

**Geological and Petrophysical Information in Geophysical  
Inversion Problems**

**DRAFT: July 19, 2016**

by

Daniel Bild-Enkin

BSc Honours, The University of Toronto, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF

**Masters of Science**

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL  
STUDIES

(Geophysics)

The University of British Columbia

(Vancouver)

August 2016

© Daniel Bild-Enkin, 2016

# **Abstract**

# Preface

# Table of Contents

<b>Abstract . . . . .</b>	<b>ii</b>
<b>Preface . . . . .</b>	<b>iii</b>
<b>Table of Contents . . . . .</b>	<b>iv</b>
<b>List of Tables . . . . .</b>	<b>vii</b>
<b>List of Figures . . . . .</b>	<b>viii</b>
<b>Glossary . . . . .</b>	<b>x</b>
<b>Acknowledgments . . . . .</b>	<b>xi</b>
<b>1 Introduction . . . . .</b>	<b>1</b>
1.1 Research Motivation . . . . .	1
1.2 Regularized Inversion . . . . .	3
1.3 Literature Review . . . . .	6
1.3.1 Mathematical methods of constraining inversion results to a given shape . . . . .	6
1.3.2 Mathematical methods of constraining inversion results to a set of litholigical units . . . . .	8
1.3.3 Implementation of constrained inversion from geological and petrophysical data . . . . .	9
1.4 Thesis Organization . . . . .	10

<b>2</b>	<b>Tools for Integrating Geological and Petrophysical Information into the Regularization of Inversions . . . . .</b>	<b>11</b>
2.1	Including Bore Hole Information in Inversion Regularization . . .	11
2.1.1	Importing Bore Hole Data . . . . .	12
2.1.2	Visualizing Bore Hole Data . . . . .	16
2.1.3	Discretizing Bore Hole Data . . . . .	18
2.1.4	Linking Lithology Information into Petrophysical Information . . . . .	19
2.1.5	Editing Bore Hole Data . . . . .	20
2.1.6	Making Constraint . . . . .	20
2.1.7	Including Surface Sample Information in Inversion Regularization . . . . .	20
2.2	Including Geological Maps in Inversion Regularization . . . . .	21
2.2.1	Preprocessing images . . . . .	22
2.2.2	Loading Images into GIFtools . . . . .	23
2.2.3	Creating a Pixel Map Legend . . . . .	24
2.2.4	Making a Geology Model from Map . . . . .	25
2.2.5	Making Constraints for an Inversion . . . . .	30
2.2.6	Inputing Fault information from Geological Maps . . . . .	31
2.3	Clustering to Create Constraints . . . . .	33
2.3.1	Clustering Algorithms . . . . .	33
2.3.2	Clustering In GIFtools and Model Builder . . . . .	35
2.3.3	Creation of Constraints . . . . .	35
2.4	Voxel-Parametric Inversion to Provide Physical Property Values for Geological Models . . . . .	36
2.4.1	Formulation of Voxel-Parametric inversion problem . . . . .	38
2.4.2	Uses of Voxel-Parametric inversion results . . . . .	39
2.5	Conclusion . . . . .	39
<b>3</b>	<b>Case Study #1 El Poma . . . . .</b>	<b>40</b>
3.1	General Overview of El Poma . . . . .	41
3.2	Overview of Deposits . . . . .	41
3.3	Discussion of the Geophysical Data Given . . . . .	41

3.4	What Information is Available . . . . .	41
3.5	Synthetic Model . . . . .	41
3.6	Blind Inversion of the Synthetic Model . . . . .	41
3.7	Determination of Magnetization Dirrection . . . . .	41
3.8	Creation of Constraints . . . . .	41
3.8.1	$\alpha$ coefficients . . . . .	41
3.8.2	Reference Models . . . . .	41
3.8.3	Weighting matrices . . . . .	41
3.8.4	Bounds . . . . .	41
3.8.5	$L_p L_q$ weights . . . . .	41
<b>4</b>	<b>Case Study #2 TKC . . . . .</b>	<b>42</b>
4.1	Overview of Deposits . . . . .	42
4.2	Discussion of the Geophysical Data Given . . . . .	42
4.3	What Information is Available . . . . .	42
4.4	Synthetic Model . . . . .	42
4.5	Blind Inversion of the Synthetic Model . . . . .	46
4.6	Determination of Magnetization Direction . . . . .	48
4.7	Creation of Constraints and Types of Data . . . . .	48
4.7.1	$\alpha$ coefficients . . . . .	48
4.7.2	Reference Models . . . . .	48
4.7.3	Weighting matrices . . . . .	48
4.7.4	Bounds . . . . .	48
	<b>Bibliography . . . . .</b>	<b>50</b>
<b>A</b>	<b>Supporting Materials . . . . .</b>	<b>53</b>

# List of Tables

Table 2.1	An example “collar file” from TKC bore holes. X and Y are UTM Easting and Northing, Z is the elevation, and Length is the bore hole total length. All units in meters . . . . .	12
Table 2.2	An example “survey file” from TKC bore holes. Depth is the position along the hole where the change in direction occurs (in meters), Azimuth and Dip are the new direction at the depth provided (degrees). The holes are the same as Table 2.1 . . . .	13
Table 2.3	An example “property file” from TKC bore holes. From and To are depth along the hole in meters and Litho is in this case a lithology unit. Other properties can be included in appropriate units. The holes are the same as Table 2.1 . . . . .	14

# List of Figures

Figure 2.1	The first Graphical User Interface (GUI) for importing bore hole data . . . . .	15
Figure 2.2	The file header GUI for importing bore hole data . . . . .	16
Figure 2.3	Visualization of TKC lithology bore hole data . . . . .	17
Figure 2.4	Visualization of El Poma susceptibility bore hole data . . . . .	17
Figure 2.5	GUI to allow the discretization of bore hole data to a given mesh	19
Figure 2.6	The geology definition of the lithological units in the TKC bore hole data. Effective susceptibility is measured in SI is (in this case) . . . . .	20
Figure 2.7	The El Poma map with fault lines (blue lines with barbs) included	22
Figure 2.8	The El Poma map with extra information removed and geological units made a single colour . . . . .	23
Figure 2.9	GUI for importing plan view image . . . . .	24
Figure 2.10	GUI for importing cross section image . . . . .	24
Figure 2.11	Example of magnetics data being viewed with a map overlaid	25
Figure 2.12	Example of a geology model created from a map with the map overlaid . . . . .	27
Figure 2.13	Example of a 2D mesh with the map overlaid . . . . .	27
Figure 2.14	Example of a 2D geology model created from a cross section map with the map overlaid . . . . .	28
Figure 2.15	GUI for adding a 2D model to a 3D model . . . . .	28
Figure 2.16	Example of a 2D geology inserted into a 3D model with the map overlaid . . . . .	29



Figure 2.17	Example of a geological definition as displayed in the GIFtools GUI, the Mean Property, and Bounds are susceptibility measurement in $\text{SI} \times 10^{-3}$ and the WS is to set the smallness weight model. . . . .	30
Figure 2.18	Example of a typical combine model dialog for a reference model	31
Figure 2.19	Example of a reference model created from a geological map .	32
Figure 2.20	The GUI for the creation of fault weights . . . . .	32
Figure 2.21	An example of fault weights that can be created with GIFtools. In this case a vertical fault (no dip) was created and the blue curtain is the low face weights that define the fault in an inversion. Only the faces that have values that are not 1 are visualized. Since the faces are set at a lower value than the rest of the face model (1) they are shown as blue. . . . .	33
Figure 2.22	Point cloud of the geological interfaces for TKC bore hole data	36
Figure 2.23	A geological model created from the data shown in Figure 2.22	36
Figure 4.1	Mesh used in this chapter along VTEM data, and outer bound of PK (blue) and HK (red) . . . . .	43
Figure 4.2	Geology model defined by geology data sets . . . . .	44
Figure 4.3	effective susceptibility result from voxel-parametric inversion	44
Figure 4.4	Normalized misfit of parametric voxel inversion . . . . .	45
Figure 4.5	Forward modeled synthetic data contaminated with Gaussian noise . . . . .	46
Figure 4.6	Blind inversion result assuming a magnetization direction to be the same as the earths field in the location of TKC. (Red cells show the extent of the HK unit in the synthetic model, only cells with a effective susceptibility of greater than $4.0348\text{E-}03$ are shown) . . . . .	47
Figure 4.7	Blind inversion result assuming a magnetization direction to be the magnetization direction of the synthetic HK unit. (Red cells show the extent of the HK unit in the synthetic model, only cells with a effective susceptibility of greater than $4.0348\text{E-}03$ are shown) . . . . .	49

# Glossary

This glossary uses the handy `acronym` package to automatically maintain the glossary. It uses the package's `printonlyused` option to include only those acronyms explicitly referenced in the  $\LaTeX$  source.

<b>MOF</b>	Model Objective Function
<b>NRM</b>	Natural Remanent Magnetization
<b>GIF</b>	Geophysical Inversion Facility
<b>GUI</b>	Graphical User Interface
<b>FCM</b>	Fuzzy C-Means
<b>MVI</b>	Magnetization Vector Inversion

# Acknowledgments

Thank those people who helped you.

Don't forget your parents or loved ones.

You may wish to acknowledge your funding sources.

# Chapter 1

## Introduction

### 1.1 Research Motivation

In mineral exploration there are many forms of information that can be used to determine the location of an economic deposit. These can be divided broadly into geological and geophysical data. Geological data refers to the study of the rocks in a region through surface samples, bore holes, and an understanding of how rock units interrelate under the surface. Geophysical data refers to recovered measurements of some field that is related to the physical properties of the rocks that will aid in the understanding of the deposit. For exploration to be as effective as possible, we need to find ways of integrating the geological and geophysical information that produce exploration vectors to the target. One of the major tools in using geophysical data to create geologically significant interpretations is inversion.

The overarching goal of geophysical inversion is to recover distributions of physical properties in the ground to aid in mineral exploration. To be useful to this end the spacial distribution of the physical property (the geophysical model) needs to both fit the geophysical data and match existing geological interpretations.

Since geophysical inversions are by their nature non-unique (because of data uncertainty and there typically being many more model parameters than data), *a priori* information needs to be added to provide a model that matches the geology

of a deposit. Much work has been done to create a mathematical framework to allow the inclusion of geological and petrophysical information into geophysical inversions (for example Li and Oldenburg (1996)). However, an area where more work must be done is the creation of tools to take the petrophysical and geological data in the forms that are generally provided and create usable constraints that can be applied to inversions.

The research in this thesis will attempt to do exactly that: provide new tools in an integrated framework that will allow the incorporation of *a priori* information into geophysical inversions. The inclusion of *a priori* information in inversions is not novel. Many researchers, especially at the Geophysical Inversion Facility (GIF), before me have used the mathematical framework to add geological and petrophysical information to inversions (for example Lelievre 2009, Phillips 2001, Farquharson et al. 2008).

In addition, Williams (2008) develops a software package to create constraints for inversions from a wide array of data types. What is lacking in the previous research is the link between the creation of inversion constraints with the processing of data and the running of inversions. By integrating the tools I create in this thesis into the framework of GIFtools and Model Builder I attempt to provide this link.

GIFtools and Model Builder are a suite of tools whose origins date back to Williams (2008), but which have been sufficiently updated that they deserve further treatment. The goal now, as in Williams (2008), is to create a set of Graphical User Interface (GUI) tools that make the running of GIF inversion codes simple and easy.

GIFtools was a set of GUI tools that made the running of GIF inversion codes (and managing the many files and parameters they require) more straight-forward. GIFtools has evolved into a software environment that allows the running of advanced inversion. It maintains the original purpose of helping run GIF inversion codes, and has incorporated many inversion codes that were not originally supported by the older GIFtools environment. It also now includes significant data visualization and quality control, mesh creation (in both 2D and 3D tensor meshes as well as 3D octree meshes), and model visualization and editing functionality.

Model Builder, first described and implemented in Williams (2008), is a set of GUI tools that facilitate the creation of constraints from geological and petrophysical data for GIF inversion codes. It largely consisted of a work flow for the

incorporation of said data into reference models, bounds, and weights that could be used in GIF inversion codes.

While the tools created in Williams (2008) were very powerful, the interface was not always straightforward to use. This was largely due to the interface’s lack of flexibility. Each part of the process occurred in a strict order, meaning that if the user made an error the whole process had to be restarted. In addition the program lacked visualization support and the ability to do quality control on the geological and petrophysical data that was to be incorporated.

By integrating these two sets of tools into one, I make the incorporation of *a priori* information into inversions much more expedient, thus encouraging greater uptake by industry. GIFtools and Model Builder allow users do quality control on data, create constraints, and run inversions within the same software framework. The interface by which *a priori* information can be incorporated has been much updated from Williams (2008), and tools to incorporate new forms of data into inversions have been added.

