QUANTITATIVE INTERPRETATION OF AIRBORNE
GRAVITY GRADIOMETRY DATA FOR
MINERAL EXPLORATION

by

Cericia D. Martinez
© Copyright by Cericia D. Martinez, 2015

All Rights Reserved
A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Geophysics).

Golden, Colorado
Date ______________________

Signed: ______________________
Cericia D. Martinez

Signed: ______________________
Dr. Yaoguo Li
Thesis Advisor

Golden, Colorado
Date ______________________

Signed: ______________________
Dr. Terry Young
Professor and Head
Department of Geophysics
ABSTRACT

In the past two decades, commercialization of previously classified instrumentation has provided the ability to rapidly collect quality gravity gradient measurements for resource exploration. In the near future, next-generation instrumentation are expected to further advance acquisition of higher-quality data not subject to pre-processing regulations. Conversely, the ability to process and interpret gravity gradiometry data has not kept pace with innovations occurring in data acquisition systems. The purpose of the research presented in this thesis is to contribute to the understanding, development, and application of processing and interpretation techniques available for airborne gravity gradiometry in resource exploration. In particular, this research focuses on the utility of 3D inversion of gravity gradiometry for interpretation purposes. Towards this goal, I investigate the requisite components for an integrated interpretation workflow. In addition to practical 3D inversions, components of the workflow include estimation of density for terrain correction, processing of multi-component data using equivalent source for denoising, quantification of noise level, and component conversion. The objective is to produce high quality density distributions for subsequent geological interpretation. I then investigate the use of the inverted density model in orebody imaging, lithology differentiation, and resource evaluation. The systematic and sequential approach highlighted in the thesis addresses some of the challenges facing the use of gravity gradiometry as an exploration tool, while elucidating a procedure for incorporating gravity gradient interpretations into the lifecycle of not only resource exploration, but also resource modeling.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xx</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>xxi</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Gravity gradient tensor</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Summary of 3D inversion</td>
<td>6</td>
</tr>
<tr>
<td>CHAPTER 2 TERRAIN AND BATHYMETRY DENSITY ESTIMATION IN GRAVITY GRADIOMETRY USING SPATIAL STATISTICS</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Terrain Effect in Gravity Gradiometry</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Density Estimation</td>
<td>13</td>
</tr>
<tr>
<td>2.3.1 Field Data Illustration</td>
<td>14</td>
</tr>
<tr>
<td>2.3.2 Moran’s $I$ and Geary’s $c$</td>
<td>14</td>
</tr>
<tr>
<td>2.3.3 Local $I$ and $c$ Statistics</td>
<td>17</td>
</tr>
<tr>
<td>2.3.4 Global $I$ and $c$ Statistics</td>
<td>20</td>
</tr>
<tr>
<td>2.4 Validation of Methodology</td>
<td>22</td>
</tr>
<tr>
<td>2.5 Test Case with Dipping Dyke and Coincident Topography</td>
<td>26</td>
</tr>
<tr>
<td>2.5.1 Component Density Estimates</td>
<td>30</td>
</tr>
<tr>
<td>2.5.2 Synthesis</td>
<td>39</td>
</tr>
</tbody>
</table>
4.3.2 \( T_{xy} \) and \( T_{uv} \) inversion ........................................ 89

4.4 Vredefort Conceptual Interpretation ........................................ 89

4.5 Vredefort Summary ....................................................... 93

4.6 Kauring Background ..................................................... 95

4.7 Kauring Observed and Calculated Datasets ............................... 96
  4.7.1 Observed Ground Gravity Data .................................... 97
  4.7.2 Observed Gravity Gradient Data .................................. 97
  4.7.3 Equivalent Source Conversion .................................... 97

4.8 Kauring Inversion of Observed Data .................................... 101

4.9 Kauring Inversion of Converted Data .................................. 103
  4.9.1 \( T_{xy} \) and \( T_{uv} \) data ............................................. 103
  4.9.2 Five component tensor data ....................................... 104

4.10 Kauring Summary ....................................................... 105

4.11 Acknowledgments ....................................................... 108

CHAPTER 5 3D INVERSION OF AIRBORNE GRAVITY GRADIOMETRY
DATA IN MINERAL EXPLORATION: A CASE STUDY IN THE
QUADRILÁTERO FERRÍFERO, BRAZIL ........................................ 109

5.1 Introduction .............................................................. 109

5.2 Deposit geology and gravity gradiometry survey ....................... 112
  5.2.1 Geology ........................................................... 112
  5.2.2 Gravity gradiometry survey ..................................... 115

5.3 Inversion results ........................................................ 117
  5.3.1 Single Component Inversion (\( T_{zz} \)) .......................... 120
  5.3.2 Three-component Inversion (\( T_{xz}, T_{yz}, \) and \( T_{zz} \)) .... 123
5.3.3 Five-component Inversion \((T_{xx}, T_{xy}, T_{xz}, T_{yy}, \text{and } T_{yz})\) ........................ 124

5.3.4 Six-component Inversion \((T_{xx}, T_{xy}, T_{xz}, T_{yy}, T_{yz}, T_{zz})\) ........................ 126

5.3.5 Summary ......................................................... 127

5.4 Conclusions ...................................................... 129

5.5 Acknowledgments .................................................. 130

CHAPTER 6 LITHOLOGIC CHARACTERIZATION USING AIRBORNE GRAVITY GRADIENT AND AEROMAGNETIC DATA FOR MINERAL EXPLORATION: A CASE STUDY IN THE QUADRILÁTERO FERRÍFERO, BRAZIL .......................... 131

6.1 Introduction ...................................................... 131

6.2 Deposit Geology and Geophysical Data .............................. 135

6.2.1 Gravity gradient data ........................................... 136

6.2.2 Aeromagnetic data ............................................. 138

6.2.2.1 Regional-residual separation ............................... 138

6.3 Lithologic Mapping ............................................... 139

6.3.1 Density contrast and susceptibility models ...................... 142

6.3.2 Lithology assignment using generic constraints ................ 144

6.3.3 Lithology assignments using geologic constraints ............... 152

6.3.4 Discussion ...................................................... 157

6.4 Conclusions ...................................................... 160

6.5 Acknowledgments .................................................. 162

CHAPTER 7 QUANTITATIVE INTEGRATION OF GEOPHYSICAL MODELS IN MINERAL RESOURCE MODELING .............................. 163

7.1 Introduction ...................................................... 163

7.2 Krigeing and Cokrigeing ......................................... 165
| Figure 2.1 | Nettleton’s method of correlating gravity anomaly corrected with various densities with the topography. From Milsom and Eriksen. | 12 |
| Figure 2.2 | Observed gravity gradient data. | 15 |
| Figure 2.3 | Calculated terrain effect displayed using a density of 1.0 g/cc. | 15 |
| Figure 2.4 | Topography over the entire survey area with black box outlining the smaller area of interest. | 16 |
| Figure 2.5 | Local Moran’s $I$ for $T_{zz}$ using various density values to remove the terrain effect. Color displays the value of $I$ at each location. | 18 |
| Figure 2.6 | Blue represents values of 0, indicating a lack of or no spatial autocorrelation. Red represents values of 1, indicating spatial autocorrelation. | 20 |
| Figure 2.7 | Global proportioned Moran’s $I$ statistic of all six components plotted against their corresponding density value on the abscissa. | 22 |
| Figure 2.8 | Global proportioned Geary’s $c$ statistic of all six components plotted against their corresponding density value on the abscissa. | 23 |
| Figure 2.9 | Anomalous gravity gradient data with terrain effect removed with a density of 2.84 g/cc. | 24 |
| Figure 2.10 | (a) Top of the dipping dyke is centered beneath the elongated hill at a depth from 53m down to 20m. (b) Top of dipping dyke is centered within the elongated hill at a depth from 61m down to 25m. (c) Dipping dyke is offset from center of the elongated hill by 30m to the east at a depth from 53m down to 20m. | 27 |
| Figure 2.11 | Observation height with locations denoted in white for (a) North trending lines with 370 observation locations (b) East trending lines with 370 observation locations and (c) North-east trending lines with 781 observation locations. | 29 |
Figure 2.12  Histograms of density estimates for the average of trace component estimates only for four groupings: (upper left) density estimates for all permutation; (upper right) density estimates for permutations on dyke beneath the hill; (lower left) density estimates for permutations on dyke within the hill; (lower right) density estimates for permutations on dyke offset from the hill. The red line indicates the true background density value.

Figure 2.13  Histogram of density estimates for each component. The red line indicates the true background density value.

Figure 2.14  Density estimates for each individual component plotted versus average density for all six components. The colors indicate line direction where black dots are estimates from east-west lines, red dots are estimates from north-south lines, and green dots are north-west trending flight lines.

Figure 2.15  Density estimates for each individual component plotted versus average density for all six components. The colors indicate the location of the dyke with respect to the north trending hill. Blue dots are estimates where the dyke is directly beneath the hill, cyan dots are estimates where the dyke is within the hill, and pink dots are estimates where the dyke is offset to the east of the hill.

Figure 2.16  Density estimates for each individual component plotted versus average density for all six components. The colors indicate the spatial statistic used. Blue indicates Moran's $I$ and red indicates Geary's $c$.

Figure 2.17  Density estimates for each individual component plotted versus average density for all six components. The colors indicate type of interpretation of the spatial statistic. Along the estimated density curves: green indicates minimum of the Moran's $I$ curve, pink indicates maximum of the Geary's $c$ curve, black indicates the minimum of the derivative of $I$ curve, gray indicates the maximum of the derivative of $I$ curve, yellow indicates the minimum of the derivative of $c$ curve, cyan indicates the maximum of the $c$ curve.

Figure 2.18  Density estimates for each individual component plotted versus average density for all six components. The colors indicate the distance interval used to generate the weighting matrix for the calculation of Moran's $I$ and Geary's $c$. 0m-40m are gray, 39m-71m are yellow, 70m-101m are black, 100m-140m are blue, 139m-181m are cyan.
Figure 2.19 Density estimates for each individual component plotted versus average density for all six components. All density estimates are plotted with gray crosses. The colored dots indicate the spatial statistic where blue indicates Moran’s $I$ and red indicates Geary’s c.

Figure 2.20 Density estimates for each individual component plotted versus average density for all six components. Only estimates with a mean across all components less than 2.75 g/cc are shown. The colored dots indicate the spatial statistic where estimates from the dyke beneath the hill are blue, within the hill are cyan, and offset from the hill are pink.

Figure 2.21 Histogram of density estimates for each component for scenarios where the dyke either beneath or within the hill and the average density across all components is less than 2.75 g/cc. The red line indicates the true background density value.

Figure 2.22 Histograms of density estimates for the average of trace component estimates only for four groupings: (upper left) density estimates for all permutation; (upper right) density estimates for permutations on dyke beneath the hill; (lower left) density estimates for permutations on dyke within the hill; (lower right) density estimates for permutations on dyke offset from the hill. The red line indicates the true background density value.

Figure 3.1 Gravity gradient data for the synthetic problem contaminated by gaussian noise with zero mean and 0.85 Eo standard deviation. The panels from left to right and top to bottom are $T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, $T_{yz}$, Observation height, and $T_{zz}$.

Figure 3.2 (a) GCV curve as a function of the regularization parameter. The minimum GCV point defines the optimal regularization value. (b) The Tikhonov curve of the equivalent source construction for the synthetic data. The filled circle denotes the position of the GCV optimal solution on the Tikhonov curve and GCV curve.

Figure 3.3 Denoised data produced by the optimal equivalent source layer according to GCV criterion. Note that all high frequency features have been removed.

Figure 3.4 The difference between observed and denoised data from equivalent source using GCV criterion.

Figure 3.5 Denoised data using the L-curve criterion. The result is consistent with that produced by the GCV criterion.
Figure 3.6  (a) The Tikhonov curve of the equivalent source construction for the synthetic data. Note the L shape of the curve and its distinct corner. (b) The curvature of the Tikhonov curve as a function of the regularization parameter. The maximum curvature point defines the corner point of the Tikhonov curve. .................................................. 58

Figure 3.7  Five-components of observed data from Leghorn survey area along with calculated $T_{zz}$. ................................................................. 59

Figure 3.8  The contour lines show the bathymetry of the area surrounding the data area. The color overlay is the observed $T_{zz}$ component prior to bathymetry correction. .............................................................. 60

Figure 3.9  The bathymetry correction calculated assuming a density contrast of 1.0 g/cm$^3$ between the seafloor and water column. ................................. 61

Figure 3.10  The bathymetry corrected data. Note that the large-scale variations have largely disappeared from the data maps and they are dominated by small-scale variations. No clear signal is visible. ............................... 62

Figure 3.11  Denoised data from equivalent source processing using L-curve criterion. . . 63

Figure 3.12  The difference between the bathymetry-corrected data and the denoised data obtained from equivalent source processing using an L-curve criterion. The difference map is representative of data errors and noise within the observed data. .............................................................. 64

Figure 3.13  The top panel shows the equivalent source recovered from the bathymetry-corrected data. The middle panel is the regional source estimated using a second-degree polynomial. The difference between the two defines the residual source from which the final residual anomaly will be calculated. .............................................................. 66

Figure 3.14  The residual gradiometer data extracted by removing a regional field from the equivalent source layer. .............................................................. 68

Figure 3.15  Observed gravity gradient data from part of the McFaulds Lake Area dataset in the Ring of Fire. From left to right, the panels display Observation Height, $T_{xy}$ and $T_{uv}$. .............................................................. 69

Figure 3.16  The constructed equivalent source layer from observed $T_{xy}$ and $T_{uv}$. . . 69

Figure 3.17  $T_{zz}$ calculated from the equivalent layer generated from the observed $T_{xy}$ and $T_{uv}$ components for the McFaulds Lake data. .............................. 70
Figure 3.18 Accurate gravity gradient data for the synthetic problem. The panels from left to right and top to bottom are $T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, $T_{yz}$, and $T_{zz}$. 73

Figure 3.19 Equivalent source layer constructed from the six noisy gravity gradient components of Figure 3.1 according to an L-curve criterion. 74

Figure 3.20 Histogram of noise estimates ($\hat{\sigma}_k^d$) for the $k$ simulated datasets. The average value for each component is summarized in Table Table 3.1. 75

Figure 3.21 Denoised gravity gradient data from part of the McFaulds Lake Area dataset. From left to right, the panels display $T_{xy}$ and $T_{uv}$. 77

Figure 4.1 Image of Vredefort Impact Structure with $T_{xy}$ overlaid. Black outline indicates data area that was inverted. Image generated using Google Earth. 81

Figure 4.2 Geologic map of the Vredefort impact structure showing main geologic units; from Henkel and Reimold. 82

Figure 4.3 Cartoon illustrating impact structure. There is a structural high in the central region of the crater with concentric fault features radiating outward. 83

Figure 4.4 Cartoon illustrating impact structure. The central portion illustrates uplift of the mantle with concentric fault structures radiating outward. 84

Figure 4.5 Digital elevation model within survey area. 85

Figure 4.6 Gravity gradient components: a) $T_{xy}$, b) $T_{uv} = (T_{xx} - T_{yy})/2$. 86

Figure 4.7 Equivalent source layer obtained from $T_{xy}$ and $T_{uv} = (T_{xx} - T_{yy})/2$. 86

Figure 4.8 a) L-curve showing data misfit plotted against model complexity b) Curvature of the L-curve plotted as a function of the regularization parameter. 87

Figure 4.9 $T_{zz}$ calculated using an equivalent source layer generated from the $T_{xy}$ and $T_{uv} = (T_{xx} - T_{yy})/2$ components. 87

Figure 4.10 (a) Side and (b) top volume rendered image of 3D density contrast model constructed from $T_{zz}$ with cells between 0.15 and -0.1 g/cc removed. Values are in g/cc. 90
Figure 4.11  (a) Side and (b) top volume rendered image of 3D density contrast model constructed from the two observed components with cells between 0.15 and -0.1 g/cc removed. Values are in g/cc.

Figure 4.12  Cross section through Vredefort impact structure constructed from gravity and magnetic data, modified from Henkel and Reimold.

Figure 4.13  Slice through 3D recovered model from inversion of observed components and correspondence to geologic section, modified from Henkel and Reimold.

Figure 4.14  Sketch of structural uplift interpreted by the central density features.

Figure 4.15  Sketch of faulting or deformation features in the upper crust.

Figure 4.16  Regional gravity around the Kauring Test Site.

Figure 4.17  Ground gravity in the airborne gravity gradient area of the Kauring Test Site.

Figure 4.18  Airborne gravity gradient data acquired over the Kauring test site by the Falcon system. Left: $T_{xy}$; Right: $T_{uv}$.

Figure 4.19  Equivalent source layer generated using ground gravity data.

Figure 4.20  The five independent gravity gradient components along with $T_{zz}$ calculated using the equivalent source layer from ground gravity data. The tensor components have been filtered with a 300m, 4th-order Butterworth filter.

Figure 4.21  Calculated Falcon components, $T_{xy}$ (left) and $T_{uv}$ (right), from the equivalent source layer in Figure 4.19.

Figure 4.22  Volume rendered image of the 3D density contrast model (g/cc) constructed from ground gravity data with values less than 1 g/cc removed.

Figure 4.23  Volume rendered image of the 3D density contrast model (g/cc) constructed from observed Falcon components with values less than 1 g/cc removed.

Figure 4.24  Volume rendered image of the 3D density contrast model (g/cc) constructed from converted Falcon components with values less than 1 g/cc removed.
Figure 4.25  Volume rendered image of the 3D density contrast model (g/cc) constructed from converted five-component data with values less than 1 g/cc removed. .................................................. 104

Figure 4.26  Plan sections through volumes at a depth of 80m. Inverted models from (A) ground gravity, (B) observed Falcon, (C) converted five component data, (D) converted Falcon. Color is from -1 g/cc (blue) to 2.4 g/cc (red). .................................................. 106

Figure 4.27  Plan sections through volumes at a depth of 180m. Inverted models from (A) ground gravity, (B) observed Falcon, (C) converted five component data, (D) converted Falcon. Color is from -1 g/cc (blue) to 2.4 g/cc (red). .................................................. 107

Figure 5.1  Geologic map of the Quadrilátero Ferrífero, with solid black lines indicating where the Minas Series is present; modified from . The survey area is outlined in red. .................................................. 113

Figure 5.2  Terrain corrected gravity gradient data covering a 4 km by 5 km area with survey topography (f) showing the location of the data subset. From left to right and top to bottom the plot panels display (a) $T_{xx}$; (b) $T_{xy}$; (c) $T_{xz}$; (d) $T_{yy}$; (e) $T_{yz}$; (f) topography; (g) $T_{zz}$. The location of the cross section of Figures Figure 5.5 and Figure 5.6 is overlain on $T_{zz}$ in panel (g). .................................................. 115

Figure 5.3  The importance of removing the terrain effect is illustrated through the visual anomalies of the $T_{zz}$ component before and after the correction is applied. (a) Observed $T_{zz}$; (b) Terrain effect calculated for $T_{zz}$ component; and (c) $T_{zz}$ after the terrain effect is removed. .................................................. 116

Figure 5.4  $T_{zz}$ difference map between the observed and predicted data from single component recovered model. .................................................. 121

Figure 5.5  (a) Geologic cross section running northwest (left) to southeast (right) and created using drill hole data. The location of this section is marked on Figure 5.2g. The pink is hematite with surrounding light and dark blue layers representing soft and hard itabirite, respectively. (b) Slice through $T_{zz}$ model corresponding to the geologic cross section; from 0 g/cc to 1.4 g/cc. The high density contrast correctly identifies hematite as a near surface body with the appropriate dip direction. Itabirite layers are defined by the dispersive positive density contrast radiating out from the dense hematite region. .................................................. 122
Figure 5.6  Single cross section comparison of four recovered models from inversion of (a) one-component, (b) three-components, (c) five-components, and (d) six-components. These sections correspond to the geologic cross section of Figure 5.5a. From the improvements seen in the recovered models from (a) one- to (c) five-components, the constraints from including additional components provide secondary information that contributes to a more accurate geologic structure. The inclusion of the sixth component (d) does not significantly enhance the ore delineation or quality of the recovered model. . . . . . . . . . . . . . . . . . . . . . 125

Figure 5.7  (a) Plan slice through five-component model at 1059m elevation with cross sections (b) - (f) denoted by white lines; (b-f) Cross sections from left (northwest) to right (southeast) through the five-component model. Dotted black line shows elevation of plan slice in (a). . . . . . . . . . . . 128

Figure 5.8  Volume rendered image of the 3D density contrast model constructed from five-component inversion with density contrast less than 1.0 g/cc removed. The high density bodies trending northeast to southwest agree with the occurrence of iron formation within the Gandarela Syncline. The semi-transparent topographic surface is shown above the 3D density contrast volume. . . . . . . . . . . . . . . . . . . . . . 129

Figure 6.1  Geologic map of the Quadrilátero Ferrífero, with sold black shading indicating where the Minas Series is present; modified from Dorr. The survey area is outlined in red. . . . . . . . . . . . . . . . . . . . . . 134

Figure 6.2  Terrain corrected gravity gradient data covering a 4 x 5 kilometer area with survey topography (f) showing the location of the data subset. From left to right and top to bottom, the plot panels display (a) $T_{xx}$; (b) $T_{xy}$; (c) $T_{xz}$; (d) $T_{yy}$; (e) $T_{yz}$; (f) topography; (g) $T_{zz}$. The location of the cross section of Figures Figure 6.9 and Figure 6.12 is overlaid on $T_{zz}$ in panel (g). . . . . . . . . . . . . . . . . . . . . . 137

Figure 6.3  Subset area total-field magnetic anomaly prior to removal of the regional field. . . . . . . . . . . . . . . . . . . . . . 138

Figure 6.4  Residual magnetic anomaly obtained from inversion-based regional-residual separation. . . . . . . . . . . . . . . . . . . . . . . 140

Figure 6.5  (a) 3D density contrast model. Cells with density contrast less than 1.13 g/cc are removed; Distance is in meters. (b) 3D susceptibility model. Cells with susceptibility less than 0.15 SI are removed. . . . . 143
Figure 6.6  (a) Scatter plot of susceptibility model versus density model. Each point represents the same cell location from the density and susceptibility models. (b) Contour lines illustrating cell concentration for a specified binning window. .................................................. 146

Figure 6.7  (a) Logarithmic scatter plot of susceptibility model versus density model ($\kappa$ vs. $\rho$) with background density of 2.67 g/cc added back in, showing cluster trends in the higher density range. (b) Logarithmic scatter plot of susceptibility model versus density contrast model ($\kappa$ vs. $\Delta \rho$), showing trends in the lower density range. .................................................. 147

Figure 6.8  (a) Scatter plot of density model versus susceptibility model with lithologic units color coded. (b) Volume rendered image of the 3D lithologic model using the assignments displayed in the scatter plot. .............................................. 148

Figure 6.9  (a) Geologic cross section generated from bore hole data showing itabirite in blue and hematite in pink. (b) Cross section through density contrast model; units are g/cc. (c) Cross section through susceptibility model plotted on a logarithmic color scale; units are SI. .................................................. 150

Figure 6.10  (a) Hematite distribution corresponding to lithologic model from generic constraints. (b) Cross section through lithologic model from generic constraints corresponding to the color scheme identified in Figure 6.8(a). .................................................. 151

Figure 6.11  (a) Scatter plot defining lithologic units with no susceptibility constraint on hematite distribution (pink). (b) Overestimate of hematite distribution when no susceptibility constraints are applied. .................................................. 153

Figure 6.12  The 2D geologic section in the top panel is sampled at discrete locations, shown in the middle panel, that corresponds to the cell centers, bottom panel, extracted from the recovered density contrast model. Every model cell coinciding with the cross sectional area now has a lithologic unit associated with the density value of that cell. .................................................. 154

Figure 6.13  (a) Geologic units associated with the 2D cross sectional areas extracted from the density and susceptibility model overlaid on the scatter plot of the two physical property models. (b) For visual clarity, the same data is plotted on a logarithmic scale using the density contrast model. .................................................. 156

Figure 6.14  (a) Scatter plot defining lithologic units with geologic density and susceptibility constraints. Hematite is pink. (b) Hematite distribution corresponding to lithologic model using geologic constraints. (c) Cross section through lithologic model using geologic constraints corresponding to the color scheme identified in Figure 6.8(a). .................................................. 158
Figure 6.15  A plan-section from 3D lithologic models at an elevation of 1150 m. Left: partial lithology model showing hematite distribution where susceptibility constraints are not used. Middle: generically constrained lithology model showing distribution of all assigned lithologic units. Right: geologically constrained lithology model showing distribution of all assigned lithologic units.

Figure 7.1  View of the 3D geologic model constructed by Krahenbuhl and Hitzman. The willemite lithology of interest is represented by the pink unit.

Figure 7.2  Synthetic generation of zinc content for 35 random locations at an approximate elevation of 224 m. The black dots represent values close to 0. Color units are zinc content.

Figure 7.3  True simulated distribution of zinc content based on geologic model of Figure 7.1.

Figure 7.4  Density contrast (g/cc) model resulting from inversion of synthetic six-component gravity gradient data contaminated with gaussian noise. All cells below 0.25 g/cc are removed to show recovery of the high density willemite.

Figure 7.5  Top: Empirical and modeled semi-variogram for zinc content. A spherical model is used with no nugget effect, a sill of 1.5, and a range of 150. Middle: Empirical and modeled semi-variogram for density contrast depth slice. A gaussian model is used with no nugget effect, a sill of 1.5, and a range of 200. Bottom: Empirical and modeled cross-variogram for zinc content and density contrast depth slice. A gaussian model is used with a nugget effect of 0.7, a sill of 0.7, and a range of 150.

Figure 7.6  (a) zinc content estimated from kriging of the 35 locations. (b) Kriging variance for the estimates in the left panel.

Figure 7.7  Density contrast (g/cc) used as a secondary variable for cokriging.

Figure 7.8  (a) zinc content estimated from cokriging of the 35 locations and density contrast of Figure 7.7. (b) Kriging variance for the estimates in the left panel.

Figure 7.9  Absolute difference between the estimates obtained from kriging and cokriging. Colorbar is zinc content.
Figure 7.10  (a) zinc content estimated from cokriging of the 35 locations and density contrast locations greater than 10% of the maximum recovered density contrast value. (b) Kriging variance for the estimates in the left panel. . 177

Figure 7.11  Top: Empirical and modeled semi-variogram for zinc content. A spherical model is used with no nugget effect, a sill of 1.5, and a range of 150. Middle: Empirical and modeled semi-variogram for density contrast depth slice. A gaussian model is used with no nugget effect, a sill of 1.3, and a range of 150. Bottom: Empirical and modeled cross-variogram for zinc content and density contrast depth slice. A gaussian model is used with a nugget effect of 0.2, a sill of 1.2, and a range of 150. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 178
**LIST OF TABLES**

Table 2.1  Table showing mean density estimate using $I$ statistics for various Gaussian noise levels dictated by $\sigma^2$ with zero mean. The true density value is 3.21 g/cc. .................................................. 25

Table 2.2  Table showing mean density estimate using $c$ statistics for various Gaussian noise levels dictated by $\sigma^2$ with zero mean. The true density value is 3.21 g/cc. .................................................. 25

Table 3.1  Noise estimates for the synthetic problem. True noise levels are given for the observed ($\sigma^o$) and denoised data ($\sigma^d$). Noise estimates from equivalent source denoising are given for the observed data ($\hat{\sigma}^o$). Mean noise estimates for the denoised data based on simulation are $\hat{\sigma}^d_k$ while the direct calculation are give as $\hat{\sigma}^d$. Noise estimates $\hat{\sigma}^o_k$ are for the known level of gaussian error used to generate the synthetic data realizations. .... 76

Table 3.2  Noise estimates for McFauld’s Lake subset based on equivalent source construction and simulation. .................................................. 78

Table 5.1  Table 1: RMS differences (in Eo) of the observed and predicted data for the four cases discussed. Shading denotes components that were not inverted and the recovered model was used to calculate the predicted data for differencing with the observed data. These values are the standard deviation of the residual map. In practice, we use these values as an estimate for the standard deviation of the data errors in the inversion process. .................................................. 121

Table 6.1  Summary of constraining inversion parameters for obtaining the 3D density contrast and susceptibility models. ................................. 142

Table 6.2  Specified density and susceptibility bounds and associated lithologic units using average petrophysical values. ................................. 149

Table 6.3  Total volume of hematite from three different lithology models. .... 160
Thank you to my advisor Dr. Yaoguo Li who has dared me, tested me, inspired me, and encouraged me. I am grateful for the opportunity to have learned from such a teacher and could not ask for a better mentor. Thank you to my committee members Dr. Murray Hitzman, Dr. Kadri Dagdelen, Dr. Paul Sava, and Dr. Richard Krahenbuhl for their guidance and thought provoking questions. It has truly been a pleasure to learn from and interact with such an esteemed group of researchers. Thank you to Vale and specifically Dr. Marco Braga and Dr. Dionisio Carlos for providing gravity gradient data used throughout this thesis. Thank you to Dr. Mark Dransfield and CGG for providing gravity gradient data used in this thesis as well. Thank you to Geoscience Australia, the Kauring Test Range sponsors, the Ontario Geological Survey, and Natural Resources Canada for making data used within this thesis publicly available.

Thank you to Dr. Terry Young, Michelle Szobody, and Dawn Umpleby for ensuring that I and all students are given every opportunity to learn and succeed here at Colorado School of Mines. Thank you to the Professors and faculty who have (knowingly or unknowingly) challenged me, championed me, and enlightened me over the years: Dr. Misac Nabighian, Dr. Dave Hale, Dr. Tom Davis, Dr. Gary Olhoeft, Dr. Mike Batzle and Dr. Mandy Hering. Thank you to Dr. Tom Boyd, Jahi Simbai, and Debra Lasich for providing graduate students with the tools to succeed and for helping me learn how to navigate the waters of leadership and progress. We have argued, celebrated, and commiserated together, thank you to my fellow CGEM students both past and present.

