Application of Physical Properties Measurements to Lithological Prediction and Constrained Inversion of potential field data, Victoria Property, Sudbury, Canada

by

Omid Mahmoodi

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy (PhD) in Mineral Deposits and Precambrian Geology

The Faculty of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

(This sheet will be inserted by Graduate Studies; please leave a blank page)

Thesis Defence Committee/Comité de soutenance de thèse

Laurentian Université/Université Laurentienne

Faculty of Graduate Studies/Faculté des études supérieures

Title of Thesis
Titre de la thèse: Application of Physical Properties Measurements to Lithological Prediction and Constrained Inversion of potential field data, Victoria Property, Sudbury, Canada

Name of Candidate
Nom du candidat: Omid Mahmoodi

Degree
Diplôme: Doctor of Philosophy

Department/Program
Département/Programme

Date of Defence
Date de la soutenance: November 30, 2015

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:

Dr. Richard S. Smith
(Supervisor/Directeur de thèse)

Dr. Douglas k. Tinkham
(Committee member/Membre du comité)

Warren Hughes
(Committee member/Membre du comité)

Dr. Ernst Schetselaar
(External Examiner/Examinateur externe)

Dr. Kalpdrum Passi
(Internal Examiner/Examinateur interne)

Approved for The Faculty of Graduate Studies
Approuvé pour la Faculté des études supérieures
Dr. David Lesbarrères
M. David Lesbarrères
Acting Dean, Faculty of Graduate Studies
Doyen intérimaire, Faculté des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

Abstract

In recent years the number of near-surface deposits has decreased significantly; consequently, exploration companies are transitioning from surface-based exploration to subsurface exploration. Geophysical methods are an important tool to explore below the surface. The physical property data are numerical data derived from geophysical measurements that can be analyzed to extract patterns to illustrate how these measurements vary in different geological units. Having knowledge of links between physical properties and geology is potentially useful to obtain more precise understanding of subsurface geology.

Firstly, down-hole density, gamma radioactivity, and magnetic susceptibility measurements in five drillholes at the Victoria property, Sudbury, Ontario were analyzed to identify a meaningful pattern of variations in physical property measurements. The measurements grouped into distinct clusters identified by the fuzzy k-means algorithm, which are termed 'physical log units'. There was a meaningful spatial and statistical correlation between these physical log units and lithological units (or groups of lithological units), as classified by the geologist. The existence of these relationships suggests that it might be possible to train a classifier to produce an inferred function quantifying this link, which can be used to predict lithological units and physical units based on physical property data. A neural network was trained from the lithological information from one hole, and was applied on a new hole with 64% of the rock types being correctly classified when compared with those logged by geologists. This misclassification can occur as a result of overlap between physical properties of rock types. However, the predictive accuracy in the training process rose to 95% when the network was trained to classify the physical log units (which group together the units with overlapping properties).

Secondly, lithological prediction based on down-hole physical property measurements was extended from the borehole to three-dimensional space at the Victoria property. Density and magnetic susceptibility models were produced by geologically constrained inversion of gravity and magnetic field data, and a neural network was trained to predict lithological units from the two physical properties measured in seven holes. Then, the trained network was applied on the 3D distribution of the two physical properties derived from the inversion models to produce a 3D litho-prediction model. The lithologies used were simplified to remove potential ambiguities due to overlap of physical properties. The 3D model obtained was consistent with the geophysical data and resulted in a more holistic understanding of the subsurface lithology.

Finally, to extract more information from geophysical logs, the density and gamma-ray response logs were analyzed to detect boundaries between lithological units. A derivative method was successfully applied on the down-hole logs, and picked the boundaries between rock types identified by geologists as well as additional information describing variation of physical properties within and between layers not identified by the geologist.

Keywords

Down-hole physical properties, fuzzy k-means clustering, neural network, classification, constrained geophysical inversion, litho-prediction model, derivative analysis, layer boundaries

Co-Authorship Statement

This manuscript is composed of six chapters of which four (Chapters 2 to 5) are research papers. Chapter 2 has been published in the Journal of Applied Geophysics. Chapter 3 has been submitted to the Journal of Applied Geophysics. Chapter 4 and 5 are going to be submitted to peer-reviewed journals.

The first paper (Chapter 2) is co-authored by Dr. Richard S. Smith; the second paper (Chapter 3) is co-authored by Richard S. Smith and Douglas K. Tinkham; the third paper (Chapter 4) is co-authored by Richard S. Smith and Bill Spicer; and the fourth paper (Chapter 5) is co-authored by Dr. Richard S. Smith. My supervisor, Dr. Smith put together the original research proposal, and guided the research by his support and advice. All papers and the thesis are edited by Dr. Smith, Dr. Tinkham, and Warren Hughes.

Acknowledgments

First and most, I am thankful to my thesis supervisor, Dr. Richard S. Smith for his invaluable support and guidance. Since the very early days of my research at Laurentian University he helped me to adapt to a new situation and get on track. He was an incredible supervisor spending plenty of time with students listening to their ideas and questions carefully. Weekly appointments to discuss progress made on the research, prompt responses and comments on my drafts and his editorial role in my project are highly appreciated.

I would also like to acknowledge the helpful discussions and advice provided by my committee, Warren Hughes, and Dr. Dougls K. Tinkham. I am also thankful to the external and internal examiners, Dr. Ernst Schrtselaar and Dr. Kalpdrum Passi for their valuable input during my defense.

This research was part of an Industrial Research Chair (IRC) financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Vale, Sudbury Integrated Nickel Operations, a Glencore company, KGHM International, Wallbridge Mining and the Centre for Excellence in Mining Innovation (CEMI). In addition to the financial support, I would also like to acknowledge in particular Bill Spicer and John Everest from KGHM International, Warren Hughes from Sudbury Integrated Nickel Operations, a Glencore company, Sean Dickie from Vale, Joshua Baily from Wallbridge Mining for their meaningful discussions and tips during semi-annual IRC meetings.

All datasets used in this project were provided by KGHM International. I would like to express my appreciation to their valuable input to this research. I am also grateful to KGHM International Ltd and the staff, in particular, John Everest and Bill Spicer, for their support during the time I worked there as a Student Geophysicist. I was fortunate to work closely with Bill Spicer and have his continuous scientific and technical support during my research.

I would also like to thank Vince Gerrie and Chris Drielsma from DGI Geosciences for their valuable guidance about the measured down-hole physical property data. Moreover, Aaron Davis, a geophysicist from CSIRO provided a meaningful contribution to this research.

I am grateful to the all staff of the Earth Sciences Department for their helpful advices and comments during my seminar talks. In particular, I would like to thank Dr. Bruno Lafrance, the Dean of Earth Sciences Department for giving me the opportunity to lecture an undergraduate-level course and obtain remarkable experiences.

For the past three years, I have been very lucky to make incredible friends in Sudbury; in particular Remy Poulin, Christopher Beckett-Brown, Craig Stewart, Evan Hastie and my fellow geophysicists Olaniyan Oladele, Frederic Guacher, Yongxing Li, Tom Naprstek, Michal Kolaj, Reza Mir, and Nikolas Gazo. I am more than grateful for my best friend, Christoph Schaub. A smart geophysicist whose efforts to teach me ice skating, hockey, sailing, and fishing, will not be forgotten. I will always remember his key to success is to "keep your stick on the ice".

More than anything in the world, I am thankful to my beloved parents, Gholam Reza and Afsaneh for always being there to support me through their endless encouragement and unwavering love. I would not make this journey through to this special day without you. I would like to thank my brother, Abbas who has been a role model for me since I opened my eyes to the world. I would also like to thank Shokooh for being the most supportive and kindest sister ever.

I am certain, I could not complete my research without Anahita's support and encouragement. You have always been the reason to remain positive and motivated even through difficulties. I am more than grateful to you.

Table of Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

1.1 Research motivation

Geophysical logs are high resolution data reflecting variation of physical characteristics of rocks surrounding a borehole. These down-hole measurements can add to the information we can extract from an expensive borehole. Geophysical logs become critical for subsurface lithological characterization when the core recovery is poor or zero (Benaouda, et.al, 1999 and Qi and Carr, 2006). The main objective of this thesis is to understand the links between variations in the physical properties and how they correlate with changes in the lithological units at the Victoria property located in Sudbury, Ontario. Ideally, we could use this knowledge to characterize lithological variation based on geophysical logs measured in a new hole; or create a 3D lithological model from a subsurface distribution of physical properties that is derived from ground or airborne geophysical measurements. The Victoria property was selected as a case study to evaluate the applicability of the down-hole geophysical analyses in a complex igneous/metamorphic environment since previous similar published works have mainly focused on sedimentary environments.

Success in this research can benefit an exploration project at different stages: 1) the geophysical logs can be measured during or immediately after drilling a new borehole; while, lithological core-logging takes a much longer time. Analyzing the down-hole physical properties measurements can provide valuable information on the geological/geophysical characteristics of

the borehole. Such a quickly produced image of the borehole gives an initial insight into the subsurface geology, which can be a basis to plan further analyses and can help guide or focus the geologist during the lithological logging process. Perhaps in the future, the physical logs could be used to automate the more mundane part of manual geological logging and allow the geologist to focus on the more interesting parts. 2) When core-logging has been undertaken, the geophysical logs can be used to predict lithological units. The predicted lithological variation derived from the geophysical logs could be used by the geologist to check their manually generated geological logs. Discrepancies between the two logs may draw attention to certain parts of the log and the geologist can then re-check the cores and modify the lithological log. In addition, the high resolution of the down-hole survey motivates using these data to enhance the accuracy of the location of lithological boundaries. 3) In a more advanced stage of exploration, the geophysical logs can effectively contribute to convert 3D petrophysical models to a 3D litho-prediction model. Down-hole physical property measurements can constrain the inversion of geophysical field data to produce petrophysical models and train a classifier to predict lithological units in a 3D space.

The results in this work are not only interpreted to demonstrate the applications explained in the previous paragraph at a specific study area (the Victoria property), but also to provide a general framework describing critical factors which might be useful for similar studies.

1.2 Down-hole physical property measurement

Geophysical borehole logging, also known as downhole geophysical surveying or wire-line logging is the branch of geophysics in which sensors are lowered down the hole to infer near-

continuous in-situ physical properties of the sequences of rock penetrated by the borehole (Lau, 1998, Keary, et al., 2002; Granek, 2009; Hearst et al., 2000). Killeen (1997) provided a general introduction to data acquisition methods and the applications of downhole physical properties measurements in mineral exploration.

Physical properties measurements are categorized into passive or active methods. In passive methods, the response emanating naturally from the surrounding rock is measured; however, in active methods, the response is measured from surrounding rocks as a result of emitted energy through an artificial source. Multi-probe in-hole measurements can be used to acquire a large amount of in-situ data, rapidly (Lau, 1998 and Keary, et al., 2002). Among the various geophysical logs, only three logs which are used in this thesis are described here:

- **Natural Gamma Logs**

The natural gamma-ray tool is equipped with a scintillation detector to measure gamma rays naturally emitted by the rocks. The number of counts of gamma rays detected is an indication of the radioactive elements, particularly K, U, and Th in the surrounding rock. The natural gamma log can be used to distinguish lithology with different concentrations of radioactive elements, and alteration zones. Typically, natural gamma log shows high values for shale, igneous and metamorphic rocks including felsic minerals, porphyritic zones and altered rocks; on the other hand, low values correspond to sandstone and calcareous rocks (Lau, 1998 and Keary, et al., 2002). The simplest tools measure all gamma rays between 0.1 and 300 MeV. These are the tools use to acquire the data analyzed in this project. More sophisticated tools can measure gamma-rays in three additional energy windows, the K windows (from 1.37 to 1.57 MeV), the U window (from 1.66 to 1.86 MeV) and the Th window (from 2.40 to 2.80 MeV).

- **Density Logs or Gamma-Gamma Logs**

The density or gamma-gamma log is used to estimate formation bulk density in g/cm^3. An artificial gamma ray is emitted from a ^{60}Co or ^{137}Cs source mounted on the probe towards the formation and gamma rays back-scattered from the surrounding rocks are received by the scintillation detector. Elastic collisions between gamma rays emitted toward the rock and electrons in the rock results in Compton scattering, which is the release of an electron and scattering of the gamma ray with a lower energy. The gamma radiation measured by detector is inversely proportional to the electron density (electron density index) which is a function of the bulk density of the formation. The denser the rocks, the more gamma rays are scattered by the Compton effect, and consequently the counts of gamma ray measured by detector are decreased. The gamma-gamma log can be used to identify lithology, porous formation, bed thickness, water content, compaction, alteration associated with compaction and/or cementation and the presence of heavy elements such as base metals, iron, and magnesium (Lau, 1998 and Keary, et al., 2002).

- **Magnetic susceptibility log**

The magnetic susceptibility log can be measured using a down-hole probe (in dry holes or holes filled with non-conductive fluid) to estimate the formation magnetic susceptibility in $SI \times 10^{-3}$. The magnetic susceptibility corresponds to the quantity of ferromagnetic magnetic minerals such as magnetite and pyrhotite (Lau, 1998 and Keary, et al., 2002). A magnetic susceptibility log can be also used to locate hydrothermal alteration since magnetite can be oxidized to a non-magnetic mineral, hematite (Hrouda, et al., 2009).

Density, magnetic susceptibility, gamma-ray response, resistivity and fluid resistivity logs were available for seven holes at Victoria property. However, the resistivity and fluid resistivity were

excluded from the analyses because there is a lack of meaningful correlation between rock type and variation in these logs.

1.3 Physical characterization of lithological units

Analyzing geophysical logs can help to understand the links between the in-situ physical properties data and the geology. Geological properties such as lithology, permeability, porosity, water and hydrocarbon saturation, formation dip, thickness, and temperature may be derived from geophysical borehole logging (Keary, et al., 2002). The principles developments have been in sedimentary environment mainly for hydrocarbon exploration, where logging is well established (Baldwin et al., 1990; Wong et al., 1995; Farmer and Adams, 1998; Benaouda et al., 1999; Qi and Carr, 2006; Maiti et al., 2007; Ojha and Maiti, 2013). In this research the primary aim is to understand the lithology in a hard rock environment (the Victoria property) based on down-hole physical property measurements. The geological complexity of an igneous/metamorphic environment distinguishes this work from previous works which have mostly been conducted on sedimentary environment.

The extent of the contrast of physical properties between rock types is the key to success in rock-type characterization (Perron et al., 2011 and Mwenifumbo and Mwenifumbo, 2012). Due to complete or partial overlap of the physical properties that rock types may have, it is not always possible to differentiate between rock types on the basis of a limited set of physical properties measurements (Rabaute, et al., 2003 and Garcia et al., 2011). In Chapters 2 and 3, I describe the similarity of physical properties of some rock types and show that the heterogeneity of the rock composition increases the overlap between physical properties of different rocks. This overlap

impedes the complete differentiation between rock types based on their physical properties and confuses the classification. The physical properties can be used to categorize the lithological units, but the overlap of physical properties between two rocks brings confusion to the prediction of rock types.

I experimented with classifying the rocks into less-ambiguous physical log units rather than the classification used by the geologist. Sometimes physical properties do not have a direct relationship with rock type, so there is not necessarily consistency between physical units and the lithological units determined by geologists. Lithological classification is mainly determined based on grain size and composition; while, geophysical logs are responsive to physical properties of rocks. Moreover, lithological classification is usually undertaken by more than one geologist, which can introduce inconsistencies due to variabilities in how different experts apply subjective criteria. The term "physical log units" can be described as homogenous intervals of one or more rock types with consistent physical properties. (Benaouda, et al., 1999 and Perron et al., 2011).

In Chapter 3 both lithological and physical unit prediction are performed to explain the strength and weakness of each method. I found that the physical units were usually related to one or more lithological units and that the inconsistencies identified in certain locations might be used as a guide to where the geologist should look more carefully at the rocks.

1.4 Pattern recognition techniques

Conventional exploratory statistical techniques, such as graphical representation of data using histograms, box and whisker plots, cross plots, or the analysis of average and variance, were

employed by Reed (1997), Killeen et al. (1997), McDowell et al. (1998), McDowell et al. (2004), and Vella and Emerson (2009) to analyze physical properties measurements. Traditional multivariate statistical techniques of supervised learning such as discriminant analysis (Bosch et al., 2002; Benaouda, et.al, 1999; Sakurai and Melvin, 1988; Eberle, 1992), maximum likelihood (Mehta, et al. 1990), and Gaussian classification (Delfiner et al., 1984) have also been successfully applied to geophysical logs to infer lithology or lithofacies. The use of modern pattern recognition techniques has been advocated to quantitatively link the geological and geophysical data, as they are felt to reduce the subjective bias in analyses compared to conventional techniques (Granek, 2011; Williams and Dipple, 2007). In addition, multivariate pattern recognition techniques allow multiple physical properties to be included in the analyses at once (Rabaute, et al., 2003; Qi and Carr, 2006; Williams and Dipple, 2007; Garcia, et al., 2011; Granek, 2011). Two classes of pattern-recognition techniques, supervised and unsupervised classification, are used to analyze the data in this thesis (Theodoridis and Koutroumbas, 2003).

Unsupervised classification methods, typically referred to as "clustering techniques", analyze unlabeled data to extract the pattern of variation in the data. These techniques are used when the distribution of data is important regardless of the class to which the measurements belong. The data are grouped in a number of clusters so that the data within each cluster have consistent characteristics and are different from other clusters (Theodoridis and Koutroumbas, 2003). These clusters are called physical units in this work. The fuzzy k-means algorithm is a clustering technique that provides outputs which can describe a fuzziness of classification due to natural overlap between physical rock properties for different lithological units. This technique has been successfully applied to down-hole geophysical logs mostly in sedimentary (softrock)

environments. Rabaute, et al. (2003) used the fuzzy k-means to identify the stratigraphy through clustering of physical-property measurements. Garcia, et al. (2011) applied the fuzzy k-means clustering and principal component analysis to down-hole geophysical logs in a sedimentary environment to identify relatively homogenous zones. The fuzzy k-means clustering was employed by Ojha and Maiti (2013) to identify physical units characterizing sedimentary rocks. They implemented the clustering technique on down-hole logging data measured during the International Ocean Discovery Program (IODP) to identify geological characteristics where core recovery is poor or zero. Application of clustering techniques to analyze physical properties measurements from a complex igneous/metamorphic environment (i.e. the Victoria property) is an interesting challenge which has not been well demonstrated in previous published works.

Chapter 2 discusses the results of application of a fuzzy k-means clustering technique on geophysical logs for five boreholes from the Victoria property. Physical characterization of lithological units is further explained by comparing the obtained physical units and corresponding lithological units. A workflow including the processing and interpretation steps in this Chapter can guide future use of fuzzy k-means clustering in other down-hole physical properties analyses.

Supervised classification refers to the techniques in which training data are labeled or classed and a classifier is designed to simulate the relationship between data and their corresponding class (Theodoridis and Koutroumbas, 2003). In physical properties analyses these methods can be used for prediction purposes. The classifier is trained using physical properties measurements and rock type information in a drillhole or parts of a hole, and then the trained classifier can be applied to a new hole or those parts lacking core to predict the rock types (Benaouda, et.al, 1999 and Qi and Carr, 2006). The neural network technique is a robust non-linear classifier

successfully applied to down-hole logging data to predict lithofacies and lithological units mostly in sedimentary environments (Baldwin et al., 1990; Wong et al., 1995; Farmer and Adams, 1998; Benaouda et al., 1999; Qi and Carr, 2006; Maiti et al., 2007; Ojha and Maiti, 2013). Application of this method to igneous and metamorphic environments like the Victoria property might be complicated as they can be more complex than sedimentary environments. In Chapter 3, neural networks are trained from data in one borehole and used to predict lithological and physical units of a second borehole. The predictions worked reasonably when used to predict lithology, but much better when used to predict the physical log units that we defined in Chapter 2. The study of Chapter 3 allows a comparison of the two types of predictions and provides an understanding of the strengths and weaknesses of the two approaches.

1.5 Three-dimensional litho-prediction model

Analyzing the geophysical logs by pattern recognition techniques provides a robust tool to predict simplified lithological units in a one-dimensional drillhole. I was interested in seeing if this approach can also effectively be used in three-dimensional space to produce a 3D geological model based on physical properties defined in 3D space. To do so, the first step is to produce distribution of physical property within the volume of interest. These petrophysical models can be obtained from 3D inversion of potential field data. This is complementary to recent work in how 3D inversion of potential field data can contribute to the difficult task of mineral exploration (Williams and Dipple; 2007; Williams, 2008; Spicer et al., 2011; Louro and Mantovani; 2012; Wang et al., 2012; Ribeiro et al., 2013; Zhou, et al., 2015).

In this thesis, among several inversion programs GRAV3D and MAG3D developed by Geophysical Inversion Facility of University of British Colombia (UBC-GIF) were chosen for gravity and airborne magnetic data inversion at the Victoria property. These programs allow for incorporating prior geological information as constraints to guide the inversion towards a result which is consistent with the known geology (Li and Oldenburg, 1996, 1998). The geology at Victoria is fairly well known, and this information was provided by KGHM International in the form of a 3D geological model which is used as the reference model for the gravity and airborne magnetic data inversion. Then, the 3D density and magnetic susceptibility models are introduced to a neural network trained using the down-hole data. The output of the network is prediction of the simplified lithological units in 3D space (a 3D litho-prediction model).