The interfaces that aid data and model viewing, quality control, and the creation of constraints is not merely a convenience. At the current level of inversion techniques, each step in the process of recovering an inversion result depends greatly on the expertise and interpretation of a geophysicist user. Without the tools to view, compare, edit, and create, these users are not able to fully exploit the information provided by geophysical data and inversion to the field of mineral exploration and other applications.

## 1.2 Regularized Inversion

In the general case, geophysical inversion involves the solving of a system that is defined by some forward operator that maps from a given model to predicted data,

$$\mathbf{d} = \mathbb{F}[\mathbf{m}], \quad (1.1)$$

where  $\mathbf{d} \in \mathbb{R}^N$  is the geophysical data ( $N$  is the number of data collected),  $\mathbf{m} \in \mathbb{R}^M$  is the discretized model that describes the distribution of some physical property in the ground ( $M$  is the number cells in the earth model), and  $\mathbb{F}$  is the forward operator that mediates between them. In the context of this thesis,  $\mathbf{d}$  is either magnetic

or gravity data measured with a ground or aerial survey,  $\mathbf{m}$  is either a magnetic susceptibility or density model, and  $\mathbb{F}$  is the magnetic or gravitational forward operator which has the convenient property of being linear. Since we are interested in recovering the model  $\mathbf{m}$ , we are interested in finding the inverse of  $\mathbb{F}$ ,

$$\mathbf{m} = \mathbb{F}^{-1} \mathbf{d}. \quad (1.2)$$

Unfortunately the inversion of  $\mathbb{F}$  is far from trivial as the problem is ill-posed. Firstly since there are usually more model parameters than data ( $M > N$ ) there are an infinite number of possible distributions of the physical property that will predict the same observed data. Secondly the system is unstable; that is, a small amount of error in the measurements can lead to large changes in the recovered model. To recover a model despite these difficulties, the problem is regularized by adding *a priori* information in the form of a Model Objective Function (MOF) or regularization functional. Once the problem is regularized, it is solved by minimizing the objective function,

$$\phi(\mathbf{m}) = \phi_d + \beta \phi_m \quad (1.3)$$

where  $\phi$  is the objective function,  $\phi_d$  and  $\phi_m$  are the data and model objective functions respectively, and  $\beta$  is a trade-off parameter that scales between them.  $\phi_d$  is defined as a least squares (or  $L_2$ ) formulation,

$$\phi_d = \|\mathbf{W}_d (\mathbb{F}[\mathbf{m}] - \mathbf{d}^{obs})\|^2 \quad (1.4)$$

$$\mathbf{d}^{obs} = \mathbb{F}[\mathbf{m}^{true}] + \mathbf{e} \quad (1.5)$$

where  $\mathbf{d}^{obs}$  is the observed data, that is the true data contaminated with noise  $\mathbf{e}$ , and  $\mathbf{W}_d$  is the data weighting matrix with the diagonal entries equal to the reciprocal

of each datum's standard deviation,

$$\mathbf{W}_d = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\sigma_N} \end{bmatrix}, \quad (1.6)$$

where each  $\sigma_i$  is that datum's assigned standard deviation. Meanwhile  $\phi_m$  can take many forms. In Li and Oldenburg (1996) it is defined in the continuous formulation as,

$$\begin{aligned} \phi_m = & \alpha_s \|\mathbf{w}_s [(\mathbf{m} - \mathbf{m}_{ref})]\|^2 + \dots \\ & \alpha_x \|\mathbf{w}_x \mathbf{G}_x [(\mathbf{m} - \mathbf{m}_{ref})]\|^2 + \dots \\ & \alpha_y \|\mathbf{w}_y \mathbf{G}_y [(\mathbf{m} - \mathbf{m}_{ref})]\|^2 + \dots \\ & \alpha_z \|\mathbf{w}_z \mathbf{G}_z [(\mathbf{m} - \mathbf{m}_{ref})]\|^2. \end{aligned} \quad (1.7)$$

The first term in Equation 1.7 promotes smallness; that is, the model must be close in value to the reference model  $\mathbf{m}_{ref}$  which is some first guess at the cell values of the area being inverted. The next three terms promote smoothness by penalizing large derivatives in the model, determined by the discrete gradient operator  $\mathbf{G}$  in each direction. There are many parameters in  $\phi_m$  that the geophysicist can use to fine-tune the recovered model.  $\alpha_s, \alpha_x, \alpha_y$ , and  $\alpha_z$ , scale the contribution of smallness and smoothness in each spatial direction.  $\alpha$  values can be used to broadly determine the length scales in each direction of a recovered model. The  $\mathbf{w}_s, \mathbf{w}_x, \mathbf{w}_y, \mathbf{w}_z$  parameters allow finer control. They weight the model's smallness and smoothness variable across its extent, allowing the user to specify certain regions as smoother than others or demanding that the recovered model more closely match the reference model in areas where the user is more certain of the reference model's validity.

In Equation 1.7 it is assumed the the norm of the regularization is  $L_2$ :

$$\|\mathbf{v}\|_2 = \left( \sum_{i=1}^N v_i^2 \right)^{\frac{1}{2}}. \quad (1.8)$$



This formulation can be generalized to an  $L_p$  norm that takes the form

$$\|\mathbf{v}\|_p = \left( \sum_{i=1}^N |v_i|^p \right)^{\frac{1}{p}} \quad (1.9)$$

where  $p$  is some positive number. Minimizing a objective function with a lower values of  $p$  promote more sparsity in the vector  $\mathbf{v}$ .

The last term in Equation 1.3 is  $\beta$  which determines the degree to which the model fits the data or obeys the regularization. Assuming that the error in  $\mathbf{d}^{obs}$  is Gaussian and independent with the standard deviations in  $\mathbf{W}_d$ ,  $\phi_d$  will be a random variable with a  $\chi^2$  distribution and an expected value of  $N$ , the number of data Oldenburg and Li (2005). Given this expected value, beta can be iteratively decreased until the misfit is sufficiently near  $N$ .

Having discussed the structure of regularized inversion I will now discuss previous research on including geological and petrophysical information in inversion using regularization.

## 1.3 Literature Review

### 1.3.1 Mathematical methods of constraining inversion results to a given shape

#### Smoothly varying recovered models

There has been much research following the lines of Li and Oldenburg (1996) and Li and Oldenburg (1998). The methods used in these two papers are described in Section 1.2. The advantage of using least-squares methods for the regularization of inverse problems is that the objective function is convex, continuous, and differentiable, which greatly aids the implementation of the optimization. Reference models and weighting matrices allow for the incorporation of geological information, making smallness and smoothness more or less significant in different parts of the model. Li and Oldenburg (2003) extend the method by also implementing upper and lower bounds that can be specified for each model cell allowing the user to make hard constraints on the value of the model.

These methods recover smoothly varying models with broad distributions of the physical property. Sometimes the geological context indicates that the model should vary sharply and the anomalous body should be compact. In other words, either the model or its derivative should be sparse. There has been a great deal of research on using sparsity-promoting norms to achieve compact or sharply-varying models.

### **Sharply varying recovered models (Blocky and Compact)**

Instead of regularizing by smallness and smoothness, Last and Kubik (1983) regularize by compactness, essentially demanding that the anomaly be as small as possible while still fitting the data. They use an  $L_0$  norm on the smallness component and do not use the smoothness constraints. Portniaguine and Zhdanov (1999) extend Last and Kubik (1983) by adding what they call a minimum gradient support functional. The effect of this support functional is a  $L_0$  norm on the smoothness terms instead of the smallness term as in Last and Kubik (1983).

Rudin et al. (1992) and Vogel and Oman (1998) propose total variation methods in the context of de-noising and de-blurring images, in other words the use of  $L_1$  norms to regularize, instead of  $L_2$  norms as in Li and Oldenburg (1996) and Li and Oldenburg (1998). Since minimizing  $L_1$  norms promotes sparsity, regularizing by them will have a comparable effect (blocky models with sharp boundaries) as the method used by Last and Kubik (1983) and Portniaguine and Zhdanov (1999). Total variation has been used more specifically in the context of geophysical inversion, such as with Guitton (2012).

Farquharson and Oldenburg (1998) also report ways of achieving sharp contrast by implementing non- $L_2$  norms such as Ekblom and Huber norms. Fournier (2015) implements a method of minimizing the general  $L_p$  (smallness) and  $L_q$  (smoothness) norms for any  $p$  and  $q$  (typically values between 0 and 2) allowing a user to specify the degree of compactness or blockiness of a recovered model in different spatial directions.

### **Recovered models with dipping anomalies**

The formulation in Equation 1.7 allows a great deal of control of the way the model varies along the three cardinal directions. However, the geometry of a deposit does not always align with any of the cardinal directions and diagonal structures are preferred. Li and Oldenburg (2000) extend the method of Li and Oldenburg (1996) and Li and Oldenburg (1998) by rotating the model objective function to allow for linear features in the recovered model to be in a direction not in line with the mesh grid. Lelièvre and Oldenburg (2009) generalize the methods in Li and Oldenburg (2000) to the 3D case.

Guillen and Menichetti (1984) extend the method described in Last and Kubik (1983). Instead of minimizing the volume of a deposit, the authors minimize its moment of inertia. By specifying an axis of rotation to determine the moment of inertia, they put dip information into the regularization in addition to having the sharp contrasts and compact models as in Last and Kubik (1983). Barbosa and Silva (1994) and Barbosa and Silva (2006) extend the method even further allowing multiple axes of rotation. The second paper also describes a GUI to interactively test the fit of various axes of rotation.

Chasseriau and Chouteau (2003) create a very general method of biasing the inversion algorithm towards anomalies of almost any shape by weighting the smallness term with a covariance matrix of the model, i.e., a matrix with the covariance of each cell versus every other cell in the model. The covariance matrix can be generated from bore hole or surface sample data or from a synthetic initial model.

### **1.3.2 Mathematical methods of constraining inversion results to a set of lithological units**

In addition to the deterministic inversions described above, much research has been done on stochastic inversion. Bosch et al. (2001) directly invert for lithologies. The authors forward model physical properties by a probabilistic relation of the physical property to the lithology. New lithology distributions are created using a pseudo-random walk. *A priori* information is included partially in the probabilistic model that links the lithology to the physical properties but also in the initial probability distribution of the lithology model. Guillen et al. (2008) implement the

method described in Bosch et al. (2001) in 3D.

Attempts have also been made at combining stochastic and deterministic methods. One particularly successful line of inquiry are Fuzzy C-Means (FCM). Paasche et al. (2006) uses FCM clustering of recovered models to derive membership functions of model cells in several clusters. The clusters are then used with *a priori* porosity data to create a likely porosity of each cluster and a porosity model is created from these results.

Instead of clustering an inversion after the fact to achieve the effect of a cooperative inversion without inversion both datasets concurrently like Paasche et al. (2006), Sun and Li (2015) use the FCM function as an extra term in the model objective function. This allows them to simultaneously invert slowness and density (from travel time and gravity data) by linking them through the FCM clusters. It also allows them to guide the FCM cluster physical properties in a way that allows the integration of petrophysical measurements of geological units without directly forcing where the units are in space.

### **1.3.3 Implementation of constrained inversion from geological and petrophysical data**

Phillips (2001) uses bore hole densities and susceptibilities to bound a gravity and a magnetic inversion. Farquharson et al. (2008) use density bore hole logs to create a reference model for a gravity inversion. They demonstrate the effect of having many bore holes versus only a few. Williams (2008) provides the most extensive review of this subject. He creates a software package to integrate a several types of geological and petrophysical data including bore hole, surface sample, geological ESRI shapefiles, and geological models both in the form of voxel models and 3D domain models. Many of the tools that he created are integrated into the work that I present. He then uses these tools to create detailed susceptibility and density constraints and applies them in inversions. Finally, Lelievre (2009) discusses the use of susceptibility samples and bore holes in constraining a synthetic magnetic inversion example. He also shows the use of orientation information of linear features as a geological constraint.

As stated in the research motivation, what is lacking in previous implementations is an integrated environment where data, models, inversions, and constraints

can be developed. By implementing the creation of constraints in such an environment, I make the creation of constraints faster even in non-trivial contexts with multiple sources of information and multiple forms of constraints.

## 1.4 Thesis Organization

In this thesis I describe the methods used to create the tools I contributed to GIFtools and give examples of their use. Chapter 2 will discuss the tools I have created. It describes the types of information that GIFtools and Model Builder can integrate into an inversion and discuss how they can be used to constrain an inversion result. I discuss how sample information (bore hole and surface sample data) can be used to set reference models and bounds. I also discuss geological maps and how Model Builder incorporates information from both cross section and plan view maps into the regularization of inversions. I also discuss parametric style inversions using a standard mesh and forward operator. Instead of allowing all cells to vary freely I constrain them to some number of units that must all be constant across the unit. Finally in Chapter 2 I discuss the use of clustering multiple inversion results to create non-trivial face weights in addition to reference models and bounds.