I’d like to thank my family who have always encouraged and supported me in my pursuit of higher education. Thank you to my mom, Antoinette, for always being a safe haven and continually reminding me what really matters in life. Thank you to my brother Santiago for our many conversations on life, love, and relativity. Thank you to my brother Javier,
for correcting me, taking care of me, and being a calming force in my life. Thank you to my padre, James, for being my cheerleader, my bedrock, and the shoulders that I stand on. Lastly, thank you to Zach Aman for being my person and the pillar that has held me up during my graduate education. Thank you for believing in me and instilling in me the tenacity and patience that have made pursuing this dream possible.
CHAPTER 1
INTRODUCTION

The advent of new or improved geophysical technologies and data acquisition platforms requires equal innovations in the processing and interpretation of the data. Gravity gradiometry is one such geophysical data type that has experienced technological advances in recent years [4–6]. Although gravity gradient instrumentation has been around for over a century [7, 8], the quantity and quality of the measurements were less than ideal. With the commercialization of classified instrumentation, the acquisition of high-quality gravity gradient measurements is now routinely available [9, 10]. The tools with which exploration geophysicists are able to analyze and interpret the data have lagged behind the ability to acquire data.

The need for gravity gradiometry interpretation tools is further accentuated by the current development of next-generation instrumentation which are expected to yield high-quality data with lower noise levels [11, 12]. Additionally, future instrumentation will not be subject to restrictions currently imposed on commercial instrumentation [13].

With such a short history compared to other well established geophysical methods, there is much to be done in understanding the utility of gravity gradiometry and best practices when used as an exploration or evaluation tool. Specific aspects of the gravity gradiometry data lifecycle need to be investigated and further related back to the practical assimilation into a geologic framework. With the understanding that data are pre-processed prior to delivery to a geophysicist, the work here builds upon current capabilities to properly handle multi-component data.

The relevance of any data integration or utilization relies on an understanding of the applied processing and prior treatment of the data. To use 3D inversion as an interpretation tool, application of terrain correction and obtaining error estimates are essential steps.
Chapter 2 and Chapter 3 address aspects of these two processing steps. Inversion algorithms can be designed to invert any combination of components measured, however map based interpretation techniques often rely on specific components. Chapter 4 addresses the ability to properly convert observed multi-component data to other components. After appropriate processing has been performed, Chapter 5 shows that 3D inversion can be used as an interpretation tool for multi-component data. Building upon the utility of 3D inversion as an interpretation tool, Chapter 6 illustrates how physical property distributions derived from multiple geophysical data types can be used for lithology differentiation. In an effort to integrate geophysically derived observations into exploration or mining workflows, Chapter 7 presents a method to incorporate geophysical models into resource evaluation.

In order to address the integrated use of products derived from gravity gradient data as summarized above, the chapters of this thesis detail practical aspects of processing and interpretation.

In Chapter 2, a technique to estimate the density value for terrain correction is discussed. Unlike gravity data, gravity gradient data only requires one correction: the terrain correction. The algorithm and digital elevation model used to calculate the terrain correction have been well addressed by other researchers. The background density, which is essential to the terrain correction, has not been well discussed. As the terrain correction (and associated background density) essentially removes signal from the observed data, there are implications on data interpretation. With this chapter, spatial statistics are exploited in an effort to identify a quantitative way to estimate the density value to be used for the terrain correction.

In Chapter 3, a method to denoise data and estimate data errors is developed using equivalent sources. While applied filtering and processing steps seek to enhance the gravity gradient signal, there are errors associated with the processing that should be considered during the interpretation stage. A simulation technique is also presented which can be used to estimate errors present in denoised or converted data.
In Chapter 4, the utility of component conversion is examined. With the basis for component conversion outlined in Chapter 2, the work presented here focuses on information content within converted components. Inversion based comparisons are used to validate the component conversion methodology.

In Chapter 5, a case study on inversion of airborne gravity gradiometry for iron ore exploration in the Quadrilátero Ferrífero, Minas Gerais, Brazil is presented. The work describes inversion of various component combinations to identify the presence of iron formation. Through comparison to a geologic section, the contribution of additional components to the recovered model is studied.

In Chapter 6, a follow-up case study on the integrated interpretation of airborne gravity gradient and magnetic data at the same site in the Quadrilátero Ferrífero is presented. The density contrast model resulting from inversion of gravity gradient data is combined with the susceptibility model from magnetic inversion in order to begin characterizing lithology.

In Chapter 7, a method is developed for integrating the quantitative nature of geophysical inverse models into the resource workflow beyond exploration. At the resource evaluation stage, quality density or lithology models can add value to the modeling process. Although there is generally more information available at the evaluation stage, there are areas within and around the target with little to no information. At a minimum, the spatial trends within the geophysical models can potentially be utilized.

These chapters assume some familiarity with gravity gradiometry and geophysical inversion. In the next two sections, the theory behind gravity gradiometry and the framework for inverting gravity gradient data are summarized.

1.1 Gravity gradient tensor

Systems currently deployed in exploration rely on parallel sets of accelerometers on a plate known as a gravity gradiometer instrument, or GGI [9, 10]. The distance between the accelerometers and the orientations of multiple GGIs within the gradiometer varies depending on the specific instrument. The combination of gravity gradient tensor components that
are collected varies as well.

Acquisition and pre-processing parameters required by current systems are controlled by the U.S. Department of State International Traffic in Arms Regulations [13]. Regulations prohibit distribution of raw data and all data available to geophysicists have been line leveled and low pass filtered as part of the internal system pre-processing. For land surveys, flight height must be higher than 30m and areal extent must be smaller than 20,000 square miles. Marine surveys must not exceed the same areal extent and must be collected within 200 miles of the nearest coast line.

The gravity gradient is a measure of the spatial rate of change of the gravitational acceleration. The gravitational acceleration, $\vec{g}$, is the gradient of the gravitational potential, $\Phi$, as shown below,

$$\vec{g} = \nabla \Phi = \left( \frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z} \right) = (g_x, g_y, g_z) \quad (1.1)$$

The gradient of the acceleration, $\vec{g}$, gives the gravity gradient tensor field, $\mathbf{T}$, [14] defined as

$$\mathbf{T} = \nabla \vec{g} = \nabla \nabla \Phi = \begin{bmatrix} \frac{\partial^2 \Phi}{\partial x^2} & \frac{\partial^2 \Phi}{\partial x \partial y} & \frac{\partial^2 \Phi}{\partial x \partial z} \\ \frac{\partial^2 \Phi}{\partial y \partial x} & \frac{\partial^2 \Phi}{\partial y^2} & \frac{\partial^2 \Phi}{\partial y \partial z} \\ \frac{\partial^2 \Phi}{\partial z \partial x} & \frac{\partial^2 \Phi}{\partial z \partial y} & \frac{\partial^2 \Phi}{\partial z^2} \end{bmatrix} = \begin{bmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{bmatrix} \quad (1.2)$$

The gravitational potential is a smooth function and satisfies Laplace’s equation in source free regions. Therefore, the gravity gradient tensor is symmetric and traceless. The smooth nature of the potential leads to symmetry, i.e., $T_{xy} = T_{yx}$, $T_{xz} = T_{zx}$, and $T_{yz} = T_{zy}$; and the Laplacian property leads to the zero trace, i.e., $T_{xx} + T_{yy} + T_{zz} = 0$ [15]. Consequently, there are only five independent components at any observation point. We remark that, in theory, the maps of five components on the observational surface are linearly dependent on each other and one component carries all the information. However, the availability of measurements only at discrete observation locations and the presence of noise in the data mean that the extra component maps do carry additional information. Thus, practical applications utilize as many components as available.
Gravity gradiometry is a valuable alternative to typical vertical gravity, \( g_z \), measurements in some applications. The commonly sparse collection of vertical gravity measurements limits the lateral resolution and extent to which a target or structure can be characterized. Compared to ground and airborne gravity, gravity gradiometry can potentially characterize the target body to a greater extent. The increased data density makes gravity gradiometry preferable to ground gravity. The higher signal-to-noise ratio makes gravity gradiometry superior to airborne gravimetry. A gravity gradient survey has improved lateral resolution over a gravity (ground or airborne) survey with the same observation locations due to the contribution from the horizontal components on how the gravity field is changing laterally. The high-frequency content in gravity gradient measurements is due to the decay of the field inversely proportional to the distance cubed. The decay rate generally results in increased signal contribution from shallower targets and relatively smaller signal contribution from deeper targets. For this reason, the resolution of near surface features is increased relative to that by vertical gravity.

The unit of measure for the gravity gradient is the Eotvos (Eo), named for Loránd Eötvös\(^1\), where 1 Eo is equal to 0.1 mGal/km, or \( 10^{-9} \text{1/s}^2 \). Similar to gravity data, gravity gradiometry data require several reduction steps to remove contributions unrelated to the underlying geologic signal. The terrain is one of the major contributors to the measured signal particularly in areas of high topographic relief. For this reason, removing the terrain effect is an important step in reducing the data to an interpretable form. Removing the terrain effect leaves only gravity gradient signal from local sources in theory, which is better for modeling or inversion for the anomalous masses of interest. For this reason, care must be taken in selecting a representative density value of the topographic region.

As discussed above, each component of the gravity gradient data contains similar information about the subsurface density distribution. However, when gridded they visually tend to present different first order information. For example, the components with horizontal

---

\(^1\)Reference to Eötvös’s name varies in the literature. His given Hungarian name and title is Vásárosnaményi Bárón Eötvös Loránd. Commonly referred to in German as Roland, Baron von Eötvös. [16]
derivatives \((T_{xx}, T_{xy}, \text{and} T_{yy})\) tend to highlight the lateral density boundaries more readily, whereas the vertical gradient, \(T_{zz}\), provides a more intuitive anomaly map, making it the most widely used for visual interpretation. For this reason, map-based interpretations relying on image enhancement tend to utilize varying components for different purposes. Prevalent methods thus far for interpreting various components of gravity gradient data are focused on enhancing edges, lineaments, and using invariants [17, 18]. An alternative, quantitative technique that can assist interpretation is 3D inversion, which is not crucially dependent on the selection of a particular component. We utilize this approach in this paper and invert the data to construct 3D distributions of density contrast and compare the recovered models to known geology to understand the potential for geologic interpretations based on the recovered 3D density model.

1.2 Summary of 3D inversion

We use the 3D potential-field inversion algorithm developed by Li [19] and Li and Oldenburg [20] to invert the gravity gradient data. This algorithm is implemented in a program library Software [21]). We adopt a right hand Cartesian system where \(x\) is northing, \(y\) is easting, and \(z\) points vertically down. The algorithm assumes a set of contiguous rectangular prisms each with a constant density contrast. The model is obtained via the inverse solution of the forward modeling equation

\[
G \vec{m} = \vec{d}
\]

where \(\vec{m} = [\rho_1, \rho_2, ..., \rho_m]^T\) is the model vector containing the density contrasts in cells, \(\vec{d} = [d_1, d_2, ..., d_n]^T\) is the data vector, and \(G\) is the sensitivity matrix that quantifies the contribution of cells in the model to the gravity gradients at observation points. The inverse solution is obtained using Tikhonov regularization [22] by minimizing a total objective function with bound constraints,

\[
\min \phi = \phi_d + \mu \phi_m \quad \text{subject to} \quad \rho_{\min} \leq \rho \leq \rho_{\max}
\]
where $\mu$ is a regularization parameter, $\phi_d$ and $\phi_m$ are data misfit and model objective functions, respectively, $\bar{\rho}_{\text{min}}$ and $\bar{\rho}_{\text{max}}$ contains the lower and upper bounds, respectively, on the density contrast to be inverted. The ability to incorporate lower and upper bounds on the recovered density contrast is useful here since we expect a large positive density contrast associated with the iron formation.

The data misfit in equation 1.4 is defined by

$$
\phi_d = \left\| W_d (d_{\text{obs}} - d_{\text{pre}}) \right\|^2 
$$

(1.5)

where $d_{\text{obs}}$ is the observed data, $d_{\text{pre}}$ is the predicted data, and $W_d$ is a data weighting matrix that contains on its diagonals the reciprocal of the standard deviations of errors in the data. The model objective function in equation 1.4 is defined by,

$$
\phi_m (\rho) = \int \int \int_V \left[ w(z)(\rho - \rho_0) \right]^2 dv 
$$

$$
+ L_x^2 \int \int \int_V \left[ \frac{\partial w(z)(\rho - \rho_0)}{\partial x} \right]^2 dv 
$$

$$
+ L_y^2 \int \int \int_V \left[ \frac{\partial w(z)(\rho - \rho_0)}{\partial y} \right]^2 dv 
$$

$$
+ L_z^2 \int \int \int_V \left[ \frac{\partial w(z)(\rho - \rho_0)}{\partial z} \right]^2 dv 
$$

(1.6)

where $L_x$, $L_y$, and $L_z$ are the length scales defining the relative smoothness of the resulting model in each of the Cartesian directions. Varying the value of these length scales helps emphasize the recovered structural continuity in a specific direction. Prior knowledge of the geologic structure can be incorporated through these length scales. For example, if we know that the geologic structure is more laterally extensive in the east-west direction than the north-south direction, we would make $L_y > L_x$ (where $y$ is easting and $x$ is northing).

The sought model is denoted by $\rho$, and $\rho_0$ is a reference model through which a priori geologic structure can be incorporated into the inversion. The function $w(z)$ is a weighting function of the form $w(z) = (z + z_0)^{-3/2}$ used to counteract the decay of the kernel function and prevent the accumulation of density occurring at shallow depths. Alternatively, a 3D
weighting function based on distance to data locations or composite sensitivity can also be used in the presence of rugged surface topography and a highly undulating observation surface.

The user-specified bounds are implemented using the logarithmic barrier method [23, 24]. It converts inequality constraints into a barrier function and is included as a part of the objective function. The objective function in equation 1.4 then becomes,

$$\phi(\lambda) = \phi_d + \mu \phi_m - 2\lambda \sum_{i=1}^{M} \left[ \ln\left( \frac{\rho_j - \rho_{j}^{\text{min}}}{\rho_{j}^{\text{max}} - \rho_{j}^{\text{min}}} \right) + \ln\left( \frac{\rho_{j}^{\text{max}} - \rho_j}{\rho_{j}^{\text{max}} - \rho_{j}^{\text{min}}} \right) \right]$$

(1.7)

where $\lambda$ is a barrier parameter and $\rho_{j}^{\text{min}}$ and $\rho_{j}^{\text{max}}$ are the lower and upper bounds, respectively, on the unknown density contrast $\rho_j$ in $j^{th}$ cell. The third term is the logarithmic barrier function and it can be conceptually thought of as a barrier at the user-supplied bounds that prevents the model from obtaining a density value outside the given bounds. A sequence of non-linear minimizations is solved to approach the solution to equation 1.4 as $\lambda$ approaches zero.

The results in this paper do not incorporate variable length scales or a non-zero reference model into the inversion. We only carry out blind inversions to understand the model quality that can be recovered in areas where little geologic information is available.

Storage of the full sensitivity matrix and multiplication of the sensitivity matrix to vectors require a large amount of memory and CPU time. The utility of wavelet compression [20] of the sensitivity matrix enables the algorithm to invert large gradiometry datasets. This is accomplished by storing a sparse representation of the sensitivity matrix and through sparse multiplication with the model vectors in the wavelet domain.

With the wavelet compression of the forward mapping defined in equation 1.3, and the optimization formulation described above, we have an efficient and practical 3D inverse methodology for the gravity gradiometry problem.
In gravity gradiometry, the observed data must be reduced in order to obtain the anomalous response due to target geologic features. One of the most important corrections that must be applied to the observed data is the terrain correction or bathymetric correction. This correction removes the response of the terrain (or bathymetry) from the observed data. In order to remove the contribution of the topographic surface from the overall measured signal, a representative density value is typically selected to calculate the terrain response to be removed. Though removal of this correction from gravity and gravity gradient data is a well-established step in the data processing workflow, there has been little work describing methods by which to estimate the density to use. In this work, the feasibility of using spatial statistical methods to estimate the representative or average density that sufficiently removes the effect from observed gravity gradient data is examined. The gravity gradient data are treated as geostatistical data. Measures such as Moran’s $I$ and Geary’s $c$ are used to quantify the spatial autocorrelation of a given geographic data set. The $I$ and $c$ statistics can be calculated for either a global or local measure of spatial autocorrelation. The level of spatial autocorrelation of the desired anomalous data is expected to vary according to the density value and the distance window used to assign weights.

2.1 Introduction

Though removal of the terrain effect from gravity and gravity gradient data is a well-established step in the data processing workflow, there has been little work describing methods by which to estimate the density to use. The standard way of estimating the density for terrain effect removal relies on profiles through the observed and topography data. There
have been other methods proposed over the past couple of decades that allow the density value to vary within the survey area.

In the context of inversion, removing the terrain effect is an efficient way to reduce computation time but is not necessarily a requirement. Inversion algorithms can be designed to invert the observed data prior to terrain correction. In practice most algorithms require the data to be reduced to the anomaly (i.e. terrain corrected data). Map-based interpretation likewise assume the gravity gradient data have been terrain corrected.

The algorithm and requirements on the resolution of the digital elevation model used to calculate the terrain correction are crucial. Chen and Macnae [25], Kass and Li [26], Dransfield and Zeng [27], and Davis et al. [28] have studied various aspects of calculating the terrain effect assuming a constant background density. For consistent processing of gravity gradient data, the modeled terrain effect should be filtered according to acquisition and prior processing steps before applying the terrain correction. This ensures high-frequency content that has been previously filtered out are not re-introduced into the data. Applying a terrain correction is not strictly error free and the terrain correction should be considered a source of error in gravity gradient anomaly data.

The density value used in conjunction with the terrain correction can highlight different geologic features and should not be considered trivial. [29] illustrates that two different terrain correction densities draw out different geologic domains within a single survey area. If the geologic feature of interest is coincident with the terrain, it becomes more important to select a representative density so as to not over- or under-correct for the terrain effect. In areas where the regional geology is not changing rapidly, a single density is sufficient to remove the terrain effect.

Often, the terrain correction is applied using a constant density although a laterally varying density function could also be used. To use a laterally varying density, a sufficient understanding of the change in geology should be available to support the density variation and depth extent of the terrain model. If applied improperly or with a lack of geologic
knowledge, a laterally varying density could unintentionally suppress an anomaly of interest. With a laterally varying density, the quantitative interpretation of models may require additional steps depending on the algorithm and whether or not the density varies laterally or step-wise.

The recovered models from inversion of the gravity gradient anomaly result in a representative density contrast model. To obtain an absolute density model, the terrain correction density is added to the recovered cell values. With a constant terrain density, the quantitative interpretation of resulting models is then straightforward. In practice, when the target is not coincident with the topography the exact value is less important and there is a small range of constant density values that can be considered sufficient to remove the terrain effect.

There have been two methods that have been commonly used to estimate a single density value and relate to the methodology proposed here. The two main methods using profiles are Nettleton’s Method [30] and a similar method proposed by Parasnis [31]. Nettleton’s Method correlates the observed data profiles using various density values with the corresponding profile of the terrain as seen in Figure 2.1. The correlation of the observed gravity profile and the topography is carried out with varying density values. The density value that provides a profile with minimum correlation to the topography is considered the representative density value. It should be noted that when the target of interest is coincident with the topography, the validity of this method is much decreased. Parasnis uses a similar approach by specifying two points along the profile that are unaffected by the anomalous source. With assumptions about the geologic features present, the slope of the line between the two points is the magnitude of the representative density that should be used. Carlos [32] has proposed a method that builds upon Parasnis’ method but uses $T_{zz}$. The optimal density is that which minimizes the covariance between the anomalous data and the topography.

These two methods have proved useful over the years. The density value selected by these methods are highly dependent on the user and their interpretation of the profiles. Some work has been done on modifying these methods and generating least squares algorithms to
automatically select the optimal density value. Despite these efforts, the standard operating procedure for selecting the density value still relies on correlation of gravity and topographic profiles with user interaction.

Figure 2.1: Nettleton’s method of correlating gravity anomaly corrected with various densities with the topography. From Milsom and Eriksen [1].

In this chapter, the feasibility of using statistical methods to estimate the representative or average density that sufficiently removes the terrain effect from observed gravity gradient data will be examined. The observed gravity gradient data is treated as geostatistical data. The intent is to exploit the fact that there is a spatial dependence in the observed data. The level of spatial autocorrelation of the desired anomalous data is expected to vary according to the density value.

First, the terrain effect will be described. Then the use of Moran’s $I$ and Geary’s $c$ to estimate the optimal density will be detailed and illustrated through a field example. Following are a validation study on the use of these spatial statistics for density estimation and investigation of a difficult case where the target geology is coincident with topography.
2.2 Terrain Effect in Gravity Gradiometry

Unlike vertical gravity measurements, gravity gradients do not require numerous corrections to the measured data in order to obtain the anomaly values. The topography is one of the strongest contributors to the measured gradient values, particularly in areas of high relief. For this reason, removing the terrain effect is an important step in reducing the data to an interpretable product. Furthermore, the density value used to remove the terrain effect from the data is crucial. With a density that is too high, signal from the target may be removed while a density that is too low fails to fully remove the terrain effect. For this reason, care must be taken in selecting the proper representative density value of the topographic region.

The gravity gradient at an observation location, \( r_0 \), from a volume source, \( V \), can be calculated according to equation 2.1 where \( \gamma \) is the gravitational constant, \( \rho \) is the average density of the volume, and \( r \) is the source location.

\[
T_{kl} = \gamma \rho \int_V \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_l} \frac{1}{|r - r_0|} \, dv
\] (2.1)

It can be seen that the terrain effect can be treated as a linear combination of discrete sources with the given topographic surface. The contribution to the measured gravity gradient from the terrain at an observation location, \( d_i \), can be calculated as in Equation 2.2. For further details on calculation of the terrain effect, see Kass and Li [26].

\[
d_i = \gamma \sum_{j=1}^{M} \rho \left( \int_{\Delta V_j} \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_l} \frac{1}{|r - r_0|} \, dv \right)
\] (2.2)

2.3 Density Estimation

The premise of this work is that the spatial dependence of the gravity gradient data may be used to assist in estimating the optimal density value for removal of the terrain effect. Various statistical methods were initially explored, but Moran’s \( I \) and Geary’s \( c \) show the
most potential for estimating the density value. In the next section, a field dataset will be introduced that will carry through the description of using the two autocorrelation statistics for estimating density. Then, a detailed discussion on the calculation of both Moran’s \( I \) and Geary’s \( c \) statistics will follow.

### 2.3.1 Field Data Illustration

To illustrate the method, a set of field gravity gradient data will be used throughout the description of the autocorrelation statistics. The gravity gradient data were collected in August-September 2005. The 93 \( km^2 \) survey was acquired with 100 m line spacing trending northeast-southwest at roughly 32 degrees from the north. We use a smaller 20 \( km^2 \) subset of the collected data.

The geologic feature of interest runs through the middle of the data parallel to the long axis of the survey area and is coincident with topographic highs in the area, making the density value used to remove the terrain effect an important parameter. The observed gravity gradient components are shown in Figure 2.2. The altitude of the survey is shown in the bottom left panel. The calculated terrain effect is shown in Figure 2.3. An image of the topography is given in Figure 2.4 with the black box outlining the smaller area of interest. Comparing only the observed gravity gradient data and the topography in the area illustrates the strong contribution in the measurements from the terrain itself. Comparing the observed gravity gradient data and the terrain effect further illustrates the substantial signal in the observed data from the terrain.

### 2.3.2 Moran’s \( I \) and Geary’s \( c \)

Moran’s \( I \) and Geary’s \( c \) are statistics that measure the spatial autocorrelation of a given geographic data set. The \( I \) and \( c \) statistics can be calculated for either a global or a local measure of spatial autocorrelation. The first step to estimating the average density is with a Local Indicator of Spatial Autocorrelation (LISA) using the Moran’s \( I \) and Geary’s \( c \) statistics. As can be seen in the data images of Figures Figure 2.2 and Figure 2.3, the
Figure 2.2: Observed gravity gradient data.

Figure 2.3: Calculated terrain effect displayed using a density of 1.0 g/cc.
data has spatial correlation that may be utilized to explore estimation of the optimal density value.

Using a LISA rather than a global estimate of Moran’s $I$ and Geary’s $c$ is helpful in characterizing the local variations of the data with varying density values and gives a qualitative view of the effectiveness of this statistics. The data example introduced in the previous section is used where the terrain effect is calculated with an initial density of 1.0 g/cc that can be scaled accordingly. The spatial correlation of the residuals will vary based on the magnitude of the removed terrain effect. The relationship between the observed data ($\vec{d}$), terrain effect ($\vec{t}$), and residuals ($\vec{r}$) is shown in Equation 2.3.

$$\vec{r} = \vec{d} - \rho \vec{t} \quad (2.3)$$

The local Moran’s $I$ and Geary’s $c$ statistics are calculated according to Equations 2.4 and 2.5, respectively [33]. The $s_j$ and $s_i$ are the (x,y)locations of the gravity gradient attribute.
denoted by $z$. $\bar{z}$ is the estimated mean value and $s^2$ is the estimated standard deviation of the data map. The $W_{ij}$ matrix is a data weighting matrix defining the spatial connectivity of the points.

$$I(s_i) = \frac{n}{(n-1)s^2} \left( z(s_i) - \bar{z} \right) \sum_{j=1}^{n} W_{ij} \left( z(s_j) - \bar{z} \right)$$  \hspace{1cm} (2.4)

$$c(s_i) = \frac{1}{s^2} \sum_{j=1}^{n} W_{ij} (z(s_i) - \bar{z})^2$$  \hspace{1cm} (2.5)

A global measure of the Moran’s $I$ and Geary’s $c$ statistics are obtained by summing the local $I(s_i)$ and $c(s_i)$ values. The global measure of Moran’s $I$ is given by equation 2.6 and Geary’s $c$ is give by equation 2.7. The $w..$ is the sum of all weights, $W_{ij}$, and scales the global statistic accordingly.

$$I = \frac{n}{(n-1)s^2w..} \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \left( z(s_i) - \bar{z} \right) \left( z(s_j) - \bar{z} \right)$$  \hspace{1cm} (2.6)

$$c = \frac{1}{2s^2w..} \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \left( z(s_i) - z(s_j) \right)^2$$  \hspace{1cm} (2.7)

The two statistics assume the data set, $z$, has a constant mean and constant variance. Interpretation of the $I$ and $c$ statistics require the expected value for each statistic. The expected value for Moran’s $I$ is given by equation 2.8 while the expected value for Geary’s $c$ is given by equation 2.9. A site is considered similar to its surrounding sites if $I > E[I]$ and the spatial autocorrelation is positive. If $I < E[I]$ then neighboring sites are not similar. For Geary’s $c$, sites are considered similar if $c < E[c]$ and dissimilar if $c > E[c]$.

$$E[I] = \frac{-1}{n-1}$$  \hspace{1cm} (2.8)

$$E[c] = 1$$  \hspace{1cm} (2.9)

2.3.3 Local $I$ and $c$ Statistics

The local $I$ and local $c$ statistics were calculated at each of the 1664 data locations for 201 density values between 2 g/cc and 4 g/cc for each of the six components. Results will
Figure 2.5: Local Moran’s $I$ for $T_{zz}$ using various density values to remove the terrain effect. Color displays the value of $I$ at each location.
be shown for the $T_{zz}$ component only, though calculations were carried out for the other five components as well.

The local Moran’s $I$ for a few of the density values used is shown in Figure 2.5. The density value used to remove the terrain effect is displayed at the top of each subplot. The red colors are higher values of $I$, indicating spatial correlation (sites are similar) while the blue colors are lower values of $I$ indicating a lack of spatial correlation (dissimilar sites). In moving from 2 g/cc to 4 g/cc, different features are highlighted by the LISA’s. At 2 g/cc, two main features are identified as being correlated while near 4 g/cc there are numerous elongated features. Visually, anomalous features can be identified as running north-east to south-west through the domain. The shape of the features and the optimal density value are not best determined by simply inspecting the plots.

As a next step, we explore how the $I$ and $c$ values compare to the expected values of the local Moran’s $I$ and Geary’s $c$ statistics. The expected values of the local versions of $I$ and $c$ are given in Equations 2.10 and 2.11, respectively. The $W_i$ is the sum of the weights relating to the $i$th point and $n$ is the number of observations.

\[
E[I_i] = \frac{-W_i}{n - 1} \tag{2.10}
\]

\[
E[c_i] = \frac{2nW_i}{n - 1} \tag{2.11}
\]

The $I$ and $c$ statistics calculated for the $i$’th observation are then compared to these local expected values of the $I$ and $c$ statistics for the $i$’th observation. If $I_i > E[I_i]$ then the index of that observation is assigned a 1 to indicate positive spatial autocorrelation, and is 0 otherwise. If $c_i < E[c_i]$ then the index of that observation is assigned a 1 to indicate positive spatial autocorrelation and is 0 otherwise.

The comparison and assignment of the assignment of 1 for spatial autocorrelation as compared to the expected value is shown in Figure 2.6 for the $T_{zz}$ component $I$ statistic. The red values correspond to 1 and indicate positive spatial autocorrelation while the blue values correspond to 0. As expected, the spatial autocorrelation becomes negative over the
area of interest as the density value becomes too large. With a feel for how the spatial correlation is changing as a function of the density value, we turn our attention to a global measure of spatial autocorrelation.

![Diagram](image)

Figure 2.6: Blue represents values of 0, indicating a lack of or no spatial autocorrelation. Red represents values of 1, indicating spatial autocorrelation.

2.3.4 Global $I$ and $c$ Statistics

Since the LISA statistics tend to be more for qualitative inference rather than quantitative conclusions, attention is turned to global measures of Moran's $I$ and Geary's $c$. Rather than
calculating new global statistics, the fact that the local and global measures are proportional is exploited. The LISA $I$ and $c$ statistics for each density in the range used are summed for a single measure of similarity or dissimilarity between the sites.

The proportional measure is to be scaled by the sum of the weights to obtain the global measure. The scaling of the measure by the sum of the weights was not performed here, as the focus is not the $I$ and $c$ value itself but rather the trend of the statistics across the density range. Scaling by the sum of the weights bulk shifts the curve and has implications for interpreting against the expected values, which is not performed here. Figure 2.7 shows the globally proportional $I$ statistic plotted as a function of the density value, with Figure 2.8 showing the $c$ statistic similarly plotted.