1.6 Layer boundary detection

Geophysical logs are high resolution data which reflect the variation of physical properties down the hole. They can be considered as a capable means to identify changes in the lithology based on changes in the physical property (Mwenifumbo and Mwenifumbo, 2012). Lithological boundaries can be associated with high-frequency variations in the physical properties, and distinguishing these rapid variations in noisy data is difficult. Cowan and Cooper (2003) and Cooper and Cowan (2009) suggested using a technique to process the log data so as to produce a smooth and unchanged log through homogenous units and sharp edges at the boundaries between individual units. Since these layer boundaries correspond to inflection points in the measured properties log, computing the second derivative of the log and then identifying zero values was proposed as an effective method to detect the sought-for layer boundaries. Chapter 5 shows a novel application of a derivative analysis developed by Davis and Christensen (2013). Their technique was developed for use on a single physical property log, but in this thesis the

technique has been extended to work on both the density and gamma-ray response measurements.

1.7 Thesis arrangement

This manuscript is composed of six chapters of which four (Chapters 2 to 5) are research papers. Chapter 2 has been published in the Journal of Applied Geophysics. Chapter 3 has been submitted to the Journal of Applied Geophysics. Chapter 4 and 5 are going to be submitted to peer-reviewed journals.

Chapter 2 discusses the application of the fuzzy k-means technique to divide down-hole physical property measurements into classes with consistent physical characteristics. Chapter 3 describes lithological and physical log parameter prediction based on down-hole physical property measurements using a neural network. In Chapter 4, the links between lithological units and physical properties is employed by a neural network to convert two 3D petrophysical models derived from inversion to a 3D litho-prediction model. Chapter 5 provides the results of lithological boundary detection by applying a derivative analysis on two geophysical logs. Finally, the results of this work and suggestions for future work are summarized in Chapter 6.

1.8 References

Baldwin, J., Bateman, A., and Wheatley, C. (1990). Application of neural network to the problem of mineral identification from well logs. The Log Analyst, 31, 279–293.

Benaouda, D., Wadge, G., Whitmarsh, R. B., Rothwell, R.G. and MacLeod, C. (1999). Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the Ocean Drilling Program, Geophysical Journal International, 136, 477-491. DOI: 10.1046/j.1365-246X.1999.00746.x

Bosch, M., Zamora, M., Utama, W. (2002). Lithology discrimination from physical rock properties. Geophysics 67(2), 573-581.

Cooper, G. R. J., Cowan, D., (2009). Blocking geophysical borehole log data using the continuous wavelet transform. Exploration Geophysics 40(2), 233–236.

Cowan, D., Cooper, G. R. J., (2003). Wavelet analysis of detailed drillhole magnetic susceptibility data, Brockman Iron Formation, Hamersley Basin, Western Australia. ASEG Extended Abstracts 2003 (January (2)), 1–4.

Davis, A., and Christensen, N., (2013). Derivative analysis for layer selection of geophysical borehole logs. Computer and Geosciences 60, 34-40.

Delfiner, P, Serra, O., and Peyret, O. (1984). Automatic determination of lithology from well logs: 59th Annual Technical Conference and Exhibition, Society of Petroleum Engineers, paper SPE 13290.

Eberle, D. (1992). Interpretation of multi-method geophysical borehole logging data from KTB Oberpfalz VB using multivariate statistical analyses. Scientific Drilling, 3, 16-26.

Farmer, R., and Adams, S. J. (1998). Facies recognition using neural networks. New Zealand Petroleum Conference. Queenstown.

Garcia, M. H., Rabaute, A., Yven, B., and Guillemot, D. (2011). Multivariate and spatial statistical analysis of Callovo-Oxfordian physical properties from lab and borehole logs data: Towards a characterization of lateral and vertical spatial trends in the Meuse/Haute-Marne Transposition Zone. Physics and Chemistry of the Earth, 36, 1469–1485.

Granek, J. (2011). Computing geologically consistent models from geophysical data. BC, Canada: University of British Coloumbia.

Hearst, J. R., Nelson, P. H., and Paillett, F. L. (2000), Well logging for physical properties, John Wiley and Sons, Chichester.

Hrouda, F., Chlupacova, M., and Chadima, M. (2009). The use of magnetic susceptibility of rocks in geophysical exploration. Terraplus.

Keary, P.; Brooks, M.; and Hill, I., (2002). An introduction to geophysical exploration. 3rd Edition, Blackwell Scientific Publications, Oxford, 262 p.

Killeen, P. (1997). Borehole geophysics: Exploring the third dimension. In A. Gubins (Ed.), Exploration 97: Fourth Decennial International Conference on Mineral Exploration, (pp. 31–42).

Lau, K.C. (1998). A review of downhole geophysical methods for ground investigation. Technical Note No. TN 4/98, Geotechnical Engineering Office, Hong Kong, 67 p.

Li, Y., and Oldenburg, D. (1996). 3-D inversion of magnetic data. Geophysics, 61, 394–408.

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

Louro, V.H.A. and Mantovani, M.S.M. (2012) 3D inversion and modeling of magnetic and gravimetric data characterizing the geophysical anomaly source in Pratinha I in the southeast of

Brazil. Journal of Applied Geophysics, 80, 110-120.
http://dx.doi.org/10.1016/j.jappgeo.2012.01.013

Mehta, C., H., Radhakrishnan, S., Srikanth, G. (1990). Segmentation of well logs by maximum-likelihood estimation. Mathematical Geology, 22(7), 853-869.

Maiti, S., Tiwari, R. K., and Kumpel, H. J. (2007). Neural network modelling and classification of lithofacies using well log data: a case study from KTB borehole site. Geophysics Journal International, 169, 733–746.

McDowell, G. M., King, A., Lewis, R. E., Clayton, E. A., and Grau, J. A. (1998). In-situ nickel assay by prompt gamma neutron activation wireline logging. SEG Annual Meeting. New Orleans, Louisiana.

McDowell, G., Fenlon, K., and King, A. (2004). Conductivity-based nickel grade estimation for grade control at Inco‟s Sudbury mines. SEG Annual Meeting. Denver, Colorado.

Mwenifumbo, C., and Mwenifumbo, A. (2012). Borehole geophysical logging in the Flin Flon Mining Camp. Geological Survey of Canada, Open File 6547.

Ojha, M., and Maiti, S. (2013). Sediment classification using neural networks: An example from the site-U1344A of IODP Expedition 323 in the Bering Sea. Deep-Sea Research II.

Perron, G. F., Phillips, N., Williston, C., Gerrie, V., and Everest, J. (2011). 3D litho-prediction model from in-situ physical rock property logging and constrained potential fields data inversion. 31st Gocad Meeting. Nancy.

Qi, L., and Carr, T. R. (2006). Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas. Computers and Geosciences, 32, 947–964.

Rabaute, A., Yven, B., Chelini, W., and Zamora, M. (2003). Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements. Geophysical Research.

Reed, L., Taner, M., Elliott, B., and Killeen, P. (1997). A comparison of physical property borehole logs with geology, mineralogy and chemistry in a borehole at Les Mines Selbaie, northwestern Québec, Canada. In A. Gubins (Ed.), Exploration 97: Fourth Decennial International Conference on Mineral Exploration, (pp. 1043–1048).

Ribeiro, V. B., Louro, V. H., and Mantovani, M. S. (2013). 3D Inversion of magnetic data of grouped anomalies—Study applied to São José intrusions in Mato Grosso, Brazil. Journal of Applied Geophysics, 93, 67-76.

Sakurai, S., and Melvin, J. (1988). Facies discrimination and permeability estimation from well logs for the Endicott field: 29th Annual Symposium, Society of Professional Well Log Analysts, Proceedings, F1-F18.

Spicer, B., Morris, B., and Ugalde, H. (2011). Structure of the Rambler Rhyolite, Baie Verte Peninsula, Newfoundland: Inversions using UBC-GIF Grav3D and Mag3D. Journal of Applied Geophysics, 75, 9–18.

Theodoridis, S., and Koutroumbas, K. (2003). Pattern recognition, (second edition). New York: Academic Press.

Vella, L., and Emerson, D. (2009). Carrapateena: Physical properties of a new Iron-oxide copper-gold deposit. Journal Article ASEG Extended Abstracts, 1 – 13.

Wang, G., Zhu, Y., Zhang, S., Yan, C., Song, Y., Ma, Z., et al. (2012). 3D geological modeling based on gravitational and magnetic data inversion in the Luanchuan ore region, Henan Province, China. Journal of Applied Geophysics, 80, 1-11.

Williams, N. (2008). Geologically constrained UBC-GIF gravity and magnetic inversions with examples from the Agnew-Wiluna greenstone belt, Western Australia. Vancouver, BC,: University of British Columbia.

Williams, N., and Dipple, G. (2007). Mapping subsurface alteration using gravity and magnetic inversion models. In B. Milkereit (Ed.), Exploration 07: Fifth Decennial International Conference on Mineral Exploration, (pp. 461-472).

Wong, P., Jian, F., and Taggart, L. (1995). A critical comparison of neural networks and discriminant analysis in lithofacies, porosity and permeability predictions. Journal of Petroleum Geology, 18(2), 191-206.

Zhou, J., Zhang, X., and Xiu, C. (2015). Lithological characterization and its application based on three-dimensional structure-coupled joint inversion of gravity and magnetic data. International Journal of Geosciences, 6, 230-237.

Chapter 2

2 Clustering of down-hole physical properties measurements at the Victoria property, Sudbury for the purpose of extracting lithological information

2.1 Abstract

Down-hole density, gamma radioactivity, and magnetic susceptibility measurements in five drillholes at the Victoria property (located in the south range of the Sudbury basin) were analyzed to identify homogenous physical units. The fuzzy k-means clustering algorithm was used for unsupervised classification of the data. Four main physical units were identified in boreholes with distinct physical characteristics. Three of them were differentiated mainly based on different gamma ray and density values, and the fourth one was characterized by high magnetic susceptibility. Physical units were compared with rock types logged by geologists to determine which rock types corresponded to physical units. We found that there was a meaningful spatial and statistical correlation between physical units (characterized based on their physical properties measurements) and lithological units as indicated by rock types at the Victoria property. However, not all rock types could be uniquely identified by the statistical classification, but a set of similar groups could be identified. Hence, identifying a group of rock types described by each physical unit can be used to translate physical data to/from lithological data. Alternatively, the physical log units could be used as a quality control procedure to check the geological logs, or to highlight areas where more careful logging or other investigation would be warranted.

2.2 Introduction

Multi-probe in-hole measurements can be used to acquire a large amount of in-situ data, rapidly. Analyzing downhole physical properties measurements can help to understand the links between this in-situ physical properties data and the geology. This approach is more likely to succeed when physical rock property contrasts exist between the different lithological units (McDowell et al., 1998; Granek, 2011; Perron et al., 2011; Mwenifumbo and Mwenifumbo, 2012). But rock types may have complete or partial overlap of physical properties. In this case, it is not always possible to differentiate between rock types on the basis of physical properties (Rabaute, et al., 2003 and Garcia et al., 2011). Physical properties are classified to produce homogenous physical units with consistent physical characteristics. The obtained physical units are not necessarily consistent with the lithological units determined by geologists, so there will not always be a direct relationship between physical properties and rock type (Benaouda, et al., 1999).

Killeen (1997) provided a general introduction to data acquisition methods and the applications of downhole physical properties measurements in mineral exploration. In the same volume, Killeen et al. (1997) introduced a method to analyze physical properties measurements to show that geophysical logging can successfully aid geological logging and provide a means of quantitatively determining layer boundaries. It was suggested by Reed et al. (1977) that the physical properties measurements can provide complementary information to that provided by visual and chemical logging. McDowell et al. (1998) employed prompt gamma neutron activation logging to provide a crude estimate of nickel assay. In a later study, McDowell et al. (2004) also probed the relationship between nickel grade and inductive conductivity logs to differentiate nickel-rich from nickel-barren sulfides zones. Physical properties measurements have also been used to differentiate a cover sequence from the basement rocks that hosts copper-

gold mineralization (Vella and Emerson, 2009). Analyses of physical properties measurements have been remarkably beneficial to ongoing Integrated Ocean Drilling Program (IODP) (Inwood et al., 2013 and Ojha and Maiti, 2013) to identify subsurface geological settings, particularly where core is not fully recovered. Inwood et al. (2013) utilized statistical techniques to classify downhole physical properties measurements obtained during IODP to indicate facies variations. Interested readers are recommended to look up for published material by the Borehole Research Group of Leicester University and Borehole Research Group of the Lamont-Doherty Earth Observatory.

To reduce the subjective bias in interpretation and to quantitatively link the geological and geophysical data, pattern recognition techniques are frequently used (Granek, 2011 and Williams and Dipple, 2007). The fuzzy k-means algorithm has been successfully used to identify stratigraphy and volcanic history through clustering of physical properties measurements (Rabaute, et al., 2003). Garcia, et al. (2011) applied the fuzzy k-means clustering and principal component analysis to down-hole geophysical logs and geomechanical measurements of core to identify relatively homogenous zones that could potentially be used to store radioactive wastes. They also identified inhomogeneities of physical units by lateral and vertical changes in the physical properties. Ojha and Maiti (2013) implemented the fuzzy k-means clustering on downhole logging data measured during IODP to identify physical units characterizing sedimentary rocks where core recovery is poor or zero.

Fuzzy k-means clustering of borehole logging data has primarily been applied to sedimentary (softrock) environments. This study focuses on identification of homogenous physical units down five boreholes from the Victoria property on the south range of the Sudbury Basin, an example of a hardrock environment. Downhole density, gamma ray response, and magnetic

susceptibility measurements were used since they are physical properties found to be reliable and responsive to changes in lithology. The data are analyzed by fuzzy k-means clustering to identify physical log units within each hole. Careful analyses on the observed physical units and their corresponding rock types are carried out to define the links between the physical and the lithological units. The degree of homogeneity of rock types in terms of physical properties can also be assessed based on statistical analyses and clustering results. A meaningful relationship between physical units and lithological units can introduce clustering as a reliable approach to characterize rock types at the Victoria property based on physical properties measurements.

2.3 Data set used

- Down-hole physical properties measurements

The gamma ray response, magnetic susceptibility, and density measurements with a vertical resolution of 20 cm, collected from five boreholes at the Victoria, were used in this work. Figure 2-1 shows the location of the collar of four boreholes on the geological map of the Victoria property. The fifth borehole, FNX1182, is located to the north beyond the area covered by this map.

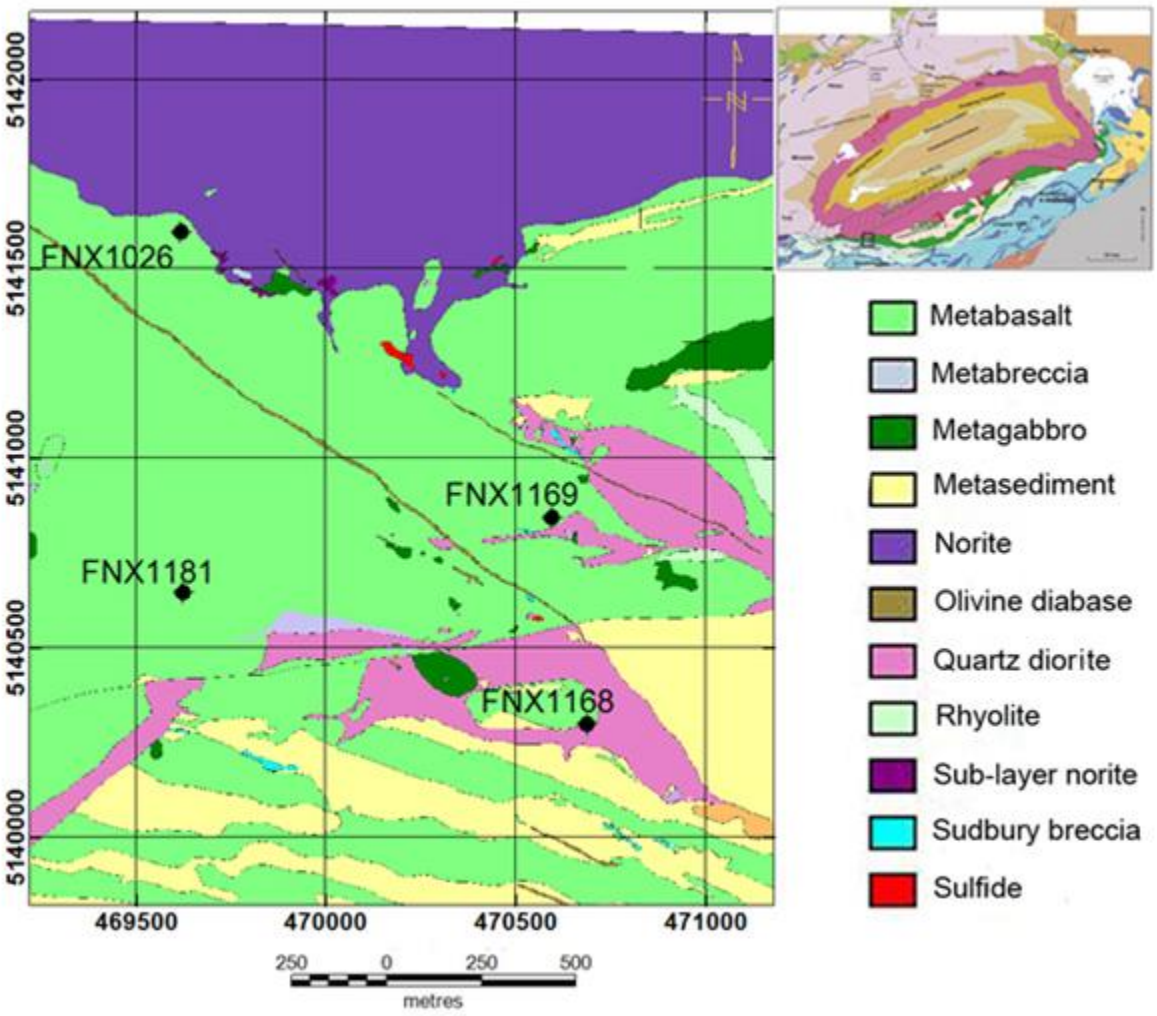


Figure 2-1: The location of four boreholes on the geological map of the Victoria property. The study area is marked by the small black rectangle in the top right figure.

All boreholes were theoretically NQ sized, but the caliper log was measured to implement on-site calibration to correct for borehole size. The fluid type filling the holes was water. A winch and steel cable system was used to lower a multi-parameter digital logging probe down the holes and measure the data. Prior to classification, data preprocessing was implemented to produce a suitable dataset. Due to different sampling intervals for each tool, the weighted average of

adjacent measurements were computed every 20 cm to resample all the logs to a uniform sample interval. The susceptibility and gamma data did not show a normal distribution so they were normalized through a logarithmic transformation. Data were subject to a low-pass filter to remove high frequency variations and increase the signal-to-noise ratio. Finally, data were normalized between 0 and 1 for clustering purposes; otherwise the results of clustering are mainly controlled by those data with a higher variance.

2.4 Methodology

- Fuzzy k-means clustering

Clustering is an unsupervised classification technique in which data are divided into clusters based on the variables measured at each data point (Lofts, 1993 and Rabaute et al., 1997). The fuzzy clustering algorithm allows data points to belong to more than one cluster and the degree of membership of each datum to each cluster is defined by a membership value for that cluster: the higher the membership value, the more strongly the datum belongs to the cluster (Zadeh, 1965 and Ruspini, 1969).

In fuzzy *k*-means clustering *n* data are divided into *p* clusters based on *q* variables. The fuzzy membership value of the i^{th} datum, m_{ik}, $i=1,\ldots,n;\ k=1,\ldots,p$) is obtained through minimization of the objective function *J*:

$$J = \sum_{i=1}^{n}\sum_{k=1}^{p} m_{ik}^{\emptyset}\, d_{ik}^2 \quad , \qquad\qquad d_{ik}^2 = \sum_{v=1}^{q} (x_{iv} - c_{kv})^2 \quad ,$$

where data are represented as vectors x_i which is the vector of i^{th} data. Each vector is composed of q elements and x_{iv} is the value of the v^{th} variable measured at the i^{th} datum. The fuzzy exponent is represented by ϕ. The term c represents the position of the centroid of the cluster. It is a vector with q elements so that c_{kv} is the average of the v^{th} variable of all the points weighted by their degree of membership to cluster k. The term d_{ik} represents the distance between the i^{th} data and the centroid of the k^{th} cluster (Bezdek, 1981 and McBratney and DeGruijter, 1992). Initially fuzzy membership values of data are randomly assigned between 0 (lowest membership) to 1 (highest membership) and the total of membership values for each data must add up to 1. The fuzzy membership values are recalculated to minimize the objective function. As a result, the centroid of clusters is relocated when the fuzzy membership values change. Therefore, a new vector of centroids is used in the objective function at each iteration. This procedure continues until the objective function is minimized. The outputs of the algorithm for each datum are the membership values, the confusion index (defined below), and the position of centroids of clusters. Interpreting the output requires careful analyses of these values to determine to which cluster each datum belongs (Bezdek, 1981; McBratney and DeGruijter, 1992; Rabaute et al., 2003).

The confusion index (CI) output by the fuzzy k-means clustering indicates how well a sample is classified. CI ranges between 0 and 1, and is calculated by:

$$CI\ (i) = 1 - (mf - mf_2),$$

where mf and mf_2 are highest and second highest memberships for each sample. If one membership for a cluster stands out, a sample more likely belongs to it, so mf is close to one and mf_2 is close to zero and the confusion index approaches zero. While, small differences between

the top two memberships for clusters creates confusion in the classification of the sample, and the confusion index approaches one (Burrough and McDonnell 1998).

