figure 5.4 Original data of $2km$ gravity anomaly grid of Canada for Matagami region

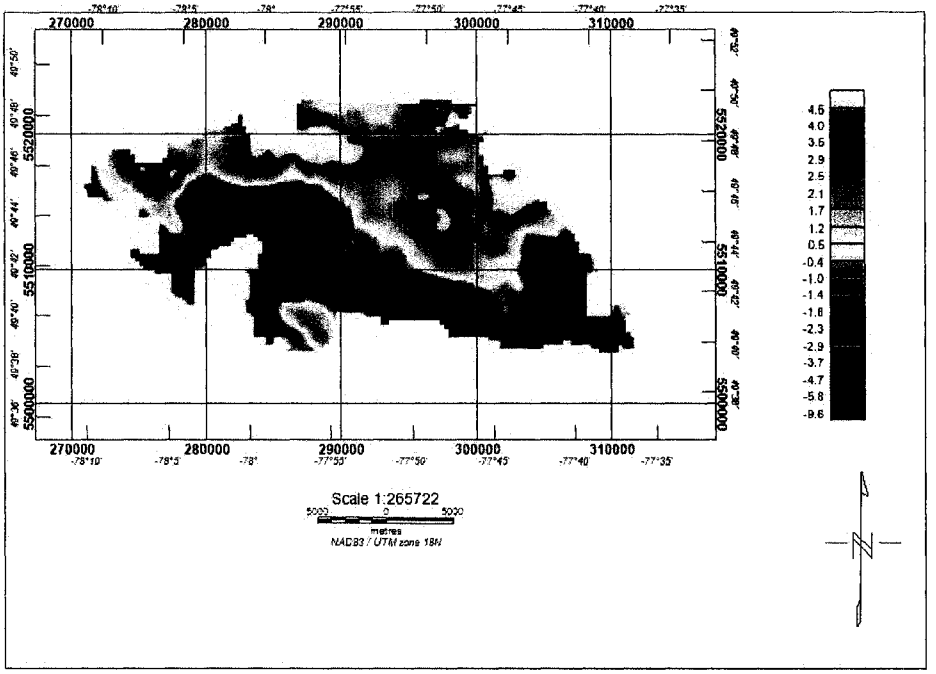

Figure 5.5 Matagami Residual field obtained from smoothing

5.2.2.2 Upward Continuation

Upward continuation transforms the potential field measured on one surface to the field that would be measured on another surface farther from all sources. As we will see, this transformation attenuates anomalies with respect to wavelength; the shorter the wavelength, the greater the attenuation. Upward continuation can be used to attenuate the shallow-source anomalies in order to emphasize deeper sources (Blakely, 1995).

Jacobsen (1987) proposed to use the upward continuation operator to build a convenient, standard family of separation filters. In fact, he designed some Wiener filters, which minimize the inevitable separation errors, from previous statistical source models suggested formerly. A major problem with this approach is to choose the

optimum height for continuation. Some people make this choice by inspection or comparing gravity anomalies upward continued to different height. Unfortunately, both of these methods don't have objective criteria to give the proper height for upward continuation.

It should be mentioned that the optimum filter for extraction of the regional is:

$$H_{reg}(\mathbf{k}) = exp(-2kz_0) \tag{5.1}$$

Where k is the norm of the wave-vector $\mathbf{k}$. In fact, the equation is a wavenumber expression for upward continuation to the height $2z_0$ above the measurement plane. Here, z_0 is depth to source and $2z_0$ is the height of upward continuation.

Upward continuation can efficiently separate the regional and the residual fields, if only we have a good estimate for optimum height. In the following, we will present four methods to calculate the height that can be used by upward continuation. Based on the results of these methods, we choose the height of $10km$ for upward continuation. The regional anomaly is obtained by upward continuation at $10km$ of the Bouguer anomaly in order to invert data on a grid of $5km$ thick. Subtracting the regional from the Bouguer anomaly results in the residual anomaly (Figure 5.6) ranging between -6.2 to 6.5 mGal. It can be seen that the residual anomaly obtained from upward continuation has a good correlation with the residual field obtained from smoothing.

In the end, we used the residual field obtained from upward continuation as the input to our inversion algorithm. Although, we can also use the method of gravity data transformation by kriging (Marcotte and Chouteau, 1993).

In this part, we explain and compare four methods to calculate the optimum height for upward continuation. These four methods are: Euler deconvolution (Thompson,

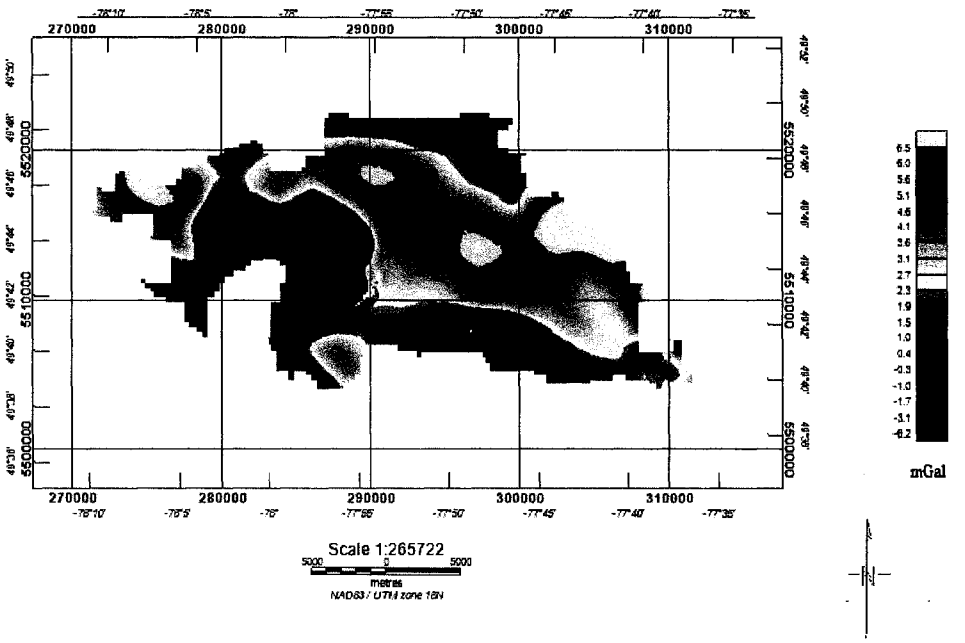

Figure 5.6 Residual anomaly obtained from upward continuation ($10km$)

1982), DEXP (Depth from Extreme Points) (Fedi, 2007), Radially averaged power spectrum and an empirical method by Zeng et al. (2007).

A. Radially averaged power spectrum

In general, power spectra of gravity data can be roughly divided into three segments. The part at the low-frequency (long wavelength) end of the spectrum with a steep slope in power is termed regional, that is, due to sources that are deep or broad. At high frequencies (short wavelengths), the residual part has a flatter slope and is due to relatively shallow sources. Spector and Grant (1970) showed that a single straight line fitted to a part of the spectrum corresponds to a single average depth. At very high frequencies, the spectrum is dominated by the effects due to

measurement errors, digitizing errors, etc. The spectral power of the random errors is independent of the frequency, and thus the spectrum flattens out at the high frequency end (white noise). When considering a grid that is large enough to include many sources, the log spectrum of this data can be interpreted to determine the statistical depth to the top of the sources.

The depth of an ensemble of sources can be determined by measuring the slope of the energy (power) spectrum and dividing it by 4π using the following expression: $h = -s/4\pi$. Where h is the depth and s is the slope of the log (energy) spectrum. In the following figures, the energy is normalized by subtracting the log of the average spectral density.

The Figure 5.7 which was plotted using Oasis Montaj (Geosoft), illustrates the reduction in energy with increasing the wavenumber. Based on this figure, the regional depth is around $2km$ and the residual depth is around $5km$.