Examining the trend of the $I$ statistic, five of the six components show high $I$ values towards the ends of the density range with minimum $I$ values occurring near the middle of the density range. The $T_{xy}$ component does not overtly exhibit this behavior as the other components do. The $c$ statistic is opposite of the $I$ statistic and reaches a maximum within the density range, trending towards lower values on the edge of the range. Again, the $T_{xy}$ component does not overtly exhibit this general behavior. In terms of estimating a density to remove the terrain effect, the residuals are expected to have lower spatial correlation and clustering than that of the observed data.

In the context of estimating a density to remove the terrain effect, it is expected that the residuals have lower spatial correlation and clustering than that of the observed data since the observed data are representative of the topography. The lowest measure of $I$ may indicate a lack of spatial correlation and correspond to the density that sufficiently removes the highly correlated terrain effect. Similarly, a higher $c$ value indicates a lack of spatial correlation. With this in mind, the density value corresponding to the lowest $I$ and highest $c$ value for each $T_{ij}$ component is extracted.

The density value corresponding to the lowest $I$ measure and highest $c$ measure for each $T_{ij}$ component is extracted. The densities corresponding to the lowest $I$ statistics are
Figure 2.7: Global proportioned Moran’s I statistic of all six components plotted against their corresponding density value on the abscissa.

\[ T_{xx} = 2.39, \ T_{xy} = 3.67, \ T_{xz} = 2.58, \ T_{yy} = 3.04, \ T_{yz} = 2.68, \text{ and } T_{zz} = 2.87 \text{ giving an average value of } 2.8717 \text{ g/cc.} \]  

The densities corresponding to the maximum \( c \) statistics are \( T_{xx} = 2.48, \ T_{xy} = 3.54, \ T_{xz} = 2.63, \ T_{yy} = 2.84, \ T_{yz} = 2.74, \text{ and } T_{zz} = 2.67 \text{ giving an average value of } 2.8167 \text{ g/cc.} \)  

These two average values seem to be somewhat consistent with each other, suggesting an average value of 2.84 g/cc. The terrain corrected gravity gradient components are shown in Figure 2.9 where the anomalous features can be seen trending northeast-southwest through the domain.

2.4 Validation of Methodology

With a general methodology outlined, the quality of the estimate can be investigated with a synthetic problem. The synthetic problem consists of four features: two dipping dykes with density of 0.7 g/cc and 1.0 g/cc, a linear feature with density 0.5 g/cc, and an
Figure 2.8: Global proportioned Geary’s $c$ statistic of all six components plotted against their corresponding density value on the abscissa.

ellipsoidal feature with density of 0.1 g/cc. This model is arbitrary and was constructed to give overlapping signals. It is noted that these values are density contrast. The density value is obtained by adding in the average density value we are trying to estimate (the background, or average density of the volume). The terrain effect and anomaly data arising from the volume and a synthetic topographic surface is calculated according to Equation 2.2. The terrain effect is added to the anomaly data and contaminated with gaussian noise at various levels.

The process of estimating the density using globally proportional statistics of $I$ and $c$ are carried out on the components to explore the effect of noise on the average density estimate. One density value was tested with the results summarized in Table 2.1 and Table 2.2. A Monte Carlo method using 100 samples was employed to obtain the statistics shown in the Tables. For each Monte Carlo sample, a gaussian noise vector based on the given standard deviation was generated using a random number generator. This noise vector was taken as
Figure 2.9: Anomalous gravity gradient data with terrain effect removed with a density of 2.84 g/cc.
the base noise for each component to simulate correlated noise. To each individual gaussian noise value, a single random value with mean 0 and standard deviation of 1 was added in order to introduce minor fluctuations into the noise level from data point to data point as well as from component to component.

Table 2.1: Table showing mean density estimate using $I$ statistics for various Gaussian noise levels dictated by $\sigma^2$ with zero mean. The true density value is 3.21 g/cc.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma = 5$</th>
<th>$\sigma = 10$</th>
<th>$\sigma = 15$</th>
<th>$\sigma = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{xx}$</td>
<td>3.17</td>
<td>3.17</td>
<td>3.17</td>
<td>3.16</td>
</tr>
<tr>
<td>$T_{xy}$</td>
<td>3.21</td>
<td>3.22</td>
<td>3.21</td>
<td>3.21</td>
</tr>
<tr>
<td>$T_{xz}$</td>
<td>3.17</td>
<td>3.18</td>
<td>3.19</td>
<td>3.19</td>
</tr>
<tr>
<td>$T_{yy}$</td>
<td>3.20</td>
<td>3.20</td>
<td>3.20</td>
<td>3.19</td>
</tr>
<tr>
<td>$T_{yz}$</td>
<td>3.23</td>
<td>3.24</td>
<td>3.24</td>
<td>3.24</td>
</tr>
<tr>
<td>$T_{zz}$</td>
<td>3.19</td>
<td>3.19</td>
<td>3.19</td>
<td>3.20</td>
</tr>
<tr>
<td>mean</td>
<td>3.195</td>
<td>3.2</td>
<td>3.2</td>
<td>3.1983</td>
</tr>
</tbody>
</table>

Table 2.2: Table showing mean density estimate using $c$ statistics for various Gaussian noise levels dictated by $\sigma^2$ with zero mean. The true density value is 3.21 g/cc.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma = 5$</th>
<th>$\sigma = 10$</th>
<th>$\sigma = 15$</th>
<th>$\sigma = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{xx}$</td>
<td>3.15</td>
<td>3.15</td>
<td>3.16</td>
<td>3.18</td>
</tr>
<tr>
<td>$T_{xy}$</td>
<td>3.20</td>
<td>3.23</td>
<td>3.21</td>
<td>3.20</td>
</tr>
<tr>
<td>$T_{xz}$</td>
<td>3.16</td>
<td>3.17</td>
<td>3.18</td>
<td>3.19</td>
</tr>
<tr>
<td>$T_{yy}$</td>
<td>3.20</td>
<td>3.20</td>
<td>3.19</td>
<td>3.19</td>
</tr>
<tr>
<td>$T_{yz}$</td>
<td>3.22</td>
<td>3.23</td>
<td>3.23</td>
<td>3.24</td>
</tr>
<tr>
<td>$T_{zz}$</td>
<td>3.18</td>
<td>3.18</td>
<td>3.18</td>
<td>3.18</td>
</tr>
<tr>
<td>mean</td>
<td>3.185</td>
<td>3.1933</td>
<td>3.1917</td>
<td>3.1967</td>
</tr>
</tbody>
</table>

From Table 2.1 it is seen that the Moran’s $I$ density estimate for each component is relatively consistent, regardless of the noise level. Comparing the estimates across components, the $T_{xy}$, $T_{yy}$, and $T_{yz}$ consistently estimate a higher density value than the $T_{xx}$, $T_{xz}$, and $T_{zz}$ components do. The same can be said of the Geary’s $c$ density estimates. The resultant mean values from the Geary’s $c$ Monte Carlo simulation are consistently 0.1 to 0.3 g/cc smaller than than the actual 3.21 g/cc density value used. The resultant mean values from the Moran’s $I$ Monte Carlo simulation are consistently 0.1 to 0.2 g/cc smaller than the
actual value. It must be reiterated that these are the results of 100 independent realizations and that the representative value will change from sample to sample.

Based on the observations above, it seems that the method is robust for this particular type of noise. In general, this method estimates the density value relatively well. In terms of the field dataset, a value of 2.84 g/cc for the terrain correction may be a plausible value though it may appear high.

2.5 Test Case with Dipping Dyke and Coincident Topography

While the two examples above illustrate the effectiveness of the technique, for the real data example there is some question as to whether or not the topography and coincident target anomaly are being separated to an acceptable degree. To address this question, we utilize a second synthetic scenario. A dipping dyke coincident or semi-coincident with an elongated hill serves as a test case for further investigating how robust each of the spatial statistics are for estimating the optimal density. The three dyke scenarios are shown in Figure 2.10 where the dyke has a density contrast of 1.0 g/cc with a constant background. The variations on an elongated hill with dipping dyke orientation are as follows:

- North trending hill with dipping dyke beneath the hill (Figure 2.10(a))
- North trending hill with dipping dyke within the hill (Figure 2.10(b))
- North trending hill with dipping dyke offset from the hill (Figure 2.10(c))

Each of the dyke orientations have three survey configurations for a total of nine variations on the observed gravity gradient data. The three survey configurations are north trending lines (Figure 2.11(a), east trending lines (Figure 2.11(b), and north-east trending lines (Figure 2.11(c) over the elongated hill. The east and north trending surveys contain 370 observation locations while the northeast trending survey contains 781 observation locations. In summary, the nine datasets used to further understand spatial statistics for density estimation in the difficult scenario of coincident target geology and topography are:
Figure 2.10: (a) Top of the dipping dyke is centered beneath the elongated hill at a depth from 53m down to 20m. (b) Top of dipping dyke is centered within the elongated hill at a depth from 61m down to 25m. (c) Dipping dyke is offset from center of the elongated hill by 30m to the east at a depth from 53m down to 20m.
• North trending lines over dipping dyke beneath hill
• North trending lines over dipping dyke within hill
• North trending lines over dipping dyke offset from hill
• East trending lines over dipping dyke beneath hill
• East trending lines over dipping dyke within hill
• East trending lines over dipping dyke offset from hill
• North-east trending lines over dipping dyke beneath hill
• North-east trending lines over dipping dyke within hill
• North-east trending lines over dipping dyke offset from hill

The observed data are obtained by calculating the terrain effect using a constant background density of 2.54 g/cc and adding it to the forward modeled response of the dyke model. Gaussian noise with zero mean and standard deviation equal to 20% of the average data range for all components is added to each gradient component. For each scenario enumerated in the previous list, 25 data realizations are generated (i.e. 25 random noise sets) and subjected to the density estimation methodology previously described. The average density estimate from the 25 realizations for each component is taken as the estimated background density.

Density estimates are obtained for both Moran’s $I$ and Geary’s $c$ statistics by identifying the minimum and maximum of the curves and their derivatives. This results in six estimates for each scenario as both the minimum and maximum density along the derivative curves are selected for the analysis. The derivative curves are utilized in order to better identify densities along the curve corresponding to changes in the overall autocorrelation of each component map.
Figure 2.11: Observation height with locations denoted in white for (a) North trending lines with 370 observation locations (b) East trending lines with 370 observation locations and (c) North-east trending lines with 781 observation locations.
Averaging the density estimates across components for all cases indicates that the distribution is right skew (Figure 2.12, upper left plot). The average component density estimates for each dyke location are shown in the other three histograms of Figure 2.12. It is clear that an average of all component density estimates may not provide the optimal density value for the difficult case where topography is coincident with the geology of interest.

![Histograms of density estimates for the average of trace component estimates only for four groupings: (upper left) density estimates for all permutation; (upper right) density estimates for permutations on dyke beneath the hill; (lower left) density estimates for permutations on dyke within the hill; (lower right) density estimates for permutations on dyke offset from the hill. The red line indicates the true background density value.](image)

Figure 2.12: Histograms of density estimates for the average of trace component estimates only for four groupings: (upper left) density estimates for all permutation; (upper right) density estimates for permutations on dyke beneath the hill; (lower left) density estimates for permutations on dyke within the hill; (lower right) density estimates for permutations on dyke offset from the hill. The red line indicates the true background density value.

For this reason, examining the ability of individual components to estimate the optimal density is the subject of the next section.

### 2.5.1 Component Density Estimates

To further understand the ability of individual components to estimate the optimal background density, the influence various parameters have on estimating the optimal density
value is simultaneously examined. In addition to the location of the dyke and direction of
the survey, the maximum distance used to generate the weights ($W_{ij}$) for the statistics is
examined. While selecting the density associated with the minimum and maximum of the $I$
and $c$ curves has proven useful, it does not provide insight into the densities at which spatial
autocorrelation is changing. For this reason, density estimates based on the derivative curves
will also be used.

To summarize, the variation of density estimates are based on permuting the:

- Direction of flight line (east
  - East-West
  - North-South
  - North-East

- Location of dyke
  - Beneath
  - Within
  - Offset

- Spatial Statistic
  - Moran’s $I$
  - Geary’s $c$

- Statistic Interpretation
  - Minimum of $I$ curve
  - Maximum of $c$ curve
  - Minimum of $I$ derivative curve
  - Minimum of $c$ derivative curve
- Maximum of $I$ derivative curve
- Maximum of $c$ derivative curve

- Weights distance
  - 0m-40m
  - 39-71m
  - 70m-101m
  - 100m-140m
  - 139m-181m

The estimated density for each component is plotted against the average density value for all six components. The color denotes the attribute permutations. In total, there are 540 possible permutations based on these attributes.

Histograms of the 540 density estimates for each component are given in Figure 2.13. From the histograms, the performance of each component is readily obvious. Overestimation of the actual density value of 2.54 g/cc (red line) occurs for all components. For $T_{xz}$ and $T_{yz}$, all density estimates are higher than the actual terrain correction density value. $T_{zz}$ estimates tend toward the actual density value.

In all plots that follow, the estimated density for each component is plotted against the average estimated density across all components for all parameter permutations previously summarized. The actual density value of 2.54 g/cc is demarcated on both axis.

The density estimates colored according to line direction are plotted in Figure 2.14. It can be seen that based on line direction, no single component outperforms any other component in estimating the actual density value.

The density estimates colored according to the location of the dyke relative to the north trending hill are plotted in Figure 2.15. The actual density value of 2.54 g/cc is demarcated
Figure 2.13: Histogram of density estimates for each component. The red line indicates the true background density value.
Figure 2.14: Density estimates for each individual component plotted versus average density for all six components. The colors indicate line direction where black dots are estimates from east-west lines, red dots are estimates from north-south lines, and green dots are north-west trending flight lines.
on both axis. Again, it can be seen that based on dyke orientation, no single component outperforms any other component in estimating the actual density value.

Figure 2.15: Density estimates for each individual component plotted versus average density for all six components. The colors indicate the location of the dyke with respect to the north trending hill. Blue dots are estimates where the dyke is directly beneath the hill, cyan dots are estimates where the dyke is within the hill, and pink dots are estimates where the dyke is offset to the east of the hill.

The density estimates colored according to the spatial statistic used are plotted in Figure 2.16. Blue indicates Moran’s $I$ and red indicates Geary’s $c$.

The density estimates colored according to the spatial statistic used are plotted in Figure 2.17. Examining the type of interpretation used leads to some interesting observations.
Figure 2.16: Density estimates for each individual component plotted versus average density for all six components. The colors indicate the spatial statistic used. Blue indicates Moran’s $I$ and red indicates Geary’s $c$. 
In general, the derivative interpretations are more self-consistent across all components. Using the minimum of Moran’s $I$ (green dots) and maximum of Geary’s $c$ (pink dots) over-estimates the background density for all components.

Figure 2.17: Density estimates for each individual component plotted versus average density for all six components. The colors indicate type of interpretation of the spatial statistic. Along the estimated density curves: green indicates minimum of the Moran’s $I$ curve, pink indicates maximum of the Geary’s $c$ curve, black indicates the minimum of the derivative of $I$ curve, gray indicates the maximum of the derivative of $I$ curve, yellow indicates the minimum of the derivative of $c$ curve, cyan indicates the maximum of the $c$ curve.

The density estimates colored according to the range interval used to calculate the inverse distance weighting matrix for the statistic estimates are plotted in Figure 2.18. That is, only observation location pairs falling within the specified separation distance are used in the density estimation process.
Observation locations having a euclidean distance separation between 100m and 140m (blue dots) seem to overestimate density values for all components, though there are a handful of underestimates. Observation location separation from 70m to 101m (black dots) and 139m to 181m (cyan dots) result in density estimates that span the entire range. The short separation range of 0m - 40m (gray dots) and 39m-71m (yellow dots) appear to provide the lowest density estimates across all components. Overall, there does not seem to be a single distance range that consistently produces the true density.

Figure 2.18: Density estimates for each individual component plotted versus average density for all six components. The colors indicate the distance interval used to generate the weighting matrix for the calculation of Moran’s $I$ and Geary’s $c$. 0m-40m are gray, 39m-71m are yellow, 70m-101m are black, 100m-140m are blue, 139m-181m are cyan.
2.5.2 Synthesis

Based on the observation of Interpretation Type, the density estimates corresponding to the derivative of the $I$ and $c$ curves provide density estimates in the lower density range. In Figure 2.19, the density estimates from the derivative curves are plotted according to statistic ($I$ or $c$) with all other density estimates plotted with gray crosses. From this, it becomes readily apparent that interpretation on the derivative curves provides more consistent density estimates for individual components.

![Figure 2.19: Density estimates for each individual component plotted versus average density for all six components. All density estimates are plotted with gray crosses. The colored dots indicate the spatial statistic where blue indicates Moran’s $I$ and red indicates Geary’s $c$.](image-url)
Narrowing down the range of interest, Figure 2.20 shows only estimates with a mean less than 2.75 g/cc across all components colored by location of the dyke. Estimates from the dyke beneath the hill are blue, within the hill are cyan, and offset from the hill are pink. From this plot, it is clear that the average density estimates closest to the true density are those corresponding to the scenario where the dyke is offset from the hill. However, this statement is not true when examining individual components.

Figure 2.20: Density estimates for each individual component plotted versus average density for all six components. Only estimates with a mean across all components less than 2.75 g/cc are shown. The colored dots indicate the spatial statistic where estimates from the dyke beneath the hill are blue, within the hill are cyan, and offset from the hill are pink.

Having previously observed that $T_{xz}$ and $T_{yz}$ tend to overestimate the true density, in the following discussion the focus is on the other four components. When removing the
scenarios where the dyke is offset from the hill, $T_{xy}$ consistently underestimates the true density while $T_{zz}$ consistently overestimates the true density. Histograms of the density estimates corresponding to the two scenarios (beneath and within the hill) are given in Figure 2.21 for each component. It can be seen that the average $T_{yy}$ estimate coincides with the true density. Likewise, $T_{xx}$ tends to identify the proper density range for these difficult scenarios as well.

![Histograms of density estimates](image)

Figure 2.21: Histogram of density estimates for each component for scenarios where the dyke either beneath or within the hill and the average density across all components is less than 2.75 g/cc. The red line indicates the true background density value.

Taking these observations into consideration, histograms of the average density based on estimates for the trace components only are given in Figure 2.22 (for all attribute permu-
While the histograms exhibit a bimodal distribution, in each case the average of the trace component estimates the true density (red line) better than the average density estimate across all components (Figure 2.12).

![Histograms of density estimates for the average of trace component estimates only for four groupings](image)

Figure 2.22: Histograms of density estimates for the average of trace component estimates only for four groupings: (upper left) density estimates for all permutations; (upper right) density estimates for permutations on dyke beneath the hill; (lower left) density estimates for permutations on dyke within the hill; (lower right) density estimates for permutations on dyke offset from the hill. The red line indicates the true background density value.

Overall, when broken down by component, no single attribute examined appears to be controlling the estimated density across all permutations. $T_{zz}$ density estimates tend towards the true density in all cases. The mid-range of 100m-140m for weights appears to overestimate the density, while both longer and smaller distance weights outperform this single
distance range. Interpretation on the derivative curves provide consistent density estimates
for individual components. For cases where the dyke is beneath or within the hill, $T_{xx}$ and
$T_{yy}$ obtain density estimates closest to the true density estimates. An average of only the
trace components has a higher frequency of estimating the true density value.

2.6 Conclusions

The density used in correcting the terrain or bathymetric effect in gravity gradient data is
an important yet neglected problem. Here, a statistical approach to estimating the density
to be used was presented. The methodology capitalizes on the spatial dependence of the
data. Statistical measures called Moran’s $I$ and Geary’s $c$ were used to evaluate the auto-
correlation of the anomalous data. The local measures of Moran’s $I$ and Geary’s $c$ do not
provide a direct means of estimating the density value, but are useful in qualitatively viewing
and understanding the spatial autocorrelation present in the data. A globally proportional
measure of the Moran’s $I$ and Geary’s $c$ statistics were determined to be a possible way of
estimating the representative density value. A representative density value is obtained when
Moran’s $I$ achieves a low when plotted as a function of density. Similarly, Geary’s $c$ achieves
a high for the representative density value.

The general applicability of the technique was explored using two synthetic examples
where the data was contaminated with various levels of Gaussian noise. For the four source
synthetic example, using both Moran’s $I$ and Geary’s $c$ statistics the representative density
value was estimated to be 3.19 g/cc or 3.20 g/cc, respectively. These estimates are close to
the true value of the 3.21 g/cc. For the more difficult case of a coincident topographic hill
and dipping dyke, it is seen that some components are more suited for this method.

When this technique was applied to a field dataset with coincident topography, the
technique estimates an average density of 2.84 g/cc for the survey area. Based on the
results of the dipping dyke synthetic example, a density of 2.84 g/cc for the field dataset was
estimated using the $I$ and $c$ curves rather than the derivative curves so may be overestimating
the optimal density value.
CHAPTER 3
DENOISING AND PROCESSING OF GRAVITY GRADIENT DATA USING AN EQUIVALENT SOURCE TECHNIQUE

The inherent relationship among different components of gravity gradiometer data requires applied processing to be consistent among the components. This restricts the applicability of some traditional potential-field processing techniques, but highlights opportunities for methods uniquely suited for such data sets. The equivalent source technique is one such method. We apply fast equivalent source construction to three aspects of gravity gradient data processing. First, multi-component data are denoised and estimates of the incoherent errors are obtained for the observations. Second, we illustrate the ability to reliably convert observed data to alternative components or perform regional residual separation. Lastly, we propose a method that can be used to estimate errors associated with the denoised or converted component data. Through both synthetic and field examples, the effectiveness of equivalent source processing for denoising and noise estimation will be illustrated.

3.1 Introduction

Although there are five independent components in the tensor at each point, it is important to note that they are all linearly related to the underlying gravitational potential. Given a complete map of the potential in a horizontal plane, the corresponding gravity gradient tensor is uniquely defined by a set of linear transformations (e.g., 34). This emphasizes the linear dependency among different components, and highlights two important aspects of gradient data processing. First, the five independent components are measuring a common underlying signal. Therefore, we should be able to use the inherent relationship among them to extract the coherent signal by removing noise that do not obey this relationship. Secondly, it requires that any processing applied to the data must be consistent among components. Otherwise, the resultant data are no longer meaningful. It follows that we must
use a “common model” related to the underlying gravitational potential for the purpose of data processing.

A natural choice of the common model is an equivalent density layer that satisfies all components of the observed gradient data. The equivalent source technique [35] is a well-established method that can be used to process gravity gradient and other potential-field data. The technique is particularly suited for multi-component data as it can reproduce any combination of components simultaneously using a common equivalent layer, thus maintaining the inherent relationship and signal content across multiple components. The constructed equivalent layer can then be used for a variety of data processing purposes. In the potential-field community, equivalent source processing has been used to interpolate irregularly spaced data [36, 37], calculate the field on alternative observation surfaces [38, 39], filter regional trends [40], and to merge multiple datasets [41, 42].

In what follows, we build upon previous work and illustrate the effectiveness of equivalent source to denoise data and provide a mechanism with which to estimate data errors. Estimating data errors in gravity gradient data has been discussed by Lane [41] and Boggs and Dransfield [42]. For the purpose of denoising gravity gradient data, Lyrio et al. [43] applies a wavelet-based method and While et al. [44] presented an approach that analyzes the power spectra and inherent relations. Pajot et al. [45] utilize gravity and gravity gradient observations together and apply a least-squares inversion to suppress data error. More recently, Yuan et al. [46] exploit the Laplace equation and apply linear inversion to remove incoherent signal using Fourier series. The latter two techniques assume that the observation locations lie within a horizontal plane and are regularly gridded, requiring additional processing prior to data denoising. Along the lines of Barnes and Lumley [47], we utilize an equivalent source technique to estimate errors in observed data.

While many have proposed methods with which to denoise multi-component data, few have explored the error levels present in the denoised data itself. It should be noted that the clean component data derived from any denoising technique are not strictly error free, nor
do they have the same noise characteristics as the observed data. The last section of this work seeks to address the gap in understanding errors in denoised data.

We first outline the procedure for constructing an equivalent density layer from multi-component gradient data. We then apply the technique to estimating and removing noise as well as a regional trend from the Leghorn marine field dataset. We then illustrate the ability of the equivalent layer to be used for converting observed components to any alternative component. Lastly, we propose a method to estimate the level of errors present in the denoised data produced by equivalent source processing and illustrate through both a synthetic example and a field dataset from McFaulds Lake, Ontario Canada.

\section*{3.2 Equivalent source construction}

We begin by describing the equivalent source construction. Assume that the observed gradient data are located at irregularly spaced points in 3D space. Framing the equivalent source construction with this assumption enables the algorithm to be practical and generally applicable. Let the observed data be denoted by

\[
\vec{d} = (T_{1xx}, T_{1xy}, T_{1xz}, T_{1yy}, T_{1yz}, \ldots, T_{pxx}, T_{psy}, T_{pxz}, T_{pyy}, T_{pyz})^T
\]

where \(p\) is the number of observation locations. The actual number of data \(N\) is equal to \(p \ast n_c\) where \(n_c\) is the number of components to be processed simultaneously.

A variety of equivalent sources can be formulated, but the most natural choice for our problem is a layer with density varying laterally and placed at some distance below the observation surface. We discretize the density layer into a set of contiguous vertical prisms and assume each prism has a constant, but unknown, density value. Choosing a suitable ordering, the density values of the \(M\) prisms can be collected into a model vector:

\[
\vec{\rho} = (\rho_1, \ldots, \rho_M)^T
\]
This vector of density contrasts values, together with the depth, thickness, and horizontal sizes define the equivalent source layer.

Since the gravity field and its derivatives are linearly related to the density value of a body with given geometry, the gradient data produced by the equivalent source are related to the density values by a linear system of equations

\[ \vec{d} = G \vec{\rho} \] (3.3)

where \( G \) is the coefficient matrix, whose elements \( g_{ij} \) define the contribution of the \( j \)'th cell density value to the \( i \)'th datum. The datum can be any one of the observed components at a given observation location.

With a suitably chosen equivalent-source layer, \( G \) can be a square and non-singular matrix. Equation 3.3 can then be inverted directly to construct the equivalent source layer. However, this is not advisable for two reasons. First, it would be difficult to design such a source layer without complicated variation in the horizontal sizes of the prisms. Second, we do not desire to reproduce the observations exactly for there is invariably noise in the data. Instead, we aim to find an equivalent source that has certain regularity and is consistent with the coherent signal in all components in a least-squares sense. Imposition of the regularity avoids overfitting the data by restricting the complexity of the source layer and allows the estimation of data errors based on established numerical methods in geophysical inverse theory. We will return to these methods and how they can be used as a mechanism to estimate data noise, but first describe the implementation of equivalent source construction as a regularized inverse problem.

Since we are attempting to estimate the errors and have no knowledge about the magnitude of noise, we define the data misfit using a simple \( L_2 \) measure:

\[
\phi_d(\vec{d}) = \sum_{i=1}^{N} (d_i^o - d_i)\]

\[= \left( G\vec{\rho} - \vec{d}^o \right)^T \left( G\vec{\rho} - \vec{d}^o \right) \] (3.4)
Here \( d^p_i \) represent the data predicted from the equivalent source and \( d^o_i \) are the observed data. Next, we define a model regularity measure of the equivalent source layer. The simplest measure is the flatness

\[
\phi(\rho) = \int \int \left( \frac{\partial \rho}{\partial x} \right)^2 dxdy + \int \int \left( \frac{\partial \rho}{\partial y} \right)^2 dxdy
\]

(3.5)

Evaluating this functional using a finite difference approximation over the source mesh yields its discrete representation

\[
\phi_m(\rho) = \tilde{\rho}^T \left( D_x^T D_x + D_y^T D_y \right) \tilde{\rho} = \tilde{\rho}^T W^T W \tilde{\rho}
\]

(3.6)

where \( D_x \) and \( D_y \) are the finite difference operators in the x- and y-directions respectively, and they collectively define the model weighting matrix, \( W \).

The equivalent source layer is then constructed by solving the minimization problem

\[
\phi = \phi_d + \mu \phi_m
\]

(3.7)

where \( \mu \) is a trade-off parameter that determines how well the data are reproduced and how complex the constructed equivalent source is.

The solution of the above minimization problem is given by the following system of equations:

\[
(G^T G + \mu W^T W) \tilde{\rho} = G^T \tilde{d}^o
\]

(3.8)

For a typical problem in gradiometer data processing, this can be a system involving a large number of unknowns and is usually not feasible to solve using a direct approach. Instead, we use an iterative solver such as the conjugate gradient (CG) technique. With an iterative solver, we do not need to form the matrix \( G^T G \) explicitly; rather, it is applied implicitly to a vector by multiplying \( G \) and then \( G^T \) to the vector. For large datasets and fine discretization of the equivalent layer, computational cost becomes increasingly important. To decrease the computational cost and time associated with constructing an equivalent
layer, wavelet compression can be applied to the regularized method described [39] or a combination of adaptive mesh and wavelet compression [48].

Once the problem is formulated in this manner, the most important parameter for an optimal solution of equivalent source is the regularization parameter $\mu$. As $\mu \to 0$, the observed data are fit as well as possible; while as $\mu \to \infty$, the data are not reproduced at all. The optimal solution is obtained when the model produces a misfit value, defined in equation 3.4, equal to its expected value. When this is achieved, the solution is the most consistent with the signal in the noisy data. The ideal value of $\mu$ would misfit the data to a level that is consistent with the noise content. Therefore, the problem of estimating the noise becomes one of estimating the optimal value of $\mu$. However, in our present problem, the noise standard deviation is the quantity we are attempting to estimate. Therefore, we apply the inverse formulation in a slightly different manner by first finding an optimal solution and then estimating the noise standard deviation from the difference between the observed and predicted data. There are a number of approaches for estimating an optimal value of the regularization parameter in a linear inverse problem without knowing the actual noise level in the data. We apply two commonly used approaches and examine their performance by estimating the noise level in gravity gradiometer data. The first is the statistical method of Generalized Cross-validation (GCV) [49] and the second is a more heuristic approach of L-curve criterion [50].