Prior to running the algorithm, two parameters need to be assigned: the optimal number of clusters and the fuzzy exponent. The optimal number of clusters is determined based on prior knowledge and minimization of three functions: the fuzzy performance index (FPI),

$$FPI = \frac{1-(p \times F-1)}{F-1}, \qquad F = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p}(m_{ik}^{\emptyset})^2 ,$$

the modified partition entropy (MPE),

$$MPE = \frac{H}{logp} , \qquad H = -\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p} m_{ik}\log(m_{ik}) ,$$

where H represents the entropy function, and finally, the separation distance (S),

$$S = \frac{J}{n\,(d_{min})^2} \quad ,$$

where d_{min} is the minimum distance between cluster centroids. Minimization of these parameters results in a maximization of the compactness and the separation of clusters (Roubens, 1982, and Xie and Beni, 1991). The procedure we adopted is to vary the number of clusters from 2 to 10 and then select the number that has a minimum. However, not all parameters always have a minimum at the same value and sometimes the best value is a local minimum, not a global minimum (Rabaute, et al., 2003). Hence, there is some subjectivity and prior knowledge and experience when selecting the optimum number of clusters.

The fuzzy exponent determines the degree of fuzziness which represents the compactness and separation of final clusters. This value can be assigned between 1 and infinity with larger values resulting in higher fuzziness in the final clustering (Bezdek, 1981; DeGruijter and McBratney, 1988; Rabaute et al., 2003). Due to the nature of physical properties data and their contrast in different rock types, Bezdek (1981) and DeGruijter and McBratney (1988) chose a value of 2 as a proper value for the fuzzy exponent, and we have followed their suggestion. The program FUZME2.1 (Minasny and McBratney, 2000) was used to classify the data.

2.5 Results

2.5.1 Borehole log data

The clustering algorithm was applied separately on the data from each hole, and the results are summarized in the following sections. Due to lacking prior knowledge of the physical units that could be resolved at Victoria, the optimal number of clusters for each borehole was determined objectively based on the three mathematical functions described above. Figure 2-2 represents the values of MPE, FPI, and S for clustering of data collected from borehole FNX1182. It suggests the samples are well separated using three clusters. The same process was implemented on the data from the other boreholes. As a result, the clustering algorithm used three clusters for FNX1168 and FNX1169, and two clusters for FNX1181 and FNX1026.

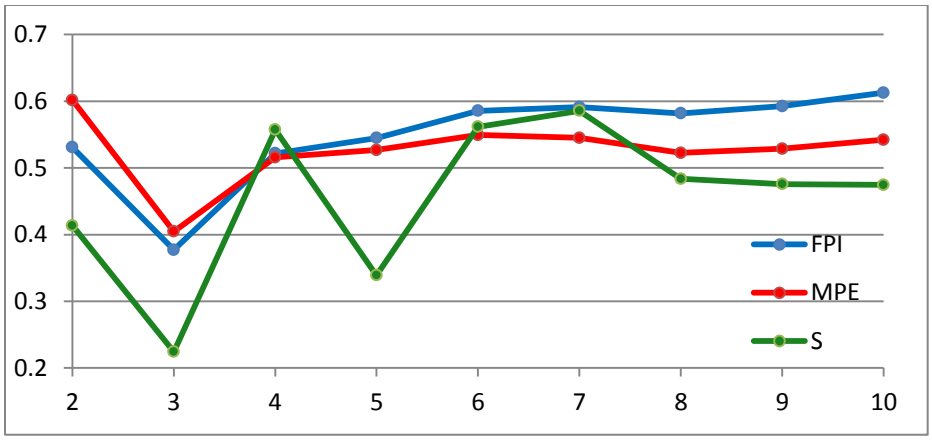


Figure 2-2- FPI, MPE, and S plots were used to determine the optimal number of clusters

The result of clustering for FNX1182 is shown in Figure 2-3. The maximum membership (left column black), defined physical units (left column colour), membership values, confusion index, physical logs, lithological unit (rock type), are presented on this plot. The following sections explain how the physical units were defined and how they relate to the rock types logged by geologists.

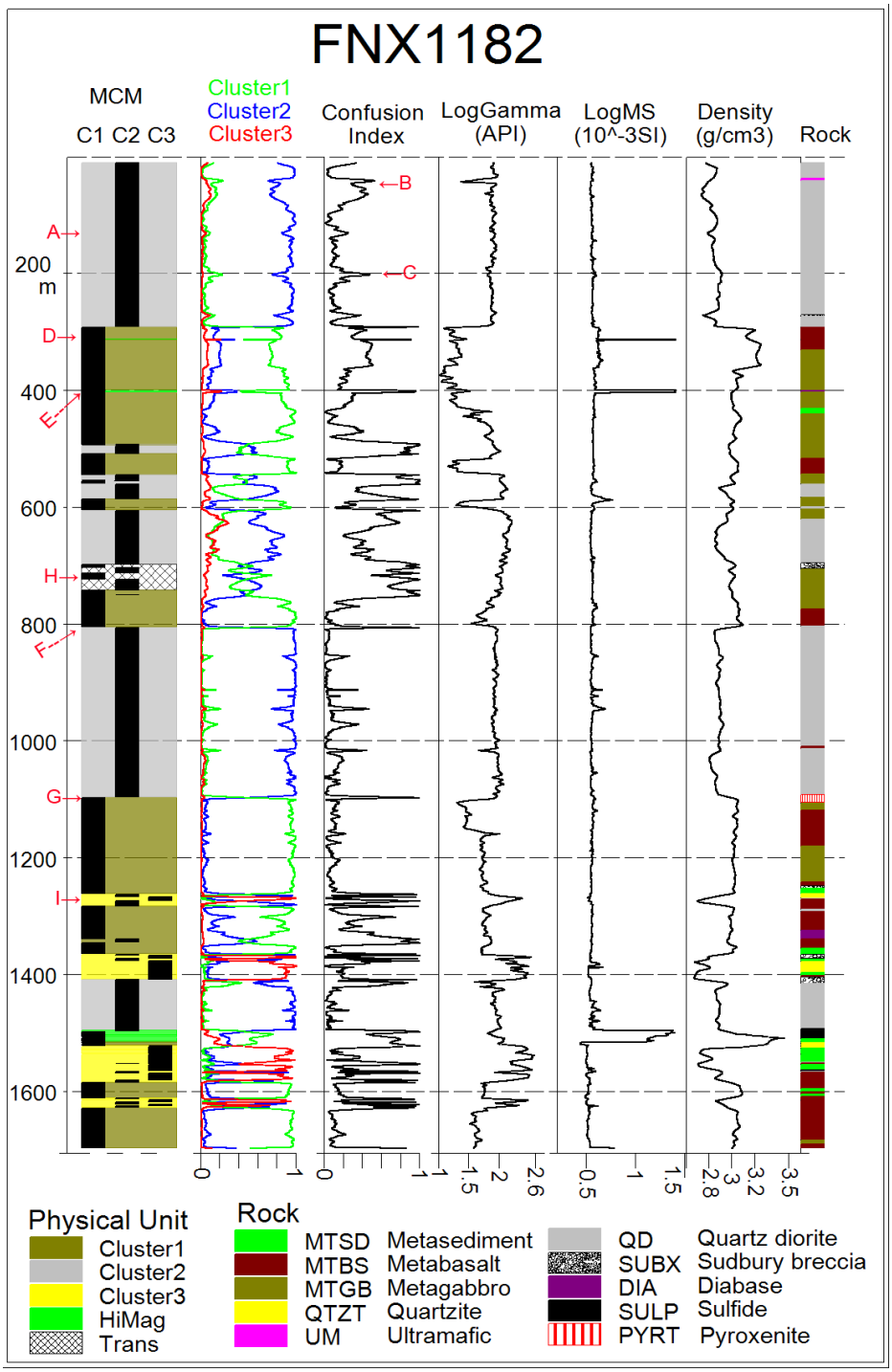


Figure 2-3- Results of clustering for FNX1182. From left to right column: modified physical units in colour and maximum membership in black; membership for clusters; confusion index; the three physical properties measurements; and rock type. The letters A to I indicate: (A) homogenous physical unit with one membership value standing out and low

confusion index; (B and C) sporadic heterogeniety within an otherwise homogeneous unit indicated by minor fluctions in the confusion index and membership values; (D and E) high magnetic unit marked by two or three similar membership values, high confusion index, and extremely high magnetic susceptibility; (F and G) sharp boundaries indicated by confusion index peaks; (H) transition zone between two physical units characterized by gradual increase and/or decrease in two or more memberships, high confusion index, and gradual change in physical properties; (I) cluster3 in the center is bracketed by two thin units of cluster2 above and below in MCM column. This zone was modified manually based on the confusion index peaks and variation in physical logs in physical unit column.

The results for FNX1169 are also shown in full in Figure 2-4 to illustrate the process of using fuzziness of the output to define physical units of the borehole. To make the manuscript short, only the defined physical units, maximum membership and rock types are represented for the other three boreholes.

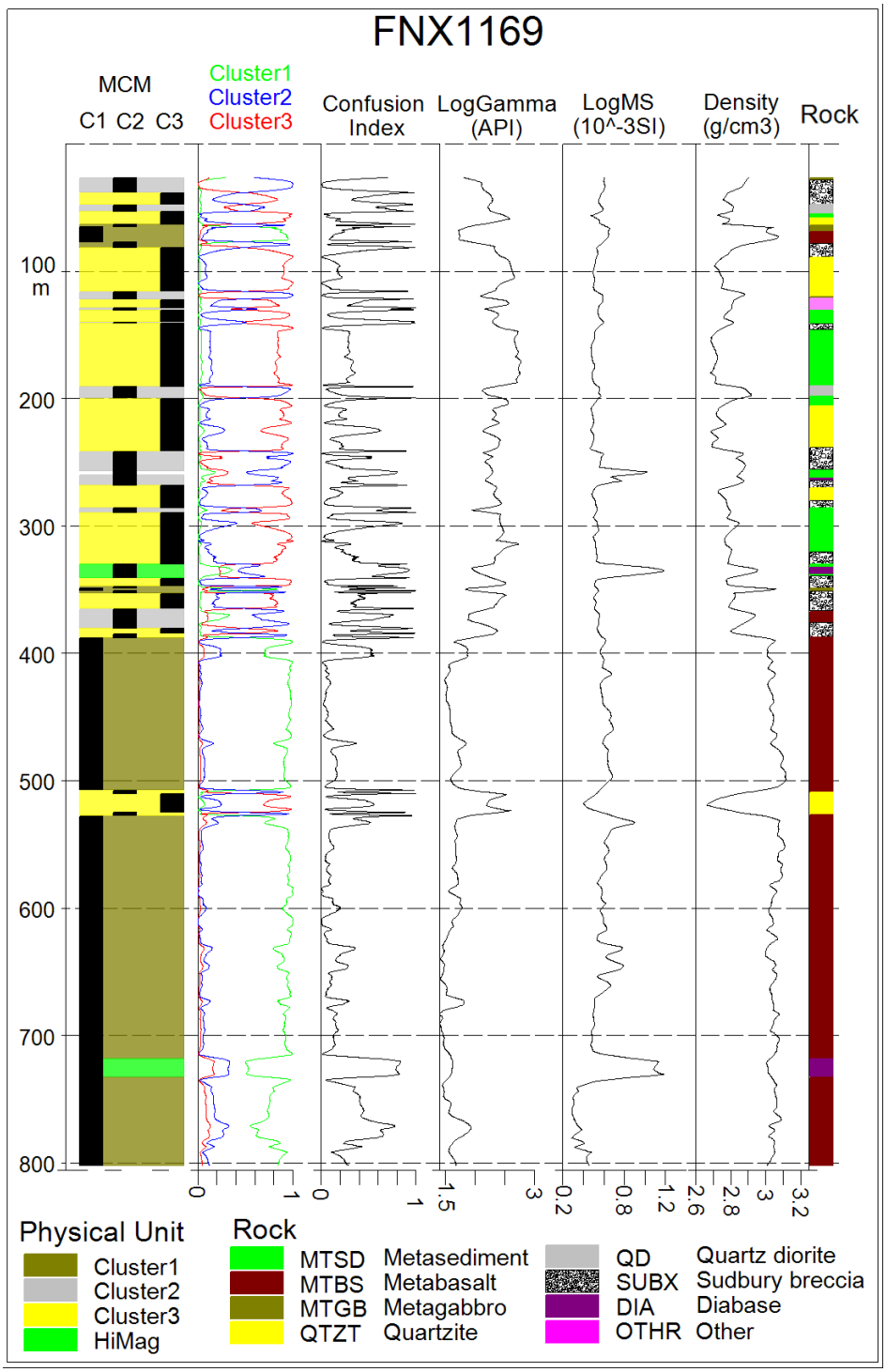


Figure 2-4- Results of clustering for FNX1169. From left to right column: modified physical units in colour and maximum membership in black; membership for clusters; confusion index; the three physical properties measurements; and rock type.

Taking into account variation in membership values can disclose important information about the geological setting along the hole. When one membership clearly stands out, it indicates homogenous formations. The top 290 m of FNX1182, which is marked by letter 'A' in Figure 2-3, forms a homogenous physical unit along which the membership of cluster2 (blue line) shows a uniform value close to unity. A relatively low confusion index for this zone brings certainty to the clustering result. In this case the classification is clearly indicated by the maximum membership value. There are, however, minor fluctuations in the confusion index that can be seen at 50 m (point B) or 200 m (point C). Because there is a lack of meaningful changes in the physical logs associated with these variations in the confusion index and membership values, these variations can be interpreted as sporadic heterogeneity within a unit. If variations in these units are of potential interest, then these locations could be further investigated. In Figure 2-4, a similar low confusion index and predominant membership for cluster1 in FNX1169 can be seen between 400-500 m and 530-700 m. Minor variation in gamma ray response around 400 and 670 m causes an increase in the confusion index, but other physical properties variations are not strong enough to change the physical unit.

A high confusion index at 400 m of FNX1182, marked by "E" in Figure 2-3, is associated with anomalous high magnetic susceptibility (MS). Observing two similar membership values (green and blue lines) at 300 m (point D) or three similar memberships (green, blue and red lines) at 400 m (point E) mean that these highly magnetic samples do not match any of the specified clusters. It is likely that the clustering algorithm did not form a separate cluster for this group of data due to insufficient number of samples, even though they have characteristic physical properties (high MS). To distinguish these units in the borehole, high magnetic susceptibility measurements can be used to manually define a "HiMag" unit which is shown in physical unit column with a green

colour. Other location where there are similar characteristics (extremely high magnetic susceptibility and confusion index, and similar low membership for three clusters) are at 1450 m on Figure 2-3 and 720 m on Figure 2-4.

A high confusion index might also be an indication of transition zones. In Figure 2-3, the confusion index peaks at 800 m (point F) and 1100 m (point G) clearly indicating the sharp boundaries between cluster1 and cluster2. But the boundary between two units is not always sharp and sometimes there is a gradual transition zone from one unit to another. This manually identified transition zone is characterized by gradual increase and/or decrease in two or more memberships, high confusion index, and gradual change in physical properties. As marked by "H" in Figure 2-3, transition zone between cluster2 and cluster1 comes to appearance at 690 m and continues down to 740 m. This transition is due to a reduction in the gamma and an increase in the density.

Gradual changes of the data identify intermediate zones when there is not a sharp transition between two units. At 1270 m of FNX1182 (point I), in MCM column cluster3 in the center is bracketed by two thin units of cluster2 above and below. However, the physical logs clearly indicate a gradual transition from cluster1 to cluster3 without a distinct cluster2 in the middle. So, the boundaries of physical units were modified manually based on the confusion index peaks and variation in physical logs. This phenomenon is better illustrated in MCM column of FNX1169 (Figure 2-4) at 520 m where two thin layers of cluster2 bound cluster3. In modified physical units, boundaries of cluster3 were shifted up and down corresponding to confusion index peaks which resulted in eliminating the thin layers of cluster2.

The same manual modifications were carried out on the results for FNX 1168, FNX1181, and FNX1026. To make the manuscript concise, only maximum membership, modified physical units, and rock type are shown for these boreholes in Figure 2-5. The centroids of the physical units (clusters 1 to 3 and HiMag) for all five boreholes are represented in physical properties space in Table 2-1.

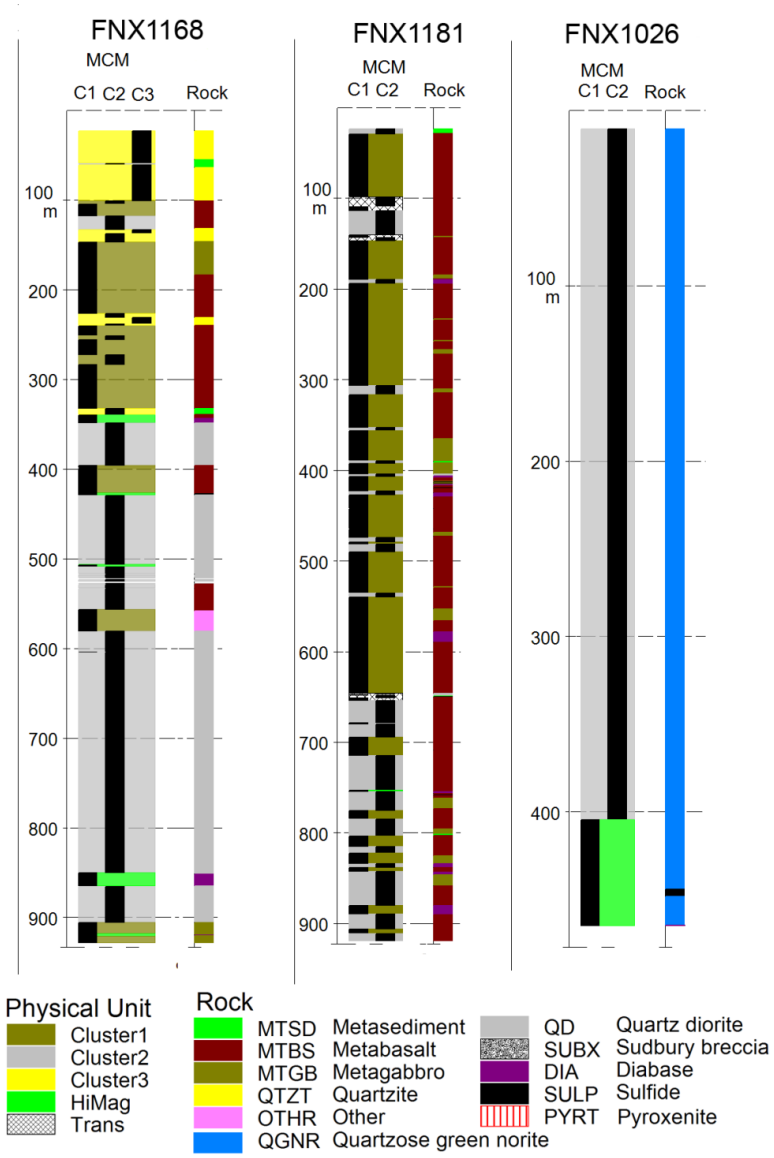


Figure 2-5- Results of clustering for FNX1168, FNX1181, and FNX1026.Modified physical units in colour and maximum membership in black; and rock type.

Table 2-1- Mean value and standard deviation of physical properties of physical units observed in each borehole.

		FNX1026	FNX1168	FNX1169	FNX1181	FNX1182
LogGamma (API)	C1	-	*1.57 ± 0.23*	*1.63 ± 0.15*	*1.63 ± 0.30*	*1.60 ± 0.24*
	C2	1.80 ± 0.07	1.87 ± 0.07	2.15 ± 0.12	*1.82 ± 0.16*	1.95 ± 0.10
	C3	-	2.34 ± 0.20	2.44 ± 0.16	-	2.28 ± 0.19
	HiMag	1.33 ± 0.18	1.58 ± 0.16	1.86 ± 0.28	1.69 ± 0.05	1.82 ± 0.25
LogMS (10^{-3}SI)						
	C1	-	0.68 ± 0.04	0.57 ± 0.12	0.56 ± 0.17	0.56 ± 0.04
	C2	0.59 ± 0.02	0.65 ± 0.04	0.59 ± 0.06	0.41 ± 0.25	0.55 ± 0.02
	C3	-	0.59 ± 0.09	0.51 ± 0.03	-	0.52 ± 0.03
	HiMag	1.05 ± 0.26	1.09 ± 0.16	1.02 ± 0.13	0.53 ± 0.10	1.28 ± 0.11
Density (g/cm^3)						
	C1	-	*3.08 ± 0.05*	*3.04 ± 0.04*	*3.06 ± 0.04*	*3.04 ± 0.06*
	C2	2.89 ± 0.03	2.95 ± 0.03	2.86 ± 0.04	*3.08 ± 0.05*	2.87 ± 0.05
	C3	-	2.83 ± 0.08	2.76 ± 0.05	-	2.81 ± 0.09
	HiMag	3.14 ± 0.09	3.04 ± 0.04	2.95 ± 0.08	3.07 ± 0.01	3.16 ± 0.17

The centroids of one specific physical unit identified in different boreholes show physical characteristics that are approximately consistent. Cluster1 is characterized by low gamma ray response and high density. Cluster2 has medium gamma ray response and medium density. Cluster3 shows high gamma ray response and low density. Magnetic susceptibility in clusters 1 to 3 varies in an irregular pattern which does not differentiate them from each other. However, extremely high values of magnetic susceptibility can be used to characterize the HiMag unit. The only exception is cluster2 of FNX1181, which has a slightly higher gamma ray value, but a similar density value to cluster1 (compare the italicized values in Table 2-1). As the values are more similar to cluster1 in other boreholes than cluster2, this cluster was considered to be cluster1 in all subsequent analyses. Unlike the HiMag unit, the transition zone is not listed in the Table 1 as a specific physical unit as it is the intermediate zone between other physical units.