In Chapter 3 I show the use of GIFtools and Model Builder in the creation of regularizations for a magnetic inversion in El Poma. El Poma is a porphyry deposit in Colombia that has magnetic properties measured from bore holes, surface samples, in addition to a geological map over the region. The region is also interesting due the large effect of remanent magnetization. I discuss a synthetic case, matching the magnetic survey, bore holes, surface samples and map to recover the anomaly. I also show the result of the inversion of the actual field data.

Finally in Chapter 4 I show an example GIFtools and Model Builder in the context of the Tli Kwi Cho Kimberlite complex in the North West Territories, Canada. In this case there have been several surveys flown over the region for electromagnetic (of which I use only the magnetic data) and gravity gradiometry data. In addition to the geophysical data there has been extensive drilling, and cross section maps have been created. I show a synthetic example inverting both gravity gradiometry and magnetic data incorporating the *a priori* information as well as using clustering of the magnetic and gravity inversion results to create further constraints.

## **Chapter 2**

# **Tools for Integrating Geological and Petrophysical Information into the Regularization of Inversions**

### **2.1 Including Bore Hole Information in Inversion Regularization**

Bore holes provide physical property measurements at depth either by providing a conduit to send geophysical instruments down hole or by recovering a core and then measuring its petrophysical properties subsequently in the lab. The inversion methods for including physical property and lithology bore hole data have been developed (e.g., Williams (2008)). In this section I discuss how GIFtools and Model Builder incorporate bore hole information into inversion constraints in a expedient and effective manner.

In addition to physical property data, bore holes can also provide qualitative rock unit information. In this section I show the linking of lithology information from bore holes with physical property information using petrophysical measurements of reasonably similar rocks. Once this linking is done, the bore hole infor-

mation can be used in the same fashion as a physical property bore hole logs.

In this section I also show the advantage of integrating the creation of inversion constraints with more general data processing tools. Once we have a set of sample data from either bore hole or surface samples loaded into GIFtools, we can then use the general data modification tools that are already provided for data quality control created both by me and the rest of the GIF team.

### 2.1.1 Importing Bore Hole Data

Bore hole data is typically stored in three separate files: the collar file, the survey file, and the property file. The collar file contains the spacial information of the very top of the hole. Table 2.1 shows an example collar file. GIFtools is sufficiently flexible that it can read both whitespace and comma delimited files. Additionally as long as there are headers for each column, it does not matter what order the columns appear in, as the correct columns for each purpose can be specified during the import process. Finally while the example shows a text hole identifier, it is also possible to identify individual holes by a numerical index.

% TKC collar data				
HOLE-ID	X	Y	Z	LENGTH
DO27-05-01	557187	7133758	418	58.52
DO27-05-02	557191	7133755	418	459.5
DO27-05-03	557165	7133682	418	230
DO27-05-04	557425	7133835	420	112.5
DO27-05-05	557425	7133835	420	99.8
DO27-05-06	557425	7133835	420	101
DO27-05-07	557425	7133835	420	218
DO27-05-08	557392	7133834	419	290
DO27-05-09	557392	7133834	419	155
DO27-05-10	557392	7133834	419	140
DO27-05-11	557400	7133913	419	374
DO27-05-12	557345	7134210	419	65

**Table 2.1:** An example “collar file” from TKC bore holes. X and Y are UTM Easting and Northing, Z is the elevation, and Length is the bore hole total length. All units in meters

The survey file provides depth, azimuth, and dip information, coding how the hole changes direction below the collar. In Table 2.2 all the holes as defined in Table 2.1 are straight and dip in different directions.

% TKC survey data			
HID	DEPTH	AZIMUTH	DIP
DO27-05-01	0	0	-90
DO27-05-01	58.52	0	-90
DO27-05-02	0	0	-90
DO27-05-02	459.5	0	-90
DO27-05-03	0.	0	-90
DO27-05-03	230	0	-90
DO27-05-04	0	180	-70
DO27-05-04	112.5	180	-70
DO27-05-05	0	200	-47
DO27-05-05	99.8	200	-47
DO27-05-06	0	80	-45
DO27-05-06	101	80	-45
DO27-05-07	0	273	-70
DO27-05-07	218	273	-70
DO27-05-08	0	265	-45
DO27-05-08	290	265	-45
DO27-05-09	0	265	-86
DO27-05-09	155	265	-86
DO27-05-10	0	348	-45
DO27-05-10	140	348	-45
DO27-05-11	0	240	-45
DO27-05-11	374	240	-45
DO27-05-12	0	230	-45
DO27-05-12	65	230	-45

**Table 2.2:** An example “survey file” from TKC bore holes. Depth is the position along the hole where the change in direction occurs (in meters), Azimuth and Dip are the new direction at the depth provided (degrees). The holes are the same as Table 2.1

Finally, the property file contains information about a given property down the hole. This property can either be a physical property (e.g. density, susceptibility, etc.) or a geological unit. The depth information of a property measurement can



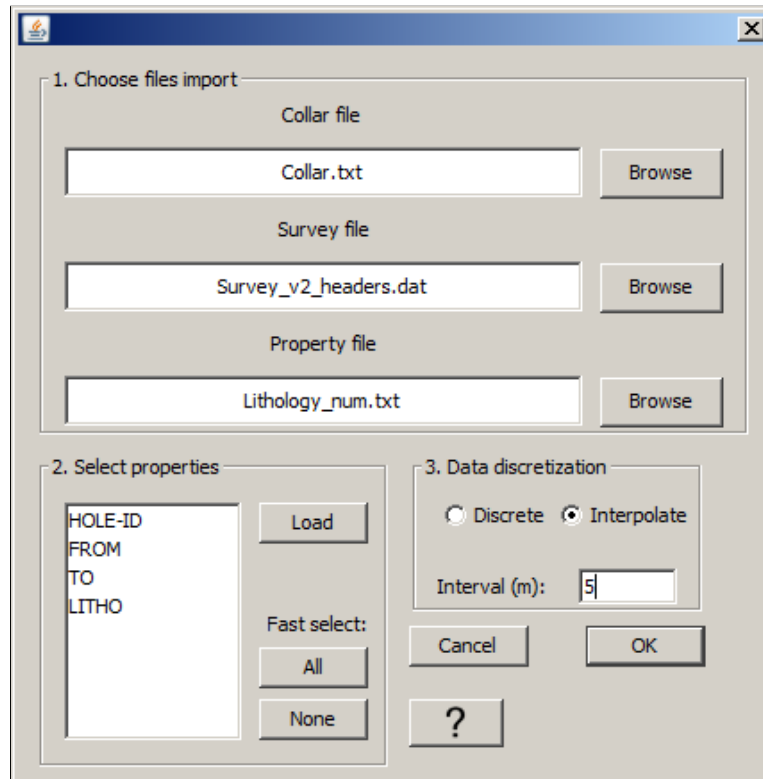
be stored as a simple depth along the bore hole, or as an interval with two depths a “from” and a “to” depth stating that a given measurement hold for the whole interval. Table 2.3 shows the second form of property file with a numerical lithology.

% TKC property data			
HOLE-ID	FROM	TO	LITHO
DO27-05-01	56.5	58.52	Kimb-1
DO27-05-02	56	459.5	Kimb-1
DO27-05-03	59	230	Kimb-1
DO27-05-04	19	63.4	Kimb-1
DO27-05-04	63.4	112.5	Kimb-3
DO27-05-05	21.8	85.8	Kimb-1
DO27-05-06	37	49.5	Kimb-1
DO27-05-06	49.5	82.9	Kimb-3
DO27-05-07	20.5	104.5	Kimb-1
DO27-05-07	104.5	131	Kimb-3
DO27-05-07	138.7	218	Kimb-2
DO27-05-08	20.8	290	Kimb-1
DO27-05-09	9	95.8	Kimb-1
DO27-05-09	95.8	117	Kimb-3
DO27-05-09	125.4	155	Kimb-2
DO27-05-10	17	100.3	Kimb-1
DO27-05-10	100.3	123	Kimb-3
DO27-05-11	44.5	223.5	Kimb-1
DO27-05-12	36	36.7	Kimb-2

**Table 2.3:** An example “property file” from TKC bore holes. From and To are depth along the hole in meters and Litho is in this case a lithology unit. Other properties can be included in appropriate units. The holes are the same as Table 2.1

The process to load bore hole data into GIFtools is as follows. Firstly the files that define the bore hole data set (collar, survey, and property) need to be provided. Secondly, the columns to be imported from the property file need to be stated. Lastly the method by which the property data is linked spatially to the drilling data (collar and survey) is stated. All of these are done by the GUI as shown in Figure 2.1. Note that the user has to set whether the depth information of the property is based on the sample location (discrete) or each point along the hole is

interpolated from the data (interpolate).



**Figure 2.1:** The first GUI for importing bore hole data

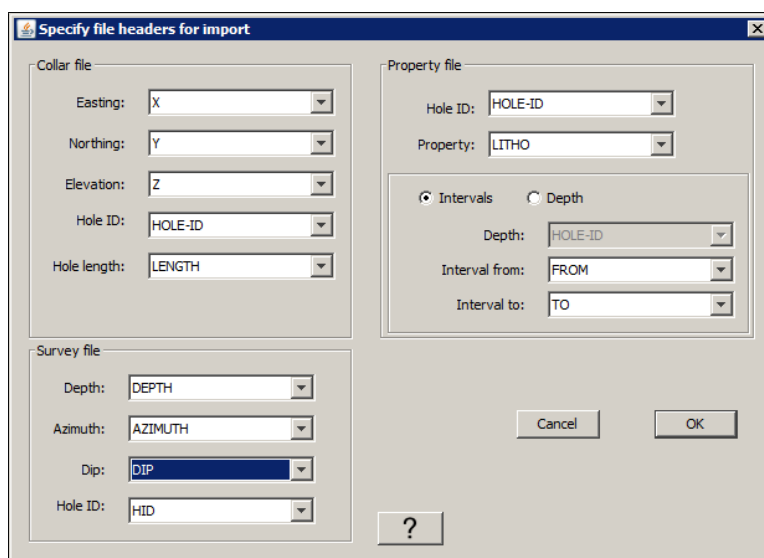
There are two methods to link property and drilling data, “Discrete” and “Interpolate”. If the “Discrete” option is selected then GIFtools simply determined the spatial location based on the depth provided in the property file. In the case that the depth is given in the form of intervals, the depth of the measurement is considered to be the midpoint of the interval.

If “Interpolate” is selected the program behaves differently depending on whether the depth information is simple depths or intervals. In both cases instead of using the depth information directly, a sample is provided at each point along the bore hole with distances between each sample defined by the sampling interval value in the first GUI (Figure 2.1).

In the case of simple depths, it is assumed that the measurement is some phys-

ical property, and given the sample depths the measurements are linearly interpolated along the bore hole. In the case of intervals it is assumed that the measurement lithologies should not be interpolated, so samples with depths that are within an interval are assigned the property given while depths outside of any range are assigned a NaN (not a number) value.

Once the files, and property location options have been set, the next step to importing bore hole data into GIFtools is the set of data columns and data headers. Since different collar, survey, and property files have different column orders it advantageous to allow the flexibility to assign columns with a given header to any property of the bore hole object in GIFtools. The GUI to assign the columns is shown in Figure 2.2



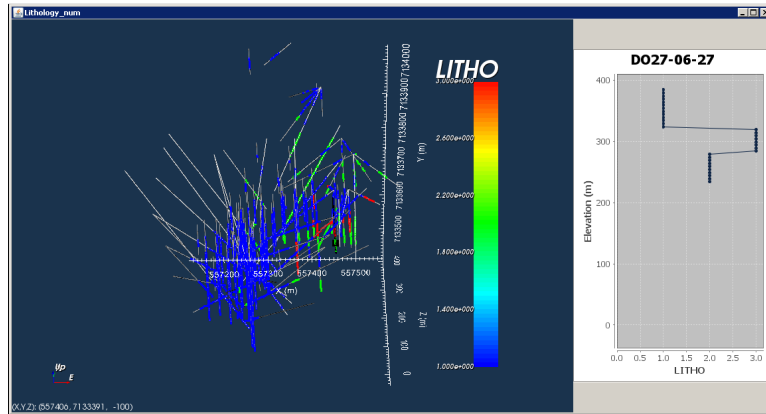
**Figure 2.2:** The file header GUI for importing bore hole data

### 2.1.2 Visualizing Bore Hole Data

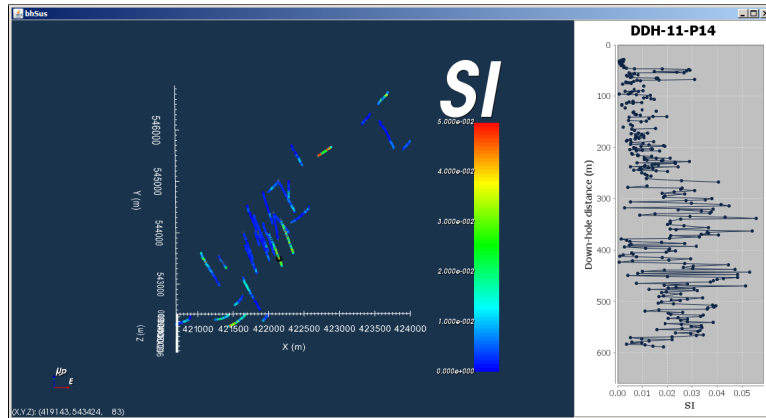
As stated in the introduction to this section, one of the advantages of integrating the Model Builder tools (including these tools for bore hole data being described) into GIFtools, a more general data visualization and quality control environment, is the ability to use those same tools for the data with which we create regularizations.