3.2.1 Generalized cross-validation (GCV) criterion

GCV is a model-based method that was originally developed in statistical literature for interpolating noisy observational data [49, 51]. Since the observed data are contaminated with noise, any interpolation should not fit each data point exactly. Instead, the interpolated data will be a smoothed version that misfits the observation by an amount that is consistent with the magnitude of the noise. The generalized cross-validation is a technique used to determine that appropriate misfit. The basic premise of the method is that the underlying signal in the data has certain cohesion while the noise is random, and that good interpolation
models derived with and without using a particular datum point should predict that datum consistently. The consistency is measured by the “leave-one-out” principle.

Let $\vec{d} = (d_1, d_2, \ldots, d_N)^T$ be the observed data. Without referring to the details of the interpolation model, let $\hat{d}_i$ be the value of the $i$’th datum predicted by the model derived from the data after leaving $d_i$ out of the model construction. The leave-one-out principle states that the total prediction error should be at a minimum when the optimal parameters are used in defining the model. That is, the optimal model should minimize the following quantity:

$$CV = \sum_{i=1}^{N} \left( d_i - \hat{d}_i \right)^2$$  \hspace{1cm} (3.9)

This is a general principle and applies to any interpolation method. We use CV to denote the quantity since it is the quantity to be minimized in the ordinary cross-validation.

The regularized inversion outlined in the preceding section can be viewed as an interpolation method that is defined by the equivalent source and the degree of smoothing is controlled by the regularization parameter. Since the controlling parameter is $\mu$, the problem becomes one of finding the value of $\mu$ that yields a minimum CV. Expressing $\hat{d}_i$ as a function of the data $(d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_N)^T$ and the regularization parameter and carrying out required algebra, a simple expression can be derived for CV as a function of $\mu$. However, a more robust measure is given by a rotation invariant representation of the CV function known as the generalized cross-validation (GCV), which is given by the following expression [49, 51] for our current problem where $\text{Trace}[]$ denotes the trace of a matrix.

$$GCV(\mu) = \frac{\| [W_d G (G^T W_d^T W_d G + \mu W^T W)^{-1} G^T W_d^T - 1] W_d \vec{d} \|^2}{\{ N - \text{Trace} [W_d G (G^T W_d^T W_d G + \mu W^T W)^{-1} G^T W_d^T] \}^2}$$ \hspace{1cm} (3.10)

The search for the optimal value of $\mu$ entails a univariant minimization and the GCV function must be evaluated many times. While the GCV function appears to be complicated and difficult to evaluate, closer inspection shows that the numerator is the data misfit function in the inversion and can be computed by performing a series of trial inversions with
different regularization parameter values. However, the more challenging task is to evaluate the trace of the matrix in the denominator. Straightforward implementation would require computing the matrix inverse \((G^TW_d^TW_dG + \mu W^TW)^{-1}\) explicitly. This is clearly impractical for the same reason that we cannot invert Equation 3.8 directly. Instead, we employ a statistical trace estimator developed by Hutchinson [52], which states that the trace of a square matrix, \(A\), is approximated by the quadratic function

\[
\text{Trace}(A) = \tilde{u}^T A \tilde{u} \tag{3.11}
\]

where \(\tilde{u}\) is a vector containing -1 and 1, each with a probability of 0.5. Thus the trace term in the denominator is approximated by

\[
\text{Trace} \left[ W_dG(G^TW_d^TW_dG + \mu W^TW)^{-1}G^TW_d^T \right] \\
\approx \tilde{u}^T W_dG(G^TW_d^TW_dG + \mu W^TW)^{-1}G^TW_d^T \tilde{u} \tag{3.12}
\]

Computation can be carried out in three steps:

1. Perform an “inversion” on the random data vector \(\tilde{u}\)
2. Apply forward modeling to the resultant “model” to compute the predicted random data vector
3. Calculate the inner product of the original and predicted random data vectors

Therefore, the GCV calculation for each value of \(\mu\) is equivalent to performing two inversions: the first inverts the observed data that are to be smoothed, and the second inverts a random data vector \(\tilde{u}\). The GCV in our algorithm is formally expressed as [20]:

\[
\text{GCV}(\mu) = \left\| \left[ W_dG(G^TW_d^TW_dG + \mu W^TW)^{-1}G^TW_d^T \right] W_d \tilde{d} \right\|^2 \\
\{ N - \tilde{u}^T W_dG(G^TW_d^TW_dG + \mu W^TW)^{-1}G^TW_d^T \tilde{u} \}^2 \tag{3.13}
\]

Using the statistical trace estimator, the implementation of GCV calculation requires no more than the basic algorithm for the equivalent source construction. Utilizing the efficient solver based upon conjugate gradient technique and fast matrix-vector multiplication, the
GCV function can be easily evaluated and its minimum identified. The corresponding regularization parameter yields an optimal equivalent source and the predicted data are the denoised data with which we can estimate data noise.

### 3.2.2 L-curve criterion

The L-curve criterion is a heuristic method first developed for least-squares problems to treat the effect of noise on the solution by estimating an optimal regularization parameter [53]. It was observed that (see Figure 3.2(b) for example) when plotted on a log-log scale, the norm of the least-squares solution as a function of the data misfit, i.e., the Tikhonov curve, exhibits a characteristic corner. As the degree of regularization decreases towards this corner point, the model norm changes very little while the misfit is being reduced. Further decrease in the degree of regularization beyond this point would result in rapid increase in the model norm with little reduction in the data misfit. As a result the curve appears to be L-shaped, and was named “L-curve” by Hansen. Hansen [50] provides a detailed analysis of the method and shows mathematically the occurrence of the corner point. Intuitively, the method can be understood by examining the behavior of the data misfit and model objective function in a least-squares problem, or more generally, in a linear inverse problem with a given regularization method.

When strong regularization is applied, the effect of noise in the data is filtered out and the least-squares solution can easily reproduce portions of the signal with little extra structure added to the model. This continues until the regularization is decreased to the point where the effect of the noise begins to affect the solution and the further decrease in misfit is achieved by reproducing the noise in the data. In order to reproduce the noise in the entire data set, there has to be a great deal of structure introduced into the model. The result is a rapid increase of model norm for small misfit reduction. The corner point represents the onset of the rapid increase in model objective function, and it therefore represents the point where the optimum misfit is achieved.
The L-curve criterion, therefore, states that the optimal solution for a given regularized inverse problem is the one that corresponds to the corner point on the Tikhonov curve. At this point, we have extracted the maximum amount of signal from the data while the equivalent source is still minimally affected by the noise in the data.

Numerically, the corner is given by the point of maximum curvature of the Tikhonov curve on a log-log plot. Both the model objective function $\phi_m$ and data misfit $\phi_d$ are functions of the regularization parameter $\mu$. Thus we have an implicit functional relation between $\phi_m$ and $\phi_d$ that is parameterized by $\mu$. Letting $\hat{\phi}_m = ln(\phi_m)$ and $\hat{\phi}_d = ln(\phi_d)$, the curvature is given by

$$\kappa(\mu) = \frac{\hat{\phi}_d \hat{\phi}_m'' - \hat{\phi}_m \hat{\phi}_d''}{\left[\left(\hat{\phi}_d'\right)^2 + \left(\hat{\phi}_m'\right)^2\right]^{3/2}}$$

(3.14)

where the prime denotes differentiation with respect to $\mu$. Application of L-curve criterion requires the solution of the equivalent source construction using a number of different regularization parameters in order to find the maximum-curvature point according to Equation 3.14. The regularization parameter corresponding to the point of maximum curvature yields the optimal equivalent source layer and an estimate of the data errors.

Equipped with these two methods to estimate the value of the regularization parameter for construction of an equivalent source layer and obtain error estimates, we next apply this method to synthetic and field data.

3.3 De-noising using equivalent source

As an illustration, we begin by applying equivalent source denoising to a synthetic data set and estimate data errors. A 3D model composed of four positive density contrast features is used to compute the data on an undulating observation surface. The data shown in Figure 3.1 are contaminated with zero mean gaussian noise and standard deviation of 0.85Eo. It can be seen that the level of noise overshadows the underlying signal. The data are calculated at 250m grid spacing over the area on the observation surface shown in the lower-left panel.
of Figure 3.1 resulting in 1085 observation locations. For this example, we calculate six-component data yielding $1085 \times 6 = 6510$ data. The equivalent layer used was discretized into 62.5m cubic cell sizes in both the northing and easting directions and extends beyond the data area approximately 6km in each direction. The equivalent layer has 160 cells in the easting and 144 cells in the northing. The equivalent source is placed at 2288m, which is the average elevation of the simulated topographic surface of the model used to generate the synthetic data.

Figure 3.1: Gravity gradient data for the synthetic problem contaminated by gaussian noise with zero mean and 0.85 $E_0$ standard deviation. The panels from left to right and top to bottom are $T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, $T_{yz}$, Observation height, and $T_{zz}$.

Figure 3.2(a) is the plot of GCV as a function of the regularization parameter while Figure 3.2(b) is the Tikhonov curve, which is the plot of the model objective function versus the data misfit on a log-log scale. The filled circle on each plot indicates the equivalent source corresponding to the GCV minimum. The corresponding value of $\mu$ from GCV is 14, which yields estimated standard deviations of $T_{xx} = 0.8158$, $T_{xy} = 0.8396$, $T_{xz} = 0.8028$, $T_{yy} = 0.8426$, $T_{yz} = 0.9020$, and $T_{zz} = 0.8617$. 

54
\[ T_{yy} = 0.8346, \quad T_{yz} = 0.8096, \quad T_{zz} = 0.7971. \] While the estimates are close to the true value of 0.85 \( E_0 \), we note that some level of coherent noise has been reproduced by the equivalent source. Figure 3.3 shows the predicted data from the equivalent source construction using the optimal regularization parameter according to GCV. This is the denoising result using an equivalent source. We can see that the denoised data exhibit few noisy features. Figure 3.4 shows the difference between the observed noisy data and the denoised data, which yields the noise estimates previously stated.

![Figure 3.2: (a) GCV curve as a function of the regularization parameter. The minimum GCV point defines the optimal regularization value. (b) The Tikhonov curve of the equivalent source construction for the synthetic data. The filled circle denotes the position of the GCV optimal solution on the Tikhonov curve and GCV curve.](image)

Figure 3.2: (a) GCV curve as a function of the regularization parameter. The minimum GCV point defines the optimal regularization value. (b) The Tikhonov curve of the equivalent source construction for the synthetic data. The filled circle denotes the position of the GCV optimal solution on the Tikhonov curve and GCV curve.

It is interesting to note the shape of the Tikhonov curve and the corresponding location of optimal solution on it. The Tikhonov curve has a pronounced corner, and it approximately coincides with the optimal solution (marked by the circle) according to the GCV criterion. This prompts us to explore the more heuristic L-curve criterion for estimating the noise.

We now apply the L-curve criterion to our problem of estimating the noise in the data shown in Figure 3.1. Using the same parameters to generate the equivalent source, the denoised data according to an L-curve criterion is shown in Figure 3.5. Note the Tikhonov curve in Figure 3.6(a) is the same as the Tikhonov curve shown with the GCV estimate. Figure 3.6(b) shows the curvature of the Tikhonov curve according to Equation 3.14. It has a
Figure 3.3: Denoised data produced by the optimal equivalent source layer according to GCV criterion. Note that all high frequency features have been removed.

well-defined positive peak flanked by two side lobes. This is a typical shape of the curvature curve. The maximum-curvature point yields a value of 10 as the optimal regularization parameter. The estimated noise standard deviations corresponding to this solution are $T_{xx} = 0.8125$, $T_{xy} = 0.8374$, $T_{xz} = 0.7959$, $T_{yy} = 0.8290$, $T_{yz} = 0.8023$, and $T_{zz} = 0.7843$. For practical purposes, these estimates are consistent with the GCV noise estimates.

It is important to note that the results from GCV and L-curve criteria are nearly identical and they both predict the magnitude of noise in the data very well. The consistency between the two results, and with the true noise level, lends confidence that either approach could work well in practical applications. However, L-curve criterion has two advantages over GCV in general. First, L-curve requires less than half the computation needed for GCV. Secondly, L-curve criterion has been demonstrated in the literature to be more robust in the presence of correlated noise [50]. We have also observed this difference when applying the L-curve criterion to other types of geophysical data. Given that the measured gradiometer data are
likely contaminated by some level of correlated noise, L-curve is likely to perform better in practice. Therefore, the L-curve approach is expected to be a more useful tool than GCV.

Having tested the algorithm on a number of field data sets available to us, we indeed empirically find that the L-curve criterion outperforms GCV for both single- and multi-component data. We have often observed that the GCV estimate of noise is smaller than that from L-curve criterion and data noise is noticeably controlling the construction of the equivalent source model. This is consistent with the observations about the two methods in the literature [50].

3.3.1 Field example: Leghorn, Gulf of Mexico

We now use the equivalent source technique to process data from the Leghorn repeat survey in the Gulf of Mexico. Five observed components are available in an area of approximately 25 km by 25 km. A total of 25 lines oriented east-west at 1 km spacing are used.
Figure 3.5: Denoised data using the L-curve criterion. The result is consistent with that produced by the GCV criterion.

Figure 3.6: (a) The Tikhonov curve of the equivalent source construction for the synthetic data. Note the L shape of the curve and its distinct corner. (b) The curvature of the Tikhonov curve as a function of the regularization parameter. The maximum curvature point defines the corner point of the Tikhonov curve.
The data are down-sampled along line to an interval of approximately 100 m, yielding 6058 data locations and 30290 observations. Figure 3.7 displays the five observed components \( T_{xx}, T_{xy}, T_{xz}, T_{yy}, T_{yz} \) as well as \( T_{zz} \) which was directly calculated from \( T_{xx} \) and \( T_{yy} \).

![Figure 3.7: Five-components of observed data from Leghorn survey area along with calculated \( T_{zz} \).](image)

The data exhibit underlying long wavelength features as well as significant high frequency content. Given that the bathymetry in the area ranges from -500 m to -2000 m, much of the high frequency component must be noise. Only extremely variable bathymetric relief or density contrast is capable of producing such high frequency data at the sea surface. This observation can be supported by examining the bathymetry correction.

Figure 3.8 shows a contour plot of the bathymetry data used to calculate the correction, with the \( T_{zz} \) component overlaid. The bathymetry correction is calculated over a 50 km by 50 km area assuming a constant density contrast of 1.0 g/cm\(^3\) between the sediments and water. The density is derived from a number of density logs in the area that indicates an average density of 2.0 g/cm\(^3\) at the mud line [54].
Figure 3.8: The contour lines show the bathymetry of the area surrounding the data area. The color overlay is the observed $T_{zz}$ component prior to bathymetry correction.
Figure 3.9: The bathymetry correction calculated assuming a density contrast of 1.0 g/cm$^3$ between the seafloor and water column.

Figure 3.9 shows the calculated bathymetry correction, illustrating smoothly varying components. Figure 3.10 is the bathymetry corrected data obtained by subtracting the field in Figure 3.9 from the observed data. A number of anomalies can be observed, but they are masked by the high frequency noise.

Next we apply the equivalent source technique to estimate noise and to generate a denoised version of the data. The equivalent source layer is composed of 200 m cubes and extends beyond the data area. The depth of the equivalent source in this example should be as deep as the highest bathymetry point. This is contrary to the usual parameterization of equivalent source, in which one would like to place the equivalent source as close to the observation points as possible in order to reproduce the high frequency content in the data. In our case, any high frequency component in the data that cannot be reproduced by a source layer on the sea floor is noise. For this reason, the layer is placed at a depth of 500 m below the sea surface. Figure 3.11 shows the denoised data using the L-curve criterion. The
Figure 3.10: The bathymetry corrected data. Note that the large-scale variations have largely disappeared from the data maps and they are dominated by small-scale variations. No clear signal is visible.
anomalies that were faintly visible in the bathymetry-corrected data are now clearly visible and the characteristic patterns of the anomaly in each data component are clearly shown. Figure 3.12 displays the estimated data error and noise present in the observed data. The standard deviation of the estimated noise for each component is $T_{xx} = 4.033$, $T_{xy} = 2.7372$, $T_{xz} = 3.1629$, $T_{yy} = 3.5788$, and $T_{yz} = 3.2260$. As expected, the data error is dominated by high-frequency variations. Looking at the denoised data, there appears to be a background trend underlying the major anomalies that are now clearly visible. As a next step, the equivalent source layer can be used to remove this long-wavelength information still present.

![Figure 3.11: Denoised data from equivalent source processing using L-curve criterion.](image)

3.3.2 Regional-residual separation

A commonly used approach for regional-residual separation is to directly estimate a regional field from the data itself. For example, a low degree polynomial approximation can be estimated from judiciously selected points in the data map. This method is easy and inexpensive to implement, however it has a number of drawbacks. First, the estimated regional
Figure 3.12: The difference between the bathymetry-corrected data and the denoised data obtained from equivalent source processing using an L-curve criterion. The difference map is representative of data errors and noise within the observed data.
field may not correspond to any field that can be produced by physically plausible source distributions. Removing such a non-physical regional field could distort the residual and negatively impact subsequent interpretation steps. Secondly, such a method is particularly unsuitable for multi-component data. Theoretically, the different gravity gradient components are linearly dependent when available over the entirety of an observational surface. However, if an estimated low-degree polynomial trend is applied to each component and a resultant regional field is removed independently, the consistency among the regional trends removed for different components is not guaranteed. The distortion introduced to the gradiometer data by such a simple regional removal can be especially severe. To avoid this problem, the regional field removed must be estimated consistently for all components. One way to ensure this condition is to calculate the regional field for each component from a single source distribution.

Both aforementioned problems can be treated effectively by using an equivalent source layer. Once an equivalent source layer is constructed, it can be separated into a smoothly varying “regional” part and a more rapidly varying “residual” part. Once this is done, a regional field can be calculated from the regional equivalent source. Since such a regional field corresponds to an actual source distribution, it is less likely to introduce distortions to the residual field, as would a purely “mathematical” regional field. When the gravity gradients are calculated from the same regional equivalent source, they provide a consistent set of regional trends for different components of the gradiometer data. When subtracted from the observed data, the resulting residual anomalies are likewise consistent.

Assuming an equivalent source layer has been constructed according to the steps outlined in the preceding sections, the remaining issue is how to separate the source into the two desired regional and residual sources. The simplest approach would be to estimate the regional source by performing a low-degree polynomial fitting to the equivalent source. The basic procedure is the same as the low-degree polynomial directly applied to observed data. We can select a number of points away from the areas of rapid change in the equivalent
source, and perform a least-squares fit. The resulting polynomial defines the regional source to be subtracted from the equivalent source. This yields the residual source and the residual anomalies are easily calculated from it.

Figure 3.13: The top panel shows the equivalent source recovered from the bathymetry-corrected data. The middle panel is the regional source estimated using a second-degree polynomial. The difference between the two defines the residual source from which the final residual anomaly will be calculated.

Returning to the Leghorn data, this approach is applied to remove the regional trend remaining in the denoised gravity gradient data. Figure 3.11 shows a number of anomalies
superimposed upon a slowly varying background. For example, the component $T_{xx}$ has a number of anomalies that displays the characteristic sequence of negative-positive-negative from south to north, which indicates an anomaly above a negative density contrast. However, the overall amplitude changes slowly in the north-south direction. The corresponding equivalent source layer shown in the top panel of Figure 3.13 displays a general trend of decreasing density contrast towards the northeast direction and there is a general negative bias. The gradual change and bias can be viewed as the regional component of the equivalent source. Performing a least squares fit to estimate a second-degree polynomial for the equivalent layer produces the regional source shown in the middle panel of Figure 3.13. After removing this regional source the resulting residual source, shown in the bottom panel of Figure 3.13, is now composed of several density anomalies in a generally zero background. The gradiometer data calculated from the residual source yields the desired residual data shown in Figure 3.14. The amplitude of the anomaly is now more balanced and the overall anomalous pattern is clearer.

3.3.3 Calculating converted components

Just as an equivalent source can be used to denoise data and remove regional trends, it can be used to calculate alternative components not measured. We refer to this as component conversion. Depending on the interpretation purpose and observed component data, it may be desirable to calculate various component combinations. Component conversion can be directly achieved utilizing the equivalent source distribution used to estimate data errors.

To illustrate, we use a gravity gradient data set from the McFaulds Lake Area in the region known as the Ring of Fire in Ontario, Canada. The flight line spacing for the survey is 250m at a heading of 135° with control line spacing of 2500m at 45°. The survey was draped with a target terrain clearance of 100m. Gravity gradient and magnetic data were simultaneously collected from January 26, 2011 to March 14, 2011. Here, we use a subset of the gravity gradient data from the McFaulds Lake survey shown in Figure 3.15. The data subset covers approximately 23.5km by 33.5km and has been downsampled to 60162
Figure 3.14: The residual gradiometer data extracted by removing a regional field from the equivalent source layer.

The equivalent layer consists of 100m cubes placed at a constant depth of 200m below the lowest topographic elevation. The layer has increasingly larger cell sizes outside of the data area. The optimal equivalent layer shown in Figure 3.16 was selected according to the L-curve criterion.

The selected equivalent source contains information from both the $T_{xy}$ and $T_{uv}$ components. Any alternative component or combination of components can be calculated using the generated equivalent layer. From the equivalent layer, $T_{zz}$ is calculated and displayed in Figure 3.17.

With the ability to denoise data and remove regional trends, qualitative interpretation can be readily performed on the denoised or converted data. Similarly, with error estimates for observed data quantitative interpretation of the observed components can be carried out. However, quantitative interpretation of denoised or converted components may require
Figure 3.15: Observed gravity gradient data from part of the McFaulds Lake Area dataset in the Ring of Fire. From left to right, the panels display Observation Height, $T_{xy}$ and $T_{uv}$.

Figure 3.16: The constructed equivalent source layer from observed $T_{xy}$ and $T_{uv}$.
additional information about the derived data and their error characteristics. The errors present in these data are not the same as the errors estimated to be in the observed data. For this reason, we must be able to estimate the error in denoised and converted components. In the next section, we discuss a method that can be used to estimate the error remaining in denoised data calculated from an equivalent source.

3.4 Estimating error level in equivalent source processed data

After an equivalent source has been used to remove incoherent noise from the gravity gradient data, we are left with what can be considered denoised gradient data. Generally, the errors associated with denoised data or converted components are not examined, though there are still errors present in this data. We clarify that while the calculation of data from the equivalent source is arguably computationally exact, propagation of observation error into the equivalent source must not be considered implausible. It is the noise constructed into the equivalent source that we refer to as the errors present in the denoised data. In the section that follows, we outline two methods to estimate the level of error present within the denoised data.

Figure 3.17: $T_{zz}$ calculated from the equivalent layer generated from the observed $T_{xy}$ and $T_{uv}$ components for the McFaulds Lake data.
Rearranging equation 3.8, the denoised data are represented by

\[
\vec{d}^d = G \left( G^T G + \mu W^T W \right)^{-1} G^T \vec{d}^o
\]  

(3.15)

The noise remaining in the denoised data can be characterized through the covariance matrix, \( C_d \), of the denoised data. In order to obtain an estimate of \( C_d \), we utilize the same inverse operator that generated the denoised data and associated error estimates obtained from the denoising process. The denoised data is a linear combination of the observed data when defining the coefficient matrix \( L \) as in equation 3.16.

\[
L = G \left( G^T G + \mu W^T W \right)^{-1} G^T
\]  

(3.16)

Given \( L \), the observed and denoised data are then related through \( \vec{d}^d = L \vec{d}^o \). Using expectation rules for random variables, the covariance matrix of the denoised data can be expressed according to equation 3.17.

\[
C_d = LC_oL^T
\]  

(3.17)

Assuming a constant variance for each observed data component, the covariance matrix \( C_o \) is simply diagonal.

Equation 3.17 provides a theoretical description of the errors in the denoised data and can be used in practice if the problem size is sufficiently small. However, modern equivalent source construction rarely uses a direct solution, and the construction of \( L \) can be prohibitively expensive. In practical applications, direct calculation of \( C_d \) is inefficient for large datasets due to the construction of \( L \).

Alternatively, a second approach to obtaining error estimates for the denoised data can be accomplished through statistical simulation. Having previously estimated observed data errors, synthetic data realizations can be generated using the denoised data to characterize \( C_d \). By applying the same denoising operation to the data realizations, an indirect estimate of the data errors can be obtained by indirectly characterizing \( L \). The simulation technique is summarized below.
1. Construct an equivalent source layer (\(\vec{\rho}^o\)) that optimally reproduces the coherent signal in the observed data (\(\vec{d}^o\)) as described in previous sections.

2. From layer \(\vec{\rho}^o\), estimate the standard deviation of noise, \(\sigma^o\), present in the observed data \(\vec{d}^o\) and calculate the denoised data \(\vec{d}^d\).

3. Generate \(k\) sets of gaussian random noise (\(\vec{n}_k\)) realizations with zero mean and standard deviation \(\hat{\sigma}^o\).

4. Generate \(k\) sets of synthetic observations using the denoised data \(\vec{d}^d\) according to \(\vec{d}_k = \vec{d}^d + \vec{n}_k\) where \(\vec{n}_k\) is a random noise vector.

5. Generate \(k\) equivalent layers and the associated denoised data, \(\vec{d}^d_k\), from the \(\vec{d}_k\) simulated data utilizing the same inverse construction that produced \(\vec{\rho}^o\).

6. From the residual data (\(\vec{r}_k\)) of the original denoised data (\(\vec{d}^d\)) and the \(k\) simulated denoised data (\(\vec{d}^d_k\)), compute summary statistics of the error still present in the simulated denoised data.

The distribution of residual vectors from \(\vec{r}_k = \vec{d}^d - \vec{d}^d_k\) provide insight into the noise level present in the denoised data through the ability of the equivalent source construction to remove known levels of noise. The two approaches outlined are first applied to the synthetic problem previously introduced in order to illustrate the technique.

### 3.4.1 Synthetic example

Through the synthetic data example, we demonstrate how to obtain error estimates for denoised data through both (a) simulation and (b) construction of \(L\).

The accurate gravity gradient data from the synthetic problem is shown in Figure 3.18.

#### 3.4.1.1 Simulation

We begin by generating gaussian random noise with zero mean and standard deviation of 0.85 Eo and add it to the accurate data which is now referred to as the observed data for
this synthetic demonstration. The observed data are plotted in Figure 3.1 and have already been subjected to equivalent source denoising as described in the denoising section. An initial noise level of $\sigma = 1$ for all components is supplied for equivalent source construction. The optimal equivalent layer illustrated in Figure 3.19 was selected according to the L-curve criterion and fits the coherent signal in the observed data. The denoised data based on this equivalent source layer is shown in Figure 3.5. The noise estimates, $\hat{\sigma}_o$, for the observed data are reproduced in Table 3.1. The regularization parameter that provided the optimal equivalent layer was found to be $\mu = 10$.

The noise estimates ($\hat{\sigma}_o$) can be compared to the true standard deviation, $\sigma^o$ of the errors calculated from the accurate data of Figure 3.18. The estimates are consistent with the true levels for the six gradient components.

We next generate $k = 100$ random gaussian noise realizations with standard deviation $\hat{\sigma}_o$ previously estimated for each individual component. Using the denoised data of Figure 3.5
we generate $k$ synthetic datasets for simulation. The initial noise level for the $k$ synthetic datasets is $\sigma_k = 1$. A single equivalent source layer is then constructed for each of the $k$ datasets using the same regularization parameter ($\mu = 10$) and inverse operation that resulted in the equivalent layer of Figure 3.19. Using the $k$ data realizations and original denoised data, summary statistics can then be generated for the simulation. From the $k$ residual maps, we calculate the mean of $k$ standard deviations to characterize the noise level present in each synthetic dataset. These estimates for $\hat{\sigma}_k$ are given in Table Table 3.1 and represent the gaussian error level used to generate each synthetic realization. From a practical standpoint, we can see that they are consistent though there appears to be some error not removed by the equivalent source.

The true error level, $\sigma^d$, still present in the denoised data is calculated using the accurate data and given in Table Table 3.1. Next we calculate the level of noise present in the $k$ denoised datasets by comparing $\vec{d}^d$ (denoised data in Figure 3.5) to each of the $\vec{d}^d_k$. These estimates are shown in Table Table 3.1 as $\hat{\sigma}_k^d$. These noise estimates are taken to be representative of the noise level present in $\vec{d}^d$ produced from the original equivalent source denoising of $\vec{d}^o$. A histogram of the $k$ estimated standard deviations for $\hat{\sigma}_k^d$ are shown in Figure 3.20 for each component. By comparing these estimates with the true level of error.
in the denoised data, we can see that they are highly consistent.

![Histograms](image)

Figure 3.20: Histogram of noise estimates ($\hat{\sigma}^d_k$) for the $k$ simulated datasets. The average value for each component is summarized in Table 3.1.

### 3.4.1.2 Direct construction of $L$

To verify the simulation approach implemented above, we explicitly construct $L$ according to equation 3.16 for this synthetic problem. This allows for a more direct calculation of the covariance matrix for the denoised data as defined in equation 3.17. As described previously, the use of wavelet compression enables the use of equivalent source to be used in larger problems. For our synthetic example above, we use wavelet compression in the generation of the equivalent layer and maintain use of it in the direct calculation of $L$.

To construct $L$, we perform $N = p * n_c$ equivalent source constructions where $p$ is the number of data locations and $n_c$ is the number of components used. For each construction, the L-curve criterion selected parameters are used with the exception of the data vector.
The value of all gravity gradient observations is set to zero except for a single observed component at a single data location. We iterate the singularity through each observation datum. As in the simulation approach, the same inverse operation is applied to the impulse data as prescribed in equation 3.15. The resulting $d\vec{d}$ data is then the $i$'th row of $L$ where $i = 1, ..., N$.

With estimated observation data covariance, $C_o$, as previously described the denoised data covariance, $C_d$, can be calculated according to equation 3.17. From the covariance matrix, we assume constant variance and calculate the standard deviation for each data point from the diagonal elements of $C_d$. Summary statistics for the direct calculation are given in Table 3.1 as $\hat{\sigma}^d$.

Table 3.1: Noise estimates for the synthetic problem. True noise levels are given for the observed ($\sigma^o$) and denoised data ($\sigma^d$). Noise estimates from equivalent source denoising are given for the observed data ($\hat{\sigma}^o$). Mean noise estimates for the denoised data based on simulation are $\sigma^d_k$ while the direct calculation are give as $\hat{\sigma}^d$. Noise estimates $\hat{\sigma}^o_k$ are for the known level of gaussian error used to generate the synthetic data realizations.