Identical physical units in different boreholes were combined to form four physical units to describe general characteristics of physical properties of boreholes. They also were used in the next step to find their relationship with lithological units. The statistical parameters of combined physical units are summarized in Table 2-2.

Table 2-2- Statistical summaries of physical properties of combined physical units.

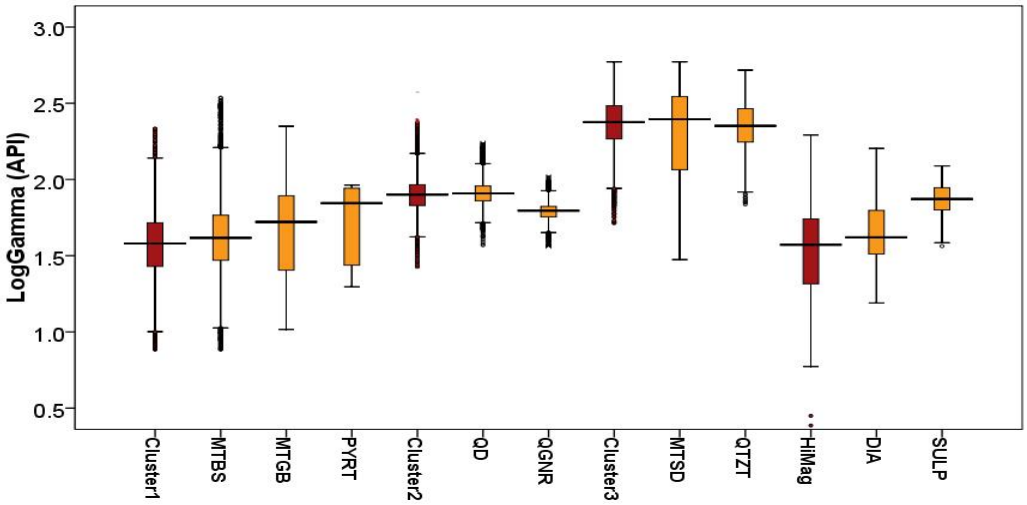
		Cluster1	Cluster2	Cluster3	HiMag
	Count	11065	8915	2628	755
LogGamma (API)					
	Mean	1.61	1.90	2.37	1.57
	Standard deviation	0.25	0.12	0.19	0.31
	Minimum	0.89	1.36	1.71	0.39
	Maximum	2.57	2.43	2.77	2.29
Log MS (10^{-3}SI)					
	Mean	0.58	0.59	0.53	1.08
	Standard deviation	0.13	0.50	0.57	0.22
	Minimum	0.01	0.41	0.39	0.48
	Maximum	1.55	1.19	0.88	1.91
Density (g/cm^3)					
	Mean	3.06	2.90	2.79	3.08
	Standard deviation	0.06	0.05	0.08	0.12
	Minimum	2.70	2.73	2.66	2.68
	Maximum	3.32	3.09	3.20	3.46

2.5.2 Correlation of physical units with rock types

The second step was to compare the physical units with lithological units (rock types) logged by geologists to find a meaningful relationship. The rock type column for each borehole is plotted on the right panel of Figures 2-3 to 2-5. Among rock types, some of them are more common in boreholes including: Metabasalt (total of 1712 m), quartz diorite (1230 m), metagabbro (570 m), quartzite (225 m), metasediment (225 m). Quartzose green norite is observed only in FNX1026 with total length of 450 m. Other lithological units occurred in short sequences are Sudbury breccia (155 m), diabase (107 m), sulfide (21 m), pyroxenite (14 m), ultramafic (3 m). Dominant

rock types have enough samples so that the statistical analysis can sufficiently characterize their physical properties; on the other hand, the limited number of samples for the minor rock types does not provide a clear characterization of the physical properties of the rocks. While interpreting the results, overlap of the physical properties of rocks should also be considered. Those rocks with similar physical properties are classified together and represented by one physical unit.

Identical rock types from different boreholes were combined, and then compared to the physical units. Figure 2-6 represents box-plots of gamma ray, density, and magnetic susceptibility measurements for the physical units (clusters) and the lithological units. The physical units and rock types are deliberately ordered in the plots based on their similarity in physical properties. The boxplots of physical units have red bars which are followed by the boxplots of rock types with orange bars. Sudbury breccia is not plotted in Figure 2-6 because it does not show consistent, distinguishable physical properties its spatial correlation with physical units is evaluated later. Ultramafic was also not plotted due to insufficient samples necessary to reliably characterize the physical properties.



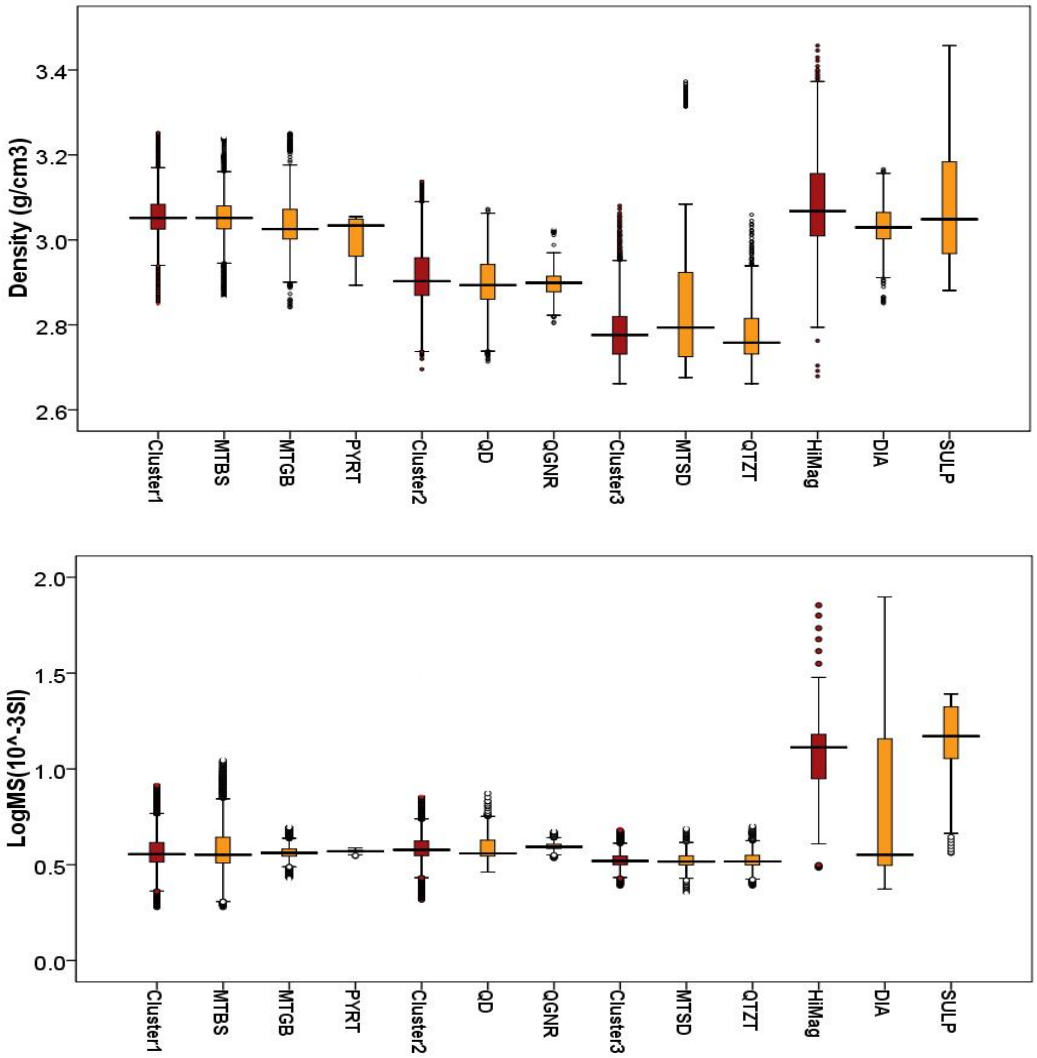


Figure 2-6- Boxplots of gamma ray (top), density (middle), and magnetic susceptibility (bottom) measurements within different physical units and rock types collected from five holes. The boxes represent the lower quartile and upper quartile. The median is represented by the horizontal line. The whiskers extending above and below the box represent the values between the lower quartile and 1.5 times the difference between the upper and lower quartiles. The outliers are shown as dots above and below. Boxplots of a physical unit (red bars) are followed by box plots of rock types (orange bars) with similar physical properties.

The boxplots of gamma ray and density measurements illustrates cluster1 has physical properties similar to metabasalt (MTBS), metagabbro (MTGB) and to a lesser extent pyroxenite (PYRT). The boxplots of gamma ray and density also indicate cluster2 closely matches the physical properties of quartz diorite (QD) and to a lesser extent norite (QGNR). The boxplots of cluster3 lies in the same range of gamma ray and density values as metasediment (MTSD) and quartzite (QTZT). Boxplot of magnetic susceptibility measurements clearly shows that the range of magnetic susceptibility values of diabase (DIA) and sulfide (SULP) fall within the range of values characterized by the HiMag zone. The rest of boxplots of magnetic susceptibility are indistinctive, which means other than diabase and sulfide, other rock types and physical units are not differentiable based on magnetic susceptibility.

Inhomogeneities of physical units and rock type could also be assessed using boxplots. As shown in Figure 2-6, broad range of gamma and density variation in metagabbro and metasediment indicates their heterogeneity of physical properties; compared to quartz diorite which is defined as a homogenous rock type with narrower range of variation.

The statistical correlation between physical unit and specified rock types is clear from the boxplots; the spatial correlation can be seen by detailed visual comparison of the clusters and rock types on Figures 2-3 to 2-5. To facilitate the comparison, Table 2-3 summarizes proportion of rocks (total length in meters and percentage) corresponding to each physical units. The information of this table agrees with the results of the statistical analyses regarding the rock types presented by each physical unit.

Table 2-3- Distribution of rock types into physical units presented in total length and percentage. Rock labels: MTBS: metabasalt; MTGB: metagabbro; PYRT: pyroxenite; QD: quartz diorite; QGNR: quartz green norite; MTSD: metasediment; QTZT: quartzite; DIA: diabase; SULP: sulfide; SUBX: Sudbury breccia.

	Cluster1	Cluster2	Cluster3	HiMag
MTBS	1567m (92%)	55 (3%)	58 (3%)	-
MTGB	472 (83%)	55 (10%)	-	3 (1%)
PYRT	9 (67%)	5 (33%)	-	-
QD	13 (1%)	1199 (97%)	10 (1%)	-
QGNR	-	394 (87%)	-	57 (13%)
MTSD	50 (22%)	11 (5%)	152 (67%)	12 (5%)
QTZT	5 (2%)	5 (2%)	214 (95%)	-
DIA	61 (56%)	2 (2%)	-	45 (42%)
SULP	-	3 (12%)	-	18 (88%)
SUBX	14 (9%)	47 (30%)	84 (54%)	2 (1%)

Significant portion of metabasalt, metagabbro, and pyroxenite are represented by cluster1. As expected, quartz diorite and norite have remarkable spatial correlation with cluster2. A minor sequence (13%) of norite (QGNR) represented by HiMag indicates the possible misclassification of the lower part of FNX1026 as norite where this part of the borehole shows anomalously high magnetic susceptibility. Quartzite is predominantly represented by cluster3. Although 67% of metasediment is represented by cluster3, partial clustering in other physical units is an indicator of the heterogeneity of this rock type. The physical behavior of diabase shows two magnetic and non-magnetic populations. The magnetic part (42%) is presented by HiMag, and the non-magnetic part (56%) has similar physical properties to cluster1.

A histogram of the magnetic susceptibility measurements of diabase in Figure 2-7 depicts a bimodal distribution, separated into magnetic and non-magnetic parts, represented by HiMag unit and cluster1. The most heterogeneous rock is Sudbury breccia which is distributed in three predominant clusters as shown in Table 2-3. Figure 2-8 shows a histogram of the density

measurements of Sudbury breccia classified by different physical units, with a trimodal distribution being evident.

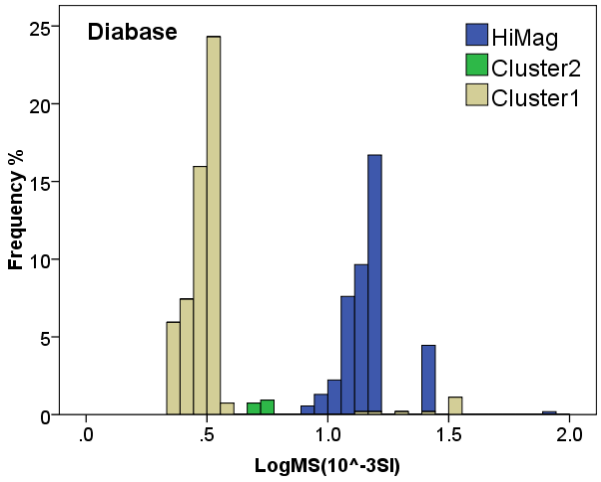


Figure 2-7- Histogram of magnetic susceptibility of diabase measured in five boreholes. Coloured bars represent percentage of diabase represented by each physical unit.

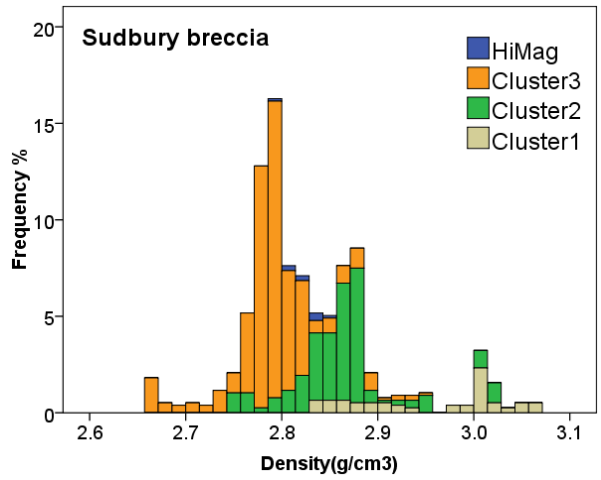


Figure 2-8- Histogram of density of Sudbury breccia measured in five boreholes. Coloured bars represent percentage of diabase represented by each physical unit

Although Table 2-3 shows remarkable spatial correlation of physical units and lithological units, significant differences may draw attention to places where a geologist might need to re-log the cores. The bottom 70 m of FNX1026 in Figure 2-5 is a good example to illustrate discrepancies: the physical properties analyses indicate a HiMag zone (diabase or sulfides), whereas the geological classification is quartzose green norite, with only a very small zone of sulfide. This suggests that some re-logging is required to modify the rock type classification. Another use of physical unit is to complete core logging when the core is missed or it is not possible for geologists to identify rock types. At 570 m of FNX1168 the rock type is not specified, but physical properties analyses suggests a rock group similar to metagabbro and metabassalt for this part of the borehole.

2.6 Workflow

The above results suggest that the following workflow for classification of borehole physical property data would be appropriate.

1) Data preprocessing (editing, filtering, taking logarithm, scaling) to produce a suitable dataset for clustering.

2) Undertaken analysis to determine the optimal number of clusters.

3) Undertake fuzzy k-means clustering with this number of clusters.

4) Undertake an initial classification of data in the hole into physical units on the basis of maximum membership.

5) Refine the classification manually in locations where the confusion index is large, placing the data into one of the following categories:

    a. If there are not significant variations in physical properties, and the physical unit is similar above and below, consider the zone of increased confusion to be a local heterogeneity.

    b. If the physical properties are unique, but in such small numbers to not warrant the creation a statistically identifiable class, create a new class and place the data in this class.

    c. If the physical properties are gradually changing spatially, classify this as a transition zone where the rock type is gradually changing from one physical unit to another. This zone is characterized by gradual increase and/or decrease in two or more memberships.

    d. If the zone is narrow and shows gradual changes in physical properties, but lies between two less confused units, then the zone might be better recognized as an intermediate zone and reclassified as one of the less confused units above or below the intermediate zone.

6) Look at the statistical analyses of the centroid (mean and standard deviation) to see if these clusters are similar to those in other holes (if already analyzed). If so bring these clusters into one group.

7) If the geological logging has already been undertaken and is available, then the physical log units and the lithological classifications could be compared either statistically, or spatially to highlight anomalous areas or areas where re-logging might be required.

8) The locations where there is higher confusion could be logged by the geologist more carefully.

2.7 Potential applications of clustering analysis

Understanding of such a relationship between physical units and rock types can help rock type characterization in new boreholes. In a borehole lacking rock-type information, clustering of physical properties measurements will provide clusters which can be related to lithological units that are known to exist within each physical unit. This information can be used in further geophysical studies like rock-type predication based on physical properties measurements or classifying potential field data inverted to 3D models to produce a lithological model from the geophysical data. On the other hand, the rock types logged by geologists can be translated to a range of physical properties and then used in constrained geophysical inversion. It also helps further geological/geophysical/geochemical activities efficiently focus on specific physical/lithological zones of interest. Finally, in case when rock types are already identified, clustering of physical properties measurements could be used to refine or check the core logging.

2.8 Conclusion

Studies involving the clustering of physical properties are successful when there is a physical properties contrast between geological zones. At the Victoria property this study was successful both in determining homogenous and heterogeneous physical log units and understanding the physical characteristics of rock types. Analyzing high resolution physical properties measurements may enhance our understanding of the geological setting of boreholes. In particular, anomalous variation in physical properties in locations where the rock type remains unchanged may draw attention to the core to recheck the log and possibly modify the rock type classification. For example, physical properties analyses identified a broad and potentially interesting HiMag zone at the bottom of FNX1026 which was missed of logged as significantly narrower in the geologists' core logging. The data could also help to predict rock type in zone where the rock is not specified (similar to the zone at 570 m in FNX1168).

Fuzzy k-means clustering was used successfully to detect the contrasts in physical properties and identify physical units. The technique is a capable tool to provide multiple outputs. Careful analysis of maximum membership, membership variation for clusters, and confusion index enhances the accuracy and reliability of the results obtained from physical properties measurements. We have outlined a workflow that allows us to use high values of the confusion index to interpret the data, specifically, to identify i) transition zones, ii) high magnetic zones and iii) intermediate zones (where we suggest modification of cluster boundaries). High confusion index is a powerful tool for identifying heterogeneity of rock types, which might not be perceived by the geologists during core logging. Heterogeneity can also be recognized as locations where there is a rapid spatial variation in memberships.

Four homogenous physical units with unique physical characteristics were identified at the Victoria property. Comparison of physical units with rock types indicated that lithological units with similar physical properties were grouped together logically. The understanding of the relationship between physical units and lithological units facilitates translation of physical data to geological information and vice versa.

2.9 References

Bezdek, J. C. (1981). Pattern recognition with fuzzy objective function algorithm. Springer US. doi: 10.1007/978-1-4757-0450-1

Benaouda, D., Wadge, G., Whitmarsh, R. B., Rothwell, R.G. and MacLeod, C. (1999). Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the Ocean Drilling Program, Geophysical Journal International, 136, 477-491. doi: 10.1046/j.1365-246X.1999.00746.x

Burrough PA, McDonnell RA (1998) Principles of geographic information systems. Oxford University Press, Oxford, pp 333.

DeGruijter, J. J., and McBratney, A. B. (1988). A modified fuzzy k-means method for predictive classification, in classification and related methods of data analysis, edited by H. H. Bock, pp. 97–104, Elsevier Sci., New York,.

Garcia, M., H., Rabaute, A., Yven, B., and Guillemot, D. (2011). Multivariate and spatial statistical analysis of Callovo-Oxfordian physical properties from lab and borehole logs data: Towards a characterization of lateral and vertical spatial trends in the Meuse/Haute-Marne

Transposition Zone. Physics and Chemistry of the Earth, 36, 1469–1485. doi: 10.1016/j.pce.2011.07.086

Granek, J. (2011). Computing geologically consistent models from geophysical data. M.Sc. thesis, University of British Coloumbia, BC, Canada.

Inwood, J., Lofi, J., Davies, S., Basile, C., Bjerum, C., Mountain, G., Proust, J.-N., Otsuka, H., and Valppu, H., (2013). Statistical classification of log response as an indicator of facies variation during changes in sea level: Integrated Ocean Drilling Program Expedition 313. Geosphere, 9:1025–1043. doi:10.1130/GES00913.1

Killeen, P.G., (1997). Borehole geophysics: Exploring the third dimension. In "Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration", edited by A.G. Gubins, 31–42.

Killeen, P., Mwenifumbo, C., Elliott, B., and Chung, C. (1977). Improving exploration efficiency by predicting geological drill core logs with geophysical logs. In "Proceedings of Exploration 97", edited by A.G. Gubbins, 713-716.

Lofts, J. C. (1993). Integrated geochemical-geophysical studies of sedimentary reservior rocks. PhD thesis, University of Leicester, Leicester, UK.

McDowell, G. M., King, A., Lewis, R. E., Clayton, E. A., and Grau, J. A. (1988). In-situ nickel assay by prompt gamma neutron activation wireline logging. In "Proceedings of SEG Annual Meeting, New Orleans, Louisiana".