B. A fast method to determine the depth and the structural index of potential fields sources

Fedi (2007) showed that potential field enjoy valuable properties when they are scaled by specific power of the altitude. The method called DEXP(Depth from Extreme Points) allows estimation of source depths, density, and structural index from the extreme points of a 3D field scaled according to specific power laws of the altitude. The scaling exponent of such laws is shown to be related to the structural index involved in Euler deconvolution theory.

This method applies to the Newtonian potential and to its nth-order derivatives and will be characterized by a high-degree of stability and accuracy in retrieving the location, the type(SI), and the excess source density. The scaling function τ is defined as the derivative of the logarithm of the potential field (any n-th vertical

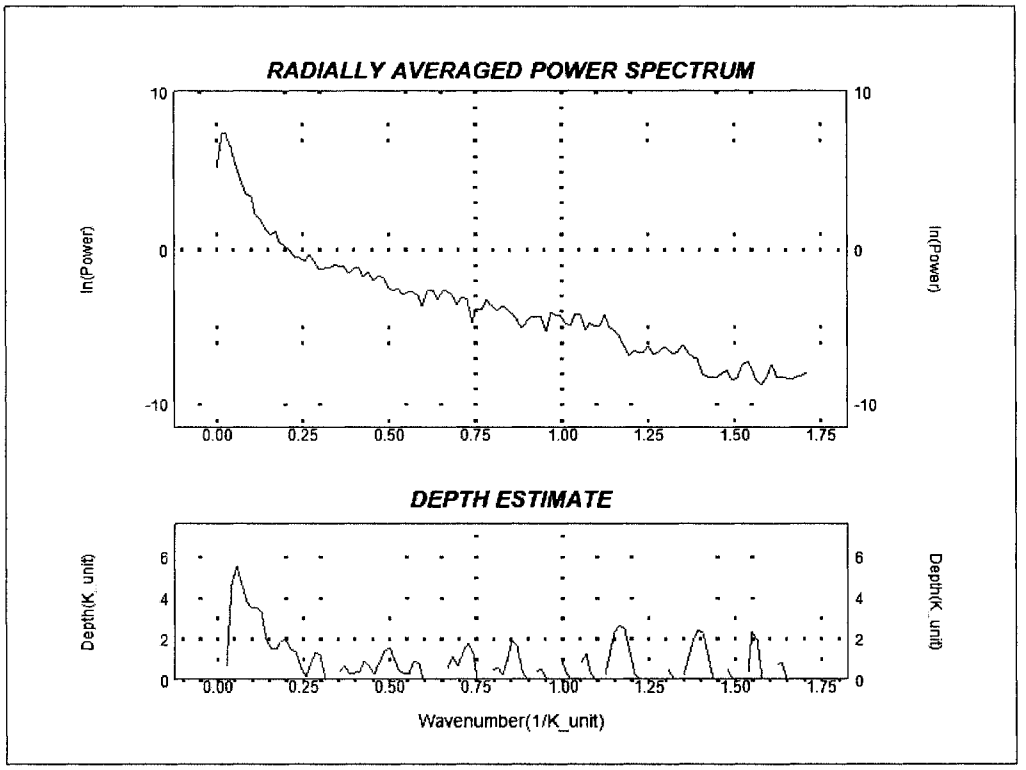

Figure 5.7 Radially Averaged Power Spectrum

derivatives) with respect to $log(z)$:

$$\tau(z) = \frac{\partial log(f_n(z))}{\partial log(z)} \tag{5.2}$$

So that at $z = -z_0$ (depth to source), the meaningful extreme point occurs for the following function:

$$W_n = z^{\alpha_n} f_n \tag{5.3}$$

The criterion of the DEXP scaling-function intercept is used to assess the scaling exponent directly from data. It follows directly from the form of the scaling func-

tion. Putting $z = \frac{1}{q}$, the scaling function may be rewritten from equation 5.2 as a function of q. For example, for a one-point,e.g, a pole source (Fedi, 2007):

$$\tau_n(q \to 0) = 2\alpha_n = -S_n \tag{5.4}$$

Hence, the intercept of τ_n versus q will give us an estimate of the structural index S_n; this allows the field to be scaled with $\alpha_n = 0.5S_n$. As we can see, this estimate does not depend on z_0 which is unlike other theories involving the structural index such as Euler deconvolution.

We have performed this method in three steps:
First, we create a 3D data volume of the potential field. In practice, this means that data are needed both horizontally and vertically. It is rare to have data sets available at different elevations. We can overcome this problem using upward continuation. The errors of upward continuation can be kept low by performing the circular convolution on a larger area than that of interest. So the input data sequences is extended to a greater length by mathematical extrapolation using maximum entropy prediction. A 3D gravity field was generated with a $100m$ step along the vertical directions extended to 5km. The process of upward continuation is shown in Figure 5.8. Second, 3D field $f(z)$ is transformed into a scaled field W_n. To do this, we used the criterion of DEXP scaling-function intercept described above to calculate α. This process is shown in Figure 5.9 where τ is sketched as a function of q. Using this figure, $\alpha = 0.03$. Third, we can determine the depth to source, which involves a search for the extreme points of the scaled potential field. For simplicity, this process for one section has been shown in Figure 5.10.

The value of α that we calculated in this method can be used by Euler deconvolution method which will be described in the next section. In fact, 2α is the structural index SI that is used by this method. The depth to source is shown in Figure 5.10.

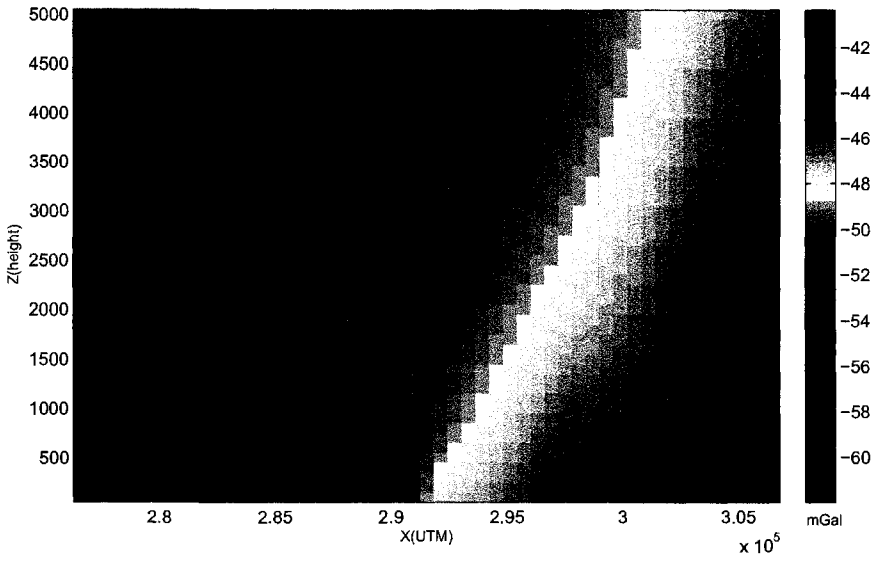

Figure 5.8 One section of 3D gravity field at $y = 5513674$; The 3D field was synthetically generated with a $100m$ step along the horizontal and vertical directions up to $5km$ with upward continuation.

C. Euler deconvolution method

Thompson (1982) suggested a method for rapidly making depth estimates from large amounts of magnetic data. The technique is based upon Euler's homogeneity relationship (hence, the acronym EULDPH). Even though applications to gravity are fewer, some authors (Keating, 1998) have applied this method to gravity data.

Euler deconvolution requires a well-founded understanding of a critical parameter, the Structural Index (SI), which characterizes the source geometry.