In this case it is very useful to be able to visualize bore hole data before it gets incorporated in reference model and bounds of an inversion.

Figure 2.3 shows the visualization of the bore hole data from TKC. Note that the data has been sampled within each unit. Figure 2.4 shows the visualization of a physical property bore hole data set from El Poma. Note that instead of the method used in Figure 2.4, each datum along the hole was given its own position, in this case, the mid point of each small interval provided. In both cases the value of the visualized property of a given hole is shown along the side of the 3D representation of all the holes.



**Figure 2.3:** Visualization of TKC lithology bore hole data



**Figure 2.4:** Visualization of El Poma susceptibility bore hole data

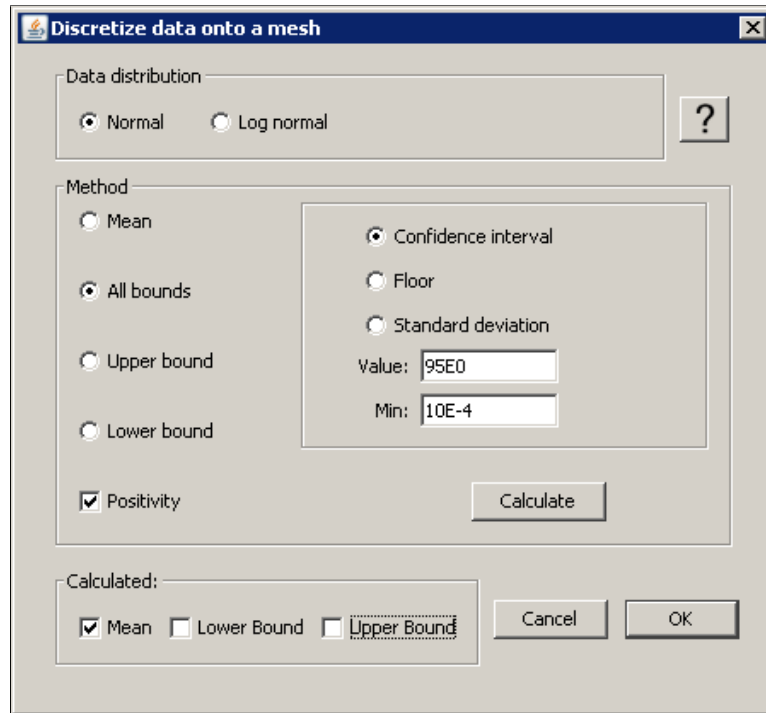
### 2.1.3 Discretizing Bore Hole Data

Before a bore hole data set can be use in the creation of inversion constraints, it needs to be discretized on the mesh that will be used in the inversion. Discretization allows the multiple data along each bore hole to be usable in the context of an inversion on a given mesh. This will allow the creation of reference models, bounds, and weights.

The process to discretized bore hole data in GIFtools and Model Builder is as follows.

- Provide mesh. If the bore data will be used with a Model Builder module to create constraints, it must have the same mesh as the Model Builder module.
- describe distribution
  - normal
  - log normal
- describe method of determining bounds
  - Confidence interval: given a number of samples in a cell and a distribution, bounds are determined from a given percent confidence interval. Where the standard deviation is zero, (if there is only one datum, or all the data are equal) he minimum value is used in place of the confidence interval.
  - Floor: Simply assigns a bound based on the mean value of each cell plus or minus the provided floor value
  - Standard Deviation: Calculates the standard deviation of the sample values in each mesh cell (given a normal or log normal distribution). The bounds are set as equal to the mean value plus some multiple of the standard deviation. In the case that the standard deviation is zero the minimum value is used instead of the multiplied standard deviation.
- positivity simply set the lower bound and mean to be at least 0

Figure 2.5 shows the GIFtools GUI that allows the user to input the discretization options.



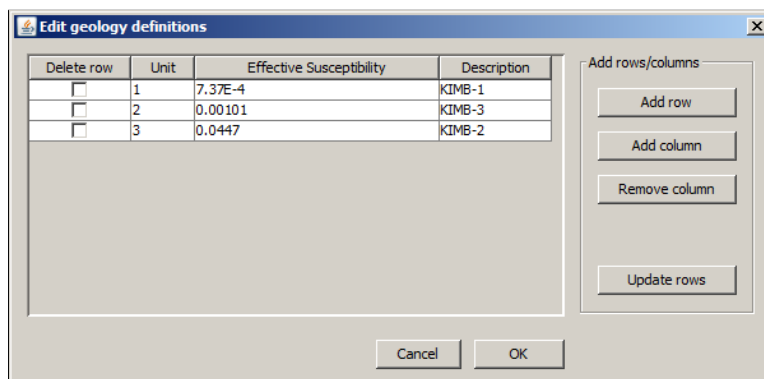
**Figure 2.5:** GUI to allow the discretization of bore hole data to a given mesh

#### 2.1.4 Linking Lithology Information into Petrophysical Information

As stated in the introduction of this section, bore hole data often consists of lithology data instead of a physical property. In these cases the lithology information needs to be converted into petrophysical information so that the bore holes are to be used to constrain inversion results. In GIFtools, linking lithology information to petrophysical data is done by what is called a geology definition. The geology definition is a lookup table that contains information of each particular geological unit's property, lower and upper bounds, and optionally the smallness weight associated with each unit.

Using the geology definition we can convert a geology model that has information about the spatial distribution of geological units but not of their physical properties into constraints that are usable by an inversion. Figure 2.6 is an example of a geology definition in the GIFtools GUI. The data in the geological definition came from lab measurements of magnetic susceptibility and remanence within

each geological unit.



**Figure 2.6:** The geology definition of the lithological units in the TKC bore hole data. Effective susceptibility is measured in SI is (in this case)

Once the geology definition is set the lithological bore hole can be treated as a physical property bore hole data set where each sample of a given unit is assigned the property in the geology definition. Multiple properties can be stored in the geology definition and the one that will be used to create a set of constraints can be changed by editing the geology definition's I/O header.

### 2.1.5 Editing Bore Hole Data

- show editing of BH discretization
- show editing of prop data for susc to eff susc

### 2.1.6 Making Constraint

- show resolve conflicts dialogue
  - steal from below section (new images necessary)

### 2.1.7 Including Surface Sample Information in Inversion Regularization

In addition to physical property measurements down hole, physical properties can also be measured on the surface, in many cases with much more ease. As with

bore hole data, much work has been done on including surface sample information in inversion regularization. Surface samples take on increased importance since in forward models the sensitivity of the data on a given cell is higher when the cell is nearer the surface. Constraining these surface cells can reduce artifacts that come from this increased sensitivity.

- importing SS
- visualize
- editing
- creation of constraints

## **2.2 Including Geological Maps in Inversion Regularization**

It is often the case that geological information is provided in the form of geological maps in either cross section or plan view. Such maps are particularly useful since they provide a great deal of information over their entire surface. Cross sections can provide information at depth and constrain a whole region often within the center of a target of interest. Plan view maps do not provide information at depth, but they do constrain the entire surface of the region being inverted. Constraining the surface of an inversion is of interest since the sensitivity of the data to the top cells is particularly high, which can lead to artifacts on the surface. For the next section the plan view model is from the El Poma case study and the cross section model is from TKC, specifically the map from (Harder et al., 2006)

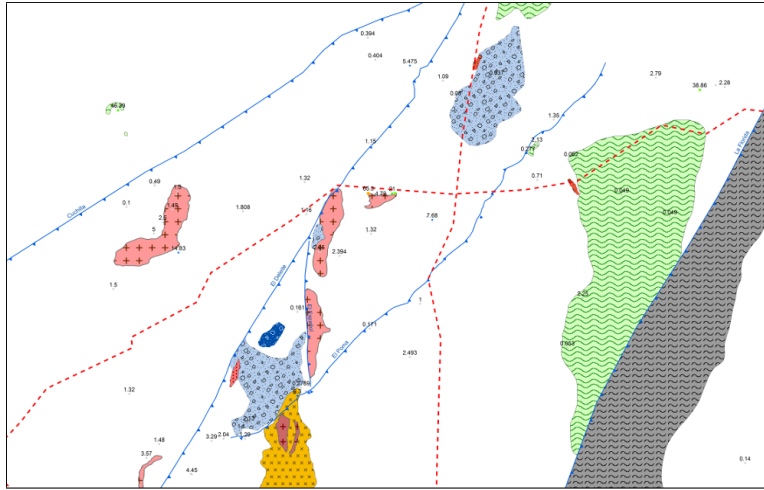
Williams (2008) discusses methods to include maps in the form of ESRI shapefiles. His method has the disadvantage of not being able to incorporate information from maps stored as pixel images. On the other hand, a method that allows the incorporation of pixel images allows the use of ESRI shapefiles since the conversion of a shapefile to a pixel image is trivial.

Below is the method I have developed to incorporate pixel maps.



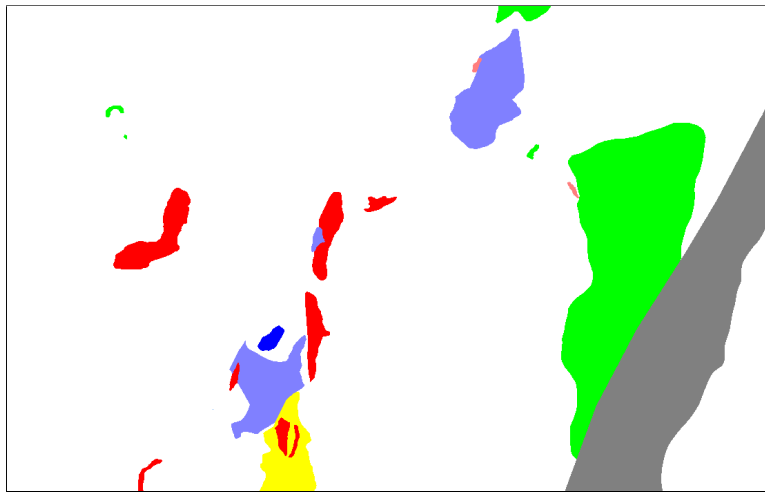
### 2.2.1 Preprocessing images

Often a geological map image will not be immediately suitable to the methods used below and some preprocessing is required. The most notable features that are undesirable in a map are geological units that are not only one or two colours and text or other annotations that could be interpreted by the program as geological information. Also map images may be of too high resolution to be efficiently used in these methods and must also be down-sampled to save computer processing time and memory.



**Figure 2.7:** The El Poma map with fault lines (blue lines with barbs) included

For example Figure 2.7 has great deal of information, (faults, magnetic susceptibility surface samples, etc.) that are not information about geological units. In addition the geological units are not a single colour polygon. The image has been edited in the GNU Image Manipulation Program (GIMP), a free image editing program, to produce Figure 2.8.

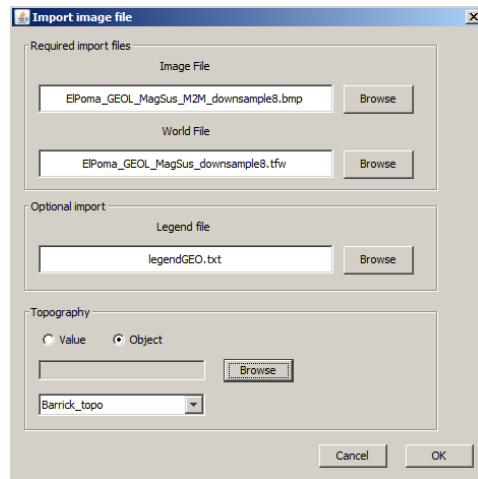


**Figure 2.8:** The El Poma map with extra information removed and geological units made a single colour

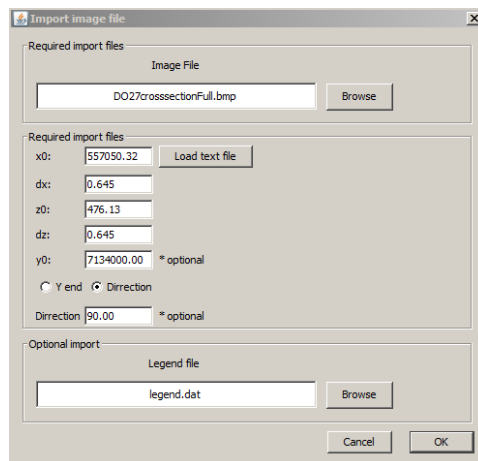
### 2.2.2 Loading Images into GIFtools

- load image into the GIFtools format (Figure 2.9)
  - Determine image format.
  - Load image using MATLAB utilities.
  - Convert image into .png style representation for faster computation.
  - Using .twf file (world file) assign location and spacial resolution to the image.
  - Assign a legend linking pixel RGB values to geological unit.
  - Assign topography (either number or GIFtools TOPOdata item) for visualization.
    - \* In the case of a cross section image, instead of topography, information for the location of the cross section in 3D or 2D space is required (Figure 2.10).