Mwenifumbo, C., and Mwenifumbo, A. (2012). Borehole geophysical logging in the Flin Flon Mining Camp. Geological survey of Canada, Open File 6547, 75 p. doi:10.4095/291534

McBratney, A. B., and DeGruijter, J. J. (1992). A continuum approach to soil classification by modified fuzzy k-means with extragrades. Soil Science, 43, 159– 175. doi: 10.1111/j.1365-2389.1992.tb00127.x

Minasny, B., and McBratney, A. B. (2000). FuzME version 2.1. Australia, Center for Precision Agriculture. New South Wales: University of Sydney. Retrieved from http://www.usyd.edu.au/su/agric/acpa

McDowell, G.M., Fenlon, K. and King, A. (2004). Conductivity-based nickel grade estimation for grade control at Inco's Sudbury mines. In "Proceedings of SEG Annual Meeting, Denver, Colorado".

Ojha, M., and Maiti, S. (2013). Sediment classification using neural networks: An example from the site-U1344A of IODP Expedition 323 in the Bering Sea. Deep-Sea Research II.

Perron, G., Fullagar, P., Pears, G., Phillips, N., Williston, C., Gerrie, V., and Everest, J. (2011). 3D litho-prediction model from in-situ physical rock property logging and constrained potential fields data inversion. In "Proceedings of 31st Gocad Meeting, Nancy".

Rabaute, A., Briqueu, L., Harvey, P. K., and Mercadier, H. (1997). Characterisation of oceanic sediments and basalts from Hole 833B, New Hebrides island arc (Vanuatu) from downhole measurements. Scientific Drilling, 6, 91-101.

Rabaute, A., Yven, B., Chelini, W., and Zamora, M. (2003). Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements. Geophysical Research, 108. doi: 10.1029/2001JB001436.

Reed, L.E., Taner, M.F., Elliott, B.E., and Killeen, P.G. (1997).A comparison of physical property borehole logs with geology, mineralogy and chemistry in a borehole at Les Mines Selbaie, northwestern Québec, Canada. In “Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral Exploration”, edited by A.G. Gubins, 1997, 1043–1048.

Roubens, M. (1982). Fuzzy clustering algorithms and their cluster validity. European Journal of Operational Research, 10, 294–301. doi: 10.1016/0377-2217(82)90228-4

Ruspini, E. H. (1969). A new approach to clustering. Information and Control, 15, 22-32. doi: 10.1016/S0019-9958(69)90591-9

Williams, N., and Dipple, G. (2007). Mapping subsurface alteration using gravity and magnetic inversion models. In "Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration" edited by B. Milkereit, 2007, 461-472.

Vella, L., Emerson, D., (2009). Carrapateena: Physical properties of a new Iron-oxide copper-gold deposit. Journal Article ASEG Extended Abstracts, 1 – 13. doi:10.1071/ASEG2009ab086

Xie, X. L., and Beni, G. (1991). A validity measure for fuzzy clustering. Pattern Analysis and Machine Intelligence, IEEE Transactions, 13, 841–847. doi: 10.1109/34.85677

Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8, 338-353.

Chapter 3

3 Supervised classification of down-hole physical properties measurements using neural network to predict the lithology

3.1 Abstract

The reliability of rock-type prediction using down-hole density, gamma ray response, and magnetic susceptibility measurements was evaluated at the Victoria property, Sudbury, ON. A supervised neural network, trained using lithological information from drillhole FNX1168, yields a predictive accuracy of 83% for the training data. Applying the trained network on drillhole FNX1182 resulted in 64% of the rock types being correctly classified when compared with the classification produced by geologists during logging of the core. The homogenous rock types, like quartz diorite, had a high accuracy of classification; while the heterogeneous rock types such as diabase were poorly classified. Overlap between physical properties of rock types caused by heterogeneity or inherent similarity in physical properties of rock types, which were verified by observing the cores, accounts for most of the misclassification. To reduce the misclassification, the network was trained from physical log units in FNX1168 derived from clustering of physical properties measurements. Four physical log units mainly represented four groups of rocks: i) quartz diorite; ii) metabasalt and metagabbro; iii) metasediment and quartzite; and iv) sulfide and diabase. The predictive accuracy in the training process rose to 95%. The trained network then was applied to predicting the physical log units in FNX1182. Given the relationships between physical log units and rock types from FNX1168, the results of physical-log-unit classification in FNX1182 were interpreted from a geological point of view. Although in ideal cases we would like to be able to extract the same classification that a geologist provides, the extraction of

physical log units is a more realistic goal. The interpretation of the lithological units from the physical log units can be compared with the geologist's classification and discrepancies or anomalies analyzed in greater detail.

3.2 Introduction

Typically when a hole is drilled, a geologist will look at the core extracted and classify the lithology or rock type as a function of depth down the hole. Rock type prediction based on log data from down-hole geophysical measurements can be considered as a potential alternative to a geologist's log when the cores are not fully recovered such as ocean drilling or drilling methods which do not provide cores such as percussion drilling in mineral exploration (Benaouda, et.al, 1999 and Qi and Carr, 2006). Physical properties logging provides a continuous set of data down the hole, which can be effectively used to improve the understanding of the geological characteristics of the hole (Killeen 1997; McDowell et al., 1998; Granek, 2011). The accuracy of rock-type characterization based on physical properties is proportional to the existent contrast of these data between rock types (Perron et al., 2011 and Mwenifumbo and Mwenifumbo, 2012). But, similarity of physical properties of rocks and the heterogeneity of the rock increases the overlap between physical properties of different rocks (Rabaute, et al., 2003 and Garcia, 2011). Overlap of physical properties between two rocks brings confusion to the prediction of rock types. We feel that it is more realistic to classify them as physical units. The term "physical log units" can be described as homogenous intervals of one or more rock types with consistent physical properties. The links between physical units and rock types can help geophysical studies to gain a better understanding of the geological setting (Benaouda, et.al, 1999 and Perron et al., 2011).

Conventional statistical techniques such as using histograms, box and whisker plots, cross plots, or the analysis of average and variance, were employed by Reed (1977), Killeen et al. (1997), McDowell et al. (1998), McDowell et al. (2004), and Vella and Emerson (2009) to extract the pattern of variation in physical properties measurements and relate them to geological setting. Recently, the multivariate pattern recognition techniques have become popular, since the conventional methods were limited in terms of the number of variables, and their ability to establish a quantitative relationship between physical properties and geological setting (Rabaute, et al., 2003; Qi and Carr, 2006; Williams and Dipple, 2007; Garcia, et al., 2011 and Granek, 2011). Supervised classification techniques can be used when the goal is to predict rock types based on the physical properties measurements. As the term 'supervised' implies, a drillhole or parts of a hole with both physical properties measurement and rock type information are used to train the classifier, and then the trained classifier can be applied to a new hole or those parts lacking core to predict the rock types (Benaouda, et.al, 1999 and Qi and Carr, 2006).

Varying composition of rock types, structures, alteration and mineralization justifies using a non-linear classifier to analyze these data, as rock types cannot be linearly separated using physical properties data. The neural network is a robust non-linear classifier successfully applied to down-hole logging data (Ojha and Maiti, 2013). Baldwin et al (1990), Wong et al. (1995), Farmer and Adams (1998), Qi and Carr (2006) and Maiti et al (2007) used the neural network to predict lithofacies based on down-hole physical properties measurements. In these works a quantitative relationship between numerical well log data and core description is simulated by the neural network, then the network is applied to the uncored wells or parts of the well lacking the core to predict lithofacies. The application of a neural network was introduced to the Ocean Drilling Program by Benaouda et al (1999) and Ojha and Maiti (2013). They used the neural network to

classify down-hole physical properties measurement to predict lithology where there is partial or zero core recovery along the hole. All of the above mentioned works have been applied in the sedimentary environment, and the prediction was considered fairly successful. Application of this method to the igneous and metamorphic environments like the Victoria property might be complicated as they are more complex than sedimentary environments due to the structural metamorphic and intrusive history.

In this research, the main objective is to compare the reliability of representing the physical properties measurements in the form of rock types and physical log units. As a first step, a neural network trained from physical properties and geological logging information collected in hole FNX1168 was used to predict the rock type in hole FNX1182. If the network was trained on the data from more than one hole, better results might be obtained, but our purpose is to test the efficacy of neural networks early in the exploration process when less data are available. The predicted rock types in FNX1182 are compared with actual rock types logged by geologists to evaluate the accuracy of the classification in the context of an igneous and metamorphic environment. With the physical properties measurements available to us, we demonstrate below that this is only moderately successful. Then rather than using the geologists' classification of rock units for training, we used physical log units in FNX1168 determined by fuzzy k-means clustering (Mahmoodi and Smith, 2015). The trained neural network was then used to predict physical units in FNX1182, with greater success. Considering the relationship between physical units and rock types are known for FNX1168, the predicted physical units of FNX1182 can be interpreted from a geological point of view. Finally, we discuss whether to represent down-hole physical properties measurements; in the form of rock type or physical log units.

3.3 Methods

The neural network is structurally comprised of an input layer, at least one hidden layer, and an output layer. Each layer has different numbers of neurons. Neurons of two adjacent layers are connected one-by-one by synaptic weights. The schematic structure of a simple three-layer neural network is shown in Figure 3-1.

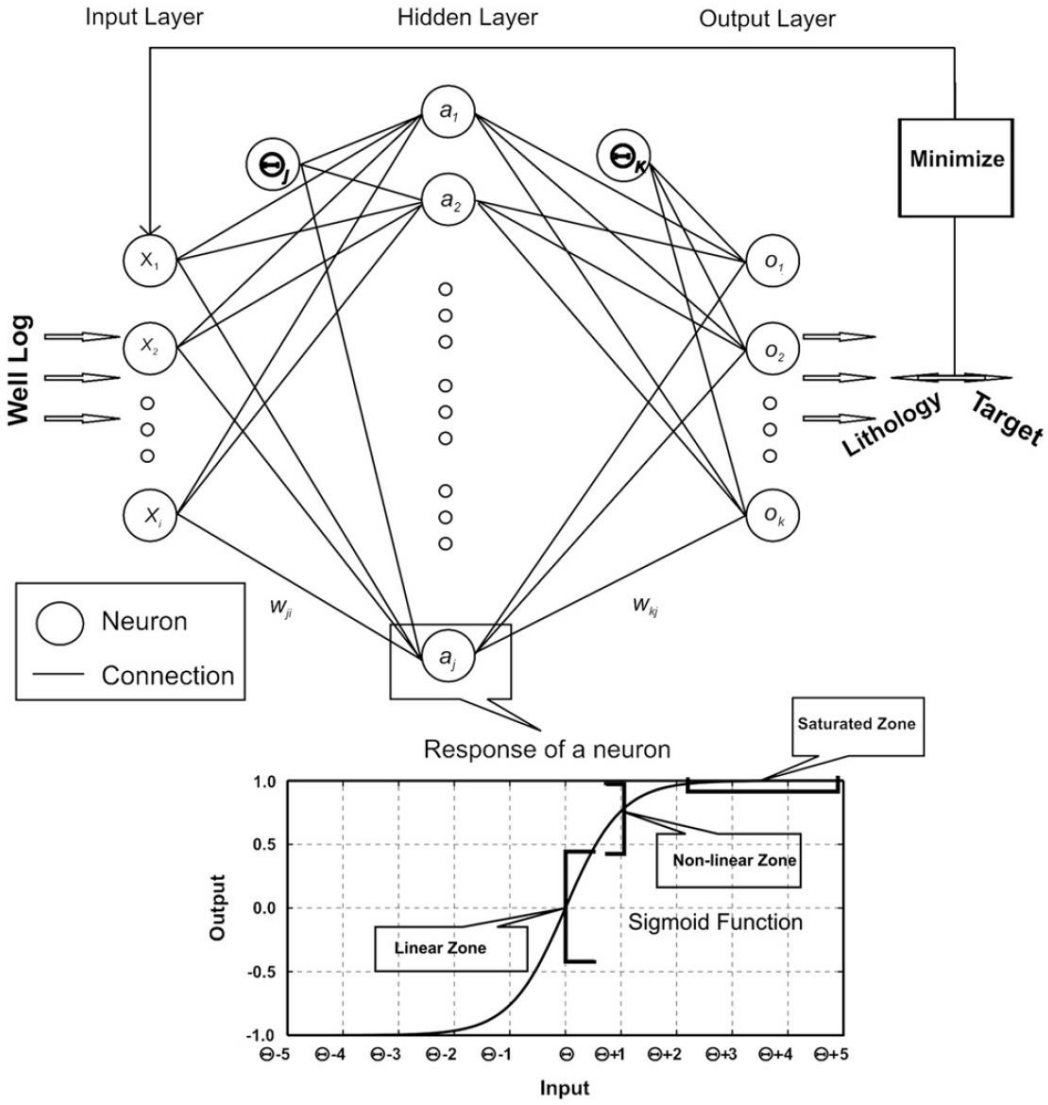


Figure 3-1- The structure of a three-layer neural network. The well log data are presented to the input layer as vectors where x_i represents the i^{th} variable in an input data vector. Each input neuron is connected to neurons of the hidden layer by synaptic weights.

Summation of input neuron multiplied by weights with bias added are fed to hidden neurons, a_1 to a_i represented by sigmoid function. The hidden neurons produce output which is input for the output layer. The sigmoid function, used as a transform function, is shown in the bottom of the figure. The bias neurons of the K^{th} layer are shown as Θ_k . The output layer has *k* neurons which is equal to the number of classes in the target vector. In the training process, the network parameters are adjusted in an iterative manner to minimize the difference between the output of network and the target vector (Reprinted from Deep-Sea Res. II, Ojha and Maiti, Sediment classification using neural networks: An example from the site-U1344A of IODP Expedition 323 in the Bering Sea, Copyright (2013), with permission from Elsevier).

The input data are represented as a data vector to the input layer. The number of neurons in the input layer equals the number of variables measured at each depth (in our case three, gamma-ray response, density and magnetic susceptibility). The number of output neurons is determined by the number of elements of the target vector; which is the number of classes in the classification problems (in our case seven rock types). The hidden neurons are computing elements of the network, which use transfer functions to generate the output. The input of a hidden neuron is a summation of bias and neurons in the previous layer multiplied by corresponding synaptic weights. Biases allow the activation function to shift to the right or left to give a desired output (Bishop, 1995; Theodoridis and Koutroumbas, 2003; Beale et al., 2014). In multilayer networks, the sigmoid function shown in Figure 3-1 is often used as the transfer function (Beale et al., 2014). There is no rigorous theoretical method available to choose the number of hidden layers and hidden neurons, and they are often determined subjectively based on trial and error. The most compact structure with acceptable performance is preferred for computational efficiency. Most classification problems can be solved by the network when one hidden layer is used. With

increase in the complexity of the relationship between input and desired output on the training data, the number of the hidden neurons should increase (Lawrence et al., 1996; Wong et al, 1998; Beale et al., 2014).

The initial weights are randomly assigned to start the training process. The main task in network training is to adjust the weights to minimize the error of the network. The error function for each iteration is a form of the difference between the actual network output and the desired or target output. Here, mean square error (*mse*) is described as the error function:

$$J = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2$$

where *t* and *a* represent desired and network output respectively for each neuron of the output layer for *N* training data. In our case *N* is the number of depths that physical properties are measured at. A gradient descent is an optimization method which is used to adjust the weights in the direction through which the most rapid decrease in error function is achieved. The adjusting term ($\Delta\omega_j^r$) which is added to the j^{th} weight of the r^{th} layer weight after iteration is obtained by the derivative (gradient) of the error function with respect to the weight. To enhance optimization efficiency, this value is multiplied by a learning rate (*lr*). The learning rate which is greater than zero and smaller than or equal to 1 controls the speed of the convergence process and how much the weights and biases can be modified at each iteration. The new estimate of the weight $\omega_{j\,(new)}^r$ is described as:

$$\omega_{j\,(new)}^r = \omega_{j\,(old)}^r - lr.\Delta\omega_j^r\,, \quad \Delta\omega_j^r = \frac{\partial J}{\partial \omega_j^r},$$

where $\omega_{j\ (old)}^{r}$ is the current weight, $\Delta\omega_j^r$ is the adjusting term, and *lr* is learning rate. The network with adjusted weights generates a new output set, and the process iterates (each iteration referred to as one epoch) until the termination criteria is fulfilled. Different criteria have been suggested to terminate the iterations, such as a threshold for the minimum performance function (cost function), the minimum decrease in the cost function in successive iterations, and the number of validation checks, which is the number of successive iterations that the cost function fails to decrease (Bishop, 1995; Benaouda et al, 1999; Theodoridis and Koutroumbas, 2003; Beale et al., 2014).

The training data set must be sufficiently large to provide enough data to train and test the network. The data set is typically randomly divided into three parts for training, validation, and testing constituting 70, 15, and 15 per cent of the available data, respectively. Testing data are used to assess generalization of the trained network which determines capability of the network to be efficiently applied to new data. If the network keeps iterating and adjusting the parameters to minimize the error, the network starts to over-fit the training data. Over-fitting occurs when the network, irrespective of generalization, tries to minimize the error of training data. In this case, the impact of random noise is incorporated into the network weights; however, the error in testing and validation rises. Validation data is used to assure that the division of data is appropriate. If the errors of testing and validation data are significantly different, it indicates poor data division (Bishop, 1995; Benaouda et al, 1999; Theodoridis and Koutroumbas, 2003; Qi and Carr, 2006; Beale et al., 2014).

The MATLAB neural network toolbox was used to train the network, and then predict rock types and physical log units. The training process was carried out with 1000 iterations, 10 validation

checks, a learning rate of 0.01, and a minimum gradient of 1×10^{-5}, as well as a performance goal of 0.03 for rock type prediction. These parameters are suggested by Lawrence et al. (1996) to avoid overfitting the network. A performance goal of 0.001 was used for physical unit prediction to take a longer time to train the network.

3.4 Data set used

The neural networks are not capable of extrapolating accurately beyond the range of data; hence, training and new data should have the same range (Beale et al., 2014). Therefore, training data and new data should be collected from one drillhole or a different hole drilled in the area with a similar geological setting. This will guarantee the consistency of the physical properties of the rocks and their relation to the rock type in the two datasets. In the current work, density, gamma-ray response, and magnetic susceptibility measured within FNX1168 and FNX1182 drilled at the Victoria property were used. Figure 3-2 shows the location of FNX1168 on the geological map of the Victoria property. FNX1182 is located to the north beyond the area covered by this map but is interpreted to intersect at depth the same geological units as appearing in FNX1168.

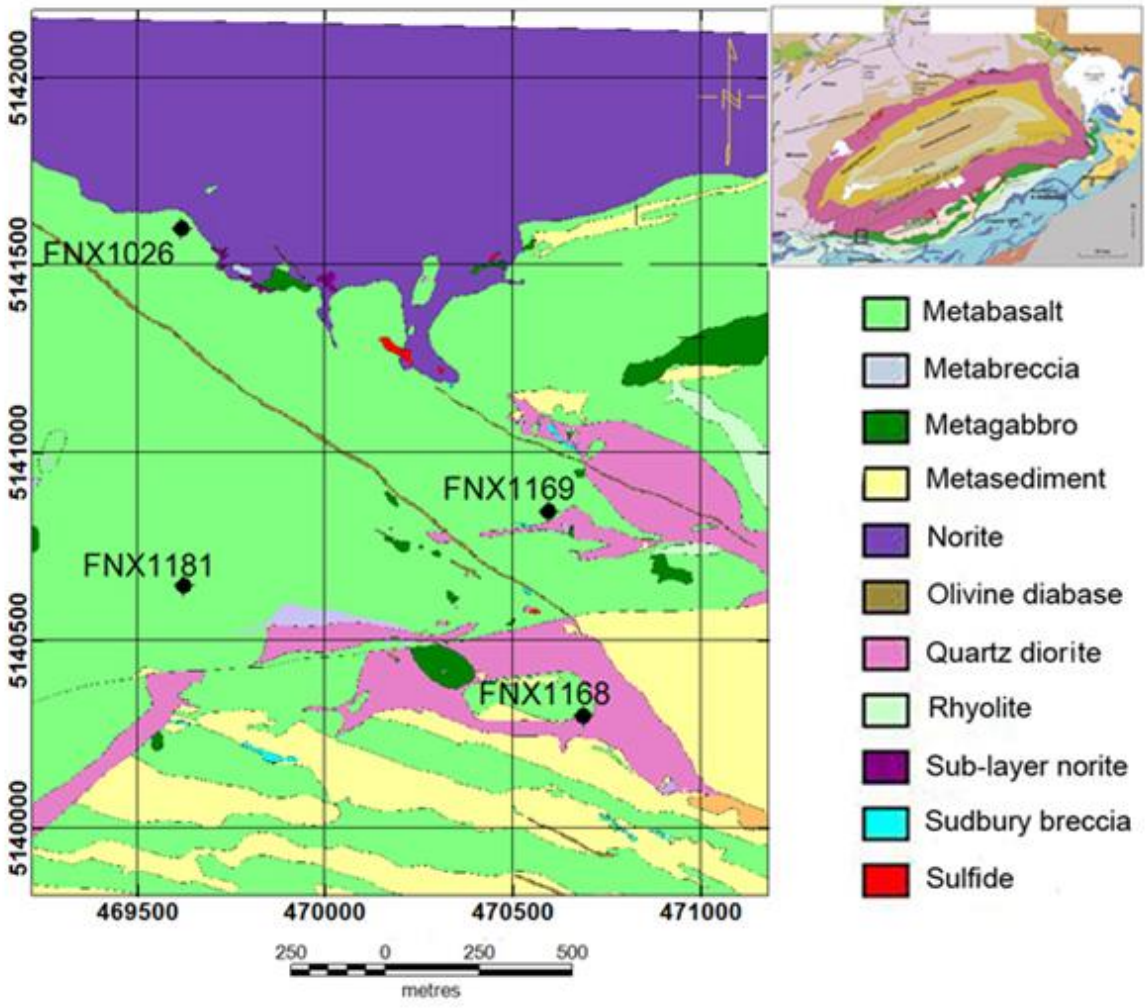


Figure 3-2- The location of FNX1168 on the geological map of the Victoria property. FNX1182 is located to the north beyond the area covered by this map. The study area is marked by the small black rectangle in the top right figure.