Strictly speaking, the function $f(\mathbf{v})$ of a set of variables $\mathbf{v} = (v1, v2, ...)$ is said to be homogeneous of degree n if $f(t\mathbf{v}) = t^n f(\mathbf{v})$. Moreover, it can be shown that if $f(\mathbf{v})$

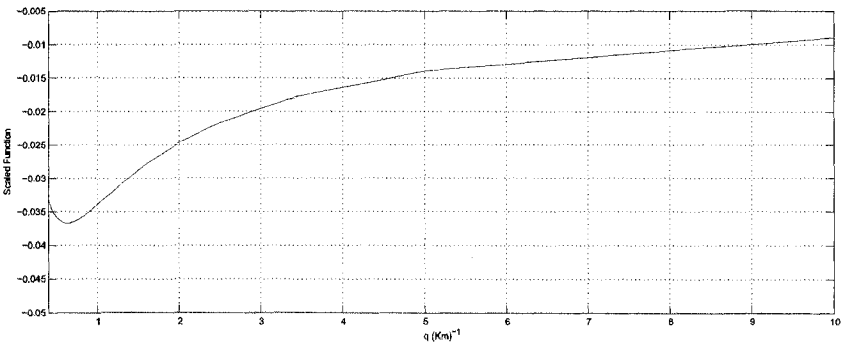

Figure 5.9 τ versus q for calculation of α

is homogeneous of degree n and differentiable, the following equation is satisfied:

$$\mathbf{v}\nabla_v f(\mathbf{v}) = nf(\mathbf{v}) \tag{5.5}$$

This partial differential equation is known as Euler's homogeneity equation or simply as Euler's equation. Suppose that $f(\mathbf{v})$ has the general functional form of $f(\mathbf{v}) = \frac{G}{r^N}$ where $r = \sqrt{(x^2 + y^2 + z^2)}$, and $N = 1, 2, 3, ..;$ G is not dependent on x, y, and z. Clearly, equation is homogeneous of order $n = -N$. For convenience we define the Structural Index (SI) as N.

In this method, we used the value of α from previous section to calculate the structural index SI. Then, the calculated value of SI is used to solve the partial differential equation above. In order to solve this equation, we used the codes written by Keating (1998) and Durrheim and Cooper (1998). One section of the area is shown in Figure 5.11. In this figure, the depth to source of different locations are shown. It shows that most depths to source are within the first $5km$.

D. Empirical method of estimating the optimum upward-continuation height

Zeng et al. (2007) proposed an empirical method based on model studies, to derive

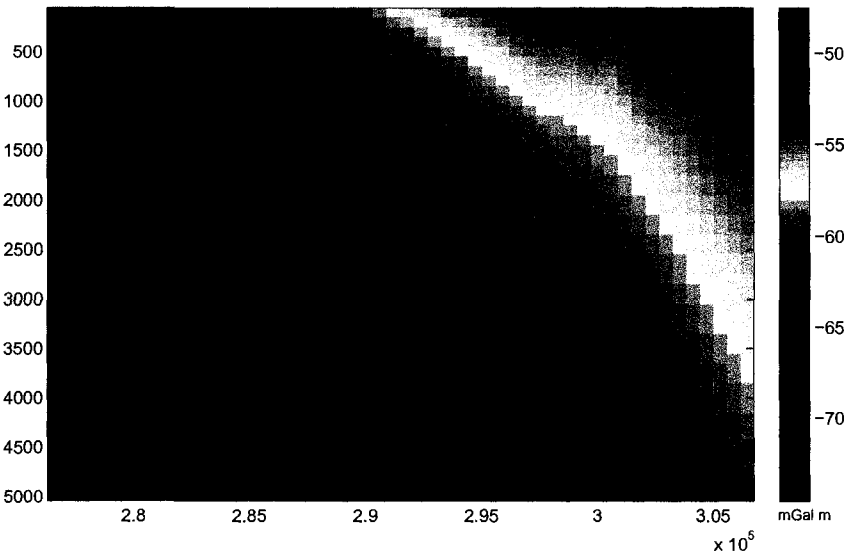

Figure 5.10 Determination of the depth to source involving a search for the extreme points of the scaled potential field; This process is done for one section at $y = 5513674$.

an optimum upward continuation height for regional-residual gravity separation. The correlation r between the regional anomaly $g_1(\Delta g_{reg})$ and each continuation g_2 can be calculated.

In reality, we don't know the real regional anomaly. So Zeng et al. (2007) suggested using correlation between the continuations of two successive heights. The observability of the maximum deflection of the curve from the chord depends mainly on a difference in depths to the anomalous bodies. The larger the difference, the more obvious the optimal height for the upward continuation to represent the regional field. However, when the method is used for sources at three or more depth levels, the plot of correlation versus the height may become complicated, and picking the point of maximum deflection from the chord may become ambiguous. We believe that this procedure is completely subjective.

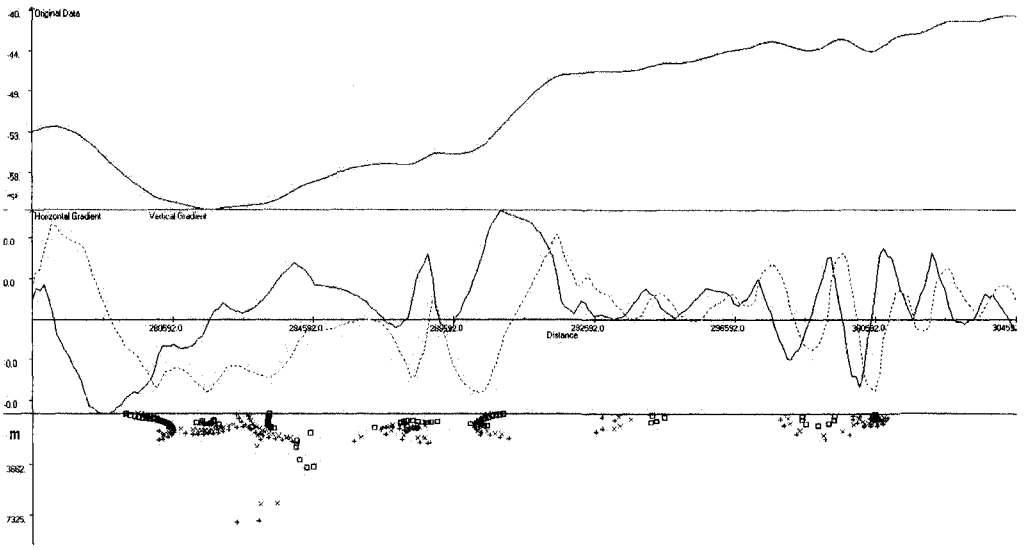

Figure 5.11 Euler Deconvolution

We perform this method using the following steps: First, we create a 3D data volume of the potential field. In practice, this means that data are needed both horizontally and vertically. As we mentioned before, data sets at different elevations are not available in our case. We can solve this problem using upward continuation. A 3D field was generated with a $100m$ step along the vertical directions extended to 5km. Second, we calculate cross-correlation between the continuations of every two successive heights. Third, we sketch the calculated correlations versus heights. The result is shown in Figure 5.12. In this figure, the optimum height, which is the position of the maximum deflection of the curve from the chord, has been shown.

Based on the results of these four methods, we choose the height of $10km$ for upward continuation. The regional anomaly is obtained by upward continuation at $10km$ of the Bouguer anomaly in order to invert data on a grid up to a depth of $5km$.