Storing a map as a GIFtools object allows its use in several ways. Notably it allows the integration of the map with models and data, allowing figures overlaying the



**Figure 2.9:** GUI for importing plan view image



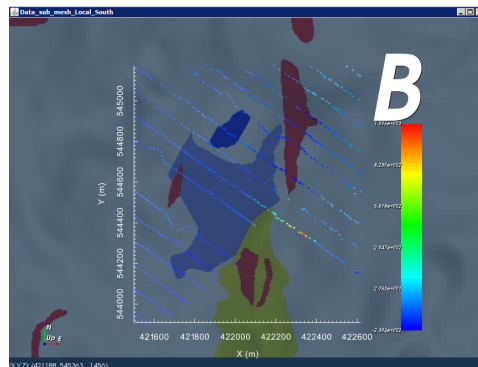
**Figure 2.10:** GUI for importing cross section image

map and data or model and allowing interpretation of the data or model with direct reference to the map (Figure 2.11).

### 2.2.3 Creating a Pixel Map Legend

Continuing on in the process of making a geological constraint.

- Find the geological unit represented of each pixel.



**Figure 2.11:** Example of magnetics data being viewed with a map overlaid

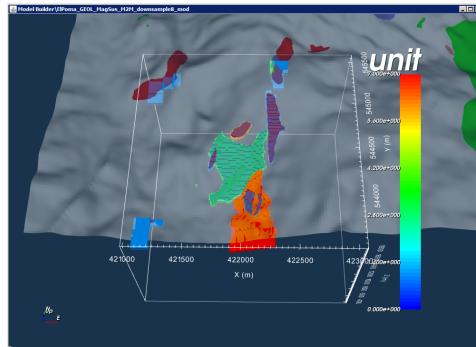
- In the .png style format as stored in MATLAB, an image consists of an “image” field, a matrix of integers, and a “map” field, which maps the image matrix to RGB value triplets.
- Each RGB triplet is compared to the legend that was provided when the image was loaded. A map field entry is considered to represent a geological unit if all three components of the RGB triplet are within a provided tolerance of any entry in the legend.
- Now that we have a relation of entries in the map field to geological units in the legend, we can assign a geological unit to each pixel in the original image simply by applying the new geological map to the image field.

## 2.2.4 Making a Geology Model from Map

### Plan View

- Provide active model. For convenience this is usually an active model already associated with a Model Builder object.
  - The active model simultaneously provides a discretized topography for the map to lay along and also a mesh (GIF 3D tensor or OcTree).
- Provide some form of depth information.

- Thickness, a certain amount of depth below topography at each point will be assigned the geological unit at each.
- Depth, the map will be used to assign a geological unit down to a fixed depth across the whole model.
- Surface, if you provide another surface below topography the cells between topography and the other surface will be assigned.
- Crop all pixels that extend outside of the mesh or that represent the background geological unit.
  - The cropping greatly speeds up the process and makes it require much less computer memory.
  - Furthermore, in the event of a mistake with coordinates the process ends almost instantly as there are few pixels to process.
- Finally the geological model is created.
  - We determine which cell of the mesh each pixel is in, including those cells below each pixel to account for thickness.
  - Each cell is assigned a geological unit based on the mode of the geological values of each pixel which colours that cell. In other words, each cell is identified with the geological unit which fills the greatest proportion of the cell.
    - \* The mode is used since each cell will be a particular unit. Since the property being mapped onto each cell by construction must represent a single geological unit, interpolation between the units will not provide the desired result.
  - The geology definition which will allow the assignment of physical properties to each geological unit. The result is shown in Figure 2.12, the continuous colour bar is not an indication of a continuous model. All model values are integers that represent geological units in the map.

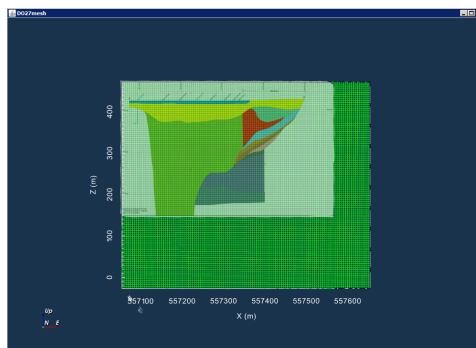


**Figure 2.12:** Example of a geology model created from a map with the map overlaid

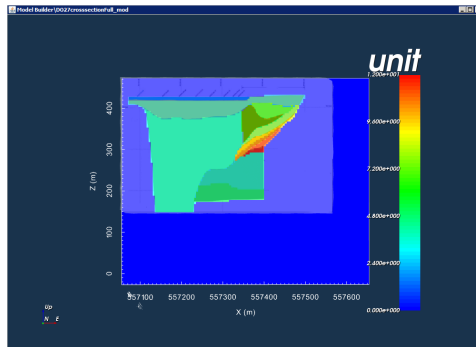
### Cross Section

The cross section case follows much the same procedure with a few exceptions. An imported cross section map is shown overlaid on a 2D mesh in Figure 2.13. Notably no parameter for the vertical extent is needed. The other notable exception is that mesh that is used is a GIF 2D mesh. The result is shown in Figure 2.14.

A 2D Geology model can be used to create constraints for a 2D inversion, it can also be used to add constraints to a 3D inversion as well. After the 2D geology model is created from the cross section map, it can be inserted into a 3D mesh (GIF 3D tensor or OcTree) given a starting and ending position or a starting position and a direction Figure 2.15, Figure 2.16.

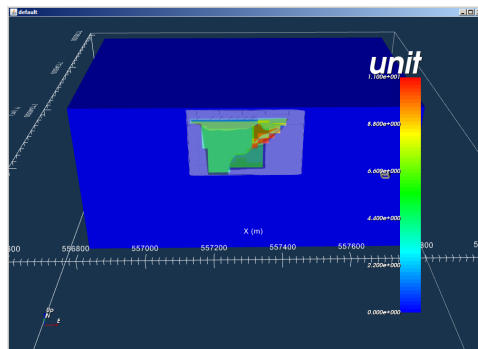


**Figure 2.13:** Example of a 2D mesh with the map overlaid



**Figure 2.14:** Example of a 2D geology model created from a cross section map with the map overlaid

**Figure 2.15:** GUI for adding a 2D model to a 3D model



**Figure 2.16:** Example of a 2D geology inserted into a 3D model with the map overlaid

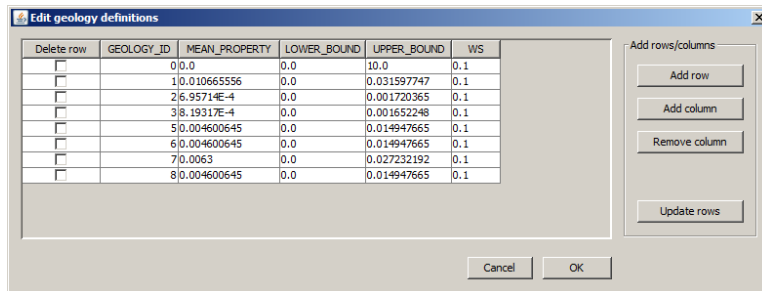


### 2.2.5 Making Constraints for an Inversion

The model that has been created is a geology model. That is, a model in which each cell represents a given geological unit. To be able to convert this model into a constraint for a geophysical inversion the link between between the geology and the petrophysics needs to be provided.

The link is stored in what is called a geology definition. In GIFtools this takes the form of a lookup table that contains information of each particular geological unit's property, lower and upper bounds, and optionally the smallness weight associated with each unit.

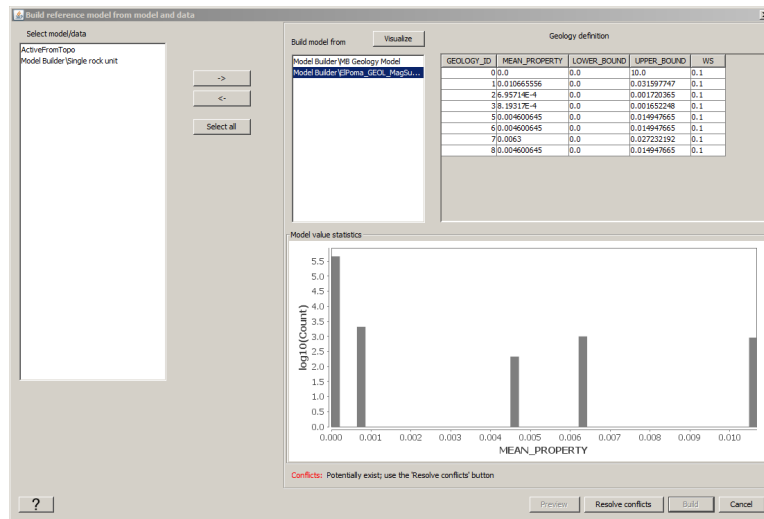
Using the geology definition we can convert a geology model that has information about the spatial distribution of geological units but not of their physical properties into constraints that are usable by an inversion. In the figures below the geological definition came from surface measurements of magnetic susceptibility within each geological unit. Figure 2.17 is an example of a geology definition in the GIFtools GUI.



Delete row	GEOLOGY_ID	MEAN_PROPERTY	LOWER_BOUND	UPPER_BOUND	WS
<input type="checkbox"/>	0	0.0	0.0	10.0	0.1
<input type="checkbox"/>	1	0.010665556	0.0	0.031597747	0.1
<input type="checkbox"/>	2	6.95714E-4	0.0	0.001720365	0.1
<input type="checkbox"/>	3	8.19317E-4	0.0	0.001652248	0.1
<input type="checkbox"/>	5	0.004600645	0.0	0.014947665	0.1
<input type="checkbox"/>	6	0.004600645	0.0	0.014947665	0.1
<input type="checkbox"/>	7	0.0063	0.0	0.027232192	0.1
<input type="checkbox"/>	8	0.004600645	0.0	0.014947665	0.1

**Figure 2.17:** Example of a geological definition as displayed in the GIFtools GUI, the Mean Property, and Bounds are susceptibility measurement in  $\text{SI} \times 10^{-3}$  and the WS is to set the smallness weight model.

Once the geology definition is provided, we can use the Combine Model Dialog (Figure 2.18) in Model Builder to create a reference model and bounds. In this case the resolution of conflicts is trivial as there is a single source of information. Less trivial examples of the creation of reference models and bounds will be discussed later. The resulting reference model is shown in Figure 2.19.



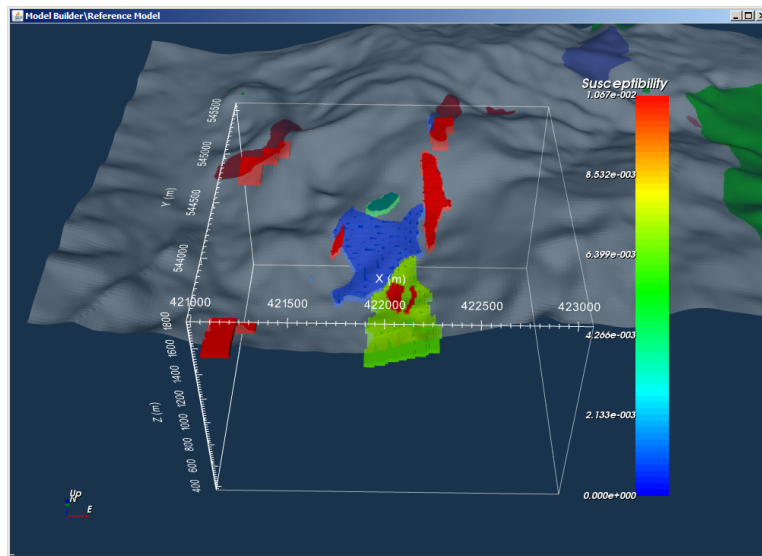
**Figure 2.18:** Example of a typical combine model dialog for a reference model

### 2.2.6 Inputing Fault information from Geological Maps

Another piece of information that can be in geological maps are fault locations. Again in the context of El Poma the map provided a whole complex of thrust faults as shown in the un-doctored map in Figure 2.7

The method used to insert faults into an inversion is as follows:

- Determine the end points of the fault.
  - GIFtools makes this easy by reporting the location of the cursor in the data viewer allowing you to find the location (including elevation) of a point along the fault.
- Using the locations provided GIFtools creates the fault weights by creating a two boxes, each with one of its sides along the fault location as defined in the GUI. By setting the value of cells in each box to one, and then taking the derivative of each of models that had a box added, two face models are created. The location of faces along the fault can be determined by taking the non-zero faces that are in common between the two models. It is then



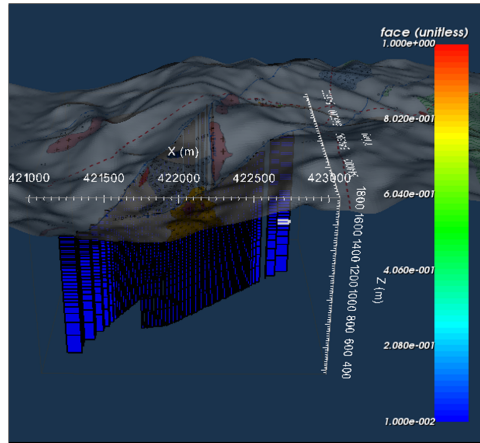
**Figure 2.19:** Example of a reference model created from a geological map

possible to set the values of the the faces that define the fault to any value desired.