The boreholes were theoretically NQ sized (diameter of 75.7 mm), but the caliper log was measured to implement on-site calibration to correct for borehole size. The fluid filling the holes was water. A winch and steel cable system was used to lower a multi-parameter digital logging

probe down the holes and measure the data. The data includes 4405 and 8484 measurement at an interval of 20 cm along the FNX1168 and FNX1182 holes, respectively.

We undertook the neural network training and classification process on the logging data from FNX1168 twice, firstly with the target vector as the rock type logged by geologists and secondly with the target as the physical log units identified by the fuzzy k-means clustering (Mahmoodi and Smith, 2015). The rock types logged by geologists in FNX1182 are only used to compare with the prediction from the trained networks. The physical properties measurements, rock type logged by geologists, and physical log units of FNX1168 are shown in Figure 3-3.

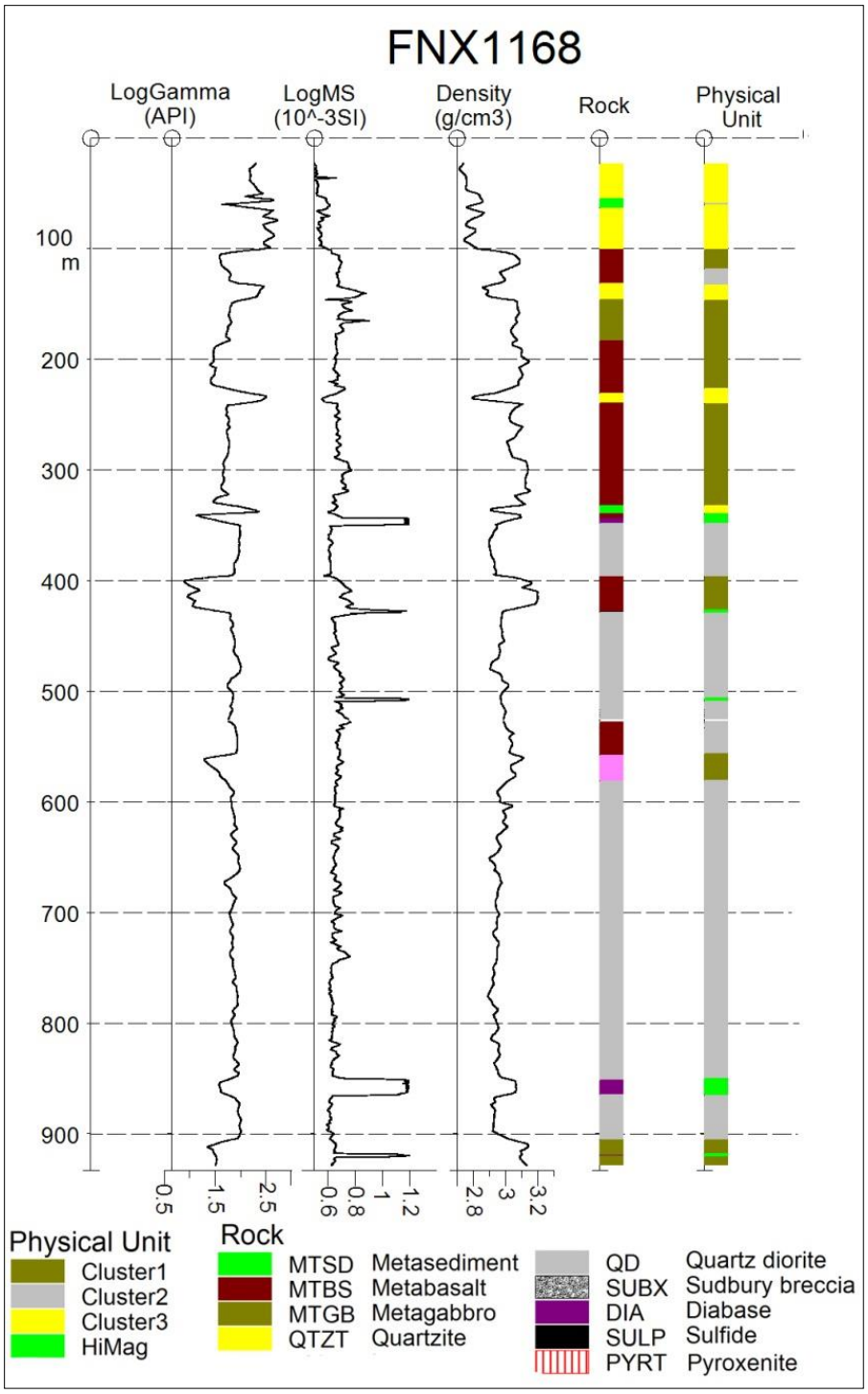


Figure 3-3- Gamma-ray response, magnetic susceptibility and density measurement in FNX1168. These properties were used to train the network to predict the rock types logged by geologists (second from right panel) or the physical log units identified by fuzzy k-mean clustering (Mahmoodi and Smith, 2015) (right panel).

The three-element input vectors representing three logging measurements at each depth were created. As is required in neural networks, each physical property should be normalized between 0 and 1. Because the magnetic susceptibility and gamma-ray response cover a large range of values, we found we could get better results when we took the logarithm of the measured value prior to normalizing the data. Each input vector of FNX1168 is then associated with a target vector identifying the rock type or physical unit belonging to each measurement point. The number of elements in the target vector equals the number of classes present in training data; i.e. seven elements for rock types and four elements for physical units. The k^{th} element of the target vector is 1 when an input vector belongs to the k^{th} class, and the rest are 0 (Qi and Carr, 2006 and Beale et al., 2014).

3.5 Results

Precise training is a very important step in the classification of data using the neural network. Great care should be taken when selecting input variables from the available data set; the number of input data should be sufficient to allow the network to be trained properly. In neural network classification, the structure of the network is of significant importance. The network with the most compact and simplest structure is generally selected to classify new data as this provides minimal error and greatest generalization. This process is not straightforward; it requires considerable experimentation and evaluation of the error of performance and generalization of the network with different structures to obtain the optimal choice (Qi and Carr, 2006). Generally the required number of hidden layers and hidden neurons increases as the complexity between input and output data increases.

3.5.1 Predicting geological units

After some experimentation, the structure of the network for rock type prediction was selected to be 3-15-7. These values represent the number of neurons in the input layer, hidden layer, and output layer, respectively. In the training process the best performance occurred at the 76th epoch and training was terminated at 86th epoch where the maximum number of validation checks was reached. Overall, the network showed 83% correlation between the predicted rocks and the actual rocks during the training process. The confusion matrix in Table 3-1 shows the accuracy of the network in classification of individual rock types.

Table 3-1-Confusion matrix for the training data in hole FNX1168. The rock types logged by geologists (desired output) and predicted rock types by the neural network are represented in rows and columns, respectively. The percentage of contribution of actual rock types in each predicted rock types are summarized in the table. The total length of logged and predicted rock types are listed in last column and last row, respectively.

		Predicted rock types							
		Quatz diorite	Meta-gabbro	Meta-basalt	Meta-sediment	Quartzite	Diabase	Sulfide	Length (m)
Actual rock types	Quatz diorite	95	0	4	0	0	1	0	455
	Metagabbro	11	1	88	0	0	1	0	59.8
	Metabasalt	18	0	81	0	1	0	0	235.6
	Metasediment	29	0	8	0	64	0	0	16
	Quartzite	3	0	0	0	97	0	0	93.2
	Diabase	4	0	3	0	0	93	0	20.4
	Sulfide	0	0	20	0	0	80	0	1
	Length (m)	492.2	1	262.6	0	102.6	22.6	0	

The homogenous rock types (rocks with spatially consistent physical properties), such as quartz diorite shows a high predictive accuracy. The accuracy for metabasalt, quartzite, quartz diorite, and diabase is remarkably high with 81, 97, 95, and 93 per cent of cases predicted correctly. On the other hand, heterogeneity in rock types results in a large range of variability of physical

properties and more overlap between rock types. Metasediment, metagabbro, and sulfide are completely misclassified, as they are predicted as other rock types with similar physical properties. For example metasediment is usually predicted as quartzite; metagabbro is predicted as metabasalt; and sulfide is predicted as diabase.

Considering the prediction of similar rock types in Table 3-1 a question arises as to why metasediment is predicted as quartzite, and not vice versa. It turns out the network shows preference for one rock type in the classification. The rock type with higher population contributes more data for training and controls the network parameters. Since quartzite is more dominant than metasediment, the network favors a choice for quartzite over metasediment; likewise, sulfide is predicted as diabase. Benaouda et al, (1999) incorporated the same number of samples from each class to avoid a bias caused by the larger populations when training the neural network. But they concluded this approach under-represents the larger classes which has an even greater detrimental impact on the results.

After the training process the parameters of the network are frozen and are then applied to data in FNX1182. The input data vector for each depth is presented to the network, and the 7-array output vector is produced. Each element in the output vector represents the predicted rock type probability, and the maximum probability determines the predicted rock type at a specific depth (Qi and Carr, 2006). The physical properties, predicted rock types, and actual rock types logged by the geologists in FNX1182 are represented in Figure 3-4.

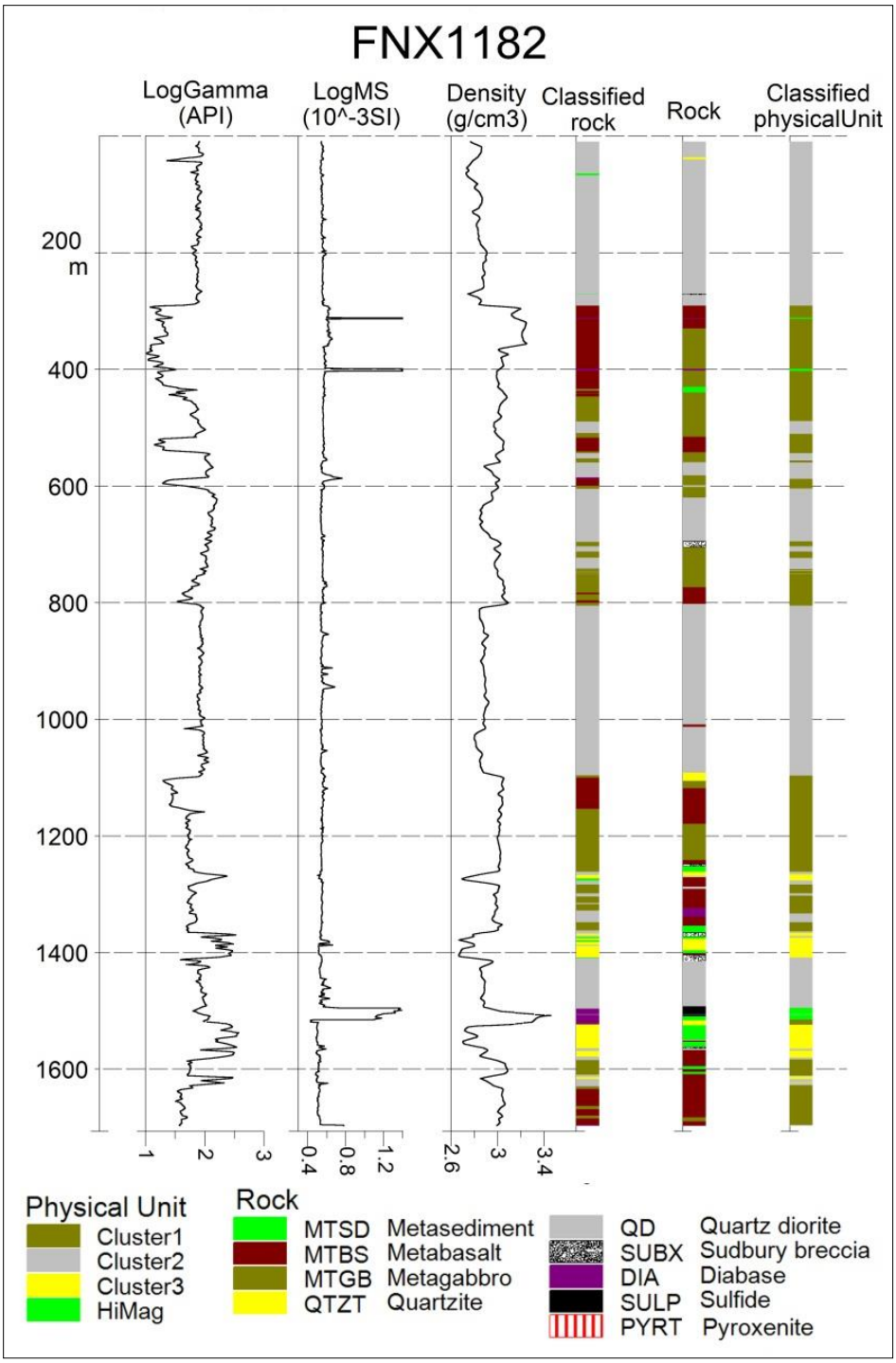


Figure 3-4- Gamma-ray response, magnetic susceptibility and density measurement in FNX1182, which were presented to the trained network. Predicted rock types (third from right panel) and physical log units (right panel) are plotted. The geologist's classified rock types (second from right panel) are shown for comparison purpose.

Comparison of the actual rock types and the predicted rock types shows they have a 64% overall correlation, which shows significant increase in the classification accuracy compared to a classification accuracy of 14%, which is what would be expected if there was a random relationship between physical rock properties and lithological units. The predictive accuracy for individual rock types in FNX1182 is summarized in Table 3-2. Three-dimensional (3D) cross plots of physical properties in actual rock types (Figure 3-5) and predicted rock types (Figure 3-6) are presented. The box plots in Figure 3-7 also depict the range of variation of physical properties (normalized between 0 and 1) in actual and predicted rock types of FNX1182. The rock types which are predicted accurately have boxplots which show similar range of variation to those for the actual rock types.

Table 3-2- The rock types logged by geologists (left column) were considered to evaluate the network accuracy in rock type prediction in hole FNX1182. The rock types predicted by the neural network are represented in columns, and the percentage of contribution of actual rock types in each predicted rock types are summarized in the table. Total length of logged and predicted rock types are listed in last column and last row, respectively.

		Predicted rock types							
Actual rock types		Quatz diorite	Meta-gabbro	Meta-basalt	Meta-sediment	Quartzite	Diabase	Sulfide	Length (m)
	Quatz diorite	98	1	0	0	1	0	0	754
	Metagabbro	20	44	35	0	0	1	0	378.2
	Metabasalt	15	35	43	1	6	1	0	349.2
	Metasediment	5	35	7	2	44	7	0	89
	Quartzite	23	1	2	8	55	13	0	33.6
	Diabase	52	25	0	0	0	23	0	20
	Sudbury breccia	40	31	0	3	27	0	0	39.2
	Sulfide	25	0	0	0	0	75	0	15.8
	Pyroxenite	47	21	31	0	0	0	0	17.8
	Length(m)	913.6	350.8	295.4	13	92	32	0	

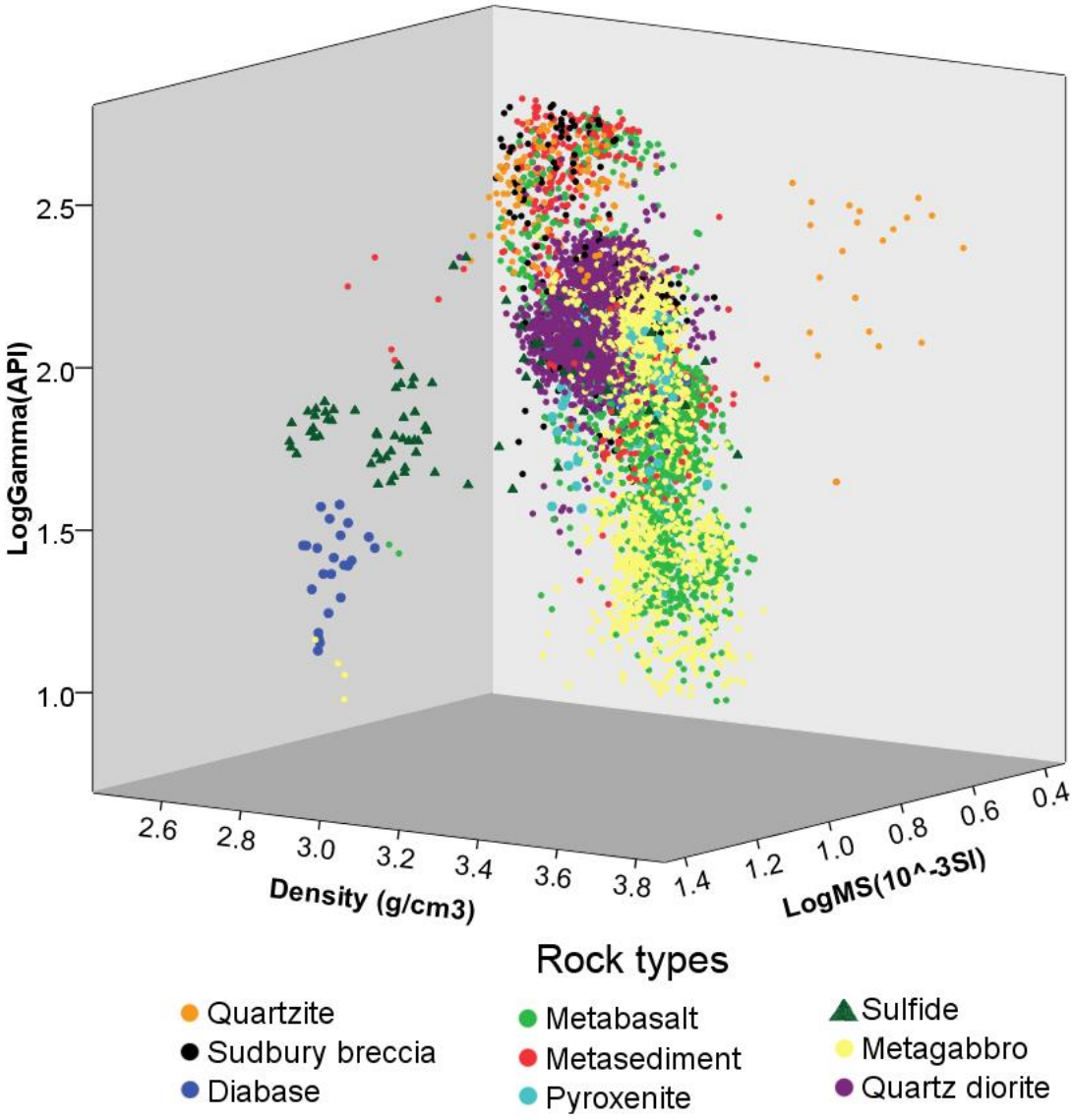


Figure 3-5- 3D cross plots of density, gamma-ray response, and magnetic susceptibility measurements for the actual rock type in FNX1182.

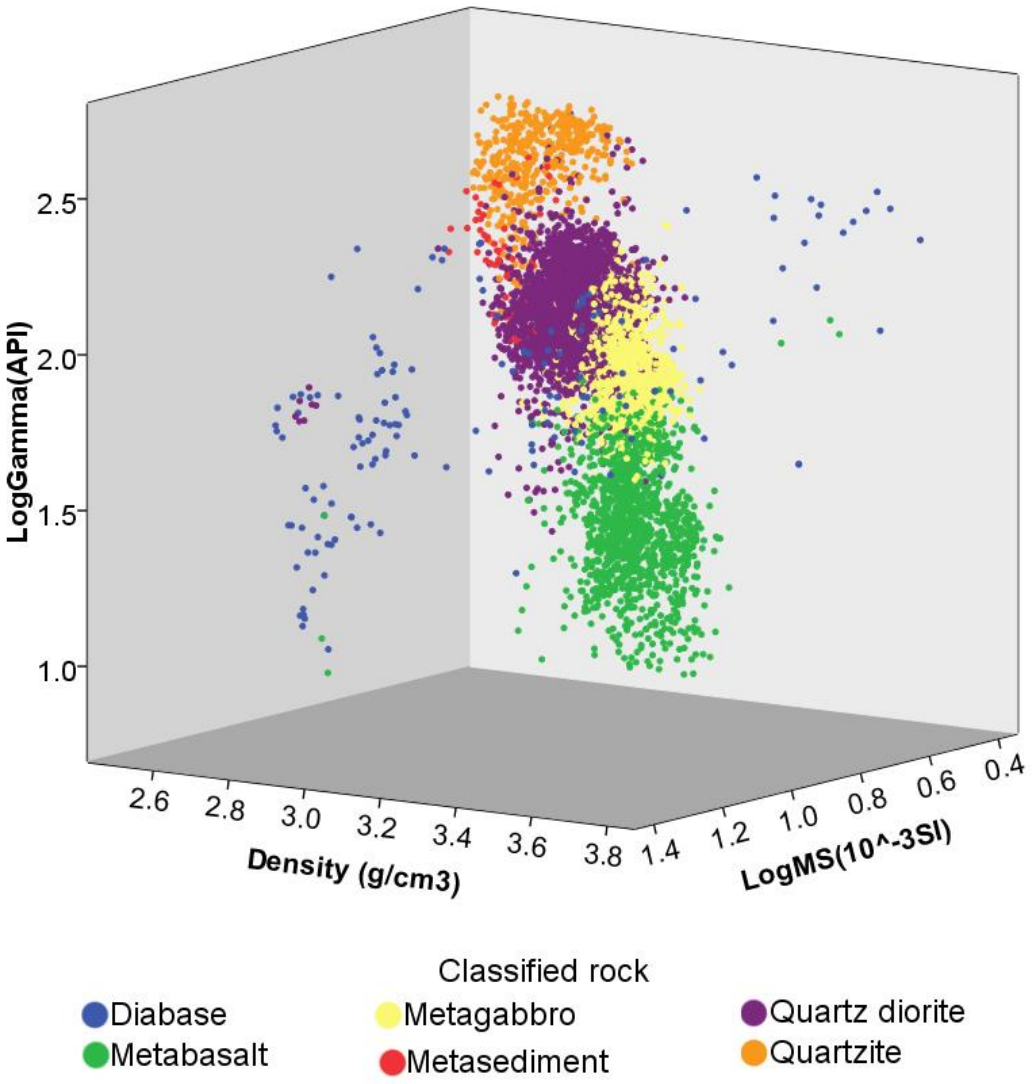


Figure 3-6- 3D cross plots of density, gamma-ray response, and magnetic susceptibility measurements for the predicted rock types in FNX1182.