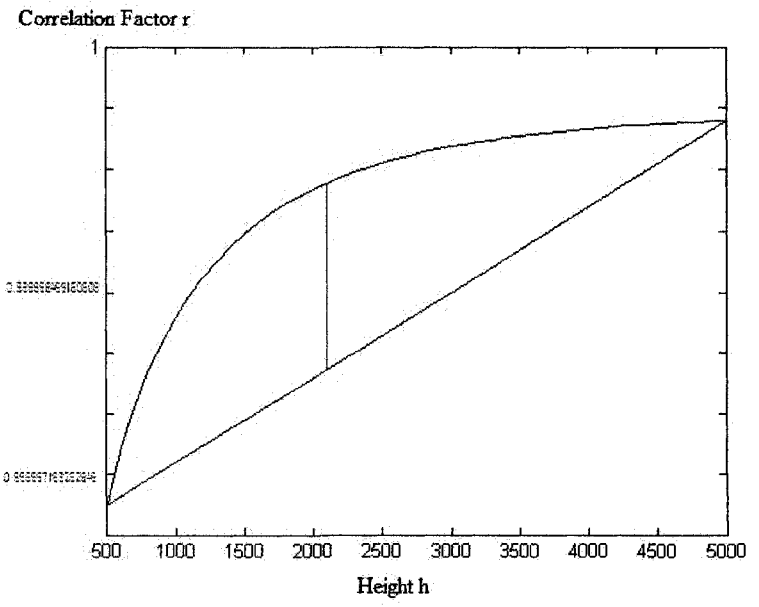

Figure 5.12 Correlations between continuations to two successive heights versus the height.

5.3 Inversion by Cokriging

Knowing the residual anomalies from the previous sections, we try to use inversion by cokriging that was discussed before in order to find the density model of our region. It should be mentioned that the residual anomaly has a mean equal to zero and therefore we can use the gravity observations to calculate the experimental gravity covariance matrix.

Geometry of Domain:
Before starting the inversion process, there are two things that should be taken care of:

- First, based on the explanations in chapter 2, we select the cell size such that it is smaller than $1.2 \times dx$ where dx is the distance between gravity

observations Boulanger and Chouteau (2001). Here, the selected cell size is equal $500m$.

- The second issue is the number of padding cells that should be added to our domain. We use the following formula for calculating the number of required pads (Boulanger and Chouteau, 2001):

$$num(pad) = \frac{d}{min(d_x, d_y)} \tag{5.6}$$

Where:

$$d \geq \frac{\sqrt{2}}{(3\sqrt{3}q)^{\frac{1}{2}}}(\frac{d_z}{2}) \tag{5.7}$$

Based on the above formula, with $q = 5\%$, we need to add two extra cells to pad our domain.

The inversion domain is divided into $n_x = 99$ by $n_y = 38$ by $n_z = 10$ cubes of dimension $500 \times 500 \times 500m$. So, the whole domain is $49.5 \times 19 \times 5km$ and the total number of prisms is $n = 37620$. The inversion will be performed using the following steps:

1. At first, we calculate the experimental gravity covariance matrix C_{gg-exp} using the gravity data.

2. Then, we use the V-V plot method that was discussed in chapter 3 with $Nlag = 50$ and we try to adjust the density model such that experimental and theoretical variograms become almost identical. The theoretical gravity covariance matrix C_{gg-th} can be calculated using the density covariance matrix $C_{\rho\rho}$ by the following formula: $C_{gg-th} = GC_{\rho\rho}G^T$. After applying this method, we conclude that the density variogram model should be anisotropic and spherical with $C0 = 105(kg/m^3)^2$, $C = 5500(kg/m^3)^2$ and $a_{45} = 6500m$,

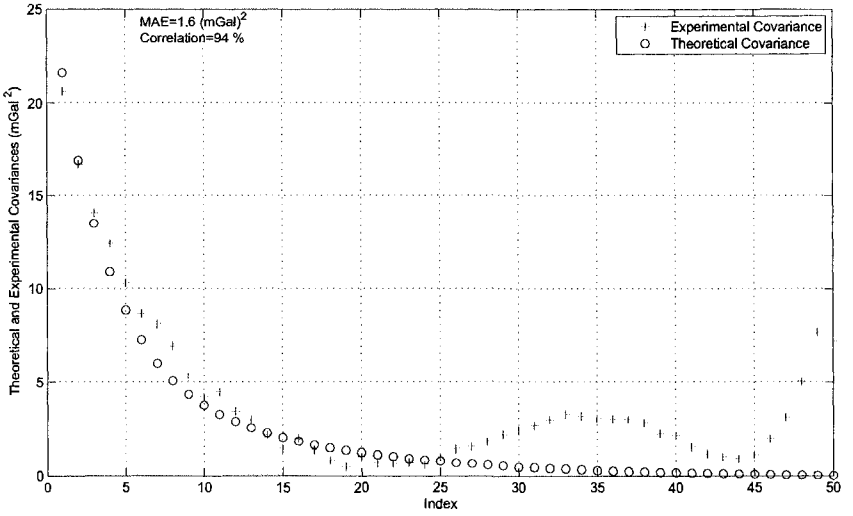

Figure 5.13 Fit of the experimental and theoretical gravity covariance matrices. Density variogram model: anisotropic and spherical with $C0 = 105(kg/m^3)^2$, $C = 5500(kg/m^3)^2$ and $a_{45} = 6500m$, $a_{135} = 7500m$ and $a_{vert} = 5000m$.

$a_{135} = 7500m$ and $a_{vert} = 5000m$. The covariance vectors v_{exp} and v_{th} are shown in Figure 5.13. It can be seen from this figure that the correlation between them is 94% and the Mean Absolute Error (MAE) is $1.6(mGal)^2$.

3. After selecting the density variogram model, we calculate the density covariance matrix and using that we calculate $C_{g\rho} = GC_{\rho\rho}$ and $C_{gg} = GC_{\rho\rho}G^T$.

   As we mentioned before, the number of prisms n is equal 37620 which is a very large number. The dimension of $C_{\rho\rho}$ is $n \times n$ and in MATLAB program, because of memory limitations, we can not store such a huge matrix. Even if we can store this matrix, the required storage and computational costs to compute $C_{g\rho}$ and C_{gg} are proportional to n^2 and become prohibitive for large n. We can solve this problem using the method proposed by Nowak et al. (2003) which is based on circulant embedding and the fast Fourier transform (FFT). The details are in chapter 3. In this method, instead of calculating $C_{\rho\rho}$

directly and saving it, we calculate its multiplication with G^T using circulant embedding and FFT. Having $C_{\rho\rho}G^T$, we can easily calculate C_{gg-th} and $C_{g\rho}$. The other problem is that, for large number of gravity data m, even using the above method, we still have the memory limitation issue with MATLAB. We solved this problem by dividing G which is an $m \times n$ matrix into smaller matrices G_i's with dimension $m1 \times n$ where $m1 < m$. Then we used the FFT method to calculate each $C_{\rho\rho}G_i^T$. Finally, we add all of the results together to have $C_{\rho\rho}G^T$. The other method to solve this problem is to use the less memory kernel described in chapter 2.

4. Knowing C_{gg-th} and $C_{g\rho}$, we can use the cokriging system equation $C_{gg}\lambda = C_{g\rho}$ to find coefficients λ and then we can find an estimate for the density distribution: $\rho = \lambda^T g$. These estimated densities are actually density contrasts. Therefore, a mean density of the crust of 2650 kg/m^3 is added as the mean to the contrasts estimated by cokriging. The 3D estimated density distribution is shown in Figure 5.14. Comparing this figure with the geological setting of the study area from Figure 5.2, we can see that the result of inversion by cokriging has a good agreement with the geology at the surface.

   We have also shown one section of the estimated densities at $y = 5515000m$ in Figure 5.15. This figure shows that using inversion by cokriging, the structure of the studied area roughly matches with the geological settings.