- faces within this box are assigned a new value that is provided in the GUI Figure 2.20.

**Figure 2.20:** The GUI for the creation of fault weights

This process can be done multiple times along several consecutive segments of multiple faults to create fault complexes that are not straight or have multiple faults, as shown in Figure 2.21



**Figure 2.21:** An example of fault weights that can be created with GIFtools. In this case a vertical fault (no dip) was created and the blue curtain is the low face weights that define the fault in an inversion. Only the faces that have values that are not 1 are visualized. Since the faces are set at a lower value than the rest of the face model (1) they are shown as blue.

## 2.3 Clustering to Create Constraints

### 2.3.1 Clustering Algorithms

In contexts where multiple data types have been collected over an area, it is often of interest to include information from one inversion result (from one data type) in the regularization of another (likely of a different data type). Much work has been done on this topic (see Section 1.3.2).

One possible way of integrating information into one inversion from another inversion result is through clustering the recovered models to create pseudo-geological models, and then using these to create reference models, bounds, and smoothness weights to constrain subsequent inversions. By pseudo-geologic I mean a geologic

type model that has discrete units and a geologic definition that links them to physical properties, but which is not the result of a geological interpretation of surface or depth lithological data. Rather pseudo-geological models in this thesis tend to be the result of clustering multiple inversion results.

Given several inversions of different data types, multiple models of different physical properties can be recovered. Assuming that each recovered model is on the same mesh (either because all inversions were performed on the same mesh or they were interpolated onto the same mesh after the fact) each cell has a value for each physical property. Since the standard representation of a model is a vector  $\mathbf{m} \in \mathbb{R}^M$  where as in Equation 1.1  $M$  is the number of cells in the discretization of the earth model, we can generalize so that  $\mathbf{M} \in \mathbb{R}^{M \times l}$  is the compiled results of  $l$  inversions. This notation allows us to say that  $\mathbf{M}$  is made out of several  $m_i$  for  $i = 1 \dots M$  and where each  $m_i$  is a row vector  $l$  long.

In each case the algorithm depends on the user providing the number of clusters for analysis. The clustering method used determines the clusters and the membership of each  $m_i$  in the clusters.

The first clustering methods implemented in Model Builder is simply a set of user defined boundaries. These allow a great deal of user control and can be very useful assuming something is known about the underground geology or the physical properties of given units. On the other hand, it can quickly be untenable to use this method if several clusters are needed or many recovered models are being clustered.

The second clustering method implemented in Model Builder is k-means clustering GAF (1984) which is based on the minimization of the following objective function,

$$\phi_{k\text{-mean}} = \sum_{j=1}^k \sum_{m_i \in S_j} \|m_i - v_j\|_2^2, \quad (2.1)$$

where  $k$  is the number of clusters and  $S_j$  are the sets of model cells  $m_i$  that make up the clustering, finally  $v_j$  are the centroids of each cluster, that is the central tendency of the data included in the given cluster  $S_j$ . The algorithm to minimize  $\phi_{k\text{-mean}}$  is a two step process firstly assigning each  $m_i$  to a cluster  $S_j$  and secondly determining a new set of  $v_j$  that better fit the data.

Finally, the third clustering algorithm implemented in Model Builder is Fuzzy C-Means (FCM) (Sun and Li, 2015). FCM is a generalization of Equation 2.1 where instead of each datum  $m_i$  being in only one cluster, each datum is assigned membership in each cluster to varying degrees allowing “fuzziness” in the classification. The objective function becomes

$$\phi_{FCM} = \sum_{j=1}^k \sum_{i=1}^M u_{ij}^q \|m_i - v_j\|_2^2, \quad (2.2)$$

where instead of the  $k$  sets  $S_j$ , membership in each cluster is represented by the membership matrix  $u \in \mathbb{R}^{M \times k}$ . Each row of  $u$  must sum to 1, in other words each datum is evenly weighted in the algorithm but may be classified as partially in each cluster. Finally  $q$ , controls the “fuzziness” of the clustering with a value of  $q = 1$  making a non-fuzzy clustering with each datum being in only one cluster and amount that a given datum can be in multiple clusters increasing as the value of  $q$  increases. A standard value of  $q$  is 2.

In the case of the clustering algorithm used in Model Builder the final step is “de-fuzzification”, that is the determining of the cluster that each model cell most fits in. Each model cell is assigned the cluster for which it has the highest membership value.

In both the k-means and the FCM case the minimization of the objective functions is done by the standard MATLAB function and notably the value of  $q$  in the FCM case is set at 2.

### 2.3.2 Clustering In GIFtools and Model Builder

- describe requirements
- show GUI
- show result

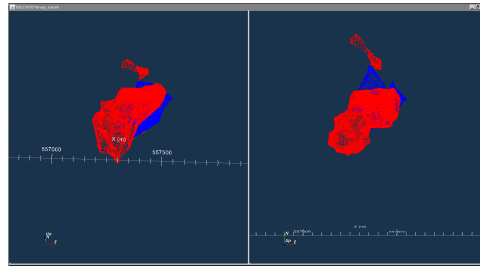
### 2.3.3 Creation of Constraints

- show creation of ref and bounds from mean cluster boundaries
- show creation of sharp bounds with weights.

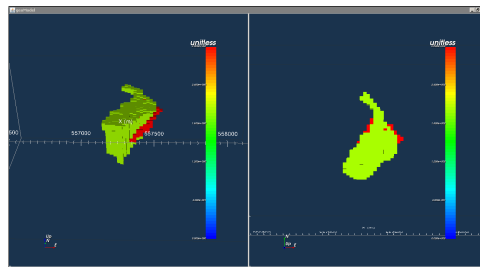
## 2.4 Voxel-Parametric Inversion to Provide Physical Property Values for Geological Models

In Section 2.3 geological type models that span the whole discretized volume are defined. These are more extensive than the geological models described in Section 2.2.4 because each cell is classified into a cluster as opposed to most cells not having any information (due to the map not intersecting the cell at all in the case of most cells).

In addition to pseudo-geological models created by clustering, it is also possible to get geological models from drill or surface geological data. In the TKC case study, one of the forms of geological data that was provided was a set of points created from the interpreted interfaces between units determined using the lithological bore hole data shown in Figure 2.3. The point cloud of the interfaces is shown in Figure 2.22 and the resultant geological model created using the “add non-convex polyhedron” command is shown in Figure 2.23.



**Figure 2.22:** Point cloud of the geological interfaces for TKC bore hole data



**Figure 2.23:** A geological model created from the data shown in Figure 2.22

In the cases listed in (the figure that shows a clustered model) and Figure 2.23,

we have acquired a distribution of geological units that extend across the whole discretized volume it becomes interesting to go further into the assignment of physical properties to the geological units in question. For a clustered pseudo-geological model the obvious method of property assignment is simply using the centroid value recovered from the clustering algorithm. In addition a standard deviation from the mean can be acquired to set the upper and lower bounds of each unit.

Another option is the direct assignment of measured physical property values to the (pseudo-)geological units. This method has the advantage of directly incorporating physical property data into the inversion constraints. Direct assignment can be done through the GIFtools GUI by editing the geologic definition of the geological model, and this method makes most sense when the model in question either is the products of geological modeling as in Figure 2.23 or there is a strong parallel between the clustered units and actual geological units.

In some cases the above method are not optimal. In the case of cluster centroids, artifacts in the inversion can lead to magnitudes in the recovered model not being quite in scale with the physical properties in the ground. Additionally this method only works if the geological model in question was derived through clustering. In the case of the direct assignment of measured physical properties to geological units, the properties often vary a great deal across the earth model. This can be due to large scale variation across the geological unit (such as by alteration) or small scale variation where, for example, one portion of a sample can be significantly more susceptible than another.

I have developed another method to assign physical properties to geological models which is well-described by the title “Voxel-Parametric inversion”. It assumes that a given geological model has correctly determined the distribution of geological units and tries to fit a given geophysical data set assuming that each unit has a single physical property value that is constant across the unit.

The method allows the assignment of physical properties to geological units based on the best fitting set of properties to a data set. This gives an advantage over the methods described above by avoiding some artifacts due to under-determined inversion and avoiding putting undue confidence in a small number of physical property measurements. The method also allows fast and efficient inversion of data without a large number of parameters since the forward problem is much less



complex and regularization beyond the setting of the geological units is unneeded.

### 2.4.1 Formulation of Voxel-Parametric inversion problem

As stated in Section 1.2 and Equation 1.1, the forward problem for which we are trying to find the inverse is formulated as follows

$$\mathbf{d} = \mathbb{F}[\mathbf{m}], \quad (2.3)$$

where as stated above  $\mathbf{d} \in \mathbb{R}^N$  is the geophysical data,  $\mathbf{m} \in \mathbb{R}^M$  is the discretized model that describes the distribution of some physical property in the ground, and  $\mathbb{F}$  is the forward operator that mediates between them.  $\mathbb{F}$  can be considered to be a matrix of size  $N \times M$  which in the case of a non-linear problem has a dependence on the model  $\mathbf{m}$  but in the case of linear problem (such as potential field problems) there is no such dependence. Equation 2.3 can be re-written

$$\mathbf{d} = \mathbf{F}(\mathbf{m})\mathbf{m}, \quad (2.4)$$

where  $\mathbf{F} \in \mathbb{R}^{N \times M}$ , and in the case of a linear problem

$$\mathbf{d} = \mathbf{F}\mathbf{m}. \quad (2.5)$$

In this formulation a geological model can be represented as a sparse matrix  $\mathbf{M}_g \in \mathbb{R}^{M \times k}$  (where  $k$  is the number of geological units) that marks the membership of each model cell in each cluster and a vector  $\mathbf{m}_g \in \mathbb{R}^k$  that has the property of each unit. Using these definitions we can define the standard property model that is represented by a geological model as

$$\mathbf{m} = \mathbf{M}_g\mathbf{m}_g, \quad (2.6)$$

where  $\mathbf{m}$  is a standard voxel model with the each value of each cell being the value assigned to geological unit of that cell. Equation 2.6 can be inserted into Equation 2.5 to form

$$\mathbf{d} = (\mathbf{F}\mathbf{M}_g)\mathbf{m}_g, \quad (2.7)$$

where  $\mathbf{FM}_g \in \mathbb{R}^{N \times k}$  is created by a sparse matrix product and is significantly smaller than  $\mathbf{F} \in \mathbb{R}^{N \times M}$ .

Voxel-Parametric inversion now simply consists of minimizing the following data objective function modified from Equation 1.2

$$\phi_{d-geo} = \|\mathbf{W}_d \left( (\mathbf{FM}_g) \mathbf{m}_g - \mathbf{d}^{obs} \right)\|^2 \quad (2.8)$$

which is significantly less computationally expensive than a full voxel inversion since the forward problem is now much smaller, and no additional regularization is required since the problem is now over determined rather than under-determined.

#### 2.4.2 Uses of Voxel-Parametric inversion results

- assign properties to geological models that fit the collected data as closely as possible
  - allows the geology constraints as in Section 2.3.3 to be more accurate
  - allows the determination of magnetization direction given a geologic or pseudo-geologic model of an magnetization anomaly
  - allows the creation of synthetic models that roughly fit data that has already been acquired

### 2.5 Conclusion

In this section I have shown the creation of constraints that are compatible with GIF inversion codes. I have created these constraints from multiple types of data: bore hole and surface sample data, maps (Cross Section and Plan View), and other inversion results (clustering and voxel-parametric inversion). I have also used different pieces of information from these forms of data to create reference models, bounds, and face weights.