As shown in Figure 3-5 quartz diorite forms a dense, well separated class which facilitates classification of this rock type. The predicted quartz diorite has similar distribution in Figure 3-6 as 98% of quartz diorite logged by geologists was predicted as quartz diorite in the neural

network classification. Such a high correlation is illustrated by a similar range of physical properties of actual quartz diorite and predicted quartz diorite in Figure 3-7.

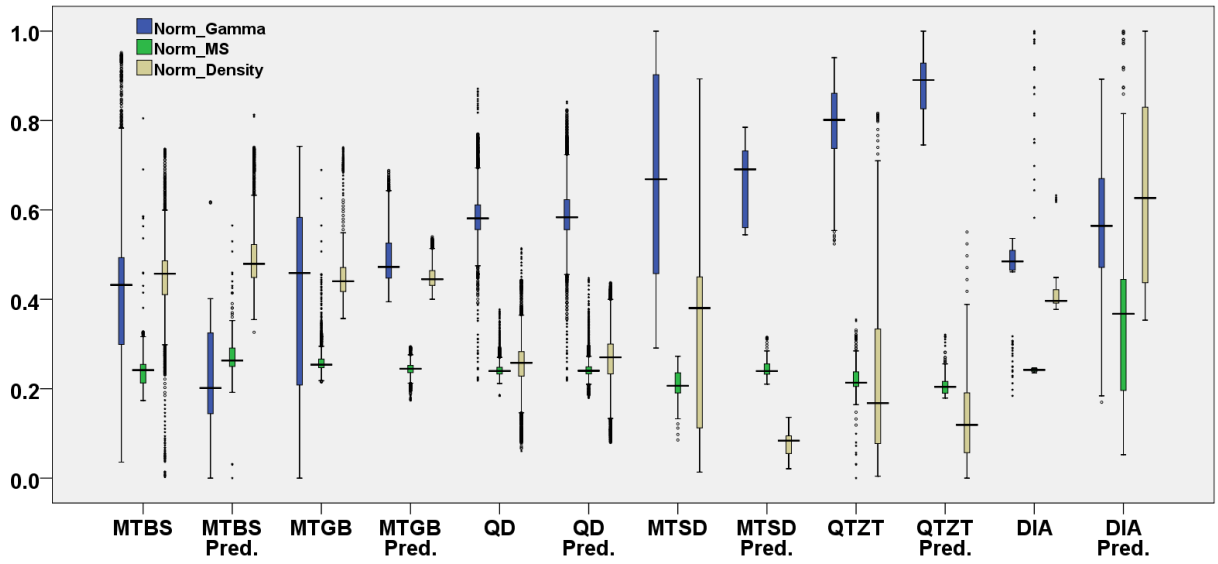


Figure 3-7- Box-plots of normalized gamma ray response, magnetic susceptibility and density measurements of rock types logged by geologists and predicted by the neural network. Labels used in the plot: MTBS: metabasalt; MTGB: metagabbro; QD: quartz diorite; MTSD: metasediment; QTZT: quartzite; DIA: diabase.

Metagabbro and metabasalt have similar physical properties illustrated with significant overlap in Figure 3-5. Table 3-2 shows that 44% and 43% of actual metagabbro and metabasalt are predicted correctly as metagabbro and metabasalt, respectively. Dissimilar distribution of these rock types in Figure 3-5 and 6 illustrates their misclassification. However, they are mutually predicted as each other, i.e. 35% of both actual metagabbro and metabasalt are predicted as metabasalt and metagabbro, respectively. It implies that the network is not capable of distinguishing between rocks with similar physical properties. The average gamma-ray response

of predicted metagabbro and metabasalt is lower than actual rocks (Figure 3-7) as the parts of these rocks with relatively high gamma-ray response (20% of metagabbro and 15% of metabasalt) are predicted as quartz diorote. Considering Figures 3-6 and 3-7, it seems metagabbro and metabasalt rocks are mainly predicted based on gamma-ray response.

High gamma-ray response and low density are the main characteristics of metasediment and quartzite in FNX1168, but Figure 3-7 illustrates they have a broader range of variation in FNX1182, which results in these rocks being predicted as metagabbro and quartz diorite. Only 2% of metasediment is correctly predicted, although similarity in physical properties of metasediment and quartzite shown in Figure 3-7 means that 44% of metasediment is predicted as quartzite. Quartzite is relatively well classified with 55% receiving the correct classification. As shown in Figure 3-6, predicted quartzite has a relatively narrower range of gamma ray response than actual quartzite; this is because 23% of quartzite with medium gamma ray is erroneously predicted as quartz diorite.

Only 23% of diabase is correctly predicted by the neural network. High density and high magnetic susceptibility are the characteristics of diabase in FNX1168. However in FNX1182, as shown in Figure 3-7, the main portion of this rock has a lower density and is surrounded in Figure 3-4 by quartz diorite and metagabbro. This means that diabase in FNX1182 have different properties from the diabase used in the training process, so are predicted to be quartz diorite (52%) and metagabbro (25%). Measurements of other rock types with high magnetic susceptibility are predicted as diabase which broadens the range of variability of physical properties of predicted diabase in Figure 3-7. Inconsistent physical properties of diabase result in inaccurate performance of the network for diabase prediction.

Due to insufficient samples in the training process and similar physical properties of sulfide and diabase, sulfide is mainly (75%) predicted as diabase. Distribution of sulfide in Figure 3-5 corresponds to predicted diabase in Figure 3-6. Sudbury breccia and pyroxenite are three rock types in FNX1182 which are absent in FNX1168. So, the trained network is not capable of predicting these rocks. Their physical properties are such that they are predicted as other rock types. Pyroxenite is mainly predicted as quartz diorite and to some extent as metagabbro and metabasalt. Sudbury breccia is the most heterogeneous rock and does not show consistent physical properties (Figure 3-5) and its contribution to different predict rock types can be seen in Table 3-2.

3.5.2 Predicting physical log units

The second approach is to represent the physical properties measurements of FNX1182 with physical log units. A 3-10-4 network was selected for physical log unit prediction in FNX1182. Physical properties measurements and identified physical log units from FNX1168 (Figure 3-3) were used to train the network. Cluster 1 has low gamma ray response and high density; cluster 2 is characterized by medium gamma ray response and medium density; cluster 3 shows low density and high gamma-ray response; and a HiMag unit is distinguished by high magnetic susceptibility. A careful comparison of the physical log units with the geological rock type from FNX1168 (Mahmoodi and Smith, 2015) concluded that cluster 1 mainly represents metabasalt and metagabbro, cluster 2 constitutes quartz diorite, cluster 3 mainly represents metasediment and quartzite, and HiMag represents diabase and sulfide.

In the training process, the best validation performance occurred at the 61st epoch where the training stopped as the minimum gradient was reached. The more rapid convergence than training the network for rock type prediction suggests the reduced complexity of the problem.

The confusion matrix for the training data in Table 3-3 shows the physical log units have a total of 95% of the samples being correctly classified. Thus, the ambiguity associated with the rock-type prediction that resulted in misclassifications is not seen with physical log units.

Table 3-3- Correlation between physical log units defined by fuzzy k-means clustering (rows) and predicted by the neural network (columns) using the trainig data set from FNX1168. The values indicate the percentage of contribution of the defined physical units in each of the predicted unit. The total length of drillhole classified to each of the units are shown in the table.

		Predicted physical unit				
Physical units identified by clustering		Cluster 1	Cluster 2	Cluster 3	HiMag	Length(m)
	Cluster1	95	3	0	2	268.4
	Cluster2	1	98	1	0	495.2
	Cluster3	1	2	97	0	109.6
	HiMag	1	4	0	95	31.2
	Length(m)	263.4	494.6	112.2	34.2	

The neural network was then applied to the physical properties measurements collected in FNX1182 and each input datum was classified as either cluster 1, 2, 3, or the HiMag unit (Figure 3-4, the right column). Figure 3-8 shows distribution of physical units in the physical properties environment.

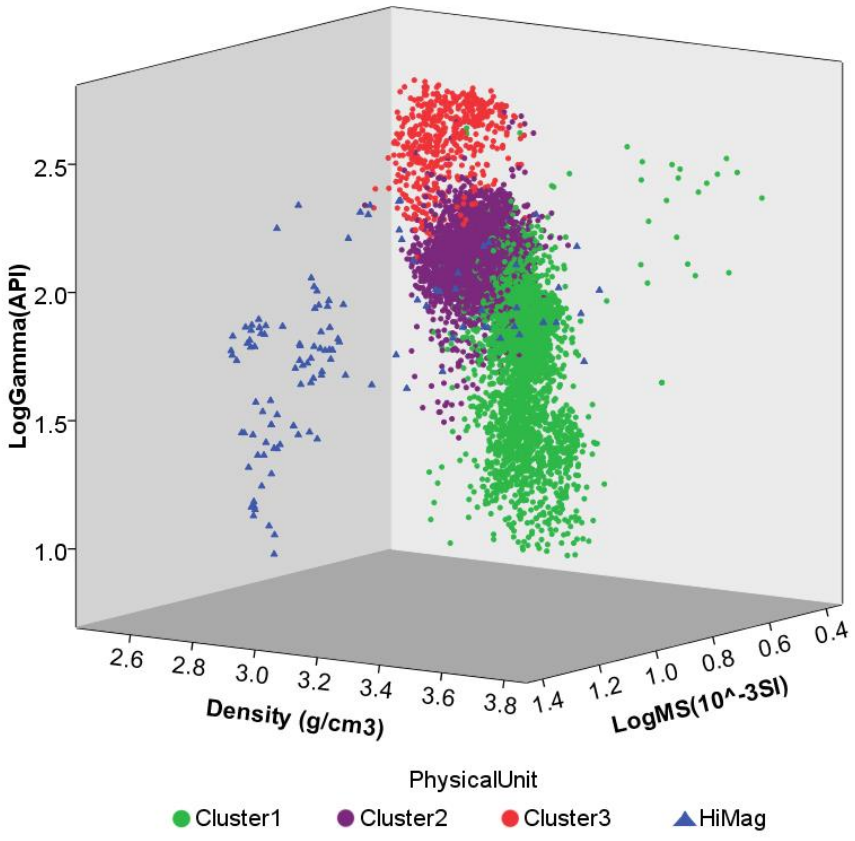


Figure 3-8- 3D cross plots of density, gamma-ray response, and magnetic susceptibility measurements for the predicted physical units in FNX1182.

In addition to being used in geophysical studies, the predicted physical units can be used in conjunction with available geological information. From analyzing FNX1168, Mahmoodi and Smith (2015) concluded that each cluster contains one or more rock types (metabasalt and metagabbro are in cluster 1; quartz diorite is cluster 2; metasediment and quartzite is cluster 3; and diabase and sulfide fall into the HiMag cluster). We assume that the same relationships apply to FNX1182.

To verify this assumption the predicted physical units were compared with the actual rock types in FNX1182 and the contribution of each rock types in the four physical units is summarized in

Table 3-4. The reliability of this approach mainly relies on our understanding of the links between physical units and rock types. Since pyroxenite and Sudbury breccia are absent in FNX1168, their relationships with physical units are unknown so the results cannot be assessed. But for the other rock types, the results mostly agreed with the expected correlation of physical units with rock types. This implies that physical unit prediction can be used as a tool to identifying classes of rock types with acceptable certainty.

Table 3-4- The percentage contribution of rock types (logged by geologists) as classified into physical log units in FNX1182.

		Predicted physical units				
		Cluster 1	Cluster 2	Cluster 3	HiMag	Length (m)
Actual rock types	Quartz diorite	1	98	1	0	754
	Metagabbro	77	22	0	0	378.2
	Metabasalt	80	12	8	0	349.2
	Metasediment	45	4	45	5	89
	Quartzite	17	13	70	0	33.6
	Pyroxenite	49	51	0	0	17.8
	Diabase	49	28	0	23	20
	Sulfide	0	16	0	84	15.8
	Sudbury breccia	34	32	34	0	39.2
	Length (m)	661	902.6	109	24.2	

Specifically, ninety eight percent of quartz diorite in FNX1182 is represented by cluster 2. Quartzite and typical metasediment with high gamma-ray response are the dominant rock types in cluster 3 with 70% and 45% contributions, respectively. Cluster 1 represents 77% of metagabbro and 80% metabasalt. Unlike in hole FNX1168, a high portion of diabase is classified into cluster 1. Sulfide was mainly represented in the HiMag cluster.

3.6 Discussion

Several factors should be considered while interpreting predicted rock types directly from physical properties or inferred from physical unit prediction.

- While training the network, the number of data belonging to each class involved in the training process influences the network performance. Because the rock types with a large population contribute more data in training, the network shows a preference for the dominant rock type rather than less populous rock types when they have similar physical properties. Thus, the significant predication of metasediment as quartzite, or sulfide as diabase can be explained, as in each case, the latter rock-type is more populous.

- Another factor is the heterogeneity of rock types which results in inconsistent physical property behavior. If the physical properties of a specific rock type are different in the training data and the prediction data, the rock type cannot be predicted correctly. The high density and high magnetic susceptibility are two distinguishing characteristics of diabase, but a remarkable portion of diabase in FNX1182 has lower density and magnetic susceptibility. This results in misclassification of diabase as quartz diorite and metagabbro. Another example of this is the misclassification of metasediment as metagabbro in FNX1182.

- Similarity in physical properties results in ambiguity. The overlap of physical properties of two or more rocks results in misclassification. Metagabbro and metabasalt are misclassified as each other in rock type prediction, and they are represented by one physical unit.

- The network is able to predict only the rock type or physical log units from which it is trained. Wise data selection for training the network is therefore critical for accurate classification. Sudbury breccia and pyroxenite were not considered as an output in the rock type prediction because they were absent in the training data.

The confusion associated with rock-type prediction is due to the overlap between the physical properties of the rock types, and this reduces the reliability of the results.

Core samples were studied to find a geological explanation for the similarities and differences in the physical characteristics of rocks. Plagioclase, amphibole, and pyroxene are the dominant minerals of both metagabbro and metabasalt, though metabasalt contains slightly higher amount of amphibole and a smaller amount of biotite compared with metagabbro. The grain size of metabasalt is finer than metagabbro. Thus, they are of broadly similar composition and mineralogy, explaining the similar physical properties of these two rocks, and the differences in grain size appears to not control the physical properties as much as the mineralogy. A detailed examination of the metabasalt and metagabbro showed patches of inconsistent grain size and the percentage of clinopyroxene observed in core samples, possibly explaining some of the heterogeneity in the physical properties measurements. Compared to metagabbro and metabasalt, quartz diorite contains less amphibole and pyroxene, and higher amounts of plagioclase and quartz. It has finer grain size, is lighter in color, and has a lower density. Occurrences of sulfide mineralization in quartz diorite will increase its density, although for low concentrations of mineralization the increase is not significant enough to make a significant change in the density log. Metasediment has a similar mineralogical composition to quartz diorite, but also contains minor garnet and iron oxide. The high gamma-ray response of the

metasediment is interpreted to be due to the relatively high concentration of radioactive minerals in the protolith (sedimentary rocks that were metamorphosed). Diabase has a fine grain size and is composed of pyroxene, amphibole, and up to 10% magnetite. Such high concentration of dense mafic minerals and magnetite increases the density and magnetic susceptibility of diabase. The portion of diabase with lower magnetic susceptibility contains less magnetite. Diabase can be mineralogically very similar to gabbro, which explains why the non-magnetic diabase is predominantly placed in cluster 1 containing metagabbro and metabasalt. Magnetite appears as disseminated mineralization. Metamorphism can result in changes in the amount and grain size of magnetite and variable metamorphism could play a role in the classification of diabase, metagabrro and metabasalt. High concentration of pyrrhotite also increases the magnetic susceptibility which corresponds to the sulfide zone on the log with high magnetic susceptibility. Pyroxenite is a rock primarily composed of the mafic mineral pyroxene which explains why it is easily misclassified as a metagabbro or a basalt (which also contain pyroxene) when there are few radioactive minerals present, or quartz diorite when there are more radioactive minerals present. The Sudbury breccia forms via cataclasis of pre-existing rocks and can form in any of the other rock types present in the area. The formation of the breccia produces a finer-grained matrix and clasts that simply reflect a mixture of the pre-existing rock types, explaining the heterogeneity and the difficulty in classifying this rock type.

If the rocks with similar physical properties are represented as a physical log unit, the problem is simplified. The significant decrease in the number of misclassified data in the process of training the network to identify physical log units illustrates greater accuracy and reliability (Table 3-3). In this case, although there are fewer types classified, interpretation can be more reliable since there is less ambiguity associated with the results.

3.7 Conclusion

Down-hole physical properties measurements can be used to extract information about geological material surrounding a bore hole. The neural network is a suitable tool to quantify the non-linear relationship between physical properties variation and their rock type. Depending on the objective, these data can be mathematically analyzed to predict geological rock type or determine physical log units in the hole.

When the physical properties measurements are used to predict rock types, three factors should be considered while interpreting the results: 1) the heterogeneity of rock types; 2) the overlap of physical properties of rock types; 3) the relative population of the different rock types in the training data. These factors control the training process. The heterogeneity causes inconsistent physical properties of rocks. The overlap is due to similarity of the physical properties of two or more rock types, and this reduces the network's ability to distinguish these rock types. The population of rock types also controls the training process as the network shows preference for dominant rock types in the output and minor rock types are predicted as the dominant rock types that have most similar physical properties.

The factors mentioned above were verified by analyzing the physical properties data, and observing the minerals and rock types in the corresponding core samples. The observed similarity of the physical properties between metagabbro and metabasalt comes from their identical to similar composition and mineralogy. Patches of inconsistent grain size and variant percentage of clinopyroxene in the metagabbro and metabasalt account for some portion of the observed heterogeneity in the physical properties measurements. Metasediment and quartzite at Victoria have anomalous gamma-ray response because both are produced by metamorphism of sedimentary rocks which is interpreted to contain a higher concentration of radioactive minerals.

The increase in the density and magnetic susceptibility of rocks can be related to the concentration of denser mafic mineral and/or sulfide minerals in rock types. High concentration of magnetite in diabase is one reason for its high magnetic susceptibility and density. A combination of the information derived from the classification results and core sample description establishes a reliable link between geology and physical properties at the Victoria.

In this work, the network was trained from data from FNX1168 to obtain the appropriate network to predict rock types and physical units in FNX1182. Predicting the physical units was simpler compared with rock type prediction, requiring only 10 hidden neurons compared with 15. When predicting the rock type, an accuracy of 83% was achieved in the training process, which is reasonable considering the complexity of this problem. However, the trained network only had 64% of accuracy in predicting rock types in FNX1182. There was an accuracy of 95% in training the network on physical log units (defined by a fuzzy k-means statistical classification), which implies an enhanced reliability and less ambiguity of the results.

In geological or geophysical studies, physical properties can be used to predict either rock type or physical units. Using a neural network to predict rock type instead of physical units gives more detailed information, however the tradeoff is less accuracy and greater ambiguity. On the other hand, when predicting the physical log unit, there is greater accuracy, but the ambiguity is still not completely resolved, as many different rocks can be classified as a single physical log unit depending on the exact factors controlling the geophysical properties. However, the physical log units do define identifiable and unique units with geological significance from a physical properties perspective and these can be useful for geophysical studies in forward and inverse modelling exercises.

3.8 References

Baldwin, J., Bateman, A., and Wheatley, C. (1990). Application of neural network to the problem of mineral identification from well logs. The Log Analyst, 31, 279–293.

Beale, M., Hagan, M., and Demuth, H. (2001). Neural network toolbox User's guide R2014a. Natick: The Math Works Inc.

Benaouda, D., Wadge, G., Whitmarsh, R. B., Rothwell, R., and MacLeod, C. (1999). Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the Ocean Drilling Program. Geophysical Journal International , 136, 477-491.

Bishop, C. (1995). Neural networks for pattern recognition. New York: Oxford University.

Farmer, R., and Adams, S. J. (1998). Facies recognition using neural networks. New Zealand Petroleum Conference. Queenstown.

Garcia, M. H., Rabaute, A., Yven, B., and Guillemot, D. (2011). Multivariate and spatial statistical analysis of Callovo-Oxfordian physical properties from lab and borehole logs data: Towards a characterization of lateral and vertical spatial trends in the Meuse/Haute-Marne Transposition Zone. Physics and Chemistry of the Earth, 36, 1469–1485.

Granek, J. (2011). Computing geologically consistent models from geophysical data. BC, Canada: University of British Coloumbia.