   In Figure 5.16, we have shown three horizontal layers of estimated densities at $z = 0$, $z = 2500$ and $z = 5000$ m. For comparison, we have also shown the geology of our domain. The approximate borders of the studied area has been shown with black lines. As it can be seen from this figure, the first layer shows an acceptable match with the geology, i.e., lower densities belong to granite and felsic rocks and higher densities belong to mafic rocks. Also, by increasing the depth, there is lack of resolution with inversion. By comparing

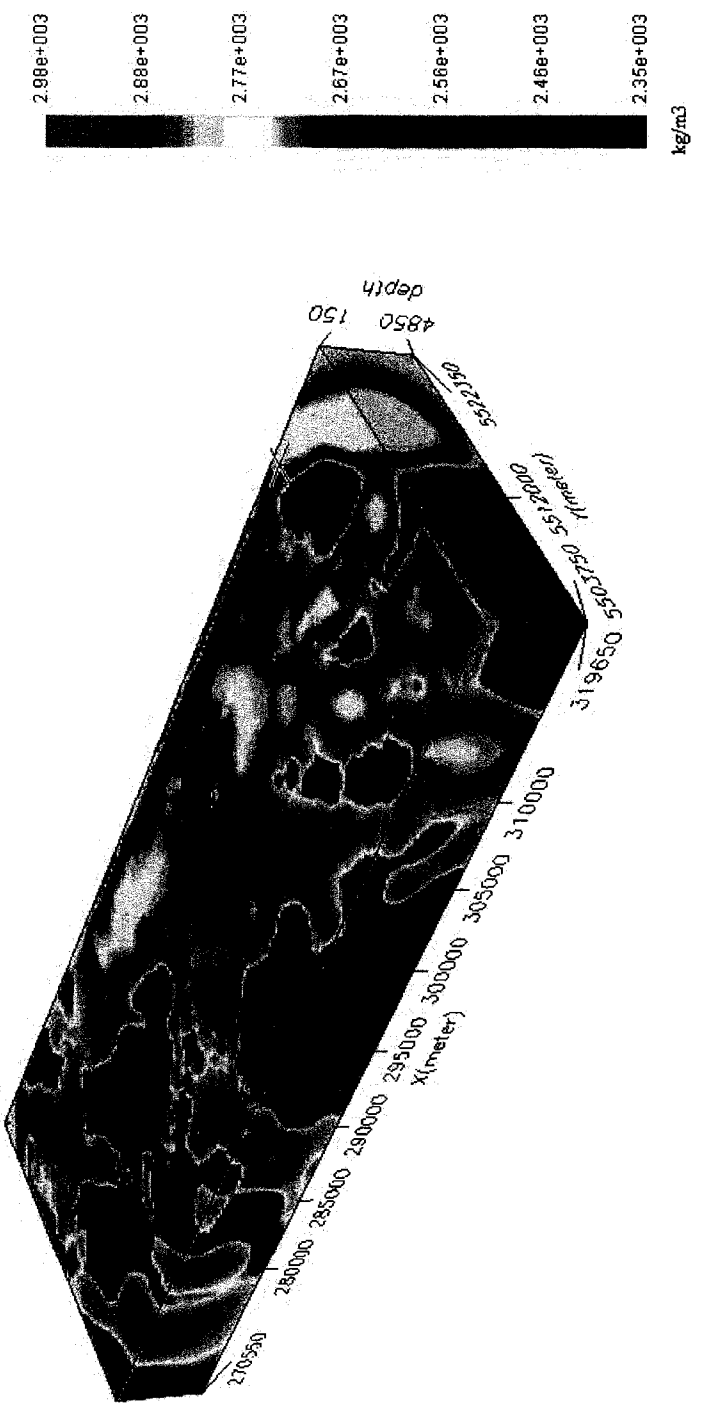

Figure 5.14 The 3D estimated density distribution using inversion by cokriging.

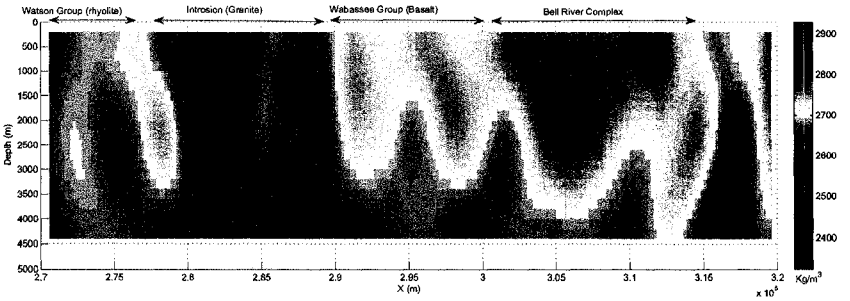

Figure 5.15 The estimated densities at section $y = 5515000m$ using inversion by cokriging.

the first horizontal layer of estimated densities, Figure 5.16; part (a), and residual gravity map in Figure 5.6, we can see that gravity and density in the first layer have good correlation with each other.

In order to investigate the inverted densities generated by inversion with cokriging, we separate them into two different ranges. First, we consider the densities between 2700 and 3000 kg/m^3. The volumes in the domain displaying this range of densities are shown in Figure 5.17. This figure can be a representative of the mafic rocks, especially the Wabassee group. Also, we consider the densities between 2300 and 2640 kg/m^3. The volumes with densities in this range are shown in Figure 5.18. This figure can be representative of the felsic rock.

5. After finding the density distribution, we can calculate the gravity data that it generates. The gravity anomaly is obtained by multiplying the geometric matrix (G) by the inverted density vector (ρ). In Figure 5.19, we compare this calculated gravity with our input (observed) gravity. It can be seen from this figure that these two gravity data are identical with MAE equal $5.77 \times 10^{-11} mGal$.

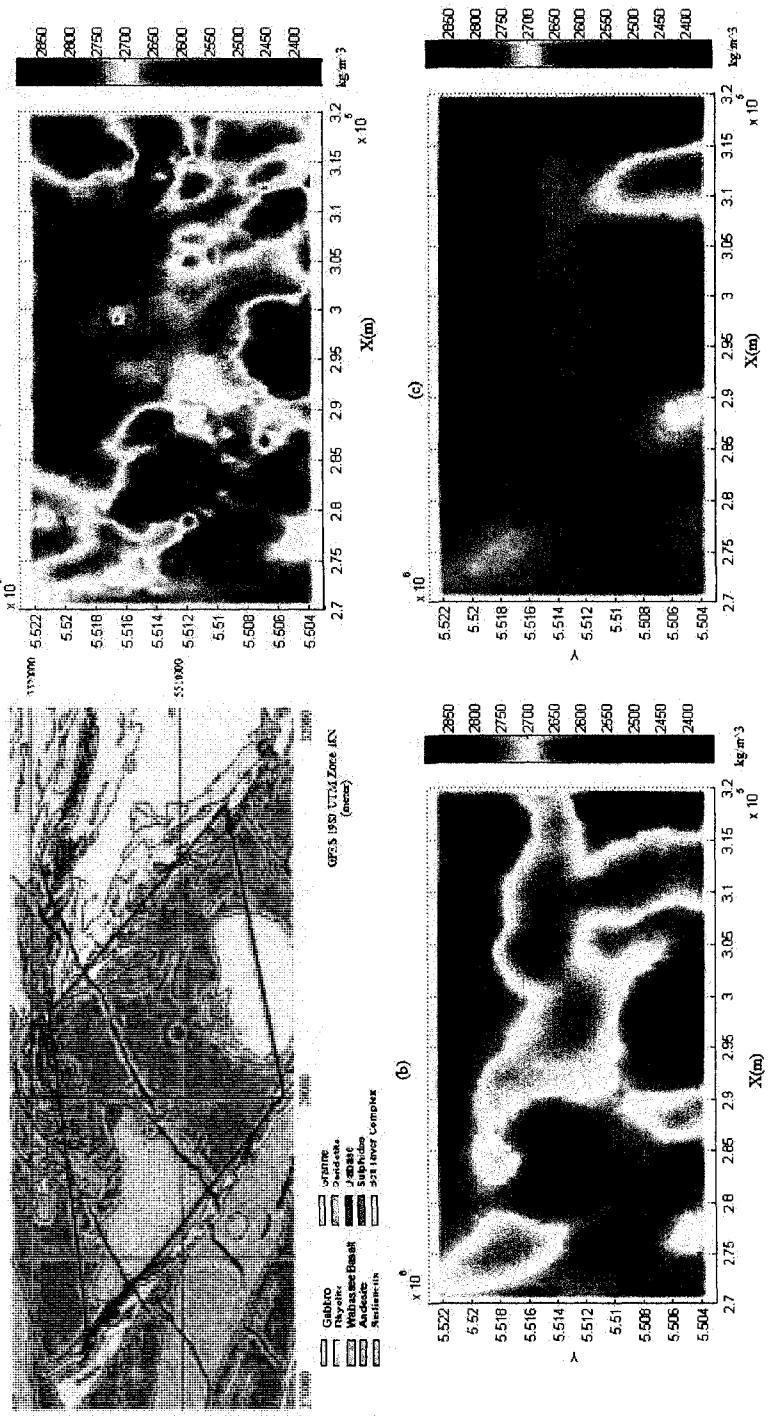

Figure 5.16 Comparing the geology with the estimated densities in horizontal layers at (a) $z = 0$, (b) $z = 2500$ and (c) $z = 5000$ m.