<table>
<thead>
<tr>
<th></th>
<th>$T_{xx}$</th>
<th>$T_{xy}$</th>
<th>$T_{xz}$</th>
<th>$T_{yy}$</th>
<th>$T_{yz}$</th>
<th>$T_{zz}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True Noise</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^o$</td>
<td>0.8358</td>
<td>0.8535</td>
<td>0.8295</td>
<td>0.8560</td>
<td>0.8297</td>
<td>0.8383</td>
</tr>
<tr>
<td>$\sigma^d$</td>
<td>0.1626</td>
<td>0.1023</td>
<td>0.1763</td>
<td>0.1466</td>
<td>0.1742</td>
<td>0.2157</td>
</tr>
<tr>
<td><strong>Estimates from Observed Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^o$</td>
<td>0.8123</td>
<td>0.8373</td>
<td>0.7959</td>
<td>0.8289</td>
<td>0.7996</td>
<td>0.7840</td>
</tr>
<tr>
<td><strong>Estimates from Simulation on Denoised Data (mean)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^o_k$</td>
<td>0.7921</td>
<td>0.8259</td>
<td>0.7697</td>
<td>0.8060</td>
<td>0.7726</td>
<td>0.7350</td>
</tr>
<tr>
<td>$\sigma^d_k$</td>
<td>0.1537</td>
<td>0.1059</td>
<td>0.1694</td>
<td>0.1530</td>
<td>0.1698</td>
<td>0.2270</td>
</tr>
<tr>
<td><strong>Direct Calculation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^d$</td>
<td>0.1683</td>
<td>0.1147</td>
<td>0.1865</td>
<td>0.1680</td>
<td>0.1857</td>
<td>0.2499</td>
</tr>
</tbody>
</table>

From the comparison of $\sigma^d$ and $\hat{\sigma}^d_k$, we see that the values obtained from simulation are representative of the true error levels present in the denoised data and consistent with the direct calculation approach. As direct construction of $L$ is not feasible in practice, we now demonstrate the simulation method on field data and estimate the level of noise present in
the equivalent source processed data.

### 3.4.2 Field Example: McFaulds Lake, Ring of Fire

We use the McFaulds Lake Area dataset previously described in the component conversion section. From the equivalent source layer (Figure 3.16), we calculate the denoised data displayed in Figure 3.21. Noise estimates for the observed data, \( \hat{\sigma}_o \), based on these denoised data are given in Table 3.2. Next, we generate 100 synthetic datasets using the denoised data in Figure 3.21 and random gaussian noise with zero mean and standard deviation of 5.2 Eo for both components. Using the same parameters that constructed the equivalent layer in Figure 3.16, we generate equivalent layers for the 100 data realizations.

From the equivalent layers and associated predicted data, we calculate residuals using the 100 data realizations. In Table 3.2, the mean value of the 100 noise estimates, \( \hat{\sigma}_k \), agrees with the level of gaussian noise contamination originally added. Next, we calculate residuals using the denoised data of Figure 3.21 and the 100 denoised data realizations. For each of the 100 residual maps, an estimate of the error standard deviation is obtained. The noise level estimated to be in the denoised data from this simulation is summarized by the mean of the resulting distribution of 100 standard deviations. These mean values are given as \( \hat{\sigma}_d^k \) in Table 3.2. The estimated 1 Eo error is an average over the dataset for the

![Figure 3.21: Denoised gravity gradient data from part of the McFaulds Lake Area dataset. From left to right, the panels display \( T_{xy} \) and \( T_{uv} \).](image)

From the equivalent layers and associated predicted data, we calculate residuals using the 100 data realizations. In Table 3.2, the mean value of the 100 noise estimates, \( \hat{\sigma}_k \), agrees with the level of gaussian noise contamination originally added. Next, we calculate residuals using the denoised data of Figure 3.21 and the 100 denoised data realizations. For each of the 100 residual maps, an estimate of the error standard deviation is obtained. The noise level estimated to be in the denoised data from this simulation is summarized by the mean of the resulting distribution of 100 standard deviations. These mean values are given as \( \hat{\sigma}_d^k \) in Table 3.2. The estimated 1 Eo error is an average over the dataset for the
denoised components.

Table 3.2: Noise estimates for McFauld’s Lake subset based on equivalent source construction and simulation.

<table>
<thead>
<tr>
<th></th>
<th>$T_{xy}$</th>
<th>$T_{uv}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_o$</td>
<td>5.1239</td>
<td>5.2666</td>
</tr>
<tr>
<td>Simulated Data (mean)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_k$</td>
<td>5.2101</td>
<td>5.2009</td>
</tr>
<tr>
<td>$\sigma_d^k$</td>
<td>1.0863</td>
<td>1.0251</td>
</tr>
</tbody>
</table>

3.5 Conclusions

With this work, we have outlined the entire processing stream of gravity gradient data using equivalent source from denoising to characterizing the error in denoised data. Additionally, three aspects of equivalent source processing have been highlighted: the equivalence of regularization parameter selection and estimating errors, regional-residual separation, and component conversion. The optimal equivalent source should reproduce the signal in the data and be minimally affected by the noise. For this reason, single- or multi-component data are consistently treated by applying equivalent source denoising. We have illustrated both the generalized cross-validation (GCV) and L-curve criteria to determine the optimal solution and found that both have produced good noise estimates in the theoretical examples and field examples, though the L-curve is more robust to correlated noise. Regional-residual separation using an equivalent source has the advantage that the regional trends removed from different components are consistent since they arise from a single source and represent a single regional field.

3.6 Acknowledgments

We thank Gary Young, Pat Millegan, and Jerry Hensel for many discussions on the processing of gravity gradiometer data. We would like to thank Desmond Rainsford for sharing with us his physical property information for the Ring of Fire area. We thank
the Ontario Geological Survey and Natural Resources Canada for making the McFauld's Lake survey data available for download and open use. This research was funded in part by the industry consortium “Gravity & Magnetics Research Consortium” (GMRC) at the Colorado School of Mines. The current sponsoring companies are Anadarko, Bell Geospace, BG-Group, BGP International, BP, CGG, ConocoPhillips, ExxonMobil, Gedex, Lockheed Martin, Marathon Oil, Micro-g LaCoste, Petra Energia, Petrobras, Shell, and Vale.
CHAPTER 4
COMPONENT CONVERSION

Though the gradient tensor has five independent components, any combination of the components can be used to invert for a structural representation of the subsurface. Currently, two acquisition systems are in use. One system measures the five independent components discussed in previous chapters. The second system is a single-axis gradiometer called 'Falcon' that measures two components. In this work, we use two Falcon airborne gravity gradiometry datasets. The first is acquired over the Vredefort Impact Structure in South Africa and will be used to illustrate component conversion as well as the ability of gradient data to image regional scale features. The data set consists of two components, i.e., \((T_{xx} - T_{yy})/2\) and \(T_{xy}\). We first examine the application of equivalent source technique in de-noising the two component gravity gradient data and the conversion of the measured components to the \(T_{zz}\) component. In an effort to understand how various components contribute to the resolution of the model, we examine the changes in the inversion results using the observed components and denoised \(T_{zz}\). The second dataset is from the Kauring test range in Western Australia. This dataset will illustrate similar conversions to the Vredefort example, but all gravity gradient data are converted from dense ground gravity data. Through the Kauring test site, a comparison between a high-quality ground gravity survey and airborne gravity gradient survey is possible.

4.1 Vredefort Background

Here we explore the utility of inverting single and multiple gravity gradient components for regional geologic features. The gravity gradient data is over the Vredefort impact structure in South Africa.

The Vredefort impact structure is in the Witwatersrand Basin of South Africa. An image of the impact structure with \(T_{xy}\) overlaid is shown in Figure 4.1. The impact structure is the
Figure 4.1: Image of Vredefort Impact Structure with $T_{xy}$ overlaid. Black outline indicates data area that was inverted. Image generated using Google Earth.
Figure 4.2: Geologic map of the Vredefort impact structure showing main geologic units; from Henkel and Reimold [2].
largest and one of the oldest known impact structures on earth. The structure is thought to be around 2023 Ma years old [55]. The original size of the impact structure is thought to be close to 300 kilometers in diameter. The Vredefort Impact Structure is in the heart of the Witwatersrand Basin, known for its economic gold reserves. It is believed that downfaulting of the Witwatersrand strata due to the impact helped preserve the vast economic resources within the basin today.

A geologic map of the impact region is given in Figure 4.2. The Vredefort Dome has two major structural pieces: the core and the collar. The core of the dome is composed of Archean granitoids while the collar is composed of overturned supracrustal strata [56]. At the center of the impact structure is the Vredefort Dome, which is the believed central peak or uplift due to the impact. There are currently two standing theories as to the type of impact structure exhibited within Vredefort. One opinion is that Vredefort is a complex crater. A visual cartoon of a characteristic crater with central peak is shown in Figure 4.3 with faulting segments radially outward. An alternative option to the structural interpretation is that it is a multi-ring basin. Similar to a complex crater, there is a structural high in the center termed a central uplift. The central uplift is considered to be caused by mantle uplift at depth. There are also concentric faulting structures outward from the central high. A visual cartoon of a characteristic multi-ring basin with central uplift is shown in Figure 4.4.

Figure 4.3: Cartoon illustrating impact structure. There is a structural high in the central region of the crater with concentric fault features radiating outward.
Many geophysical surveys have been conducted within the Vredefort impact structure to characterize the geologic features. Density of the core area was constrained by a seismic refraction survey across the Dome [57]. Henkel and Reimold [2] investigated the magnetic anomaly of the Vredefort Dome thought to be the result of a thermal overprint from the impact event. More recently, Muundjua et al. [58] implemented a ground magnetic survey in an attempt to characterize the geologic structure.

4.2 Vredefort Gravity Gradient Data

The gravity gradient data were acquired using the Falcon® system with north-south trending flight lines spaced 1 kilometer apart. Only a subset of the collected data is used for inversion here and covers approximately 1200 square kilometers. The digital elevation model within the survey area is shown in Figure 4.5. The survey is semi-draped with flight heights ranging from 50 to 280 meters above the topographic surface. The acquired data underwent routine proprietary processing and corrections for residual aircraft motion and self-gradient. Prior to data delivery, the data are demodulated, filtered, and leveled. The terrain effect
was removed using a density of 2.67 g/cc.

For this investigation, the two observed components are the curvature gravity gradients $T_{xy}$ and $T_{uv} = (T_{xx} - T_{yy})/2$. These two observed components are displayed in Figure 4.6. To explore what information can be extracted from the data, we calculate a third, more common component, $T_{zz}$. The $T_{zz}$ component was calculated using an equivalent source method. The equivalent source layer consisted of 1000 meter cubic cells located 1000 meters below the lowest elevation point. Both of the observed data components ($T_{xy}$ and $T_{uv}$) were used to obtain the equivalent source layer shown in Figure 4.7.

An L-curve criteria [50] was used to select the optimal equivalent source layer. The model objective function is plotted against the data misfit on a log-log scale as seen in Figure 4.8(a). This plot illustrates that models with increasingly complex model structure tend to fit the data to a higher degree than those of simple structures. The point of maximum curvature on the plot is considered to yield the optimal tradeoff between fitting the data and obtaining a reasonably realistic model. The curvature is plotted as a function of regularization parameter in Figure 4.8(b). The model selected is denoted by the pink marker on both the L-curve and
curvature plots. The forward modeled \( T_{zz} \) component from this model is shown in Figure 4.9. The equivalent source method doubles its utility by removing noise from the data due to the inability of the equivalent source layer to reproduce the random fluctuations that are characteristic of noise. For this reason, the \( T_{zz} \) component is expected to have less noise than that of the observed components.

![Figure 4.6: Gravity gradient components: a) \( T_{xy} \), b) \( T_{uv} = (T_{xx} - T_{yy})/2 \)](image)

Figure 4.6: Gravity gradient components: a) \( T_{xy} \), b) \( T_{uv} = (T_{xx} - T_{yy})/2 \)

![Figure 4.7: Equivalent source layer obtained from \( T_{xy} \) and \( T_{uv} = (T_{xx} - T_{yy})/2 \).](image)

Figure 4.7: Equivalent source layer obtained from \( T_{xy} \) and \( T_{uv} = (T_{xx} - T_{yy})/2 \).

### 4.3 Vredefort Inversion

Two models were obtained by inverting \( T_{zz} \) and \( T_{xy} \) and \( T_{uv} \). The mesh is composed of rectangular prisms with constant density contrast within each cell. The mesh discretization
Figure 4.8: a) L-curve showing data misfit plotted against model complexity b) Curvature of the L-curve plotted as a function of the regularization parameter

Figure 4.9: $T_{zz}$ calculated using an equivalent source layer generated from the $T_{xy}$ and $T_{uw} = (T_{xx} - T_{yy})/2$ components.
used has cell sizes of 500 meters in the easting and northing directions. The depth cells start with 25 meter thickness near surface to accommodate the topography and gradually increase to 500 meter thickness. Padding cells extend beyond the data area. The rectangular mesh has cell dimensions of 86 cells in the easting by 80 cells in the northing by 87 cells in depth giving a total of 598560 cells.

The two models were obtained via blind inversion of the components. A zero reference model is used with an initial model of 0.0 g/cc so that we are not assuming anything about the geologic features. The lower and upper bounds placed on the density contrast are -5.0 g/cc and 5.0 g/cc. These bounds were selected to allow for a somewhat unrestricted recovery of density contrasts while maintaining a plausible range of densities. The length scales in each direction are two times the cell size such that $L_E=L_N=1000$ meters and $L_Z=500$ meters. Topography is used in the inversion to remove model cells that lay above the ground surface.

4.3.1 $T_{zz}$ inversion

We first invert the $T_{zz}$ component to obtain a density contrast distribution. Though the $T_{zz}$ component was calculated, it serves as a base model to compare whether the addition of more components (observed or calculated) increases the quality of the model. A series of inversions using a range of values were carried out in order to select the optimal regularization parameter. Again, we use an L-curve criteria and plot the data misfit versus the model objective function [50].

For brevity, we omit the predicted data and difference data from the model shown. The range of values seen in the difference map are an indication of the noise in the data which the recovered model does not fit. The standard deviation calculated from the entire difference map is 0.96 Eo. We note that a low noise level is expected from inversion of the $T_{zz}$ data since it is calculated from the other two observed components.

A volume rendered image of the density contrast model is shown in Figure 4.10(a) and Figure 4.10(b) with all cells between 0.15 and -0.1 g/cc removed for clarity. The remaining high density contrast elucidates two dense features. The rounded feature to the west is near
the core of the Vredefort Dome while the linear feature to the east lies near the collar of the Vreederfort Dome.

4.3.2 $T_{xy}$ and $T_{uv}$ inversion

We next invert the two observed components $T_{xy}$ and $T_{uv} = (T_{xx} - T_{yy})/2$ together using the same mesh as that of the single $T_{zz}$ component. The errors in the data are estimated in the same manner as above. The estimated noise level for the $T_{xy}$ component is 5.42 Eo with a similar value of 5.53 for that of the $T_{uv}$ component.

The density contrast volume is displayed in Figure 4.11(a) and Figure 4.11(b) with all cells between 0.15 and -0.1 g/cc removed for comparison. The remaining high density contrast again identifies two main features. The two features maintain a similar density contrast to that of the $T_{zz}$ model. The two features are more compact with the high density contrast boundaries being significantly tightened to the source locations of the structures.

4.4 Vredefort Conceptual Interpretation

Modeling of geologic features within the Vredefort Dome have been performed by Henkel and Reimold [2]. The geologic modeling utilized gravity and magnetics with constraints from refraction seismic. The fitted geologic model is shown in Figure 4.12. The location of this cross section is shown on the geologic map in Figure 4.2. Along the 300 kilometer cross section, roughly 22 kilometers fall within the data area that was inverted (marked in Figure 4.9 as a black line). The approximate location of this 22 kilometer stretch is denoted on the section of Figure 4.12 as a red line. From the reconstructed section, it can be seen that there are likely massive fragments of crust in the first 15 kilometers of the subsurface. The slice through the 3D model corresponding to the geologic section is shown in Figure 4.13. The slice is taken from the recovered model from inverting the observed components.

By examining the two cross sections side by side, structural similarities are observable. The structural high represented by the red and blue crustal layers appears to correspond roughly with the location of a change in density contrast feature near the central portion of
Figure 4.10: (a) Side and (b) top volume rendered image of 3D density contrast model constructed from $T_{zz}$ with cells between 0.15 and -0.1 g/cc removed. Values are in g/cc.
Figure 4.11: (a) Side and (b) top volume rendered image of 3D density contrast model constructed from the two observed components with cells between 0.15 and -0.1 g/cc removed. Values are in g/cc.
Figure 4.12: Cross section through Vredefort impact structure constructed from gravity and magnetic data, modified from Henkel and Reimold [2].

Figure 4.13: Slice through 3D recovered model from inversion of observed components and correspondence to geologic section, modified from Henkel and Reimold [2].
the model. Conceptually mapping the structural features seen within the cross section onto
the density contrast model, the structural high is sketched in as illustrated in Figure 4.14.
The structural deformation of the upper crust modeled in the geologic section can also be
mapped onto the density contrast reversals captured in the region to the north. Mapping
these as faulted features, a more representative conceptual correspondence of the geologic
section with the 3D density model is obtained as shown in Figure 4.15.

![Figure 4.14: Sketch of structural uplift interpreted by the central density features.](image)

4.5 Vredefort Summary

With the two recovered models, we gain an understanding of the effect single and multiple
components have on the resulting inverted model structure and recovered density contrast.
The two models are overall structurally similar with small scale differences seen. By qualita-
tively comparing the recovered models, we can infer that the information contained within
the observed components has been translated to the equivalent source derived $T_{zz}$ compo-
Figure 4.15: Sketch of faulting or deformation features in the upper crust.
nent. In this example, one component gives a satisfactory model where the main features are distinguishable. Inverting the observed components delineates the same structure but with compact and continuous features.

The structural correspondence of the 3D models to the interpreted geologic structure from other geophysical and geological investigations provides validation of the component conversion and inversion technique. Further, we have demonstrated the utility of inversion for regional scale investigations using gravity gradient data. Given the lateral resolution and decay typified by gravity gradient data, large scale features such as those seen here within the Vredefort dome can be resolved.

4.6 Kauring Background

The Kauring Gravity Gradiometry Test Site offers a unique opportunity to understand the information content of various types of gravity gradient measurements and the required processing and interpretation techniques. In this paper, we utilize the measured ground vertical gravity data to simulate airborne gravity gradiometry data from several current and future systems using an equivalent-source technique. We then invert these converted data for the underlying density contrast and evaluate the performance of these inversions. The converted data and inverted models provide a reference for establishing the validity of available airborne gravity gradiometry data at the site. The major density features recovered from the optimally converted gradient data are highly consistent with those from the original ground gravity and observed gravity gradient data. With the comparable results from ground gravity and observed gravity gradiometry, we validate the use of gravity gradiometry at the test site. The consistency seen in this study demonstrates a coherent understanding of the workflow necessary to process and interpret gravity gradiometry.

With the technological advancements in gravity gradiometry instrumentation, there is a need for both understanding the information content of modern gravity gradiometry data relative to traditional gravity data, and for developing processing and interpretation tools that assess the myriad of measured gradient components and information they contain.
The Kauring Airborne Gravity Gradiometry Test Site and associated data provide a unique and important opportunity to address these two aspects and to explore the applicability of existing tools.

In this study, we focus on the available gravity data and airborne gravity gradient data to investigate the gravity gradient signal expected. We then compare converted data and the recovered density distributions to those obtained from inversion of observed data. We put the converted data through a realistic suite of processing, noise reduction, and noise estimation techniques in order to properly assess the information content and resulting density contrast models from inversion of airborne gravity gradiometry.

We first examine the expected airborne gravity gradient response due to the density structure below the test site. From the observed ground gravity data, we calculate full tensor gravity gradient data and typical components from newer gradiometer systems in development by using a regularized equivalent source method. These derived gravity gradient data are then subjected to appropriate low-pass filtering to simulate realistic processing. The converted data sets provide a reference for comparison with currently available acquired data and that of future surveys. Secondly, we interpret both the simulated and field data sets by inverting various components for 3D density contrast distributions.

The information content of these converted data sets are assessed through the recovered density models in comparison to those obtained from the inversion of observed gravity and gradient data.

4.7 Kauring Observed and Calculated Datasets

We first introduce the two observed datasets used in this study: the ground gravity and airborne gravity gradiometry. Then the application of equivalent source conversion of the ground gravity data for modeling of gravity gradient components will be discussed. Lastly, the calculated components using the equivalent source model will be presented.
4.7.1 Observed Ground Gravity Data

The ground gravity data were acquired for the explicit purpose of establishing the reference for assessing airborne gravity gradient data to be acquired at the Test Site [59]. The gravity data available from the website for the Test Site are shown in Figure 4.16. The data is dominated by a regional trend from the northeast to southwest. The data consists of sparse ground gravity from the Cunderdin regional area and dense ground gravity in the Airborne Gravity Gradiometry (AGG) test area. The regional data have station spacing of 2km while the AGG area data have a minimum of 500m station spacing. The smaller AGG area ground gravity is shown in Figure 4.17. Observation lines are oriented northeast to southwest, with progressively greater line spacing moving away from the central northeast axis of the data. The data shows a focused anomaly high in the centre and a broad linear feature to the southeast. There is also a regional field trending east-west. It is noted that there is a negative DC component in the gravity data shown in Figure 4.17.

4.7.2 Observed Gravity Gradient Data

A recently acquired Falcon gravity gradiometry survey is available over the airborne test site. The observed gravity gradient components are shown in Figure 4.18. The $T_{xy}$ component shows a linear feature trending NW to SE in the lower portion of the data in the same location and direction as in the ground gravity data. The observed components will be inverted and the recovered model compared to that of the model recovered from the converted gravity gradient components.

4.7.3 Equivalent Source Conversion

We examine the associated gravity gradiometry responses from equivalent source and their ability to characterize the high-density anomaly beneath the test site through inversion. We use a regularized equivalent source technique to convert the ground gravity data to gravity gradient components [35, 60]. The approach has the benefit of attenuating data error and producing more reliable gradient calculations.
The equivalent source layer is composed of 100-m cubic cells located 100 m below the lowest observation point. The gravity data consists of both sparse regional data and denser data within the AGG test area, so using 100 m cell sizes allows for low and high frequencies to be represented within the layer. A regularized inverse approach is used to determine the density values in the layer. An L-curve criteria [61] was used to select the optimal regularization level so that the calculated gradient components are minimally affected by the noise in the ground gravity data.

![Figure 4.16: Regional gravity around the Kauring Test Site.](image)

The constructed equivalent source layer is shown in Figure 4.19. The equivalent source method used in our processing and conversion provides two benefits. First, it attenuates data error due to the ability of the equivalent source layer to optimally misfit data according to the noise level. Secondly, it allows for the removal of a regional trend so the calculated gravity gradients are consistent with the target anomaly.

From the equivalent source layer, we first calculate the five independent gravity gradient components along with $T_{zz}$ (referred to as six-component data). In order to be consistent with realistic acquisition of gravity gradient data, a 300m, 4th-order Butterworth filter was
Figure 4.17: Ground gravity in the airborne gravity gradient area of the Kauring Test Site.

Figure 4.18: Airborne gravity gradient data acquired over the Kauring test site by the Falcon system. Left: $T_{xy}$; Right: $T_{uv}$

Figure 4.19: Equivalent source layer generated using ground gravity data.
used to approximate the low-pass acquisition filter. The application of this filter serves to maintain a realistic level of processing before the data, and resulting signal, is inverted. The filtered six-component data is shown in Figure 4.20.

Figure 4.20: The five independent gravity gradient components along with $T_{zz}$ calculated using the equivalent source layer from ground gravity data. The tensor components have been filtered with a 300m, 4th-order Butterworth filter.

We then combine these components to form the specific data quantities measured by different systems. For example, two components measured by Fugro’s Falcon system [62] are given by

\begin{align*}
T_{xy} &= T_{xy} \\
T_{uv} &= \frac{T_{xx} - T_{yy}}{2}
\end{align*}

where $T_{ij}$ is the gravity gradient component when $x$ is northing, $y$ is easting, and $z$ is down. The converted Falcon components are shown in Figure 4.21 with the same flight lines as the observed Falcon components. The $T_{xy}$ component shows a linear feature running northwest.
to southeast through the data. The major features present in these converted gravity gradient components from the ground gravity are consistent with the features seen in the observed gravity gradient data of Figure 4.18.

Figure 4.21: Calculated Falcon components, $T_{xy}$ (left) and $T_{uv}$ (right), from the equivalent source layer in Figure 4.19.

4.8 Kauring Inversion of Observed Data

The ground gravity data is first inverted for a 3D density contrast distribution. The recovered density contrast model provides a basis for understanding the information contained within the converted gravity gradient components. The mesh used is composed of rectangular prisms with density contrast being constant within each cell. The mesh uses cuboidal cells of 50 m by 50 m by 50 m within the core of the data area and with increasing cell sizes towards the edges of the survey area and at depth.

The ground gravity data was inverted using a zero reference model and no a priori information incorporated. We select the recovered model based on the discrepancy principle [63]. A volume rendered image of the recovered model is shown in Figure 4.22 with all cells less than 1 g/cc removed for clarity. Note the positive density contrast feature in the middle of the model corresponding to the gravity high.

We next invert the observed Falcon components ($T_{xy}$ and $T_{uv}$) of Figure 4.18. The density contrast distribution recovered will serve two purposes. First, we seek to establish
the common recoverable density distribution from observed ground gravity and airborne gravity gradient data. Second, we validate the recovered models from converted data with the recovered models from observed data (ground gravity and gravity gradient components).

The observed gradient components were inverted with the same inversion parameters as that of the ground gravity. A series of inversions were performed in order to explore and estimate the data errors. Through the process of choosing a regularization parameter, the noise is estimated to be $2.3 \times 10^2$ for the $T_{xy}$ component and $2.1 \times 10^2$ for the $T_{uv}$ component. The selected density distribution shown here had a misfit value close to the number of data. The density distribution selected is shown in Figure 4.23 where cells with density contrast
below 1 g/cc are removed. As with the ground gravity density distribution, the observed gravity gradient model identifies a main dense feature in addition to a southern trending feature with density contrast higher than 1 g/cc.

4.9 Kauring Inversion of Converted Data

To explore the information content of the various converted gravity gradient components, we invert $T_{xy}$ and $T_{uv}$, and the five component tensor data. The models discussed here were obtained using the same mesh discretization as that of the observed data inversions. No prior geologic knowledge of the site was incorporated into the inversion. Generic lower and upper bounds of -5.0 g/cc and 5.0 g/cc, respectively, are placed on the density contrast with a zero reference model.

For all converted data inversions, we follow the same methodology previously outlined to obtain the density distribution given. A series of inversions using a range of regularization parameter values were carried out in order to determine the optimal parameter. In determining this parameter, we also gain insight into the noise content of the data by examining what part of the data the recovered model is unable to fit. We do not reproduce the predicted or difference data maps here for brevity.

4.9.1 $T_{xy}$ and $T_{uv}$ data

We first invert the converted $T_{xy}$ and $T_{uv}$ components given in Figure 4.24. The recovered model in Figure 4.24 identifies a dense feature when cells below 1 g/cc are removed. The location and structure of this feature is consistent with both models obtained from observed ground gravity and airborne gravity gradient data.

The standard deviation calculated from the difference maps for both the $T_{xy}$ and $T_{uv}$ components is less than 1 Eo. This is a reasonable estimate since the calculated data were subjected to two levels of noise reduction through the equivalent source construction and the applied filter to simulate the acquisition filters.
4.9.2 Five component tensor data

We next invert the five independent tensor components to obtain a density contrast distribution. The standard deviation calculated from the difference maps is less than 1 Eo. A volume rendered image of the density contrast model from inverting tensor data is shown in Figure 4.25 with all cells below 1 g/cc removed in order to see the structure associated with the central anomaly. The recovered dense feature is consistent with both models obtained from observed ground gravity and airborne gravity gradient data.
4.10 Kauring Summary

The various converted gravity gradiometry inversions are consistent with the observed ground gravity and Falcon gravity gradient inversions. All models recover a positive density contrast feature associated with the observed gravity high in the ground data. Though not properly conveyed with the volume-rendered images used, all density distributions do recover less dense features trending southeast away from the main anomaly. The depth slices in Figure 4.26(a), Figure 4.26(b), Figure 4.26(c), and Figure 4.26(d) display this common feature in four recovered models. A second depth slice through these same four models is displayed in Figure 4.27(a), Figure 4.27(b), Figure 4.27(c), and Figure 4.27(d). These depth slices show similar density contrast structures across all four models at greater depth.

We have used the ground vertical gravity data at the Kauring Test Site to model airborne gravity gradiometry data and investigate the information content of various component combinations through inversion. All inversions of observed and converted data have produced similar structural models, indicating that the particular converted data components and orientation used in this study are equivalent in information content. Further, the consistently recovered features indicate that airborne gravity gradiometry is a reliable alternative to ground gravity at this site. The variations in these density contrasts are representative of the varying information contained among individual components. The use of regularized equivalent source construction proves to be a reliable tool for data conversion.