## **Chapter 3**

# **Case Study #1 El Poma**

### **3.1 General Overview of El Poma**

### **3.2 Overview of Deposits**

### **3.3 Discussion of the Geophysical Data Given**

### **3.4 What Information is Available**

### **3.5 Synthetic Model**

### **3.6 Blind Inversion of the Synthetic Model**

### **3.7 Determination of Magnetization Direction**

### **3.8 Creation of Constraints**

#### **3.8.1 $\alpha$ coefficients**

#### **3.8.2 Reference Models**

#### **3.8.3 Weighting matrices**

41

#### **3.8.4 Bounds**

#### **3.8.5 $L_p L_q$ weights**

## **Chapter 4**

### **Case Study #2 TKC**

#### **4.1 Overview of Deposits**

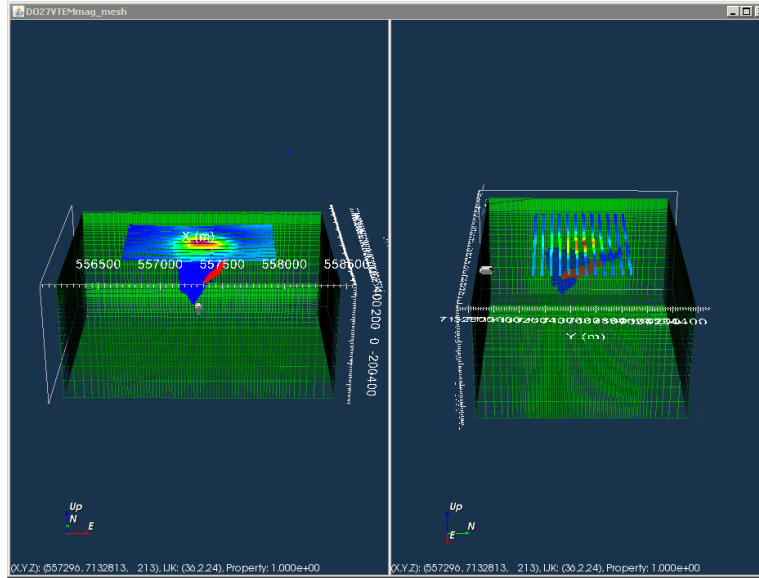
#### **4.2 Discussion of the Geophysical Data Given**

#### **4.3 What Information is Available**

#### **4.4 Synthetic Model**

Given the amount of *a priori* information that has been collected in the region we can make a fairly non-trivial synthetic model to test various forms of constraints. The primary source of information for the creation of the synthetic model are drill hole logs of the kimberlite pipe Eggleston et al. (2014). From these drill holes, a data set representing the outer surface of the main geological units (PK and HK), was created. These are shown in Figure 4.1 along with the VTEM magnetic data over top of DO27, and the mesh used throughout the rest of this section. The of Earth's field in the location has an inclination of  $83^\circ$  and a declination of  $19^\circ$ .

Using the geology unit data sets and a geological model is created Figure 4.2. Given this geology model and using a magnetization vector voxel-parametric inversion we now have a magnetization model that closely matches known geological structure and makes an attempt at fitting the data Figure 4.2. For this paramet-

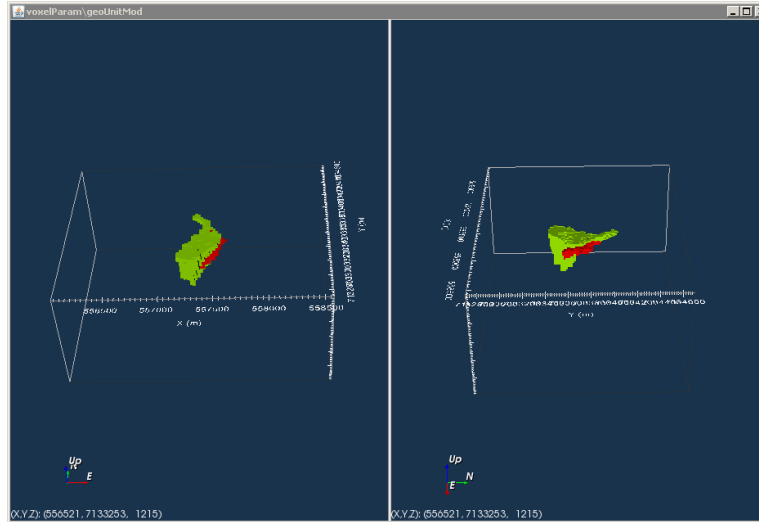


**Figure 4.1:** Mesh used in this chapter along VTEM data, and outer bound of PK (blue) and HK (red)

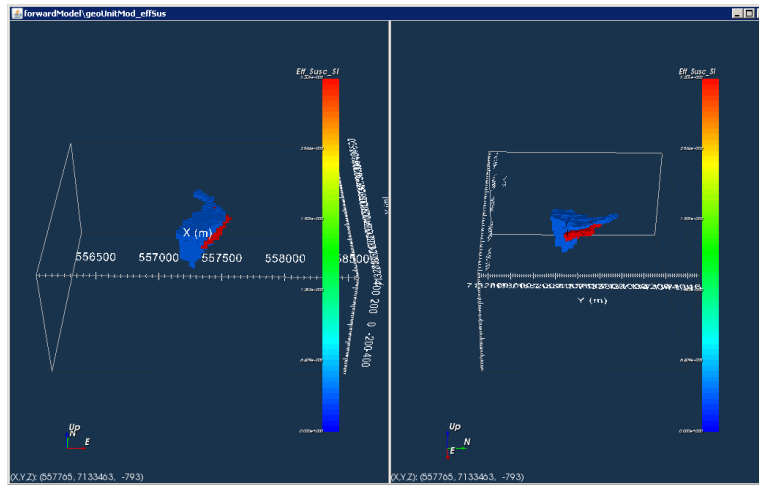
ric inversion result the PK unit has an effective susceptibility of  $3.28\text{E-}3$  (SI) and the HK unit has an effective susceptibility of  $.032$  (SI). These numbers may seem high but it is important to note that they are effective susceptibilities and not true susceptibilities.

Though we do not have oriented Natural Remanent Magnetization (NRM) measurements for TKC samples, the un-oriented NRM measurements that we do have added to the measured susceptibilities give us a rough maximum for effective susceptibilities of the geological in TKC. Given that inversion results imply that the magnetization direction in TKC is close to the earth's field in the location (either show this or cite Sarah's paper), the true values for effective susceptibilities should be near the maximum value.

Measured maximum effective susceptibility values for HK and PK samples from TKC the results roughly fit petrophysical measurements. The average of measured maximum effective susceptibility for PK is  $7.37\text{e-}3$ (SI), and for HK is  $0.045$ (SI) both greater than the recovered values but not so much greater as to imply a very different NRM direction from the direction of earth's field.



**Figure 4.2:** Geology model defined by geology data sets

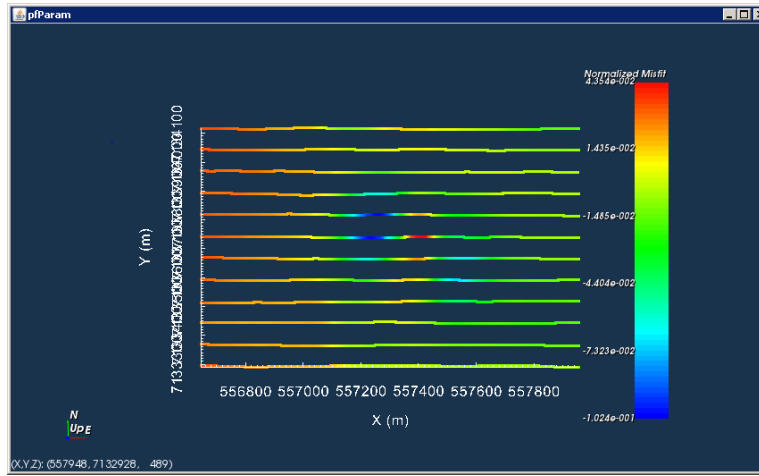


**Figure 4.3:** effective susceptibility result from voxel-parametric inversion

In terms of the recovered directions of magnetization of the units, the PK unit has an inclination of  $83^\circ$  and a declination of  $70^\circ$  (an angular difference of  $6^\circ$ ) and the HK unit has an inclination of  $53^\circ$  and a declination of  $19^\circ$  (a much larger angular difference of  $20^\circ$ ).

The fit of this synthetic model to the measured field data is shown in Figure 4.4,

the effective  $\Phi_d$  of the predicted data is 1.01E6 which is significantly higher than the expected value of 7657. This is because the inversion is over-determined rather than under-determined as is the case in most inversions. Since there are so many fewer degrees of freedom than data, the model will only fit the data so well. In addition the voxel-parametric inversion assumes that the geological model provided is completely true while in reality it is only an approximation based on drilling; it also assumes that each unit has a simple, constant property across its extent which is almost certainly untrue.



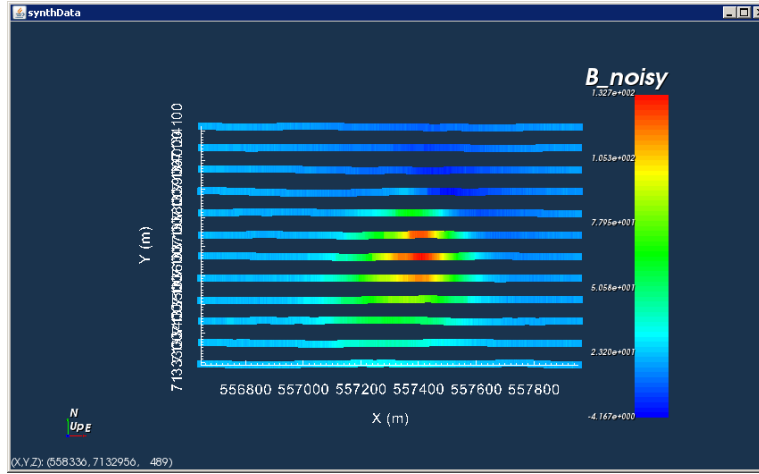
**Figure 4.4:** Normalized misfit of parametric voxel inversion

Given these caveats, the model shown in Figure 4.3 will suit nicely as a sufficiently complex synthetic model for the purposes of showing the capabilities of GIFtools and Model Builder in the performing of inversion on a complicated model with extensive *a priori* information.

To make the inversions easier to perform, the background value (which had an effective susceptibility of 1.09e-3(SI) was set to zero. This avoids having a non-zero reference model in the inversion of the data. The synthetic model with a zero background was forward modeled using the GIF Magnetization Vector Inversion (MVI) forward model code. Figure 4.5 shows the synthetic data contaminated with Gaussian noise with a standard deviation of 1 nanoTesla.

The data shown in Figure 4.5 is the data I will use for the rest of this chapter.



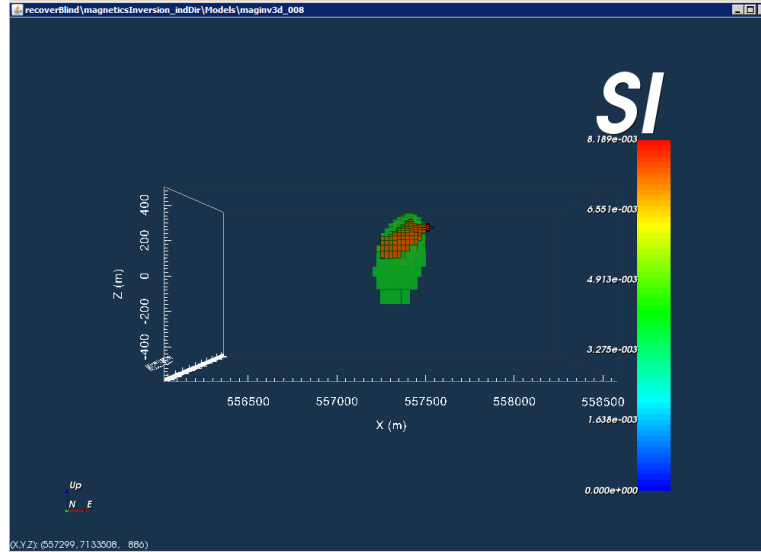


**Figure 4.5:** Forward modeled synthetic data contaminated with Gaussian noise

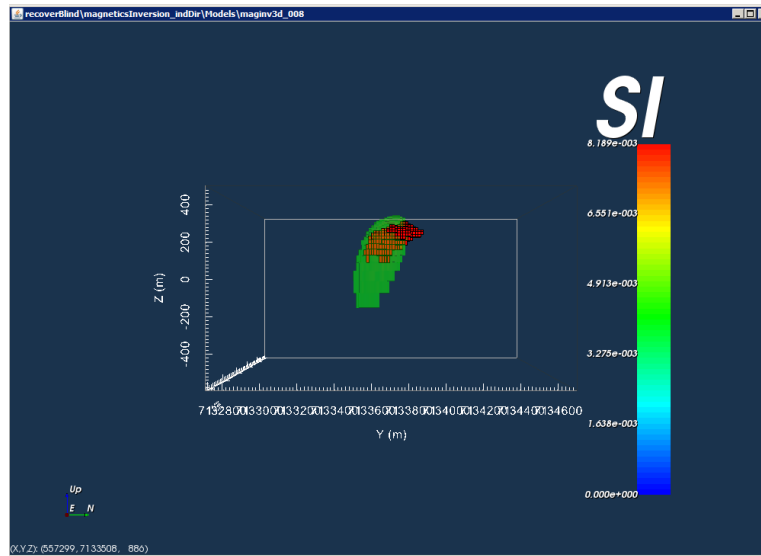
It is the predicted data that would be measured over the model shown in Figure 4.3 with a zero background. From this data I will try to recover an anomaly with a similar shape to the model that created the data using various other forms of data (boreholes, cross section maps etc.)