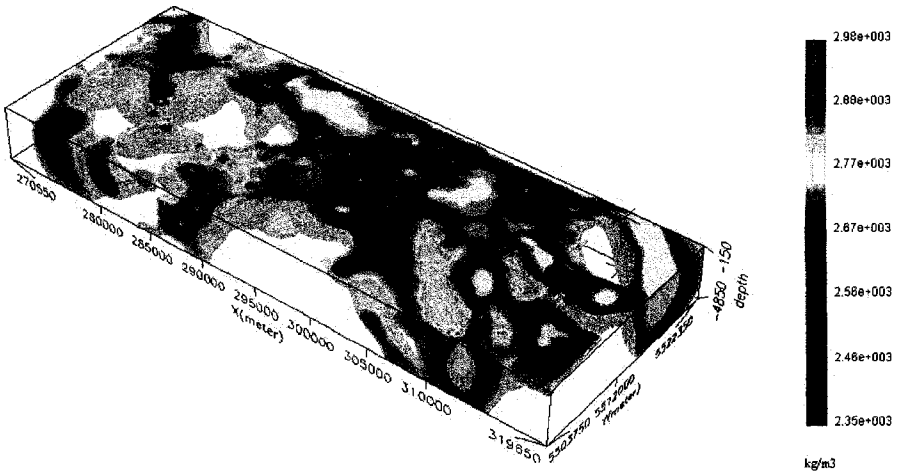

Figure 5.17 The 3D estimated densities in the range of 2700 and 3000 kg/m^3 generated by inversion with cokriging. These densities are related to mafic rocks.

5.4 Co-simulation results

As we mentioned in chapter 3, cokriging gives a smooth estimate of the density distribution. Using geostatistical simulation algorithms, we can have various reasonable solutions showing the kind of variability that can be expected from the density covariance model. We use the same density variogram model obtained during inversion by cokriging and using FFT-MA method, we provide various non-conditional density simulations (here, 100). We start post-conditioning with the cokriging of densities ρ^* obtained from observed gravities in previous parts. For each of the density simulations ρ_s, we calculate the surface gravity. Then we use the same weights λ used by cokriging method to calculate the cokriging of densities ρ_s^*. Finally, the conditional simulated densities can be written as follows:

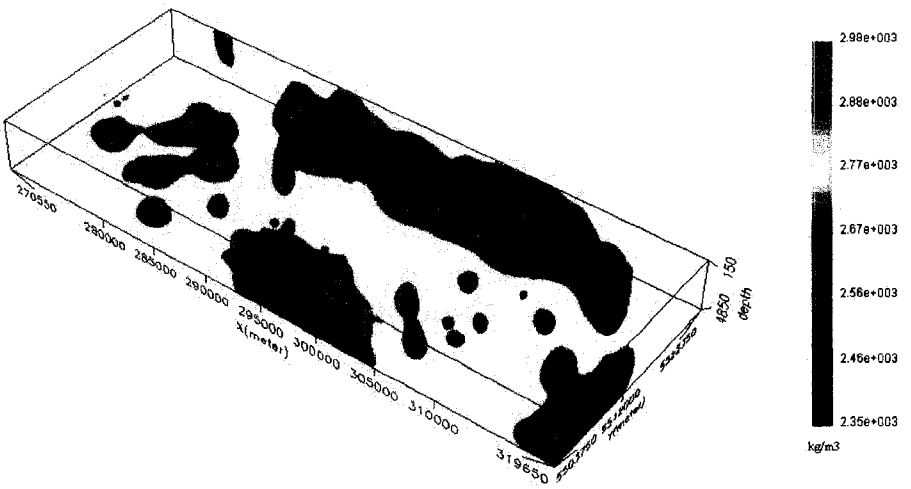

Figure 5.18 The 3D estimated densities between 2300 and 2640 kg/m^3 generated by inversion with cokriging. These densities are related to felsic rocks.

$$\rho_{sc} = \rho^* + (\rho_s - \rho_s^*) \tag{5.8}$$

At each realization, we calculate the surface gravity of ρ_{sc}. The 3D conditional simulated densities added to the mean density 2650 kg/m^3 is shown in Figure 5.20.

We also show the histogram of the inverted densities using cokriging and also the conditional simulated densities in Figure 5.21. As we saw in chapter 4 for synthetic densities, the histogram of the simulated densities had a better match with the synthetic densities, in term of the mean, variance and distribution compared with the inverted data using cokriging. Therefore, we can say that our real density distribution is closer to the simulated densities in terms of mean, variance and distribution provided we use a realistic covariance model.

As we mentioned in chapter 4, geostatistical simulation of density distributions

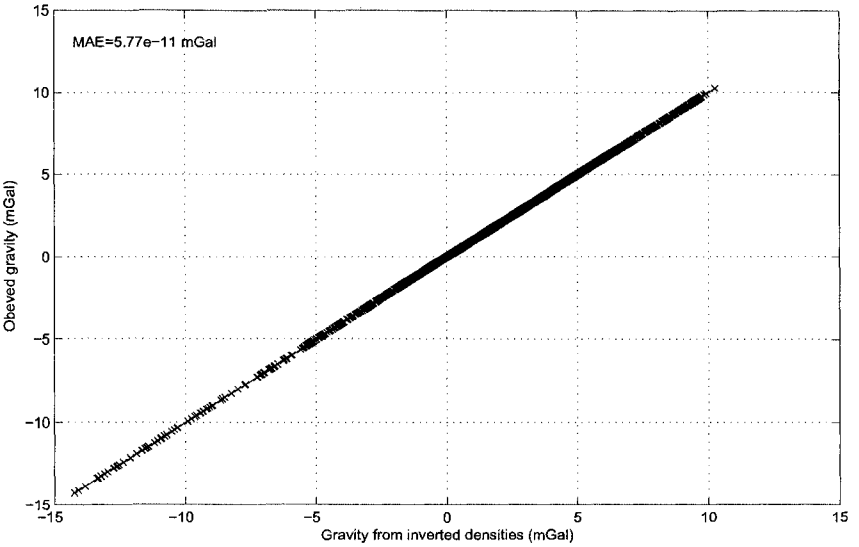

Figure 5.19 Comparing the input gravity with gravity calculated from inverted densities.

(ρ) allows computing their probability field. A probability that the density falls under a given range (threshold) can be obtained from the cumulative distribution of the simulated densities. Figure 5.22 shows co-simulated densities and inverted densities by cokriging at section $y = 5515000m$ along with the probability map of $2700 < \rho < 3000kg/m^3$ and the probability map of $2300 < \rho < 2650kg/m^3$ in this section.