Compared with the density model from airborne gravity gradient data, the models from converted gradient data appear to have slightly more small-scale circular features. The presence of these features indicate a difference in the high-frequency information content resulting from the equivalent source conversion. The agreement of the structural features in this study demonstrates the high quality of the airborne gravity gradient data and consistency in the information contained in the ground gravity and airborne gravity gradient data.
Figure 4.26: Plan sections through volumes at a depth of 80m. Inverted models from (A) ground gravity, (B) observed Falcon, (C) converted five component data, (D) converted Falcon. Color is from -1 g/cc (blue) to 2.4 g/cc (red).
Figure 4.27: Plan sections through volumes at a depth of 180m. Inverted models from (A) ground gravity, (B) observed Falcon, (C) converted five component data, (D) converted Falcon. Color is from -1 g/cc (blue) to 2.4 g/cc (red).
4.11 Acknowledgments

Thank you to Mark Dransfield and CGG (formerly Fugro) for the Vredefort and Kauring data and for helpful information in completing these studies. We thank Ned Stoltz and Richard Lane from Geoscience Australia for providing the airborne gravity gradient data at the Kauring Test Site. We gratefully acknowledge the researchers, agencies, and companies who established the Kauring test site and made the associated data available.
CHAPTER 5

3D INVERSION OF AIRBORNE GRAVITY GRADIOMETRY DATA IN MINERAL EXPLORATION: A CASE STUDY IN THE QUADRILÁTERO FERRÍFERO, BRAZIL

We present a case study of applying 3D inversion of gravity gradiometry data to iron ore exploration in Minas Gerais, Brazil. The ore bodies have a distinctly high density contrast and produce well defined anomalies in airborne gravity gradiometry data. We have carried out a study to apply 3D inversion to a 20 km² subarea of data from a larger survey to demonstrate the utility of such data and associated inversion algorithm in characterizing the deposit. We examine multiple density contrast models obtained by first inverting $T_{zz}$; then $T_{xz}$, $T_{yz}$, and $T_{zz}$ jointly; and finally all five independent components to understand the information content in different data components. The commonly discussed $T_{zz}$ component is sufficient to produce geologically reasonable and interpretable results, while including additional components involving horizontal derivatives increases the resolution of the recovered density model and improves the ore delineation. We show that gravity gradiometry data can be used to delineate the iron ore formation within this study area.\(^\text{2}\)

5.1 Introduction

The Quadrilátero Ferrífero, or Iron Quadrangle (Figure 1), is an area of significant mineral resources in the state of Minas Gerais, Brazil. The Quadrilátero Ferrífero covers approximately 7,000 square kilometers. Iron and gold are major mineral resources that have historically been produced from the area [65].

The host formations for the iron ore are the Cauê Itabirite and to lesser degrees, the overlying Gandarela and the underlying Batatal. Itabirite is a general term for oxide-facies iron formation and is the namesake of the Cauê formation. The iron ore of interest is hematite and itabirite. The high-density contrast between the iron formation and the host rock make

\(^{2}\)Content within this chapter has been published in Geophysics ([64]).
gravity methods effective tools for exploration and ore delineation in this area. However, the target iron formation is generally extensive in two dominant directions and requires wide area coverage of geophysical data. The topography of the area is rugged with the landscape composed of canyons, plateaus, and valleys. The occurrence of iron formation is commonly coincident with the rugged peaks, making ground geophysical exploration difficult in this region. In addition, the thick blanket of vegetation severely limits the mobility of equipment for both manual and vehicular surveys. Much of the remote area is inaccessible via ground for a few months out of the year due to tropical rainstorms, further impeding the use of ground geophysics. With the many obstacles that the environmental conditions impose, performing ground-based geophysics is impractical.

To overcome these difficulties, considerable effort has been put into mapping the occurrence of iron formation in the region through outcrop and other surface expressions. The iron formation is characteristically shallow. The typical exploration strategy in the area, prior to the introduction of airborne geophysics, relied on the geologic knowledge resulting from such mapping investigations and extensive drilling programs. Drill holes were located primarily based on the known surficial orientation of iron formation and the presence of canga, which is a cemented, surficial iron formation-derived layer.

Alternatively, airborne gravity gradiometry has the potential to delineate iron formation effectively over extended areas where there are difficulties in ground exploration. The method is advantageous since data can be collected over large areas in a short period of time. Gravity gradiometry has higher resolution of near surface features and a higher signal-to-noise ratio (SNR) than airborne gravity because of its ability to reject common-mode noise. Thus, airborne gravity gradiometry is an effective method for resolving the shallow dense ore bodies typical in this area.

As for many geophysical methods utilized in mineral exploration, quantitative modeling such as inversion is necessary to extract geological information from the data. With the rapid emergence of gravity gradiometry in exploration (e.g. [9, 10]) and multiple-component
datasets being acquired over large areas, the need for processing and interpretation tools is all the more important. One essential aspect is to utilize the entirety of information contained in multi-component gradiometry measurements. Gravity gradient data can be inverted to obtain a 3D density contrast distribution. In addition to characterizing the geometry and density of the target, a 3D density model can potentially be used to estimate the volume of the target; infer lithologic changes; or provide guiding information for more targeted and efficient drilling programs.

There is a plethora of work in the literature addressing different aspects of data processing and interpretation in gravity gradiometry. For instance, Chen and Macnae [25], Kass and Li [26], Dransfield and Zeng [27], and Davis et al. [28] studied various aspects of terrain correction. Inversion of gravity gradient components was developed in conjunction with airborne systems, such as the work done by Vasco and Taylor [66], Condi and Talwani [67], Li [19], and Zhdanov et al. [68]. Kirkendall et al. [69] developed $l_p$-norm inversion for recovering sharp boundaries within the context of cargo container imaging. Uieda and Barbosa [70] extended Rene’s (1986) method with seeds and Camacho’s (2000) method of growing bodies by introducing a regularizing function for compactness and using multiple seeds in the model region to guide the recovery of spatially complex density contrasts. Lee et al. [73] and Lyrio [74] examine the effects of acquisition processing and post-processing on the data and its subsequent inversion. Pilkington [75] has examined the optimal gravity gradient components to utilize for inversion. In contrast, there have been fewer case studies on the application of gravity gradiometry inversion for quantitative interpretation in mineral exploration.

In this paper, we invert full tensor gravity gradient data over an iron formation in order to generate a 3D density contrast distribution using the approach of Li [19]. We examine the utility of various gravity gradient component combinations in obtaining a 3D density contrast distribution using a field dataset from the Quadrilátero Ferrífero. We will first review the theory and acquisition of airborne gravity gradient data, the inversion methodology, and then
describe the survey area and geology of the site. We then invert the different combinations
of gravity gradient tensor components to qualitatively evaluate their information content
for resolving the high-density zone associated with the iron formation. The results show
that the inversion of airborne gravity gradiometry data can reliably characterize the iron
ore formation by imaging its spatial distribution, dip structure, and depth extent. We
conclude by discussing the practical consideration of gravity gradiometry inversion in such
applications.

5.2 Deposit geology and gravity gradiometry survey

The rugged terrain, vegetation, and weather hinder the use of ground geophysics for
exploration in this region. Previous exploration efforts have been unable to conduct ground
geophysical surveys, making airborne gravity gradiometry specifically advantageous since
the utilization of this method requires little manpower and no ground access. The near
surface ore bodies and stark contact with the host rock make airborne gravity gradiometry
well suited for iron exploration in this area.

5.2.1 Geology

The area of interest for this study, the Gandarela Syncline, lies within the Quadrilátero
Ferrífero. There are varying interpretations as to the tectonic evolution of the Quadrilátero
Ferrífero such as those given by Dorr [65], Chemale Jr. et al. [76], and Alkmim and Marshak
[77]. Although there is much discussion about the deformation sequence of the Quadrilátero
Ferrífero, it is generally accepted that regional scale folding during the Transamazonian event
is responsible for the formation of the Gandarela syncline [76]. A fluid inclusion study con-
ducted by Rosiere and Rios [78] gives evidence that the ore bodies are the result of hypogene
and supergene processes, with mineralization beginning during the Transamazonian event
and continuing in stages through the Brasiliano/Pan-African orogeny.

Dorr [65] gives a detailed description of the stratigraphy of the Quadrilátero Ferrífero and
the structure of interest, the Gandarela Syncline. The iron bearing formations occur within
Figure 5.1: Geologic map of the Quadrilátero Ferrífero, with solid black lines indicating where the Minas Series is present; modified from [3]. The survey area is outlined in red.
the Minas Series, which are composed of Precambrian metasedimentary rocks. The Minas Series presently occurs in regional synclinal features such as the Gandarela Syncline. The structurally controlled occurrence of the Minas Series is shown in Figure 5.1. The Minas Series, containing Lake Superior type iron formation, lies unconformably on the Rio das Velhas Series, which contains Algoma-type iron formation ([77]). The eastern limb of the Gandarela syncline has been overturned while the western limb remains right side up.

Within the Minas Series, the Cauê Itabirite hosts the majority of the iron mineralization. The Cauê Itabirite lies between the overlying Gandarela Formation and underlying Batatal formation. The Cauê is commonly 300-500 meters in thickness but can range from a few meters to over 1400 meters. Generally, the iron deposits occur near faults and fold zones characterized by low pore pressure and high permeability that allow for the movement of fluids. Iron formation on ridgelines is protected by canga, which is a cemented iron formation-derived layer resistant to chemical and mechanical weathering [65].

The details below on the structural occurrence and properties of the iron ore bodies can be found in Dorr [3]. The ore bodies tend to be shallow and can range anywhere from 25 to 150 m below the surface. The ore deposits follow the structure of the host formation and are generally tabular, have large strike length, and dip southeast with an approximate dip of 25°. The contact between the ore and host itabirite can be gradational or feathery in nature, but is usually abrupt. The high-grade ore typically contains an average of 66% Fe and the intermediate grade ores range between 57% and 66%. The high-grade deposits are easily differentiated from the dolomitic and quartz-rich country rock by the high-density contrast. The densities of the high-grade iron ores range from 3 g/cc for soft and porous hematite to 5 g/cc for hard and compact hematite. Variations in the soft and hard hematite are due to the ore forming process where the soft hematite is the result of weathering of preexisting iron formation.
5.2.2 Gravity gradiometry survey

The gravity gradient data were collected in August-September 2005 in the Quadrilátero Ferrífero. The 93 – km² survey was acquired with 100 m line spacing. Flight lines were oriented 32 from the north. The survey was semi-draped with flight heights ranging from 60 to 500 m above the topographic surface. The acquired data underwent routine proprietary processing and corrections for the centripetal force, self-gradient, and acceleration compensation by the acquisition company. The lines were also leveled and filtered to attenuate noise.

Figure 5.2: Terrain corrected gravity gradient data covering a 4 km by 5 km area with survey topography (f) showing the location of the data subset. From left to right and top to bottom the plot panels display (a) $T_{xx}$; (b) $T_{xy}$; (c) $T_{xz}$; (d) $T_{yy}$; (e) $T_{yz}$; (f) topography; (g) $T_{zz}$. The location of the cross section of Figures Figure 5.5 and Figure 5.6 is overlain on $T_{zz}$ in panel (g).

We use a subset of the survey area for this study. The data subset is displayed in Figure 5.2 in the tensor format. For reference, the topography within the entire survey area
is displayed in Figure 5.2f, and the black box indicates the subset area. The data subset covers approximately $20 \text{ km}^2$ in the northern part of the survey area. There is an active mine, the Cauê Mine, within this area and borehole data are available. A geologic cross section generated from drill data is also available within this area. The geologic feature of interest runs north-east through the middle of the data map parallel to the long axis of the survey area and is visually distinguishable within the gridded data map.

![Figure 5.3: The importance of removing the terrain effect is illustrated through the visual anomalies of the $T_{zz}$ component before and after the correction is applied. (a) Observed $T_{zz}$; (b) Terrain effect calculated for $T_{zz}$ component; and (c) $T_{zz}$ after the terrain effect is removed.](image)

The importance of removing the terrain effect is illustrated in Figure 5.3. The observed $T_{zz}$ subset is shown in Figure 5.3a. The presence of long wavelength features is primarily due to the terrain while small-scale variations are representative of acquisition and processing errors. As seen in the topographic map in Figure 5.2f, the terrain is rugged with an average elevation of 1100 m. The topography is spatially coincident with the presence of iron formation, with ore bodies occurring near topographic highs in this area.

With the availability of LIDAR (Light Detection and Ranging) data over the region, the terrain effect can be calculated to a high degree of accuracy. The terrain effect shown for $T_{zz}$ in Figure 5.3b is calculated using a constant density of 2.67 g/cc. To select the optimal density value that removes the topographic effects in the survey area, data maps for the
The terrain corrected $T_{zz}$ subset is shown in Figure 5.3c. From visual examination of Figure 5.3, the importance of removing the terrain effect from gravity gradient measurements becomes clear. In the anomalous data map, we can see the response from the iron formation oriented northeast to southwest near the flanks of the topographic high. We can also visually distinguish the lateral range of the ore body while depth and further structural information is not readily discernible. From these qualitative observations, it is clear that proper removal of the terrain effect plays an important role in reducing the data to interpretable anomalies. With the anomalous data, we can now proceed to inversion in order to further ascertain structural and spatial characteristics of the ore bodies that are not apparent by visual inspection of the data maps.

### 5.3 Inversion results

We examine whether gravity gradient data can be used to resolve the ore bodies and define geologic structure of the iron formation. Through three different combinations of components, we explore the extent to which additional components provide extra information about the structural elements of the target iron formation.

To demonstrate how well gravity gradient data can resolve the structure of the ore bodies, we first invert a single component, $T_{zz}$. To explore how additional components enhance the recovered source distribution, we compare models obtained from various component combinations. As described in the inversion algorithm section, the mesh is composed of rectangular prisms or cells with constant density contrast within each cell. The discretization

$T_{zz}$ component were examined with different density values used for the terrain correction. A density close to 2.67 g/cc reveals anomalies in the $T_{zz}$ map that correspond to known geologic features of interest in the area. Using small deviations from the actual mean value is not crucial since the background density is not strictly an absolute constant. The goal is to remove as much of the terrain effect as possible in order to identify signal from the target of interest. Conceptually, the final density model used for interpretation is the sum of the terrain density and the recovered density contrast.
used has cell sizes of 25 m in the easting by 25 m in the northing by 20 m in depth in the central region of the mesh and padding cells beyond the data area and at large depth, resulting in 156 cells in the easting, 241 in the northing, and 45 in depth. The second inversion uses three components ($T_{xz}$, $T_{yz}$, and $T_{zz}$), and the third uses the five independent tensor components ($T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, and $T_{yz}$). Both use the same mesh discretization as that of $T_{zz}$ inversion. A total of 3017 data locations were used to invert for the three models, resulting in 3017 data for the first inversion, 9051 data for the second, and 15085 data for the third.

The same inversion parameters, except for the regularization parameter, are used to obtain the three models. A zero density contrast reference model is used. The lower and upper bounds placed on the density contrast of all cells are 0.0 g/cc and 4.0 g/cc, respectively. These bounds were selected because the density contrast of the iron formation is entirely positive. The length scales in each direction are two times the cell size $L_E = L_N = 50$ m and $L_Z = 40$ m, requiring a relatively equal amount of smoothing in each direction. Topography is incorporated into the inversion by removing model cells that lie above the ground surface. Although the reference model and length scales can be used to incorporate the known geologic structure and lateral extent primarily in the north-south direction, we did not utilize these parameters to incorporate a priori information. Instead, we only applied the above-mentioned bound constraints. Thus, all models were obtained via blind inversion with generic inversion parameters and little a priori information.

An important step in inverting any combination of components is to obtain knowledge of the relative data errors in different components. These error estimates are required in the inversion to properly weight the contributions of different tensor components to the total data misfit. Here, we outline a general approach of using inversion to explore the data errors of each component and obtain a representative standard deviation. The data error estimation is achieved iteratively through the selection of the optimal regularization parameter for an initial estimate of data errors. In this approach, we are able to simultaneously explore the
selection of the optimal solution.

In order to estimate the errors in the data, we used the L-curve criterion [50] and relied on the statistical characteristics of the misfit data map from preliminary inversions. To do so, we start with an assumed uniform standard deviation and performed multiple inversions using a range of values for the regularization parameter. A plot of the model objective function versus the data misfit as a function of regularization parameter on a log-log scale forms the Tikhonov curve that characterizes the trade-off between data misfit and model complexity. This curve is also referred to as the L-curve when a pronounced corner point is visible, which represents the optimal balance between reproducing the signal in the data versus constructing a structurally simple model. The L-curve criterion thus states that the regularization parameter corresponding to this corner point yields the optimal regularization.

We begin by using the L-curve criterion to identify the optimal regularization range and investigate the data errors. We then refine the data misfit function and regularization parameter by examining the data residual, i.e., the difference map between the observed and predicted (or forward modeled) data. By inspecting the difference map, we are able to determine whether the model is reproducing the trends and features within the data that are representative of the geologic formation of interest. The difference map is expected to contain high frequency fluctuations and no trends relating to the target geology if we have a perfect characterization of the data errors. The Root Mean Squared (RMS) difference between observed data and predicted data by the current optimal model is calculated for individual components and used as the approximations to standard deviations in the subsequent inversion. We repeat this step several times.

In summary, while the L-curve criterion enables us to zoom into the neighborhood of regularization for the optimal model, the final regularization parameter is selected according to the data misfit in equation 1.5 that is defined based on the iteratively estimated error standard deviations. Corresponding to the optimal model, the resulting differences between the observed and predicted data are considered to best characterize the errors in the data. The
differences then allow us to calculate an estimate standard deviation of the data errors. The standard deviations estimated using the method described are given for each combination of components for the purpose of understanding the data errors in general.

5.3.1 Single Component Inversion ($T_{zz}$)

We first invert the $T_{zz}$ component. Comparison of the observed and predicted data indicates that the major signals in the observed data have been well reproduced by the density distribution in the predicted data. For brevity, we have not reproduced the plots here. The difference map, Figure 5.4, of the observed and predicted data shows primarily small fluctuations with a near zero mean, indicating that the selected model explains the major trends, i.e., the signal, in the data. The range of values seen in the difference map between the observed and predicted data is an indication of the noise in the data. The standard deviation calculated from the entire difference map is 16.5 Eo. This standard deviation is reasonable since the error at this stage has several different sources including acquisition system, aircraft movement, previous processing steps, the terrain correction, and bound constraints imposed during inversion.

From the single component recovered model, we calculate the predicted data for all unused gravity gradient tensor components. The standard deviation calculated from the difference between the observed and predicted data for the tensor components is shown in Table 5.1. The shaded cells denote the components that were not inverted. The values are the standard deviation of the difference maps. In practice, we utilize these as an estimate of the standard deviation of the errors in the data.

The recovered model identifies a series of dense and shallow bodies trending northeast to southwest through the survey area. These features are consistent with the known iron formations localized within the Gandarela Syncline. The features in the density contrast model demonstrate that the inversion algorithm imaged the iron ore formation within the Minas Series from single $T_{zz}$ gradient data alone. We compare one cross section of the recovered 3D model with a geologic cross section in Figure 5.5. The location of the cross
Figure 5.4: Tzz difference map between the observed and predicted data from single component recovered model.

Table 5.1: RMS differences (in Eo) of the observed and predicted data for the four cases discussed. Shading denotes components that were not inverted and the recovered model was used to calculate the predicted data for differencing with the observed data. These values are the standard deviation of the residual map. In practice, we use these values as an estimate for the standard deviation of the data errors in the inversion process.

<table>
<thead>
<tr>
<th></th>
<th>Txx</th>
<th>Tyx</th>
<th>Tzx</th>
<th>Tyy</th>
<th>Tyz</th>
<th>Tzz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1C</td>
<td>23.9</td>
<td>23.2</td>
<td>31.8</td>
<td>23.1</td>
<td>26.1</td>
<td>16.5</td>
</tr>
<tr>
<td>3C</td>
<td>17.5</td>
<td>16.0</td>
<td>15.9</td>
<td>16.0</td>
<td>12.4</td>
<td>22.5</td>
</tr>
<tr>
<td>5C</td>
<td>16.6</td>
<td>12.6</td>
<td>16.3</td>
<td>15.8</td>
<td>12.2</td>
<td>24.3</td>
</tr>
<tr>
<td>6C</td>
<td>15.7</td>
<td>13.0</td>
<td>17.9</td>
<td>13.8</td>
<td>13.8</td>
<td>21.4</td>
</tr>
</tbody>
</table>
section through the area is marked with a black line on the $T_{zz}$ data in Figure 5.2g. The geologic cross section (Figure 5.5a) is generated from drill hole information and shows that the main iron ore body is dipping to the southeast, parallel to general bedding in the area. The pink zone in the section is the known hematite, while the surrounding light and dark blue layers are, respectively, soft and hard itabirite. A slice through the recovered 3D density model corresponding to the geologic cross section is shown in Figure 5.5b. We have discretized the outlines of the hematite and itabirite units and displayed them as overlays on the recovered density section for direct comparison between the two sections.

Figure 5.5: (a) Geologic cross section running northwest (left) to southeast (right) and created using drill hole data. The location of this section is marked on Figure 5.2g. The pink is hematite with surrounding light and dark blue layers representing soft and hard itabirite, respectively. (b) Slice through $T_{zz}$ model corresponding to the geologic cross section; from 0 g/cc to 1.4 g/cc. The high density contrast correctly identifies hematite as a near surface body with the appropriate dip direction. Itabirite layers are defined by the dispersive positive density contrast radiating out from the dense hematite region.
The first zone of high density contrast near 200 m easting on the section is related to the gravity gradient high as seen in the $T_{zz}$ data just to the north-east of the cross section. The presence of high density contrast associated with this anomaly is expected in this cross section due to the proximity of the anomaly and the smoothing by the algorithm. The inverted model shows a major dense body dipping in the direction of bedding approximately coincident with the ore body in the geologic cross section. We observe that the depth extent (approximately 200 m) and dip (about 35°) of the recovered dense region are consistent with the iron formation indicated on the geologic section. Thus the inversion has successfully delineated the presence of the ore body with the appropriate dip and approximately correct location. The second thin hematite body seen in the geologic section is not properly imaged due to the size and geometry of the small body.

5.3.2 Three-component Inversion ($T_{xz}$, $T_{yz}$, and $T_{zz}$)

From inversion of the $T_{zz}$ component, it is clear that gravity gradient data with the given acquisition parameters can image the presence and general structure of iron formation. We next invert three components ($T_{xz}$, $T_{yz}$, and $T_{zz}$) together using the same mesh to investigate the improvement that can be gained through multiple components of the data. We choose $T_{xz}$ and $T_{yz}$ components because they intuitively complement the $T_{zz}$ component and are expected to improve the lateral resolution of the inverted density contrast model if they contain different information about the source.

Initial estimates of the errors in the data were obtained in the same manner as described for the $T_{zz}$ component by carrying out a series of inversions. However, we also utilize the estimate of the error standard deviation of 16.5 Eo from the $T_{zz}$ residuals as a guide. The final estimated standard deviations for the $T_{xz}$ and $T_{yz}$ are respectively 15.9 Eo and 12.4 Eo (shown in Table Table 5.1), while the error estimate for $T_{zz}$ has increased to 22.5 Eo. The increase is expected since the inclusion of two new components enables the identification of additional noise in $T_{zz}$ that is inconsistent with the newly added components. In general, the estimate from multiple components should be greater than that from a single component because
the estimation procedure essentially relies on the consistency of signals among different components.

A slice through the recovered model in the same position as the geologic cross section in Figure 5.5a is shown in Figure 5.6b. As with the single component $T_{zz}$ inversion, the model shows a similar dense body dipping in the direction of known bedding. Using the outlines as a comparison, it can be seen that the density contrast corresponding to the ore body in the geologic cross section appears to be better defined structurally.

Compared to the model section from $T_{zz}$ inversion, the lateral boundaries of the dense body are better aligned and the dip is more pronounced. The dense hematite core is better characterized as an elongated dipping body as indicated by the geologic section. From these observations, we can conclude that including $T_{xz}$ and $T_{yz}$ components in the inversion process tightens the lateral boundaries and location of the recovered source when compared to the single component inversion using $T_{zz}$.

### 5.3.3 Five-component Inversion ($T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, and $T_{yz}$)

Next we invert five components ($T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, $T_{yz}$) together to evaluate the utility of inverting all independent components. With the additional information at each observation location, the overall density model is expected to improve. Preliminary estimates of the errors in the data are found according to the same methodology previously described. The estimates obtained for the $T_{xz}$ and $T_{yz}$ components (16.3 Eo and 12.2 Eo, respectively) are highly consistent with those (15.9 Eo and 12.4 Eo, respectively) obtained from the three-component inversion. Individual standard deviations estimated are listed in the third row of Table 5.1.

A slice through the recovered model at the same location of the geologic cross section is shown in Figure 5.6c. We can see that the orientation and shape of the recovered body has been further refined and that the density model provides the best structural match with the ore body. The edges of the ore body and iron formation are sharper than the previous two sections and have better agreement with the expected dip of the itabirite formation.
Figure 5.6: Single cross section comparison of four recovered models from inversion of (a) one-component, (b) three-components, (c) five-components, and (d) six-components. These sections correspond to the geologic cross section of Figure 5.5a. From the improvements seen in the recovered models from (a) one- to (c) five-components, the constraints from including additional components provide secondary information that contributes to a more accurate geologic structure. The inclusion of the sixth component (d) does not significantly enhance the ore delineation or quality of the recovered model.
An improved representation of the hematite body is achieved with high density contrast aligned with the elongated outline and extending to the surface in the direction of bedding. Compared to the single- and three-component inversions, the added components significantly increase the interpretable quality of the resulting model.

5.3.4 Six-component Inversion \((T_{xx}, T_{xy}, T_{xz}, T_{yy}, T_{yz}, T_{zz})\)

For this survey, six components of the gravity gradient tensor were measured and are available for interpretation. Although the gravity gradient tensor can be fully reconstructed from five independent components, it is common to use all available information for inversion. We invert all six measured components in order to gain some insight into the information content extracted through inversion from all measured components compared to that of select components as previously described. The fourth model in our study was obtained using the same mesh, inversion parameters, and methodology as the previous three models. The estimates obtained from this six-component inversion are consistent with the earlier results as seen in the fourth row of Table Table 5.1.

The cross section through this model corresponding to the geologic section is shown in Figure 5.6d. The structural features are consistent with the previous three models. The high density contrast corresponding to the hematite is located in the same near surface location. The edges of the hematite ore are more elongated with slightly greater depth extent than that of the five-component model. The dip of the itabirite is still tightly constrained.

The inversion results using five and six components, respectively, are highly consistent and they will serve the purpose of interpretation equally well. Thus, it can be confidently stated that adding the sixth component is unnecessary. It then follows that one should not invert six components in practice because the sixth component does not bring additional information, yet it increases the computational cost of inversion by one-sixth.
5.3.5 Summary

We have compared the inversion results from the four gradient tensor component combinations in Figure 5.6. From the $T_{zz}$ model alone (Figure 5.6a), it is clear that gravity gradient data are able to resolve the iron formation structure. With two additional components, improvements are made in the overall placement and boundaries of the dense hematite ore (Figure 5.6b). When five components are inverted (Figure 5.6c), the high-density hematite is properly placed near surface with a tightening of the dipping iron formation edges. Inclusion of a sixth component (Figure 5.6d) does not improve the delineation of the iron formation.

The first-order structural elements in all cross-sections are in agreement in that there is a dense dipping body represented by the high density contrast. Incorporating additional gradient components into the inversion improves the structural representation by better defining the dip, density contrast, as well as the depth and lateral extent of the body. Using multiple components provides secondary contributions to the overall recovered source distribution and refines the structural elements identified by the single component inversion. This is particularly advantageous when data are at discrete locations. The additional components provide secondary information about how the geology is changing around a given observation location in the absence of a perfectly continuous, noise-free data map. From the progression shown in Figure 5.6, it can be seen that inversion of gravity gradiometry is a valuable tool for characterizing exploration targets.

To convey the 3D nature of the model, we display additional cross sections parallel to the previous cross sections in Figure 5.7. The cross sections and depth slice are from the five-component model. A volume rendered image of this density contrast model is shown in Figure 5.8, which displays regions with density contrast greater than 1.0 g/cc. For brevity, we have not reproduced the similar image for other inversions. However, all are highly consistent. The volume rendered image shows the shallow, dense bodies trending northeast-southwest through the survey area. The general distribution of high-density zones corresponds well
Figure 5.7: (a) Plan slice through five-component model at 1059m elevation with cross sections (b) - (f) denoted by white lines; (b-f) Cross sections from left (northwest) to right (southeast) through the five-component model. Dotted black line shows elevation of plan slice in (a).
with the known iron formation.

Figure 5.8: Volume rendered image of the 3D density contrast model constructed from five-component inversion with density contrast less than 1.0 g/cc removed. The high density bodies trending northeast to southwest agree with the occurrence of iron formation within the Gandarela Syncline. The semi-transparent topographic surface is shown above the 3D density contrast volume.

5.4 Conclusions

The ability to invert gravity gradient data for detailed deposit-scale interpretation and ore delineation has been demonstrated using a single component of a field dataset from the Quadrilátero Ferrífero. From the $T_{zz}$ component alone, we are able to delineate the iron ore formation. Incorporation of additional gradient components provides secondary information that serves to improve the overall structural quality of the recovered model. When inversion results are compared to a geologic section through the area, it is seen that all models are representative of the iron formation localized within the Gandarela Syncline. The models shown here were obtained through blind inversion of the data with little a priori knowledge incorporated into the process.

Each of the four recovered models can potentially be used for interpretation purposes. For a basic characterization of the anomalous source distribution, $T_{zz}$ alone sufficiently images the iron formation to a first order. While improvements are seen in the three-component model, structural interpretation would be more reliable using the five-component model. Investigations with other component combinations provide further insight into information
content through the analysis of data fit and model structure. The various component combinations show improvements to the defined dip, density contrast, and depth and lateral extent of the body.

In obtaining the results shown here, we have developed a procedure for processing and 3D inversion of gravity gradient data in such an application. The procedure that takes us from raw data to the density contrast models that are highly consistent with known geology requires a comprehensive set of steps from basic gravity gradient processing, terrain correction, noise characterization, and careful selection of inversion parameters.