## 4.5 Blind Inversion of the Synthetic Model

The inversion whose result is shown in Figure 4.6 assumes the the magnetization of the anomaly is in the direction of the earth's field. This is not true as discussed in Section 4.4. The inversion result highlights the position of the HK unit rather than the PK unit. This is expected given the much higher susceptibility of the HK unit. Since the inversion performed used  $L_2$  norms in the MOF the model is not compact and over estimates the size of the susceptibility anomaly, especially below the unit in the synthetic model. Additionally, while the position of the HK unit is well estimated at the depth of the true synthetic HK unit, the recovered anomaly dips quite a bit to the south, implying a very different susceptibility distribution than the true synthetic model. By using extra information from *a priori* sources and other inversions I will make the result shown match more closely with the distribution from the synthetic model shown in Figure 4.3.



(a) View from south



(b) View from east

**Figure 4.6:** Blind inversion result assuming a magnetization direction to be the same as the earth's field in the location of TKC. (Red cells show the extent of the HK unit in the synthetic model, only cells with a effective susceptibility of greater than  $4.0348 \times 10^{-3}$  (SI) are shown)

## 4.6 Determination of Magnetization Direction

One of the first pieces of information that can be inserted into the magnetics inversion is the anomalous magnetization direction. Figure 4.7 shows the result when the true magnetization direction of the synthetic HK unit is provided. The result has several advantages over that of the one shown in Figure 4.6. While the recovered model still greatly overestimates the size of the susceptibility anomaly below the synthetic HK units, anomaly no longer dips giving a better idea of the lateral position of the susceptibility anomaly. Even at the depth of the synthetic HK unit the lateral position of the anomaly is closer to the synthetic truth.

Unfortunately we usually don't have the true magnetization direction of an anomaly. For example we do not have the anomaly direction the actual TKC example. To recover the magnetization direction of the anomaly we can use MVI. Using the GIF MVI code we get the result shown in ??.

- we can perform an MVI inversion

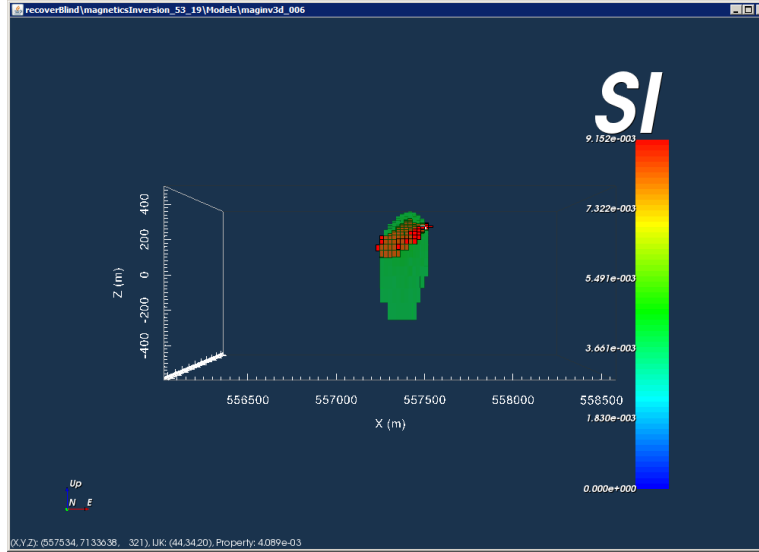
## 4.7 Creation of Constraints and Types of Data

### 4.7.1 $\alpha$ coefficients

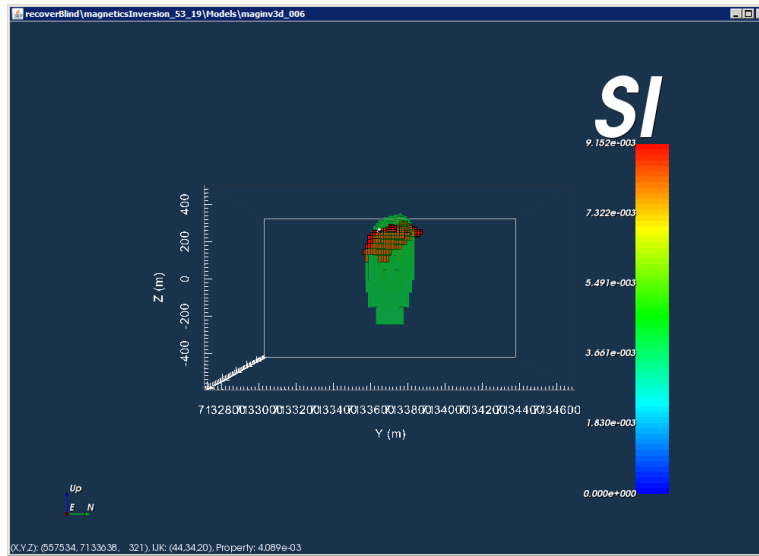
### 4.7.2 Reference Models

### 4.7.3 Weighting matrices

### 4.7.4 Bounds

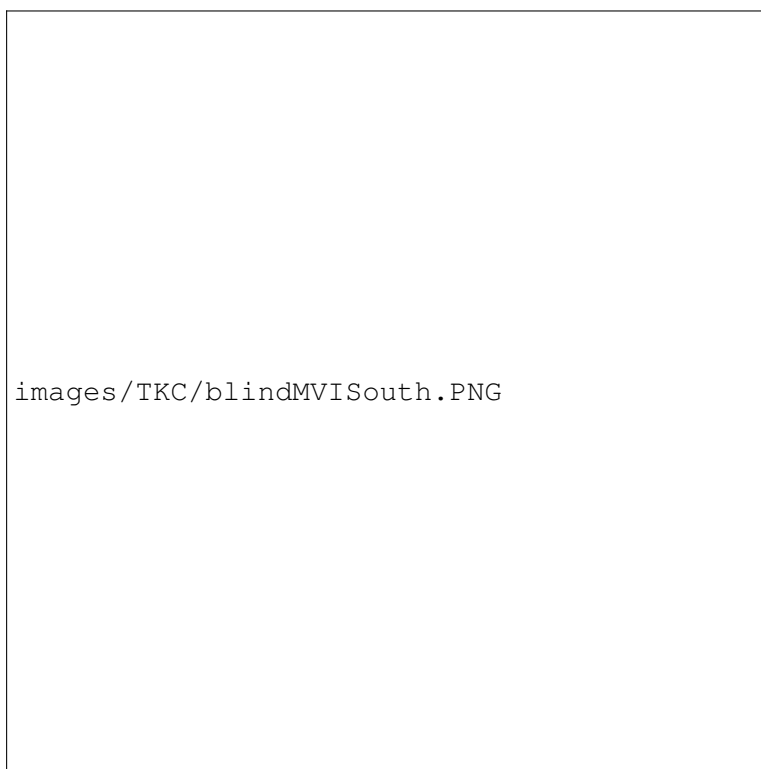


(a) View from south



(b) View from east

**Figure 4.7:** Blind inversion result assuming a magnetization direction to be the magnetization direction of the synthetic HK unit. (Red cells show the extent of the HK unit in the synthetic model, only cells with a effective susceptibility of greater than 4.0348E-03 (SI) are shown)



**(a)** View from south



**(b)** View from east

**Figure 4.8:** Blind inversion result assuming a magnetization direction to be the magnetization direction of the synthetic HK unit. (Red cells show the extent of the HK unit in the synthetic model, only cells with a effective magnetization of  $10^{-4}$  A/m or greater are shown.)

# Bibliography

- Barbosa, V. C., and J. B. Silva, 2006, Interactive 2d magnetic inversion: A tool for aiding forward modeling and testing geologic hypotheses: *Geophysics*, **71**, L43–L50. → pages 8
- Barbosa, V. C. F., and J. B. Silva, 1994, Generalized compact gravity inversion: *Geophysics*, **59**, 57–68. → pages 8
- Bosch, M., A. Guillen, and P. Ledru, 2001, Lithologic tomography: An application to geophysical data from the cadomian belt of northern brittany, france: *Tectonophysics*, **331**, 197–227. → pages 8, 9
- Chasseriau, P., and M. Chouteau, 2003, 3d gravity inversion using a model of parameter covariance: *Journal of applied geophysics*, **52**, 59–74. → pages 8
- Eggleson, T., R. SME, K. Brisebois, P. Eng, and J. Pell, 2014, Peregrine diamonds ltd. lac de gras project northwest territories, canada ni 43-101 technical report. → pages 42
- Farquharson, C. G., M. R. Ash, and H. G. Miller, 2008, Geologically constrained gravity inversion for the voisey's bay ovoid deposit: *The Leading Edge*, **27**, 64–69. → pages 2, 9
- Farquharson, C. G., and D. W. Oldenburg, 1998, Non-linear inversion using general measures of data misfit and model structure: *Geophysical Journal International*, **134**, 213–227. → pages 7
- Fournier, D., 2015, A cooperative magnetic inversion method with lp-norm regularization: PhD thesis, University of British Columbia. → pages 7
- GAF, S., 1984, Multivariate observations. → pages 34
- Guillen, A., P. Calcagno, G. Courrioux, A. Joly, and P. Ledru, 2008, Geological modelling from field data and geological knowledge: Part ii. modelling validation using gravity and magnetic data inversion: *Physics of the Earth and Planetary Interiors*, **171**, 158–169. → pages 8
- Guillen, A., and V. Menichetti, 1984, Gravity and magnetic inversion with minimization of a specific functional: *Geophysics*, **49**, 1354–1360. → pages 8
- Guitton, A., 2012, Blocky regularization schemes for full-waveform inversion: *Geophysical Prospecting*, **60**, 870–884. → pages 7

- Harder, M., C. Hetman, B. Scott Smith, and J. Pell, 2006, Geology of the do27 pipe: a pyroclastic kimberlite in the lac de gras province, nwt, canada: Presented at the Long abstracts, Kimberlite Emplacement Workshop, Saskatoon, Sask. Available at <http://www.venuewest.com/8IKC/files/14%20Harder.pdf> [accessed 18 June 2008]. → pages 21
- Last, B., and K. Kubik, 1983, Compact gravity inversion: *Geophysics*, **48**, 713–721. → pages 7, 8
- Lelievre, P. G., 2009, Integrating geologic and geophysical data through advanced constrained inversion. → pages 2, 9
- Lelièvre, P. G., and D. W. Oldenburg, 2009, A comprehensive study of including structural orientation information in geophysical inversions: *Geophysical Journal International*, **178**, 623–637. → pages 8
- Li, Y., and D. W. Oldenburg, 1996, 3-d inversion of magnetic data: *Geophysics*, **61**, 394–408. → pages 2, 5, 6, 7, 8
- , 1998, 3-d inversion of gravity data: *Geophysics*, **63**, 109–119. → pages 6, 7, 8
- , 2000, Incorporating geological dip information into geophysical inversions: *Geophysics*, **65**, 148–157. → pages 8
- , 2003, Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method: *Geophysical Journal International*, **152**, 251–265. → pages 6
- Oldenburg, D. W., and Y. Li, 2005, Inversion for applied geophysics: A tutorial: *Investigations in geophysics*, **13**, 89–150. → pages 6
- Paasche, H., J. Tronicke, K. Holliger, A. G. Green, and H. Maurer, 2006, Integration of diverse physical-property models: Subsurface zonation and petrophysical parameter estimation based on fuzzy c-means cluster analyses: *Geophysics*, **71**, H33–H44. → pages 9
- Phillips, N. D., 2001, Geophysical inversion in an integrated exploration program: Examples from the san nicolas deposit. → pages 2, 9
- Portniaguine, O., and M. S. Zhdanov, 1999, Focusing geophysical inversion images: *Geophysics*, **64**, 874–887. → pages 7
- Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms: *Physica D: Nonlinear Phenomena*, **60**, 259–268. → pages 7
- Sun, J., and Y. Li, 2015, Multidomain petrophysically constrained inversion and geology differentiation using guided fuzzy c-means clustering: *Geophysics*, **80**, ID1–ID18. → pages 9, 35
- Vogel, C. R., and M. E. Oman, 1998, Fast, robust total variation-based reconstruction of noisy, blurred images: *Image Processing, IEEE Transactions on*, **7**, 813–824. → pages 7

Williams, N. C., 2008, Geologically-constrained ubc–gif gravity and magnetic inversions with examples from the agnew-wiluna greenstone belt, western australia. → pages 2, 3, 9, 11, 21



## **Appendix A**

# **Supporting Materials**

-