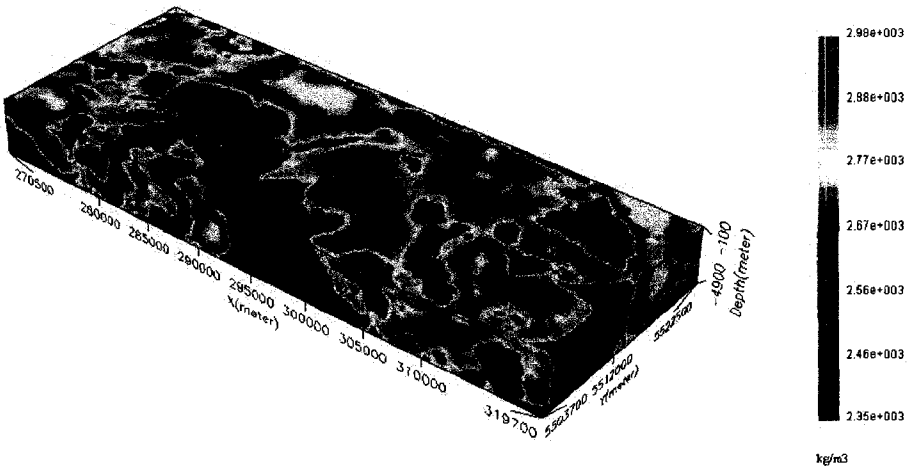

Figure 5.20 The 3D conditional simulated densities added with the mean 2650 kg/m^3.

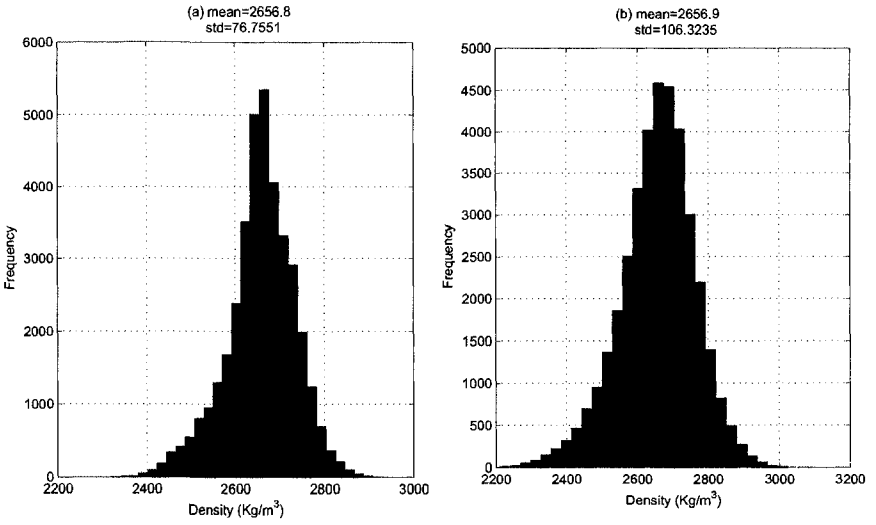

Figure 5.21 The histogram of (a) the inverted densities using cokriging and (b) the conditional simulated densities.

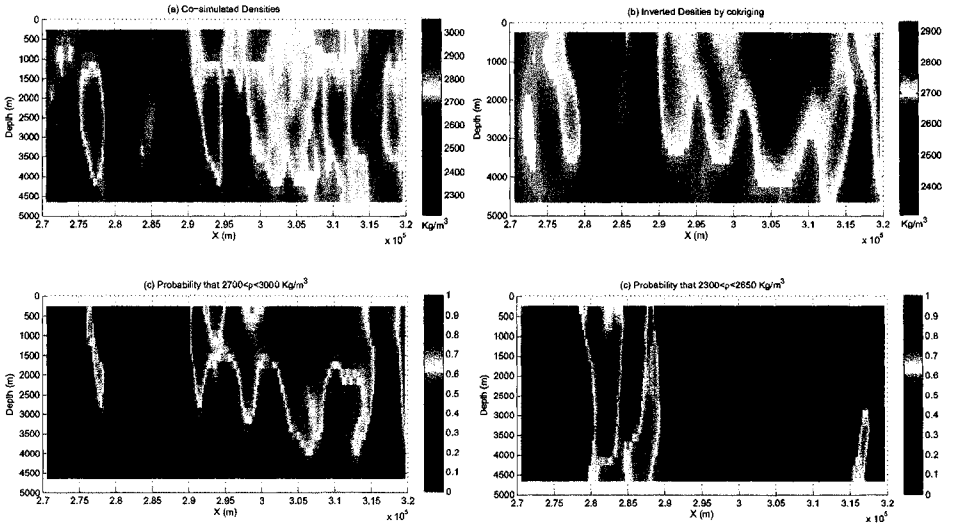

Figure 5.22 (a) Co-simulated Densities, (b) Inverted Densities by cokriging, (c) Probability map of $2700 < \rho < 3000 kg/m^3$ and (d)Probability map of $2300 < \rho < 2650 kg/m^3$. All the figures are from section $y = 5515000m$

5.5 Discussions

In this chapter, we applied the proposed method based on cokriging and co-simulation to a real case. For real Bouguer anomalies, the regional and residual fields were separated using upward continuation to a 10 km height. Then inversion was applied to the residual field. There were two important matters during inversion by cokriging:

- Obtaining a good fit between the theoretical and experimental covariances is sometimes challenging. Nevertheless, it is possible to obtain at least crude estimates of the sill and the ranges along principal geological directions. It is also possible to validate prior estimates deduced from geological knowledge.
- In the real case, the number of prisms n was very large. Therefore, using MATLAB software, the required storage and computational costs to compute

$C_{g\rho}$ and C_{gg} were prohibitive for large n. We can solve this problem using the method proposed by Nowak et al. (2003) which is based on circulant embedding and the fast Fourier transform (FFT). Another idea is to discretize the field initially with much larger blocks to find covariance at this scale. Often, it is relatively easy to deduce the approximate covariance model for smaller blocks.

We also simulated the densities using the geostatistical simulation method based on FFT-MA (Le Ravalec et al., 2000):

- We generated the histogram of the conditional simulated densities and we expect that our real density distribution is close to the simulated densities in terms of mean, variance and distribution.
- We also used the results of co-simulation to derive the probability map of $2700 < \rho < 3000 kg/m^3$ and the probability map of $2300 < \rho < 2650 kg/m^3$ for our domain.

From the results presented in this chapter, the estimated densities using inversion by cokriging and also co-simulation are in good agreement with the geology of the studied region.

CHAPTER 6

CONCLUSIONS

In this thesis, we presented an inversion method based on a geostatistical approach (cokriging) for three dimensional inversion of gravity data including geological constraints and after that co-simulation of data using covariance matrices. In the proposed method, we used the V-V plot tool (Asli et al., 2000) that estimates the covariance matrix of parameters by fitting experimental and theoretical gravity covariance matrices. Then, we explained how to add some constraints to our cokriging system to improve the performance of the inversion and consequently parameter estimation. We also described co-simulation based on the FFT moving average (FFT-MA) generator. We used this method as a geostatistical simulation algorithm to obtain various reasonable solutions in order to see the variability that can be expected from the density covariance model adopted.

We applied the proposed method on synthetic data which was simulated based on a geostatistical method and also dipping dyke. Then, we added some geological information as constraints to the system of cokriging. The use of known gradients and densities as constraints is new in this approach and can be useful in the mining industry if a 3D geological model can be built with a GIS.

Based on the results shown in chapter 4, we saw that the non-iterative proposed method was computationally very efficient. The algorithm is very flexible as it can be used for both regular and non-regular grids. Moreover, it is stable in the presence of data noise. For obtaining good fits between theoretical and experimental covariance matrices, it was necessary to consider carefully the geological information available from the domain under study. We used the method with

less memory and computation requirements for efficient calculation of the gravity-density covariance matrix and the gravity-gravity covariance matrix which is based on circulant embedding and the fast Fourier transform (FFT). It was observed that this method is very useful and has acceptable speed especially for large amount of data and large grids.