5.5 Acknowledgments

We would like to thank Vale for providing the data used in this study. We also thank the current sponsors of the Gravity and Magnetics Research Consortium (GMRC) for supporting this work: Anadarko, Bell Geospace, BGP, BG Group, BP, ConocoPhillips, Fugro, Gedex, Lockheed Martin, Marathon Oil, Petrobras, Shell, and Vale. Cericia Martinez is also supported in part by a John Poate Fellow from the Colorado School of Mines. We thank Bill Morris and Nigel Phillips for constructive feedback in the review of this paper.
We present a study on utilizing airborne gravity gradient and magnetic data to characterize an iron ore formation in Minas Gerais, Brazil. The target iron ore bodies have a distinctly high density contrast and produce well defined anomalies in airborne gravity gradiometry data. The high grade hematite iron ores are associated with low and moderate susceptibility, making magnetic data useful in distinguishing potential ore bodies from the host iron formation. The airborne gravity gradient and magnetic data over part of the Gandarela Syncline iron formation in the Quadrilátero Ferrífero are independently inverted to obtain a 3D susceptibility and density contrast model. These detailed 3D physical property distributions of subsurface features are then used for geologic characterization and interpretation purposes through lithologic associations. We outline two approaches to link the two physical property distributions and identify representative geologic units in the study area. The geologic units are then organized into a 3D lithology model in order to help characterize subsurface geologic structure and ore distribution. The lithologic models provide intuitive representation of geology and can assist in future exploration plans or in assessment of resource distribution and quality. Our study demonstrates that such approaches are feasible on the deposit scale.\footnote{Content within this chapter has been published in Interpretation (March 2015).}

6.1 Introduction

The Quadrilátero Ferrífero, or Iron Quadrangle (Figure 6.1), is an area of significant mineral resources in the state of Minas Gerais, Brazil. The Quadrilátero Ferrífero
covers approximately 7,000 square kilometers towards the eastern edge of Brazil. Iron and gold are major mineral resources that have historically been produced from the area [65].

The iron formation of interest in this area is generally tabular and known to be extensive in one dominant lateral direction, requiring wide area coverage of geophysical data. The area has rugged terrain with the landscape composed of canyons, plateaus, and valleys. The occurrence of iron formation today is coincident with the rugged peaks, making ground geophysical exploration in this region difficult. In addition, the thick blanket of vegetation severely limits the mobility of equipment for both manual and vehicular surveys. Much of the remote areas are inaccessible on ground for a few months out of the year due to tropical rain storms. With the many obstacles which the environmental conditions impose, performing ground based geophysics is impractical considering the restrictions on transporting equipment.

To overcome these difficulties, considerable effort has been put into mapping the occurrence of iron formation in the region through outcrop and other surface expressions. As currently understood, the iron formation is characteristically shallow. Typical exploration strategy in the area prior to the introduction of airborne geophysics relied on the geologic knowledge resulting from such mapping investigations and extensive drilling programs. Drill holes and previous exploration campaigns were planned primarily based on the known surficial orientation of iron formation and the presence of canga, which is a cemented, iron-formation derived surficial layer.

The improved characterization of a target underlines the intent of any geophysical investigation and combining multiple data types for interpretation can significantly contribute to that aim. Given the difficulties of ground geophysics in the Quadrilátero Ferrífero, airborne potential field data can be useful in characterizing the near surface exploration targets present in the region. With the advent of accessible systems to measure gravity gradients, it is natural to combine gravity gradient and magnetic data for interpretation purposes where applicable. The high density contrast of iron ore in this area generates distinct anomalies in
airborne gravity gradiometry data. The susceptibilities associated with the high and interme-
diate grade ores provides a secondary response through magnetic data that may contribute
to an improved understanding of the target ore bodies.

There has been much work in exploiting the potential field relationships to extract more
geologic information from gravity and magnetic data. In 1970, Kanasewich and Agarwal
explored the validity of examining the magnetization to density ratio in the wavenumber
domain as an interpretation tool. Combining magnetic and gravity derived gravity gradient
data in the spatial domain through the Poisson relation has been accomplished by Dransfield
et al. [80] and Price and Dransfield [81] in order to generate pseudolithology maps based on
the ratio of apparent susceptibility to density. Following the same methodology, gravity
gradient and magnetic observations can be directly used to generate pseudolithologic maps
as in Braga et al. [82].

Beyond utilization of data to generate apparent susceptibility and density maps in the
data space, inversion of potential-field data can be explored for lithologic and geologic char-
acterization. Much effort has been focused on incorporating a priori information into the
inverse methodology in order to identify lithologic units and their structure. In 1999, Bosch
present a method to obtain lithologic distinctions by incorporating known petrophysical,
geostatistical, and structure information into the inversion algorithm. Lane and Guillen [84]
explore inversion guided by lithologic categories with density and susceptibility properties
being ancillary information. Williams and Dipple [85] explore mineral abundance estimates
with drill data and 3D property distributions obtained by inversion of magnetic and gravity
data with geologic reference models.

While these methodologies could be advantageous in well-established exploration fields
where a priori information is abundant, they may not be ideal in scenarios with a lack
of detailed geologic information. Toward this end, [86] use 3D inversions of magnetic and
gravity data to obtain regional susceptibility and density contrast models that were used to
divide the model region into distributions of lithology classes based on a scatterplot of the
Figure 6.1: Geologic map of the Quadrilátero Ferrífero, with sold black shading indicating where the Minas Series is present; modified from Dorr [3]. The survey area is outlined in red.
In this work, we utilize a similar approach to obtain a 3D lithologic model by relating susceptibility and density distributions obtained through independent inversion of magnetic and gravity gradient data from the Quadrilátero Ferrífero. In contrast to the work by Kowalczyk et al. [86], which focused on the regional scale, our work attempts to characterize lithology on the deposit scale. We refer to this work as lithology characterization, which may contribute to differentiating multiple geologic units and geologic structure. In addition to using known generic ranges of physical property values for different lithologic units, we also develop a new approach that integrates known geologic sections with the inverted physical property models. The first approach, using generic constraints, is advantageous in that it requires minimal to no prior site-specific knowledge of the target geology, whereas the second approach, using geologic constraints, is more adaptive to a specific site.

In the following, we first describe the site geology and the gravity gradiometry and magnetic data. We then discuss the inverse procedure used to obtain the density and magnetic susceptibility models. We next present two methods that use varying degrees of prior geologic knowledge to identify lithologic units to derive 3D lithology models. We conclude the paper by comparing the results with site geology and known ore distribution and a discussion on the applicability and limitations of such an approach.

6.2 Deposit Geology and Geophysical Data

The high grade ore (hematite) typically contains an average of 66% Fe with the intermediate grade ores (itabirite) ranging between 57% and 66% (average of 63% Fe). Itabirite is synonymous with banded-iron formation, the term originating in Brazil. The high-grade deposits are easily differentiated from the dolomitic and quartz-rich country rock by the stark density contrast. The densities of the high-grade iron ores range from 3 g/cc for soft, porous hematite to 5 g/cc for hard, compact hematite. Variations in the soft and hard hematite are due to the ore forming process where the soft hematite is thought to be the result of weathering of previously formed iron formation. The susceptibility of the iron ore in this area
is generally low for hard, compact hematite, slightly higher for that of soft, porous hematite, and higher still for that of economic itabirite.

6.2.1 Gravity gradient data

The data were collected in August-September 2005. The gravity gradient and total field aeromagnetic data were acquired simultaneously and therefore share similar survey parameters. The 93 square kilometer survey was acquired with 100 meter line spacing. Flight lines are oriented 32° from the north. Three tie lines were flown perpendicular to the main survey orientation. The survey was semi-draped with flight heights ranging from 60 to 500 meters above the topographic surface.

The gravity gradient data underwent routine proprietary processing and corrections for the centripetal force, self-gradient, and acceleration compensation by the acquisition company. Before data delivery, the lines were leveled and filtered to attenuate noise. We focus on a subset of the survey area displayed in Figure 6.2 in tensor format. For reference, the topography within the entire survey area is displayed in the lower left panel of Figure 6.2 where the black box indicates the area of subset. The elevation within the area ranges from 750 meters to 1350 meters, with an average elevation of 1014 meters.

The subset data covers approximately 20 square kilometers in the northern part of the survey area. The Cauê Mine is within this area and limited borehole data are available. Geologic cross sections generated from drill data are also available within this area. The geologic feature of interest runs through the middle of the gravity gradient data map parallel to the long axis of the survey area and is visually distinguishable within the data. The iron formation is coincident with topographic highs in the area, making the density value used to remove the terrain effect an important parameter. To obtain the displayed gradient anomaly of Figure 6.2, a density value of 2.67 g/cc was used to remove the terrain effect.
Figure 6.2: Terrain corrected gravity gradient data covering a 4 x 5 kilometer area with survey topography (f) showing the location of the data subset. From left to right and top to bottom, the plot panels display (a) $T_{xx}$; (b) $T_{xy}$; (c) $T_{xz}$; (d) $T_{yy}$; (e) $T_{yz}$; (f) topography; (g) $T_{zz}$. The location of the cross section of Figures Figure 6.9 and Figure 6.12 is overlaid on $T_{zz}$ in panel (g).
6.2.2 Aeromagnetic data

The total-field magnetic data used in this study are a subset of the aeromagnetic survey corresponding to the gravity gradient subset with an original line spacing of 100 meters. Processing by the acquisition company included leveling and microleveling as well as removal of the diurnal variation and the International Geomagnetic Reference Field (IGRF). To prepare the magnetic data for inversion, the IGRF corrected data is gridded and decimated, providing 12,037 data locations for inversion. The observed magnetic data are shown in Figure 6.3. The magnetic anomalies exhibit both broad and small scale features. The local fluctuations are related to the iron formation and need to be extracted prior to inversion for a susceptibility model.

![Figure 6.3: Subset area total-field magnetic anomaly prior to removal of the regional field.](image)

6.2.2.1 Regional-residual separation

Extraction of the magnetic anomaly related to the iron formation of interest is performed by regional-residual separation via an inversion based method [87]. We briefly summarize
the steps taken to separate the local iron formation signal from underlying, deeper geology.

The subset of total-field data in Figure 6.3 is initially inverted on a coarse mesh in order to reproduce the regional features seen in the data. The coarse mesh consists of only four depth layers encapsulating the majority of the near surface volume of interest. Susceptibility is constrained to be within the range of 0 to 1 SI. Using generalized cross validation, a susceptibility model is obtained that optimally reproduces the data given the coarse discretization. The coarse mesh contains cell sizes of 250m in the easting, 250m in the northing, and 100m in depth. The mesh extends 10km in the easting and 12km in the northing to a maximum depth of 10km.

Two features of the recovered model are then exploited in order to separate signal from regional and local sources. First, it can be observed that the coarse inversion does not reproduce variations associated with the near surface geologic features of interest with the desired level of details. Second, the constructed model can be partitioned into local and regional sources in order to remove model contributions of near surface features to the predicted data.

The first observation is a natural consequence of discretization. The second observation is achieved by assigning all cells between the topographic surface and an elevation of 919 meters a susceptibility value of zero, while the rest of the model is left unchanged. Through this step, we are effectively removing the near-surface susceptibility sources from the model. The predicted data from this altered model now represents deep, regional sources contributing to the regional trends in the data. The predicted data from this altered susceptibility model is now considered the regional trend that we wish to remove from the observed data in order to obtain the response of the near surface geologic features. The calculated regional trend is subtracted from the observed data and we use this residual magnetic data (Figure 6.4) to invert for a shallow, near surface susceptibility model.

### 6.3 Lithologic Mapping

In order to perform lithologic differentiation, we first construct a density and susceptibility model by inverting the gravity gradient and magnetic anomalies, respectively. We use the
3D potential field inversion algorithm developed by Li and Oldenburg to invert the gravity gradient [19] and magnetic data [20, 88] independently. In the brief description that follows, the vector $\vec{m}$ represents density, $\vec{\rho}$, when $\vec{d}$ is gravity gradient data and susceptibility, $\vec{\kappa}$, when $\vec{d}$ is magnetic data. For the general construction outlined here, we use a right hand Cartesian system where x is northing, y is easting, and z points vertically down. The algorithm assumes a set of contiguous rectangular prisms each with a constant physical property. The model is obtained via the inverse solution of the forward modeling equation

$$G\vec{m} = \vec{d} \tag{6.1}$$

where $\vec{m} = [m_1, m_2, ..., m_m]^T$ is the model vector containing the physical property values of the cells, $\vec{d} = [d_1, d_2, ..., d_n]^T$ is the data vector, and $G$ is the sensitivity matrix that quantifies the contribution of cells in the model to the data value at the given observation locations. The inverse solution is obtained using Tikhonov regularization by minimizing a total objective function with bound constraints,
\[
\begin{align*}
\min & \quad \phi = \phi_d + \mu \phi_m \\
\text{subject to} & \quad \vec{m}_{\min} \leq \vec{m} \leq \vec{m}_{\max}
\end{align*}
\] (6.2)

where \( \mu \) is a regularization parameter, \( \phi_d \) and \( \phi_m \) are data misfit and model objective functions, respectively, \( \vec{m}_{\min} \) and \( \vec{m}_{\max} \) contains the lower and upper bounds, respectively, on the physical property distribution. The ability to incorporate lower and upper bounds on the recovered model is useful in the case of gravity gradiometry since we expect a large positive density contrast associated with the iron formation. In the magnetic case, we impose positivity on the recovered susceptibility model because we only expect positive susceptibility assuming a general zero background.

The density value used to terrain correct the gravity gradient data is added to each cell of the density contrast model to obtain density rather than density contrast. The 2.67 g/cc terrain correction density restores the background density associated with non-target geology initially removed from the data and provides a density model that general physical property values are more readily compared to. Our analyses have indicated that no significant remanent magnetization is present and total-field inversion is employed. The recovered susceptibility model from the total-field anomaly can be treated as the susceptibility value of the rock units, as we assume that the general background has a zero susceptibility.

We apply an existing method and develop a new one for assigning lithologic units based on the recovered density and susceptibility values. The first method uses general physical property values of specific rock types (in particular hematite and itabirite) to assign lithologic units from the recovered density and susceptibility models. This method requires little knowledge of the geologic structure, but assumes that the expected types of lithology are known. The second method uses a geologic cross-section generated from borehole data to identify lithologic units based on the correspondence between the lithologic units present in the cross-section and the recovered physical property values. This method requires detailed information about the geologic structure in one or more cross-sections or plan-sections and
seeks to extend that knowledge to a full 3D geologic model. Before describing the two lithology methods in greater detail, we first summarize the inversion process by which the density and susceptibility models are obtained.

### 6.3.1 Density contrast and susceptibility models

To allow for model association during lithologic mapping, the same mesh discretization is used for both the gravity gradiometry and magnetic data inversions. The mesh is composed of rectangular prisms (referred to as cells henceforth) with constant density contrast and susceptibility. Cell sizes are 25 m in the easting by 25 m in the northing by 20 m in depth in the central region of the mesh, with padding cells beyond the data area and at large depth. The rectangular mesh has dimensions of 156 cells in the easting by 241 cells in the northing by 45 cells in depth. A topographic surface is included in the inversion so that all cells above the topographic surface are excluded from the model region.

Inversion parameters used to obtain the density contrast and susceptibility models are summarized in Table 6.1. Six measured gravity gradient components ($T_{xx}$, $T_{xy}$, $T_{xz}$, $T_{yy}$, $T_{yz}$, and $T_{zz}$) were simultaneously inverted to obtain a representative density contrast model of the target subsurface geology. The density contrast model was obtained by inversion of the 18,102 data points. Little a priori information was incorporated into the inversion. A zero density contrast reference model is used with lower and upper bounds placed on all model cells of 0.0 g/cc and 4.0 g/cc, respectively. The bounds were selected because a positive density contrast is expected from the dense iron ore with respect to the less dense host rock.

Table 6.1: Summary of constraining inversion parameters for obtaining the 3D density contrast and susceptibility models.

<table>
<thead>
<tr>
<th>Data Inverted</th>
<th>$m_{ref}$</th>
<th>$m_0$</th>
<th>$m_{min}$</th>
<th>$m_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Six-component gravity gradient</td>
<td>0 g/cc</td>
<td>2 g/cc</td>
<td>0 g/cc</td>
<td>4 g/cc</td>
</tr>
<tr>
<td>total-field magnetic</td>
<td>0.0 SI</td>
<td>0.001 SI</td>
<td>0.0 SI</td>
<td>1.0 SI</td>
</tr>
</tbody>
</table>
The correlation lengths in three directions are set to be $L_x = L_y = 50$ m and $L_z = 40$ m, requiring a relatively equal amount of flatness in each direction. For further discussion on obtaining the density contrast distribution at this site, see [64]. A volume rendered image of the density contrast model selected is shown in Figure 6.5(a) with model cells below 1.13 g/cc removed for clarity.

![Figure 6.5: (a) 3D density contrast model. Cells with density contrast less than 1.13 g/cc are removed; Distance is in meters. (b) 3D susceptibility model. Cells with susceptibility less than 0.15 SI are removed.](image)

The susceptibility model was obtained by a similar inversion of the total-field data. A zero reference model was used with an initial model of 0.001 (SI). Lower and upper bounds of 0.0 and 1.0 were placed on the model to keep the recovered values within a reasonable range of positive susceptibilities. A volume rendered image of the recovered susceptibility distribution is shown in Figure 6.5(b) with model cells below 0.15 (SI) removed for clarity.
6.3.2 Lithology assignment using generic constraints

We first examine lithology characterization by using general physical property information expected based on the assumed rock types present. We use physical property values for general iron rocks from Hunt et al. [89] and appropriately scale the range to correspond with the range of recovered density and susceptibility values. In obtaining both a density contrast and susceptibility model, minimal prior information specific to this site was employed. Therefore, any association between density, susceptibility, and lithologies we can extract are based purely on the two potential field datasets and represent the minimum information content in the data in regards to the lithology. Such a result also provides an indication of the base level lithology differentiation we can accomplish without using any site-specific geological information.

We begin by examining the correlation between density and susceptibility obtained from inversions. A scatter plot of the corresponding model cell values from the susceptibility and density models of Figure 6.5 is generated to furnish lithologic unit assignments across multiple physical properties. The process of assigning a lithologic unit to a given cell in the model accomplishes the objective of lithology differentiation. The scatter plot for the susceptibility and density models is shown in Figure 6.6(a). Each point on the scatter plot denotes a single cell from the 3D model defined by the mesh discretization described earlier. Within this region, the high grade iron ores (high density) are associated with low and moderate susceptibility. To better visualize the density ($\rho$) and susceptibility ($\kappa$) relationship in areas of high point- (or cell-) concentration in the scatter plot, four contour lines are overlaid on the scatter plot in Figure 6.6(b). The contour lines show the binned number of cells in rectangular windows with a size of $\delta \kappa = 0.0028$ and $\delta \rho = 0.0119$ g/cc on the scatter plot. Through the contours, it can be seen that a significant number of cells correspond to low density and low susceptibility.

The scatter plot is characterized by a concentrated region of low susceptibility and low density cells. There are a few visible clusters extending outward from the main cloud of
points, particularly at large density values. From the two independent physical property models, visually extracting any type of correlation or clusters indicative of different rock types is difficult. The majority of the model cells represent lithologic units of low density and low susceptibility. In order to obtain a better understanding of the density and susceptibility trends, we examine two logarithmic scatter plots in Figure 6.7. Figure 6.7(a) uses the absolute density model and emphasizes trends at high density whereas Figure 6.7(b) uses the density contrast model and better highlights trends at low density.

From these initial plots, we observe that the density and susceptibility models appear to contain first- and second-order trends. An example of the trends can be seen in Figure 6.7(a) where high density tends to occur in regions of low and high susceptibility, but not in the mid-range. This is also supported by the curved clusters seen on the low susceptibility range and the spires on the high susceptibility range of Figure 6.7(a). These second-order trends demonstrate the relation of physical property values that represent a lithologic unit when taking into account the smoothing imposed by the algorithm. Since the two models were obtained independently, any lithologic relation illustrated through the crossplot will be dispersive rather than compact. With this understanding, these first and second order trends are evidence of related physical property occurrences.

Using general physical property values, we identify key lithologic units that may be representative of rock types at the site. In particular, we focus on the distribution of hematite (low susceptibility, high density), itabirite (moderately high density, moderately high susceptibility), and the host rock for the iron formation (low density, low susceptibility). These general physical property bounds are scaled to overlap with the range of values seen in the recovered models. The bounds for each of the lithologic units are given in Table 6.2. These lithologic bounds are illustrated in the scatter plot in Figure 6.8(a). The rock units of interest here are hematite and itabirite and we focus on hematite as the main unit for presenting results of the two methods. The hematite assignment is colored pink (upper left unit on the scatter plot), itabirite and similar iron formation are denoted by dark and light
Figure 6.6: (a) Scatter plot of susceptibility model versus density model. Each point represents the same cell location from the density and susceptibility models. (b) Contour lines illustrating cell concentration for a specified binning window.
Figure 6.7: (a) Logarithmic scatter plot of susceptibility model versus density model ($\kappa$ vs. $\rho$) with background density of 2.67 g/cc added back in, showing cluster trends in the higher density range. (b) Logarithmic scatter plot of susceptibility model versus density contrast model ($\kappa$ vs. $\Delta \rho$), showing trends in the lower density range.
blue (upper right and left middle unit), and the host rock is identified by brown (lower left unit).

![Figure 6.8](image)

(a) Scatter plot of density model versus susceptibility model with lithologic units color coded. (b) Volume rendered image of the 3D lithologic model using the assignments displayed in the scatter plot.

Figure 6.8: (a) Scatter plot of density model versus susceptibility model with lithologic units color coded. (b) Volume rendered image of the 3D lithologic model using the assignments displayed in the scatter plot.

With these physical property assignments, we are able to generate a 3D lithology model shown in Figure 6.8(b) that spatially corresponds to the density and susceptibility models. The geologic section given in Figure 6.9(a) provides reference for evaluating our lithologic assignments. In this geologic section, the hematite is represented by the pink units while
Itabirite is represented by the blue units. For additional comparison, the corresponding cross section from the individual density contrast and susceptibility models are shown in Figures 6.9(b) and 6.9(c), respectively. A cross section through our constructed lithologic model is shown in Figure 6.10(b). For consistency, we have maintained the same color scheme for the lithologic units in the cross plot and the lithologic model. The geologic outline of hematite and itabirite are marked on the lithologic, density, and susceptibility cross sections.

Table 6.2: Specified density and susceptibility bounds and associated lithologic units using average petrophysical values.

<table>
<thead>
<tr>
<th>Rock Unit</th>
<th>Color</th>
<th>$\kappa_{min}$</th>
<th>$\kappa_{max}$</th>
<th>$\rho_{min}$</th>
<th>$\rho_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minas Piracicaba</td>
<td>Brown</td>
<td>0.0</td>
<td>0.04</td>
<td>0.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Phyllite</td>
<td>Tan</td>
<td>0.04</td>
<td>$\kappa_{max}$</td>
<td>0.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Itab. Dolomite</td>
<td>Teal</td>
<td>0.1</td>
<td>$\kappa_{max}$</td>
<td>3.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Itab. Ocre</td>
<td>Dark Tan</td>
<td>0.02</td>
<td>0.1</td>
<td>3.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Itab. Friable</td>
<td>Light Blue</td>
<td>0.0</td>
<td>0.02</td>
<td>3.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Itab. Compact</td>
<td>Dark Blue</td>
<td>0.04</td>
<td>$\kappa_{max}$</td>
<td>3.8</td>
<td>$\rho_{max}$</td>
</tr>
<tr>
<td>Hematite</td>
<td>Pink</td>
<td>0.0</td>
<td>0.04</td>
<td>3.8</td>
<td>$\rho_{max}$</td>
</tr>
</tbody>
</table>

A cursory look at the cross sections in Figure 6.9 and Figure 6.10(b) may lead one to conclude that the density bounds from the scatter plot are more influential than susceptibility bounds. The dominant features within the lithologic section, Figure 6.9(a), appear to be largely consistent with the density model, Figure 6.9(b), with little to no likeness to the susceptibility model, Figure 6.9(c). However, this is not true in general, especially when one examines the classification in the entire 3-D volume. Classification based on density alone can lead to a gross overestimate of total ore volumes.

This is demonstrated by first examining the full 3D lithologic volume that the cross section of Figure 6.10(b) was extracted from. A 3D image of the model showing only the distribution of hematite cells is given in Figure 6.10(a). This image is a direct display of the cells classified as hematite and no thresholding is used in the volume rendering. For comparison, we then remove the susceptibility constraints on the hematite, and thus define
Figure 6.9: (a) Geologic cross section generated from bore hole data showing itabirite in blue and hematite in pink. (b) Cross section through density contrast model; units are g/cc. (c) Cross section through susceptibility model plotted on a logarithmic color scale; units are SI.
Figure 6.10: (a) Hematite distribution corresponding to lithologic model from generic constraints. (b) Cross section through lithologic model from generic constraints corresponding to the color scheme identified in Figure 6.8(a).
the lithologic unit based solely on the upper and lower bounds of density. This alternate scenario is illustrated in the scatter plot of Figure 6.11(a), where the pink hematite unit has no susceptibility constraints and encompasses the full upper part of the scatter plot. The volume rendered image of the resulting 3D hematite distribution from this alternative lithologic scheme is shown in Figure 6.11(b). Comparing the two volumes, it is visually evident that the lack of susceptibility constraints results in a more continuous and abundant distribution of hematite, Figure 6.11(b), as opposed to the more sparse hematite distribution where susceptibility constraints are used, Figure 6.10(a). The structure of this altered hematite distribution is similar to the recovered high density contrast volume, Figure 6.5(a), but has a much smaller spatial extent. However, knowledge of the site’s geology indicates that the total hematite volumes may be significantly overestimated using this second lithology scheme where susceptibility is ignored. Therefore, both recovered density and susceptibility constraints are required to generate a practical 3D lithologic model of the site.

6.3.3 Lithology assignments using geologic constraints

The second method uses a geologic section as a training section to guide lithologic assignments. The known geologic structure and corresponding recovered models are used to define regions, where the lithologic bounds can then be defined specifically for the local geology. We use the geologic cross section provided in Figure 6.9(a) to obtain physical property bounds based on the recovered density and susceptibility model cell values. This technique is illustrated in Figure 6.12 using a 2D slice from the recovered density model. The top panel of Figure 6.12 illustrates the digitized units from the geologic section. The bottom panel shows the same cross sectional area extracted from the density model at discrete locations (cell centers). The middle panel shows the geologic section from the top panel sampled at the same discrete locations as the density model (bottom panel). From this sampling, a geologic unit is now associated with each discrete element of the extracted density section. The same approach is used to assign a lithologic unit to the same discrete locations extracted from the susceptibility model.
Figure 6.11: (a) Scatter plot defining lithologic units with no susceptibility constraint on hematite distribution (pink). (b) Overestimate of hematite distribution when no susceptibility constraints are applied.
Figure 6.12: The 2D geologic section in the top panel is sampled at discrete locations, shown in the middle panel, that corresponds to the cell centers, bottom panel, extracted from the recovered density contrast model. Every model cell coinciding with the cross sectional area now has a lithologic unit associated with the density value of that cell.
For every model cell, there now exist three pieces of information (density $\rho$, susceptibility $\kappa$, and geologic unit) for the 2D cross sectional area. The geologic unit associated with the two physical properties can now be overlaid on the crossplot as in Figure 6.13(a) and Figure 6.13(b). From here, the user can now define lithologic bounds based on the grouping of geologic units and their associated recovered physical properties. In this manner, the geologic units are now defined solely by the range of recovered model values. Likewise, the physical property bounds for each lithology are now a function of the spatial geometries permitted in the geologic cross section.

The polygonal bounds are based on the localization of geologic units on the scatter plot and requires the user to identify regions delineating geologic units. The method outlined above does not provide distinct bounds within the scatter plot, but rather identifies compact regions that represent lithologic units. However, assigning lithologic units in this manner is advantageous in that it addresses an often overlooked limitation of geophysical inversion. The 3D generalized inversion formulation using an $l_2$ norm results in a smooth model regardless of the true nature of the geologic source. Since the models used are smooth, small-scale jumps between assigned lithologies are not expected because the boundary is abrupt rather than transitional. Furthermore, the recovered physical property values often bias towards the reference models and may not be consistent with the general physical property values for use in classification. With this method, however, the model bias does not pose a direct difficulty since we do not compare the recovered density and susceptibility values to the true, measured values of specific rock types. Instead, the lithologic regions are now defined by the range of recovered physical property values corresponding to known geologic units. Therefore, the known geology section acts as the template instead of a generic table of physical property values. The errors inherent in the inversions are mitigated to a large extent by using such a site specific template.

We note that the quality of the lithology model derived by this approach is highly dependent upon the reliability of the geologic cross-section used. If the geometries of the
Figure 6.13: (a) Geologic units associated with the 2D cross sectional areas extracted from the density and susceptibility model overlaid on the scatter plot of the two physical property models. (b) For visual clarity, the same data is plotted on a logarithmic scale using the density contrast model.
structure are erroneous in the geologic section, these inaccuracies will propagate into the lithologic model. Geologic sections are often an interpreted and interpolated version of the true geology through surface observation and drill hole logs, so some level of errors are to be expected. Thus, the accuracy of the lithology differentiation is highly dependent on the errors in the section used. Moreover, our study only has one cross-section available, and its representativeness may be limited. Multiple cross-sections and plan-sections will provide a better sampling of the 3D geology and should improve the quality of the resulting lithology models. Similarly, if substantial and representative petrophysical data are available, they will directly contribute to improved lithology differentiation models.

The scatter plot with lithologic assignments using the known geologic section is shown in Figure 6.14(a). The lithologic section using this approach is shown in Figure 6.14(c). The lithologic section identifies similar overall lithologic units within the area. The difference in the two techniques is noted by the change in the structure of both the hematite (pink) and itabirite (blue). The hematite body covers a larger area while the region designated as itabirite is decreased. We provide a volume rendered image of the hematite distribution across the entire lithologic model in Figure 6.14(b). As seen by the constrained distribution of hematite, using the known geologic section as a reference has a noticable impact on the characterization of this particular unit.

6.3.4 Discussion

Based on the inversions of gravity gradiometry and magnetic data, we have obtained three different lithologic models: (1) a partial lithology model identifying possible hematite distributions from density values alone (Figure Figure 6.11(b)), (2) a generic lithology model by assigning different lithologic units using general density and susceptibility bounds for different rock types known to be present in the study area (Figure Figure 6.10(a)), and (3) geology constrained lithology model using density and susceptibility bounds identified using geological cross-sections in the study area (Figure Figure 6.14(b)). The partial lithology model only identifies possible distribution of hematite because it is the unit with the highest
Figure 6.14: (a) Scatter plot defining lithologic units with geologic density and susceptibility constraints. Hematite is pink. (b) Hematite distribution corresponding to lithologic model using geologic constraints. (c) Cross section through lithologic model using geologic constraints corresponding to the color scheme identified in Figure 6.8(a).
density at this site and identifying it using highest recovered density values is reasonable. Examining a depth slice through each model provides additional insight into the differences between the lithology differentiation methods. Figure 6.15 shows the three different lithology models in a common plan-section at an elevation of 1150 m. The left panel displays only the distribution of hematite from the partial lithology model, the middle panel shows all lithologic units from the generic lithology model, while the right panel displays all lithology units from the geologically constrained lithology model. Two major observations can be made from these models.

First, the locations of the predicted hematite ore bodies are consistent among all three models, and there is a consistent refinement from the the partial model to the generic lithology model to the lithology model constrained by known local geology. For example, the two large hematite bodies in the partial model based on density alone are eliminated in the two lithology models incorporating magnetic susceptibility. The difference with the partial model is expected since the magnetic susceptibility brings in additional information and the corresponding refinement will exclude volumes that are simultaneously high density and high susceptibility, which is contradictory with the physical property of hematite. The predicted hematite in the latter two models are identical in general locations but differ in the sizes and shapes. Such consistency gives us confidence in the predicted hematite distributions. As a quantitative measure, we have calculated the total volume of predicted hematite in each lithology model. The result is shown in Table Table 6.3. As noted from visual inspection, the estimated volume decreases as more information is incorporated into the differentiation. There is a 2.7-times reduction in volume of hematite from the partial model to the generic model, while the reduction from the generic to the geology-constrained model is only 20%.

Secondly, the general distributions of other lithologic units besides hematite are highly consistent in the two models incorporating the recovered susceptibility. In particular the three different types of itabirite exhibit the same general distribution and orientation. As with the hematite, only the shape and size of each unit at the same location differ between
Table 6.3: Total volume of hematite from three different lithology models.

<table>
<thead>
<tr>
<th>Lithology model</th>
<th>Estimated hematite volume (m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial model - $\rho$ only (Figure 6.11(b))</td>
<td>$1.292 \times 10^8$</td>
</tr>
<tr>
<td>Generic model - $\rho$ and $\kappa$ (Figure 6.10(a))</td>
<td>$4.663 \times 10^7$</td>
</tr>
<tr>
<td>Geology-constrained model (Figure 6.14(b))</td>
<td>$3.765 \times 10^7$</td>
</tr>
</tbody>
</table>

the two models. Overall, the two models obtained by using two independent differentiation scheme have yielded the same lithology model to the first order. This consistency again lends credence to the lithology differentiation results by using physical property values from a generic table or from correlating with local geology. The consistency is also a verification of the validity in using generic physical property to perform lithology differentiation.

Figure 6.15: A plan-section from 3D lithologic models at an elevation of 1150 m. Left: partial lithology model showing hematite distribution where susceptibility constraints are not used. Middle: generically constrained lithology model showing distribution of all assigned lithologic units. Right: geologically constrained lithology model showing distribution of all assigned lithologic units.

6.4 Conclusions

We have investigated lithology characterization for an iron ore deposit by using airborne gravity gradient and magnetic data from the Quadrilátero Ferrífero. The gravity gradient and magnetic data are inverted to obtain independent density and susceptibility models. We have applied one existing approach and developed a new method to identify lithologic units
based on the recovered physical property models. The first method uses general density and susceptibility values for specific rock types while the second relies on the known geologic structure. Both techniques produce reasonable lithologic models. The quality and confidence in potential interpretation of the model is naturally increased as more prior knowledge is incorporated. In areas lacking ground data, the similarities of the two models illustrates the advantages of acquiring multiple geophysical datasets to assist in initially characterizing the quality of a target.

The first lithology characterization method requires little site-specific geologic knowledge and uses general physical property values of commonly known rock types. The generic physical property values result in a lithologic model consistent with a geologic cross section of the iron formation. By applying both density and susceptibility bounds, the extent of hematite in the region is better constrained compared to the hematite volume using only density bounds.

The second method relies on the availability of a geologic section to extract physical property bounds directly from the geometry of the known geology. The resulting model proves to be consistent with both the geologic section and lithology model obtained via the first method. The hematite distribution is further constrained by this geologically guided approach. Incorporating known geologic structure and utilizing recovered physical property models also alleviates the model smoothing and possible bias introduced by the inversion algorithms. The consistency between lithology models obtained from these two approaches validates the use of generic physical properties for lithology differentiation.

Utilizing models obtained from geophysical inversion has the potential to be a powerful interpretation tool, particularly combined with other geologic information. The methods presented here illustrate that it is possible to combine multiple physical property models to help characterize lithology and subsurface structure. Through the geologic sections, partial verification is possible for the deposit scale interpretation at this site. Preliminary models such as the susceptibility distribution, density distribution, and lithologic character may
prove useful in planning further exploration of an area.

6.5 Acknowledgments

We would like to thank Marco Antonio Braga and Vale for providing the data used in this study. We also thank the sponsors of the Gravity and Magnetics Research Consortium (GMRC) for supporting this work: Anadarko, Bell Geospace, BGP, BG Group, BP, ConocoPhillips, Fugro, Gedex, Lockheed Martin, Marathon Oil, Petrobras, Shell, and Vale.
Incorporation of geophysical models at all stages of the exploration and extraction of a mineral resource is often not done in a quantitative fashion. If a deposit model does incorporate knowledge based on a geophysical model, it tends to be a qualitative contribution. As geophysical tools and interpretation methods improve, so do the quality of the geophysical models and their potential to contribute to modeling a resource. While there are multiple potential approaches to integrating geophysical models into the quantitative assessment of a mineral resource, here we focus on the use of physical property models as a supplemental variable in cokriging for estimation of mineral content.

7.1 Introduction

Exploration for mineral deposits have long used geophysical methods [90]. On the contrary, the use of geophysics for mineral content estimation is less common. The use of geophysical techniques in resource classification is gaining attention in the oil and gas industry as it provides insight into structural controls and physical properties in the subsurface [91–95].

The incorporation of geophysical data in the form of well logs has been identified to assist in mineral resource classification and deposit modeling [96–98]. Integrated modeling using well logs and surface geophysical techniques to map structural contacts has also been shown to contribute to deposit modeling [99].

The 3D physical property distributions obtainable from inversion of potential field data can potentially contribute to mineral resource assessment through models for mineral content estimation. The model selected for understanding mineral content and grade distribution is
a building block for grade-tonnage curves, local and global reserve estimation, and resource classification.

In general, a model of the mineral resource grade is generated using assay data, though assay data is not always available through the entirety of the model area. Various methods exist for estimating grade throughout the model area [100–102] where the approach taken is driven by the data.

Physical property models derived from geophysical data can be incorporated in geostatistical modeling of grade distribution. Grade models often rely on the use of drill hole data and other geological information in order to develop a model of the mineral resource. Such models require a significant drill hole sample size for optimal results. In some instances, the number of drill holes available within an area can be either sparse or concentrated to intersect a specific geologic feature of interest.

Physical property models such as a density model can potentially be used to assist resource model construction in areas with limited drill data. One method of particular interest for modeling grade distribution is kriging [103]. Ordinary and simple kriging allow for modeling of a single variable, while cokriging allows for multiple supplementary variables in the estimation of a single variable [104, 105].

Economic profitability and mine planning are to some extent dependent on modeling and simulation of the mineral resource. In particular, the selective mining unit (SMU) will contribute to the determination of economic profitability. The SMU is the smallest volume of rock mined that must be identified as either waste or ore. Geophysical models may be useful in helping to identify the geologic units or physical properties associated with each SMU.

The particular definition of cross-variograms implemented in cokriging here has some unique properties that may benefit not only the mineral industry, but the petroleum industry as well. A commonly cited drawback to the use of cokriging is that (1) the variables must be sampled at the same location and (2) the variables must have a commonly defined unit.
(where the secondary variable is usually defined as a proxy for the primary variable). The use of a variance based cross-variogram can overcome both of these hinderances providing more flexibility in the use of a secondary variable in the application of cokriging. For example, in the petroleum industry a proxy for permeability can be derived from well logs or core samples. While the units may align, the actual permeability values may be representative of slightly different measurement phenomena. Rather than a direct increment relationship between core and wireline derived permeability, a scaled relationship can be used providing a correlation less susceptible to differences associated with the various measurement and conversion techniques.

In this work, we first discuss kriging, semi-variograms, and the use of cross-variograms for cokriging. An synthetic example from the Beltana Zn-Pb ore deposit in Australia will then be used to illustrate the application of cokriging for mineral content using a density model obtained from inversion of gravity gradient data.

7.2 Kriging and Cokriging

Kriging and cokriging as an estimation procedure exploit the regional spatial relationship of the data. For this work, the semi-variogram and cross-variogram are used to model the spatial variance and covariance structure used in estimation. In the discussion that follows, the observed data for which estimation is to take place is termed the primary variable. The supplementary data that can be used to assist in describing the spatial dependence structure is termed the secondary variable.

Kriging of a single (primary) variable requires modeling of a single semi-variogram to describe the spatial structure observed within the data. The use of a secondary variable in cokriging requires additionally modeling the semi-variogram of the secondary variable and the cross-variogram of the primary and secondary variable. Inclusion of auxiliary variables (tertiary, quaternary, etc.) in the cokriging procedure requires modeling of their semi-variograms and the associated cross-variograms.
In the mining community, cokriging is used with caution. As the number of variables included in cokriging increases, so does the number of semi- and cross-variograms that need to be modeled. The additional modeling requires substantial time and effort and has the potential of negatively impacting estimation of the primary variable if care is not taken to properly select variogram models.

In the following sections, cokriging is presented in the context of two variables: the primary variable (mineral content) and the secondary variable (density model). The primary variable is taken to be at discrete points while the secondary variable is taken to be a contiguous model that can be represented by discrete points or block volumes. The distinction between discrete point estimation versus an area or volume estimation is important and addressed in the discussion section. The approach taken for estimating the semi-variogram, cross-variogram, and modeling of the variograms are summarized next.

7.2.1 Semi-variogram

Many methods exist in order to generate the empirical semi-variogram from sample data. Here we consider two methods: the Matheron Estimator and the Cressie-Hawkings Estimator [106]. The Matheron Estimator is easily computed and provides unbiased estimates. The Matheron semivariogram is defined in equation 7.1

\[
\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} \{Z(s_i) - Z(s_j)\}^2
\]

(7.1)

where \(h\) is the lag distance, \(s_i\) and \(s_j\) are locations, and \(Z\) is the spatial process. Since the estimator is defined as the average of squared differences, it can be affected by outliers.

The Cressie-Hawkins estimator is defined as

\[
\hat{\gamma}(h) = \frac{1}{\left[ \frac{1}{|N(h)|} \sum_{N(h)} |Z(s_i) - Z(s_j)|^{(1/2)} \right]^4} 0.457 + \frac{0.494}{N(h)}
\]

(7.2)

Taking the square root of the differences and raising them to the fourth power diminishes the affect of outliers. Though the estimator is no longer considered strictly unbiased, the additional term makes the estimates approximately unbiased. For ordinary kriging, using
the Cressie-Hawkins estimator can be beneficial so that the presence of outliers does not overly affect semivariogram modeling. For simplicity and comparison purposes, the Matheron Estimator is used in the synthetic example that follows.

### 7.2.2 Cross-variogram

There are two common methods used to calculate the empirical cross-variogram. An in-depth discussion on the two approaches can be found in [107]. The first (and more common) empirical cross-variogram is defined in equation 7.3.

\[
\hat{\nu}_{1,2}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} \{(Z_1(s_i) - Z_1(s_j))(Z_2(s_i) - Z_2(s_j))\} \tag{7.3}
\]

This definition of the cross-variogram requires the primary and secondary variables to be sampled at the same data locations and be consistent in units from process \(Z_1\) to \(Z_2\).

An alternative definition of the cross-variogram is given in equation 7.4. This alternative approach does not require the primary and secondary variables to be sampled at the same locations.

\[
\hat{\gamma}_{1,2}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} \{Z_1(s_i) - Z_2(s_j)\}^2 \tag{7.4}
\]

The use of either of these definitions assumes that the primary variable \((Z_1)\) has constant mean and the secondary variable \((Z_2)\) also has constant mean. In order to utilize the cross-variogram in 7.3 for estimation it must be assumed that \(\hat{\nu}_{1,2}(h) = \hat{\nu}_{2,1}(h)\), though this equality is not necessarily true. No such assumption must be made on the equivalence of \(\hat{\gamma}_{1,2}(h)\) and \(\hat{\gamma}_{2,1}(h)\).

While the flexibility of the cross-variogram of 7.4 is attractive, the units of measure for \(Z_1\) and \(Z_2\) may not necessarily be equal. In order to use the cross-variogram of 7.4, the variables are first standardized (see Cressie and Wikle [107]). Equations 7.5 and 7.6 define the standardized data where \(\bar{Z}_1\) is the sample mean and \(S_1\) is the sample standard deviation for the primary variable \((Z_1)\) and \(\bar{Z}_2\) is the sample mean and \(S_2\) is the sample standard deviation for the secondary variable \((Z_2)\).
\[ z_1(s_i) = \frac{Z_1(s_i) - \bar{Z}_1}{S_1} \]  
\[ z_2(s_j) = \frac{Z_2(s_j) - \bar{Z}_2}{S_2} \]  

(7.5)
(7.6)

Since geophysical models tend to be defined over a discrete area or volume, it is ideal to use the cross-variogram defined in equation 7.4 where the locations of the sample do not necessarily have to be at the same locations as the primary variable.

### 7.2.3 Variogram Modeling

Once a representative empirical semi-variogram or cross-variogram has been obtained, a variogram model can be fit to the data in order to assist in estimation of the primary variable at unobserved locations. In the example that follows, spherical and gaussian models with a nugget effect were used to fit the empirical variograms so only those models are given here.

The spherical model is given in equation 7.7 where \( c \) is the sill and \( a \) is the range.

\[ \gamma(h) = \begin{cases} c \left[ 1.5 \frac{h}{a} - 0.5 \left( \frac{h}{a} \right)^3 \right] & \text{if } h \leq a \\ c & \text{otherwise} \end{cases} \]  

(7.7)

The gaussian model is given in equation 7.8 where \( c \) is the sill and \( a \) is the range. The gaussian model typically describes highly continuous functions.

\[ \gamma(h) = c \left[ 1 - e^{-\frac{ah^2}{\alpha^2}} \right] \]  

(7.8)

Lastly, a nugget effect model given in equation 7.9 can be added to any variogram or nested variogram model where \( c_0 \) is the displacement from the origin. The nugget effect simply bulk shifts any variogram model, giving a new representative sill of \( c_0 + c \) for any sill \( c \) as defined in the models of equations 7.7 or 7.8.

\[ \gamma(h) = \begin{cases} 0 & \text{if } h = 0 \\ c_0 & \text{otherwise} \end{cases} \]  

(7.9)
7.3 Synthetic Example Based on Beltana Zinc Deposit

The Beltana deposit is a non-sulfide zinc-lead deposit north of Adelaide in South Australia. The main willemite mineralization is structurally controlled and trends north-west. The associated hematite-dolomite alteration zone contains small amounts of zinc as well.

In 2004, a modeling study was conducted based on detailed geologic cross sections to identify geophysical methods that could potentially detect the presence of the willemite ore body [108]. A 3D view of the geologic model is shown in Figure 7.1. The gravity, DC, and IP methods were identified as potentially useful.

![Figure 7.1: View of the 3D geologic model constructed by Krahenbuhl and Hitzman. The willemite lithology of interest is represented by the pink unit.](image_url)

Using the geologic model constructed by Krahenbuhl and Hitzman, a synthetic example for cokriging can be generated. From the lithologic units identified and associated densities, a representative zinc concentration is assumed for 35 random locations for a single bench. The samples were then scaled by a percentage of the value so the relationship between the primary and secondary variables are not exact. The zinc content for the 35 locations are
shown in Figure 7.2. The generation of random locations were restricted in the easting directions as they would be in practice given the known trend of mineralization. The true distribution of zinc content from which the 35 values were derived is given in Figure 7.3.

![Figure 7.2: Synthetic generation of zinc content for 35 random locations at an approximate elevation of 224m. The black dots represent values close to 0. Color units are zinc content.](image)

To obtain the secondary variable to be used in cokriging for zinc content at unobserved locations, the gravity gradient response of the Beltana deposit is calculated. Six-components are calculated at a constant barometric 335m with 50m line spacing and 8.5m sample spacing along line. Gaussian noise with zero mean and 5 Eo standard deviation is added to the calculated data. The synthetic observations are then inverted for a 3D density contrast model that can be used as a secondary control on the spatial structure in cokriging. The density contrast model selected is shown in Figure 7.4 consisting of 20m cubes in the core area.

Next, we utilize kriging and cokriging to predict zinc content at discrete points across the bench interval. The first step is to generate the empirical variograms and fit models to describe the spatial structure. The empirical semi-variogram and model fit for the zinc content is shown in the top panel of Figure 7.5. The empirical semi-variogram and model
Figure 7.3: True simulated distribution of zinc content based on geologic model of Figure 7.1.

Figure 7.4: Density contrast (g/cc) model resulting from inversion of synthetic six-component gravity gradient data contaminated with gaussian noise. All cells below 0.25 g/cc are removed to show recovery of the high density willemite.
fit for the density contrast depth slice is shown in the middle panel of Figure 7.5. The empirical cross-variogram and model fit for zinc content and density contrast is shown in the bottom panel of Figure 7.5. The variograms exhibit an interesting phenomena of decreased variability (high spatial dependence) at large lag distances \( h \). Given the nature of the area, where beyond the willemite ore deposit both the zinc content and density contrast tend toward zero, the indication of high spatial dependence in these near zero zones is not unexpected.

Kriging is first applied using the 35 locations with zinc content values. The semi-variogram model \( \gamma_1 \) is the solid blue line shown in top the panel of Figure 7.5. Estimates of zinc content are obtained at the same locations as the secondary variable to be used in cokriging. Prediction locations do not have to coincide with the secondary variable and is done so here simply for ease of visualization. The kriging estimates are shown in the left panel of Figure 7.6(a) with the kriging variance shown in the right panel (Figure 7.6(b)).

Using the density contrast as a secondary variable, cokriging is applied to the zinc data. The density contrast values are displayed in Figure 7.7. All three variogram models shown in Figure 7.5 are used. The cokriging estimates are shown in the left panel of Figure 7.8(a) with the kriging variance shown in the right panel (Figure 7.8(b)).

7.3.1 Comparison

At first glance, two things are immediately obvious. The first is that both kriging and cokriging recover a similar trend. The second is that the kriging variance is considerably higher than that for cokriging. The values estimated from kriging are generally higher than the estimates from cokriging. There is more spatial structure in the cokriging estimates (particularly in the west) than in the kriging estimates. This is a direct result of the local near-zero variability exhibited within the density contrast model beyond the mineralization zone.

For a quantitative comparison, the true zinc content was sampled at thirty randomly generated locations. The kriging and cokriging estimates for these thirty locations were then
Figure 7.5: Top: Empirical and modeled semi-variogram for zinc content. A spherical model is used with no nugget effect, a sill of 1.5, and a range of 150. Middle: Empirical and modeled semi-variogram for density contrast depth slice. A gaussian model is used with no nugget effect, a sill of 1.5, and a range of 200. Bottom: Empirical and modeled cross-variogram for zinc content and density contrast depth slice. A gaussian model is used with a nugget effect of 0.7, a sill of 0.7, and a range of 150.
Figure 7.6: (a) zinc content estimated from kriging of the 35 locations. (b) Kriging variance for the estimates in the left panel.

Figure 7.7: Density contrast (g/cc) used as a secondary variable for cokriging.
Figure 7.8: (a) zinc content estimated from cokriging of the 35 locations and density contrast of Figure 7.7. (b) Kriging variance for the estimates in the left panel.
obtained and the root mean squared difference calculated. The RMS error for the kriging estimates is 0.2362 while for the cokriging estimates it is 0.2223. An image of the absolute difference between the kriging and cokriging estimates is shown in Figure 7.9. The most prominent feature has a maximum difference upwards of 0.3.

![Image of absolute difference between the estimates obtained from kriging and cokriging. Colorbar is zinc content.](image)

Figure 7.9: Absolute difference between the estimates obtained from kriging and cokriging. Colorbar is zinc content.

To illustrate, cokriging is performed with the secondary variable only at locations where the density contrast is greater than 10% of the maximum recovered density contrast value. The cokriging estimates when utilizing less of the secondary variable are displayed in Figure 7.10(a). The estimated zinc content across the area has further decreased with a subset of the secondary variable utilized. There is now less structure in the areas outside of the deposit trend. The variance (Figure Figure 7.10(b)) has become more consistent across the area as well. The empirical variogram and their respective models for these alternative estimates are given in Figure 7.11. For the same thirty verification locations, the mean squared error for these estimates is 0.2221.

### 7.4 Discussion

Here, we restricted the application of cokriging to point estimates using a secondary variable representative of a single depth interval (or bench). Block kriging estimation rather than point kriging can be performed where the volume (selective mining unit) that can be
Figure 7.10: (a) zinc content estimated from cokriging of the 35 locations and density contrast locations greater than 10% of the maximum recovered density contrast value. (b) Kriging variance for the estimates in the left panel.
Figure 7.11: Top: Empirical and modeled semi-variogram for zinc content. A spherical model is used with no nugget effect, a sill of 1.5, and a range of 150. Middle: Empirical and modeled semi-variogram for density contrast depth slice. A gaussian model is used with no nugget effect, a sill of 1.3, and a range of 150. Bottom: Empirical and modeled cross-variogram for zinc content and density contrast depth slice. A gaussian model is used with a nugget effect of 0.2, a sill of 1.2, and a range of 150.
mined is taken into consideration. The primary variable is considered to represent a single point location in space. In practice, this is true as assays are typically done on a very small rock sample that may not be representative of the rock composition a few meters away. The support and information effect for these samples must be considered when estimation is performed for a volume rather than a point. However, in the example above the density model can be directly used as a secondary variable for block cokriging since the variable itself is defined in terms of volume.

An alternative cokriging method such as indicator cokriging could assist in differentiating lithologic units rather than mineral content. From the combination of multiple geophysical methods, geology differentiation can be defined according to how multiple physical properties vary with each other. The resulting geologic or lithologic model must satisfy multiple physics relationships. Indicator cokriging and simulation with lithologic models as a supplemental variable could assist in identifying the SMU range where classification of a volume as ore or waste begins to change. The ore and waste classification would depend on the lithology present which in turns depends on the ability of the bulk lithologic unit to satisfy governing physics equations.

In practice variogram modeling can take into account anisotropy and directional dependencies. The example here assumes isotropic variogram models for simplicity. While the example here is in 2D, cokriging can be performed in 3D. For 3D kriging, anisotropic variogram modeling will be important to properly model realistic geologic spatial dependencies.

Lastly, we note that the sample locations used here were randomly generated while in practice the availability of locations will be more structured and potentially clustered. The structured nature of borehole locations in exploration can be reflected in the statistical summary of the data. Further, the systematic or clustered nature of sample data represent an information effect. The presence of this effect requires declustering or another means of accounting for clustered data in statistical analysis.
7.5 Conclusions

The cokriging methodology identified does not require the primary and secondary variables to be coincident. This allows for more flexibility in the use of secondary variables that may not have been observed at the same locations as the primary variable. Cokriging using geophysical models as a supplemental variable can potentially contribute to assessing a mineral resource if a spatial correlation of the two (or multiple) variables exists. Further, the variables do not need to be in comparable units. To the first order, cokriging with a geophysical model as a supplement can provide additional information on the spatial continuity of geologic features. In mining, geophysical models verified by geologic evidence can further contribute by bridging areas where there is a lack of borehole measurements to areas where there are measurements through a quantitative and spatial relationship.

7.6 Acknowledgments

Thank you to Rich Krahenbuhl for sharing the Beltana deposit model for testing of this method and helpful discussions on implementation. This research was funded in part by the industry consortium “Gravity & Magnetics Research Consortium” (GMRC) at the Colorado School of Mines. The current sponsoring companies are Anadarko, Bell Geospace, BG-Group, BGP International, BP, CGG, ConocoPhillips, ExxonMobil, Gedex, Lockheed Martin, Marathon Oil, Micro-g LaCoste, Petra Energia, Petrobras, Shell, and Vale.
The work presented in this thesis has generated a workflow for the processing and interpretation of gravity gradient data. Aspects of processing that have been discussed include terrain correction density estimation, error estimation, and reliable component conversion. Inversion based interpretation considerations have illustrated the benefit of multiple components, additional geophysical data for lithology differentiation, and the potential contribution geophysical models can play at the resource evaluation stage.

While this work has addressed specific aspects of the practical application of gravity gradiometry for resource exploration, there remains questions to be answered and nuances still that need to be explored. Before my concluding remarks, summarizing the major contributions of my research, I first outline three main areas where further research could be fruitful: the use of spatial statistics for estimating the terrain correction density, equivalent source denoising, and integration of geophysical models into resource evaluation.

8.1 Future Research Directions

I presented a method for estimating the density to be used in the terrain correction, additional effort in understanding the spatial dependence of the gradient data may be useful in developing a more robust methodology. The work in Chapter 2 provides the foundation for which to begin testing the use of spatial statistics for such a purpose. Further evaluation on estimating the terrain correction density could include:

- Analysis of simple gravity gradient anomaly autocorrelation
- Alternative spatial statistics to select optimal density value
- Ability to identify spatially varying terrain correction for map interpretation purposes
The utility of equivalent source processing has been shown to be useful not only for denoising, estimating data errors, and component conversion but also for characterizing data errors associated with denoised data. With the framework developed, there is much that can be done to further understand the effect of various equivalent source parameters on the resulting error estimates and denoised data. Research questions yet to be answered include:

- Optimal depth and discretization of flat layer
- Pros and cons of using a flat or undulating layer following topography or flight surface
- Relationship between remaining error estimates and equivalent source construction parameters (e.g. discretization, wavelet compression)

A method for incorporating geophysical models into resource evaluation has been described. While the method presented utilizes cokriging and exploits the spatial information in the geophysical models to contribute to resource evaluation, there is a strong assumption that the physical property is highly correlated with the resource occurrence. Geophysical models and mineralization may not be directly correlated, and spatial continuity and structure may not be the controlling factors. Lithologic or similar types of models derived from geophysical data may be better suited for incorporation into resource evaluation. Potential directions on understanding the practical application of such integration include:

- Identification of physical property and mineralization relationships or proxies
- Case study examining feasibility of method
- Extension of the method to 3D with a synthetic illustration.

The listing above is in no way meant to be extensive and reflects the need to further investigate the novel contributions included in this thesis. To recap, a summary of the thesis work and contributions are highlighted.
8.2 Summary

In this thesis, I investigated terrain correction density estimation, error estimation, and reliable component conversion as well as interpretation through inversion of multiple components, lithology differentiation, and the incorporating geophysical models into the resource evaluation stage. The results of my research in these areas form a workflow for inversion based interpretation of gravity gradiometry. To invert gravity gradient data, current algorithms assume the data have been terrain corrected and data error estimates are available. Map-based interpretation and some inversion formulations require specific components that may need to be calculated from observed components. Interpretation of gravity gradient data need not end at the recovery of a 3D density contrast model and integration for lithologic differentiation and resource modeling can be achieved. The workflow is a natural development from investigation of these components of gravity gradiometry processing and interpretation.

In Chapter 2, a novel method for the selection of the density to be used in the terrain correction has been outlined. While the terrain correction is a common data reduction step, few methods exist with which to identify the optimal density value. The method is predicated on the fact that the gravity gradient data have a spatial dependence that can be quantified at each point within the data map. The use of spatial statistics to identify a representative background density has been illustrated through a field dataset and validated through two synthetic examples. The difficult example of a dipping dyke with coincident topography has been used to illustrate the promise of such a method in estimating the terrain correction density.

In Chapter 3, the equivalent source technique has been shown to be a reliable way to denoise data, quantify data errors, and perform component conversion. The technique is particularly advantageous for gravity gradient data as it can reproduce any combination of components simultaneously using a common equivalent layer, thus maintaining the inherent relationship and signal content across multiple components. Additionally, a method to calculate errors associated with the data calculated from an equivalent layer (denoised data)
has been developed.

In Chapter 4, the application of reliable component conversion is illustrated through two field examples from Canada and Australia. It has been shown that any combination of components can be calculated from an equivalent source. Further, the inversion based comparison of observed and converted recovered models indicate the equivalent source technique has translated information from the observed to the converted components.

In Chapter 5, a case history on inverting full tensor gravity gradient data over an iron formation for a 3D density contrast distribution has been performed. The utility of various gravity gradient component combinations has been investigated using data from the Quadrilátero Ferrífero. The results show that the first-order structural elements are in agreement across all recovered models. It can be seen that there is a highly dense dipping body represented by the high density contrast. Incorporating additional gradient components into the inversion improves the structural representation by better defining the dip, density contrast, as well as the depth and lateral extent of the body. The improvements seen are an indicator of the supplemental information contained within the additional gravity gradient components.

In Chapter 6, the previous case history is built upon and lithology characterization is achieved by incorporating a magnetic susceptibility model for joint-interpretation with the 3D density model. Two approaches to generating 3D lithologic models have been explored by relating susceptibility and density distributions obtained through independent inversion. The methods are advantageous in that they incorporate minimal to no prior knowledge of the target geology. The first approach relies on general physical property values for iron rich rocks in order to identify density and susceptibility ranges of target formation types. The second approach relies on a geologic section, in order to identify density and susceptibility ranges that are site-specific. Through the two approaches, it has been shown that the ability to delineate iron formation can potentially be further broken down to delineate economic hematite from less economic iron formation.
In Chapter 7, a method has been presented on incorporating geophysical models into the resource evaluation stage. The use of cokriging has been identified for the ability to utilize geophysical models as secondary or tertiary information in the modeling process. From the synthetic illustration, it has been shown that observations need not be coincident for all variables included in the modeling. At a minimum, the spatial information inherent to geophysical models may prove useful to resource evaluation.

Three novel contributions from this work include estimation of terrain correction density, noise estimation, and integration of geophysical models into resource evaluation. Through both synthetic and field examples, the utility of gravity gradiometry for contributing to geologic interpretations has been demonstrated. The ability to produce useful 3D models to be used for reliable geologic interpretation requires an understanding of processing and interpretation requirements. My research has developed specific tools and methods that address these needs and facilitate interpretation of gravity gradiometry for resource applications.
REFERENCES CITED