We saw that the orientation of a dyke can be recognized with stochastic methods. Also the proposed method could provide good resolution at depth knowing some gradient constraints. These two advantages are not present in the least squares inversion method which can not provide accurate estimate for the orientation of the body and also can not show good resolution in depth.

We found out that the histogram of the simulated densities shows a better match with the synthetic densities, in term of the mean, variance and distribution compared with the inverted data using cokriging. We realized that, using geostatistical simulation, we can provide probability maps with different thresholds that can be used as a tool in the interpretation of the inversion results.

We also saw the effect of increasing the amount of information as constraints. It improved the estimation of density distribution especially at deeper depths. For the dyke model, we used some known gradients as the constraints and we saw that they had a very positive effect on inversion by cokriging and reconstruction of the density distribution contrast. For the synthetic model, we used some borehole densities as the constraints and we saw their effectiveness in estimation of the density distribution.

Finally, we applied the proposed method based on cokriging and co-simulation on real data. As a case study, we considered one part of the survey data (Bouguer anomaly) from the Matagami area in Quebec. Based on the gravity measurements

available from this region, the regional and residual fields were separated using upward continuation. Then inversion was applied to the residual field in order to estimate the density distribution in this region. Based on the results of inversion, we tried to interpret the densities and have an estimation of the lithologies that exist in this region.

We found out that obtaining a good fit between the theoretical and experimental covariances is sometimes challenging. Nevertheless, it is possible to obtain at least crude estimates of the sill and the ranges along principal geological directions. It is also possible to validate prior estimates deduced from geological knowledge.

In the real case, the number of prisms n was very large. Therefore, using MATLAB software, the required storage and computational costs to compute $C_{g\rho}$ and C_{gg} were prohibitive for large n. We recommended using the method based on circulant embedding and the fast Fourier transform (FFT) for efficient calculation of the gravity-density covariance matrix and the gravity-gravity covariance matrix.

We also estimated the densities using the geostatistical simulation method based on FFT-MA (Le Ravalec et al., 2000). We generated the histogram of the conditional simulated densities and we expect that our real density distribution is close to simulated densities in terms of mean, variance and distribution. We also used the results of co-simulation to derive the probability maps with different thresholds for the densities in our domain.

From the results presented for the real case, the estimated densities using inversion by cokriging and also co-simulation were in good agreement with the geology of the studied region.

REFERENCES

Asli, M., Marcotte, D., and Chouteau, M. (2000). Direct inversion of gravity data by cokriging. In Kleingeld and Krige, editors, *6th International Geostatistics Congress*, Cape Town, South Africa, pages 64–73.

Blakely, R. (1995). *Potential Theory in Gravity and Magnetic Applications*. Cambridge University Press.

Boulanger, O. and Chouteau, M. (2001). Constraints in 3D gravity inversion. *Geophysical Prospecting*, **49**(2), 265 – 280.

Calvert, A. and Li, Y. (1999). Seismic reflection imaging over a massive sulfide deposit at the Matagami mining camp, Québec. *Geophysics*, **64**, 24–32.

Chasseriau, P. and Chouteau, M. (2003). 3D gravity inversion using a model of parameter covariance. *Journal of Applied Geophysics*, **52**(1), 59 – 74.

Chilès, J. and Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. Wiley.

Durrheim, R. and Cooper, G. (1998). EULDEP: a program for the Euler deconvolution of magnetic and gravity data. *Computers and Geosciences*, **24**(6), 545–550.

Fedi, M. (2007). DEXP: A fast method to determine the depth and the structural index of potential fields sources. *Geophysics*, **72**, I1–I11.

Fedi, M. and Rapolla, A. (1999). 3-D inversion of gravity and magnetic data with depth resolution. *Geophysics*, **64**, 452–460.

Franklin, J. (1970). Well posed stochastic extensions of Ill posed linear problems. *J. Math*, **31**, 682–716.

Giroux, B., Gloaguen, E., and Chouteau, M. (2007). bh_tomo–a Matlab borehole georadar 2D tomography package. *Computers and Geosciences*, **33**(1), 126–137.

Gloaguen, E., Marcotte, D., Chouteau, M., and Perroud, H. (2005). Borehole radar velocity inversion using cokriging and cosimulation. *Journal of Applied Geophysics*, **57**(4), 242 – 259.

Gloaguen, E., Marcotte, D., Giroux, B., Dubreuil-Boisclair, C., Chouteau, M., and Aubertin, M. (2007). Stochastic borehole radar velocity and attenuation tomographies using cokriging and cosimulation. *Journal of Applied Geophysics*, **62**(2), 141 – 157.

Green, W. (1975). Inversion of gravity profiles by use of a Backus-Gilbert approach. *Geophysics*, **40**, 763–772.

Gupta, V. and Ramani, N. (1980). Some aspects of regional-residual separation of gravity anomalies in a Precambrian terrain. *Geophysics*, **45**, 1412–1426.

Haáz, I. (1953). Relations between the potential of the attraction of the mass contained in a finite rectangular prism and its first and second derivatives. *Geophysical Transactions II*, **7**, 57–66.

Jacobsen, B. (1987). A case for upward continuation as a standard separation filter for potential-field maps. *Geophysics*, **52**, 1138–1148.

Keating, P. (1998). Weighted Euler deconvolution of gravity data. *Geophysics*, **63**, 1595–1603.

Last, B. and Kubik, K. (1983). Compact gravity inversion. *Geophysics*, **48**, 713–721.

Le Ravalec, M., Noetinger, B., and Hu, L. (2000). The FFT Moving Average (FFT-MA) Generator: An Efficient Numerical Method for Generating and Conditioning Gaussian Simulations. *Mathematical Geology*, **32**(6), 701–723.

Li, X. and Chouteau, M. (1998). Three-dimensional gravity modeling In all Space. *Surveys in Geophysics*, **19**(4), 339–368.

Li, Y. and Oldenburg, D. (1998). 3-D inversion of gravity data. *Geophysics*, **63**(1), 109–119.

Marcotte, D. (1991). Cokriging with Matlab. *Computers & Geosciences*, **17**(9), 1265–1280.

Marcotte, D. and Chouteau, M. (1993). Gravity data transformation by kriging. In Soares, A., editor, *Geostatistics Tróia '92*, pages 249–269. Kluwer Academic Publishers Dordrecht.

Matheron, G. (1965). Les variables régionalisées et leur estimation. *Une Application de la Théorie des Functions Aléatories aux Sciences de la Nature. Masson, Paris.*

Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press.

Montagner, J. and Jobert, N. (1988). Vectorial tomography-II. Application to the Indian Ocean. *Geophysical Journal International*, **94**(2), 309–344.

Nelder, J. and Mead, R. (1965). A simplex method for function minimisation. *Computer J.*, **7**, 308–313.

Nowak, W., Tenkleve, S., and Cirpka, O. (2003). Efficient computation of linearized cross-covariance and auto-covariance matrices of interdependent quantities. *Mathematical Geology*, **35**(1), 53–66.

Piché, M., Guha, J., and Daigneault, R. (1993). Stratigraphic and structural aspects of the volcanic rocks of the Matagami mining camp, Quebec; implications for the Norita ore deposit. *Economic Geology*, **88**(6), 1542–1558.

Pilkington, M. (1997). 3-D magnetic imaging using conjugate gradients. *Geophysics*, **62**, 1132–1142.

Spector, A. and Grant, F. (1970). Statistical models for interpreting aeromagmetic data . *Geophysics*, **35**, 293.

Tarantola, A. and Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys. Space Phys*, **20**(2), 219–232.

Thompson, D. (1982). EULDPH: A new technique for making computer-assisted depth estimates from magnetic data. *Geophysics*, **47**, 31–37.

Zeng, H., Xu, D., and Tan, H. (2007). A model study for estimating optimum upward-continuation height for gravity separation with application to a Bouguer gravity anomaly over a mineral deposit, Jilin province, northeast China. *Geophysics*, **72**, I45–150.