Killeen, P. (1997). Borehole geophysics: Exploring the third dimension. In A. Gubins (Ed.), Exploration 97: Fourth Decennial International Conference on Mineral Exploration, (pp. 31–42).

Killeen, P., Mwenifumbo, C., Elliott, B., and Chung, C. (1977). Improving exploration efficiency by predicting geological drill core logs with geophysical logs. In A. Gubbins (Ed.), Exploration 97, (pp. 713-716).

Lawrence, S., Giles, C. L., and Tsoj, A. C. (1996). What size neural network gives optimal generalization? Convergence properties of backpropagation. Institute for Advanced Computer Studies. College Park: University of Maryland.

Mahmoodi, O., and Smith, R. (2015). Clustering of downhole physical property measurements at the Victoria property, Sudbury for the purpose of extracting lithological information. Journal of Applied Geophysics, 118, 145-154.

Maiti, S., Tiwari, R. K., and Kumpel, H. J. (2007). Neural network modelling and classification of lithofacies using well log data: a case study from KTB borehole site. Geophysics Journal International, 169, 733–746.

McDowell, G. M., King, A., Lewis, R. E., Clayton, E. A., and Grau, J. A. (1998). In-situ nickel assay by prompt gamma neutron activation wireline logging. SEG Annual Meeting. New Orleans, Louisiana.

McDowell, G., Fenlon, K., and King, A. (2004). Conductivity-based nickel grade estimation for grade control at Inco‟s Sudbury mines. SEG Annual Meeting. Denver, Colorado.

Mwenifumbo, C., and Mwenifumbo, A. (2012). Borehole geophysical logging in the Flin Flon Mining Camp. Geological Survey of Canada, Open File 6547.

Ojha, M., and Maiti, S. (2013). Sediment classification using neural networks: An example from the site-U1344A of IODP Expedition 323 in the Bering Sea. Deep-Sea Research II.

Perron, G. F., Phillips, N., Williston, C., Gerrie, V., and Everest, J. (2011). 3D litho-prediction model from in-situ physical rock property logging and constrained potential fields data inversion. 31st Gocad Meeting. Nancy.

Qi, L., and Carr, T. R. (2006). Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas. Computers and Geosciences, 32, 947–964.

Rabaute, A., Yven, B., Chelini, W., and Zamora, M. (2003). Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements. Geophysical Research.

Reed, L., Taner, M., Elliott, B., and Killeen, P. (1997). A comparison of physical property borehole logs with geology, mineralogy and chemistry in a borehole at Les Mines Selbaie, northwestern Québec, Canada. In A. Gubins (Ed.), Exploration 97: Fourth Decennial International Conference on Mineral Exploration, (pp. 1043–1048).

Theodoridis, S., and Koutroumbas, K. (2003). Pattern recognition, (second edition). New York: Academic Press.

Vella, L., and Emerson, D. (2009). Carrapateena: Physical properties of a new Iron-oxide copper-gold deposit. Journal Article ASEG Extended Abstracts, 1 – 13.

Williams, N., and Dipple, G. (2007). Mapping subsurface alteration using gravity and magnetic inversion models. In B. Milkereit (Ed.), Exploration 07: Fifth Decennial International Conference on Mineral Exploration, (pp. 461-472).

Wong, P., Jian, F., and Taggart, L. (1995). A critical comparison of neural networks and discriminant analysis in lithofacies, porosity and permeability predictions. Journal of Petroleum Geology, 18(2), 191-206.

Chapter 4

4 Using constrained inversion of gravity and magnetic field to produce a 3D litho-prediction model

4.1 Abstract

Geologically constrained inversion of gravity and magnetic field data of the Victoria property (located in Sudbury, Canada) was undertaken in order to update the present three dimensional (3D) geological model. The initial and reference model was constructed based on geological information from approximately 965 drillholes to constrain the inversion. As well, down-hole density and magnetic susceptibility measurements of seven holes were statistically analyzed to derive lower and upper bounds on the physical properties of the lithological units in the reference model. Constrained inversion of the ground gravity and the airborne magnetic data collected at the Vitoria property were performed using UBC-GIF programs. A neural network was trained to predict lithological units from the physical properties measured in seven holes. Then, the trained network was applied on the 3D distribution of physical properties derived from the inversion models to produce a 3D litho-prediction model. Some of the features evident in the lithological model are remnants of the constraints, where the data did not demand a significant change in the model from the initial constraining model (e.g the thin pair of diabase dykes). However other changes from the initial model are evident, for example: a larger body was predicted for quartz diorite which may be related to the prospective offset dykes; a new zone was predicted as sulfide which may represent potential mineralization; and a geophysical subcategory of metabasalt was identified with high magnetic susceptibility and high density. The litho-prediction model agrees

with the geological expectation for the 3D structure at Victoria, and is consistent with the geophysical data, which results in a more holistic understanding of the subsurface lithology.

4.2 Introduction

Inversion methods are an important means to use geophysical data to explore for deep ore bodies and investigate the geology at depth. Three dimensional inversion methods estimate a three-dimensional discrete physical property distribution which can explain the observed geophysical field data. Li and Oldenburg (1996) described a method for inverting magnetic data to estimate the magnetic susceptibility and Li and Oldenburg (1998) used gravity data to estimate the density. Since this early work, a variety of 3D inversion algorithms have been developed for mineral exploration (Zeyen and Pous, 1993; Portniaguine and Zhdanov, 2002; Chasseriau and Chouteau, 2003; Dias, et al., 2009; Fergoso and Gallardo, 2009).

When there are no constraints applied to the data the estimated models look geologically unrealistic, either having the physical property variations too shallow, or varying too rapidly. Many authors incorporate prior geological information in different forms to constrain the inversion algorithms in order to derive a physical model which satisfies the expected geology and the geophysical data. Bosch and McGaughey (2001) incorporated lithological constraints into joint inversion of magnetic and gravity data. Fullagar and Pears (2007) developed a 3D inversion program which operates on a geological/petrophysical model. It uses an adaptive mesh which enables it to account for changes in the geological surfaces during the inversion. Boulanger and Chouteau (2001) used a Lagrangian formulation to combine a selection of constraints including minimum distance, flatness, smoothness and compactness. A method

introduced by Shamsipour et al. (2014) uses structural geologic constraints in potential field data inversion within a stochastic framework.

Recently, 3D inversion has become a popular tool applied to mineral exploration problems. Williams and Dipple (2007) mapped subsurface alteration using gravity and magnetic inversion models. In a different work Williams (2008) thoroughly explained how a 3D inversion algorithm can be utilized to integrate geophysical and geological data sets to produce a 3D prediction model of the subsurface. Using 3D inversion of gravity and magnetic data, Spicer et al. (2011) found the structure of a complex deformed lithological unit which hosts massive sulfide deposits. Louro and Mantovani (2012) detected anomalous zones of mineral exploration interest based on 3D inversion models of density and magnetic susceptibility. Wang et al. (2012) implemented 3D gravity and magnetic data inversion to derive a 3D geological model. Potential gold, copper, and zinc mineralization zone were proposed by Ribeiro et al. (2013) based on 3D inversion of magnetic data of grouped anomalies investigated. Three-dimensional joint inversion of gravity and magnetic data was performed by Zhou et al. (2015) to characterize the subsurface lithology.

Among the inversion methods, the UBC-GIF algorithm developed by Li and Oldenburg (1996, 1998) can obtain more realistic results by weighting the importance of specific elements depending on their depth, or by specifying that the model must vary smoothly in some way. Constrained inversion algorithms incorporate prior geological information to derive a physical model which can satisfy the expected geology and the geophysical data (Li and Oldenburg, 1996 and 1998). The inversion package used in this study is supplied by the University of British Columbia's Geophysical Inversion Facility and allows the model to be constrained by available information in a variety of forms such as i) down-hole and surface physical property measurement, geological maps, lithological logs, or three-dimensional geological models (Li and

Oldenburg, 1996, 1998; Williams, 2008). Another approach is to use physical rock properties estimated from the chemical rock composition (Boszczuk et al., 2011).

The obtained 3D inversion models are a representation of the distribution of physical property within the volume of interest. In this paper, we show how these petrophysical models can be converted to a geological model. This process is most successful when the physical characteristics of rock types show distinct changes when the lithological units change. Converting the petrophysical models to lithology requires a link between the petrophysical data and the lithological data, which can be extracted from downhole physical property measurements (Killeen, 1997; McDowell et al., 1998; McDowell et al., 2004; Vella and Emerson, 2009; Granek, 2011; Mwenifumbo and Mwenifumbo, 2012; Mahmoodi and Smith, 2015a). A classifier which simulates the quantitative relationship between geological features and physical properties can be used to identify the geology (Rabaute, et al., 2003; Garcia, et al., 2011; Granek, 2011). A neural network is a supervised multivariate classification technique which has been successfully applied to physical log data. Neural networks have been used by Baldwin et al., (1990), Wong et al. (1995), Farmer and Adams (1998), Qi and Carr (2006), Maiti et al. (2007), Maiti and Tiwari (2009, 2010) to classify physical logs and attempt to predict litho-facies. Neural networks have also been used to predict lithological units from down-hole physical measurement taken during the Ocean Drilling Program (Benaouda et al., 1999; Ojha and Maiti, 2013). Further, Konate et al. (2015) aimed to use a neural network to simulate the complex relationship between metamorphic lithological units and physical log data. The success of lithological prediction using a neural network trained from down-hole geophysical logs at the Victoria property is described by Mahmoodi and Smith (2015b). The capability of the neural networks in predicting geological

features based on physical data motivated us to apply this tool so as to convert physical properties obtained from 3D inversion to produce a 3D lithological model.

In this work, GRAV3D and MAG3D, inversion programs developed by the UBC-GIF, are used to produce density and magnetic susceptibility models from the gravity and magnetic data collected at the Victoria property. Statistical characteristics of the lithological units derived from the down-hole measurements are incorporated into a supplied geological model to constrain the range of values that the inversion can adjust the physical property within each lithological unit. Then, a neural network is trained from down-hole density and magnetic susceptibility measurements in seven holes to predict lithological units. Finally, the trained network is applied to the three-dimensional (3D) physical properties models (density and susceptibility) with the intention of deriving a 3D lithological-unit model. Some prior knowledge from analysis of the downhole data is then used to adjust the lithological classification.

4.3 Geological context

The Victoria property is located at the western portion of the south range of the Sudbury Igneous Complex (SIC); operated by KGHM the property contains a number of historic and potential future copper-nickel-platinum group element (Cu-Ni-PGE) deposits. The Victoria deposit is categorized as an “offset dyke deposit” in the Sudbury basin deposit classification scheme used by Ames and Farrow (2007). Figure 4-1 locates Victoria which is situated at the junction of the SIC and the Worthington Offset Dyke. Norite is a dominant SIC related rock along the northern portion of the property. A cyclical repetition of sedimentary sequences with felsic and mafic volcanic rocks and gabbroic intrusive form the footwall rocks. The historic Victoria Mine site

was located in a faulted structure at the intersection zone of the main mass of SIC and the SIC related offset quartz diorite dykes. The mineralization at Victoria is mainly hosted by quartz diorite and recrystallized Sudbury breccia.

Down-hole physical property measurements from seven drillholes at Victoria (four of them are plotted on the geological map in Figure 4-1) indicated there is a strong contrast in density and magnetic susceptibility between the dominant lithological units (Mahmoodi and Smith, 2015a, b). This suggests that Victoria is an interesting site for lithological and physical unit modeling using gravity and magnetic data inversion. Among the rock types at Victoria, the diabase and the sulfide zones have a distinguishable high magnetic susceptibility and density. The magnetic susceptibility of other rock types shows remarkable variation to the point that both metabasalt and metasediment can be subcategorized into non-magnetic and magnetic categories. Due to their high magnetic content, magnetic metabasalts and magnetic metasediment have relatively high density as well. Non-magnetic metasediment shows medium to low density. Non-magnetic metabasalt generally represents medium to high density and low magnetic susceptibility. The quartz diorite and norite have medium density, but the magnetic susceptibility of norite is more varied than quartz diorite. Statistical parameters of the density and susceptibility for each lithological unit are summarized in Table 4-1.

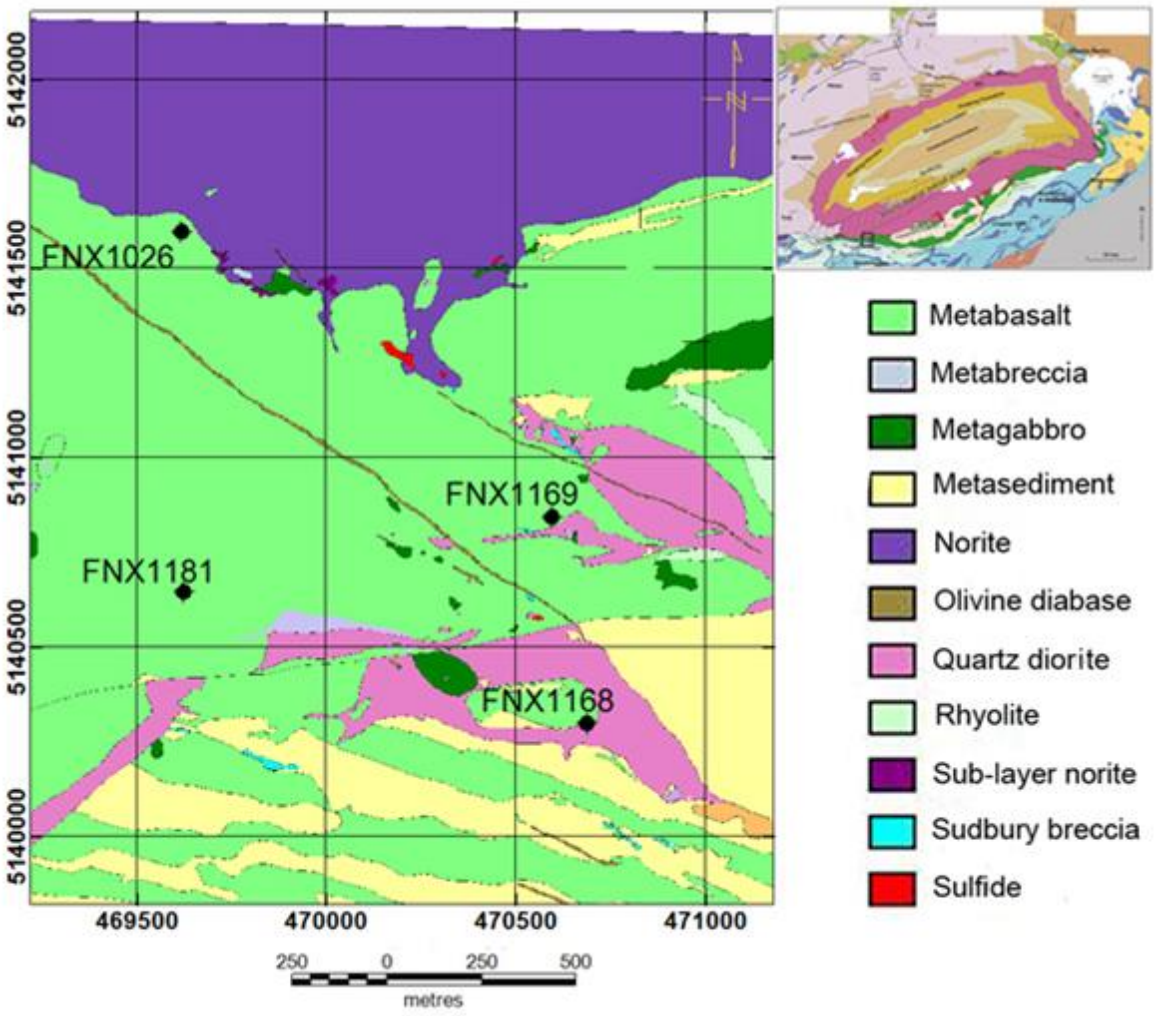


Figure 4-1- Geological map of the study area at Victoria showing the location of four boreholes, in which physical properties were measured. The study area is marked by the small black rectangle on the Sudbury Structure (top right figure). Two boreholes (FNX1182 and 700150) are located to the north, and one borehole (FNX1175) is located south beyond the area covered by this map.

4.4 Inversion

Inversion is a mathematical process which derives a distribution of physical properties to explain observed geophysical anomalies at the surface. Three dimensional inversion of gravity data

returns a 3D density contrast model, and similar inversion of magnetic data produces a 3D magnetic susceptibility model. Due to the presence of intrinsic noise in the geophysical data, reproducing the exact observed data results in a noisy model. Therefore, a proper uncertainty is attributed to the observed data to allow data to be predicted within a specified tolerance. In inversion theory, this uncertainty is ideally the standard deviation of the presumably Gaussian noise distribution in the data (Williams, 2008). In practice, the uncertainty level is set to 1 to 2% of the data range for gravity data (in mGal), and 3nT+3% to 5nT+5% of residual data range for magnetic data (in nT). These are the suggested values for a starting estimate of the noise and it is suggested that these values be adjusted during the inversion. For example, if the recovered model is excessively noisy (tightly fitted data) higher uncertainty should be used, and in a case of excessively smooth model (poorly fitted data) lower uncertainty should be used (Williams, 2008).

The non-uniqueness of the recovered model is a fundamental issue of any geophysical field data inversion. It means an infinite number of mathematically feasible but geologically-unlikely models can satisfactorily reproduce the observed data (Li and Oldenburg, 1996 and Williams, 2008). The non-uniqueness of the solution can be tackled by constraining the inversion to reproduce a solution consistent with available geological knowledge (Li and Oldenburg, 1996; Bosch, 2001; Farquharson, 2008; Williams, 2008). In the UBC-GIF software a reference model supplies the best estimate of the expected physical property distribution in the model. It allows incorporating as little or as much as geological/geophysical knowledge that is available at the time of the inversion. The reference model specifies the likely expected physical property of each cell in the model, and lower and upper bounds define the minimum and maximum value one cell is allowed to take (Williams, 2008).

The inversion algorithm developed by the UBC-GIF is based on the minimization of the model objective function and data misfit which is described as the difference between the observed data and the data predicted by the model. The model objective function (ϕ_m) is expressed as the following:

$$\phi_m(m) = \alpha_s \int_V w_s\left[w_r(z)\left(m - m_{ref}\right)\right]^2 dV + \cdots$$
$$\alpha_x \int_V w_x \left[\frac{\partial}{\partial x}\left(w_r(z)\left(m - m_{ref}\right)\right)\right]^2 dV + \cdots$$
$$\alpha_y \int_V w_y \left[\frac{\partial}{\partial y}\left(w_r(z)\left(m - m_{ref}\right)\right)\right]^2 dV + \cdots$$
$$\alpha_z \int_V w_z \left[\frac{\partial}{\partial z}\left(w_r(z)\left(m - m_{ref}\right)\right)\right]^2 dV$$

The smallness, which refers to the difference between the recovered model (m) and the reference model (m_{ref}), is measured by the first integral component. The smoothness of the difference between the recovered model and the reference model is measured by three other integrals which contain derivatives in the orthogonal *x*, *y*, and *z* directions. The alphas ($\alpha_s, \alpha_x, \alpha_y, and\ \alpha_z$) are user defined position-invariant parameters to achieve desired balance between the smallness and the smoothness in different directions throughout the model. The weighting functions $w_s, w_x, w_y,$ and w_z are position-dependent parameters which are imposed on specific cell or cell boundaries to force the recovered model to be closer to the reference model, or to control the smoothness across specific cell boundaries. By implementing the depth or distance weight function(w_r), it is possible to control the likelihood that a cell will have a physical property that has an impact on the objective function. If the depth to the j^{th} cell is z_j then one form for the depth function (w_r) is:

$$w_r^2(z_j) = \frac{1}{(z_j + z_0)^{\beta}}$$

The depth function is best matched to the fields decay with depth by adjusting the parameters z_0 and β. For gravity data β=2 is used, and $\beta = 3$ is used for magnetic data. The algorithm automatically adjusts z_0 to match the decay of the field in the center of the defined mesh. Distance weighting is recommended for surface or airborne data inversion which takes into account the actual 3D distance between the observation points and the cells (UBC-GIF MAG3D manual, 2013). For detailed information on the distance weighting see Li and Oldenberg (1996), and Williams (2008).

The data misfit between observed data (d^{obs}) and predicted data (d^{mod}) is calculated as:

$$\phi_d = \|W_d\,(d^{mod} - d^{obs})\|^2, \quad W_d = diag\left(\frac{1}{\sigma_i}\right)$$

where W_d is a diagonal matrix where the *ith* element represents the standard deviation of noise attributed to the *ith* data point. Combining both the model objective function (ϕ_m) and the data misfit (ϕ_d) functions, gives one objective function (ϕ):

$$\phi = \phi_d + \mu\phi_m,$$

where μ is the trade off parameter controlling the relative importance of the data and the model. The algorithm aims to minimize ϕ subject to the condition $\rho_{min} < \rho < \rho_{max}$ set by bounds (ρ is the recovered physical property in the inversion model, ρ_{min} and ρ_{max} are lower and upper physical property bounds). In unconstrained inversion, the software takes default models (such as uniform ground) as a reference model, an upper bound model, and a lower bound model. In constrained inversion, these models are customized to guide the inversion towards the expected

geology. The algorithm terminates once the objective function ϕ is minimized. With the assumption that the data noise is truly Gaussian and the correct data uncertainty is allowed, the target misfit (ϕ_d) equals the number of data points (Li and Oldenberg, 1996 and Williams, 2008).

Inversion using MAG3D and GRAV3D requires defining the volume of interest, which is discretized into cells, each of which represents a volume where the physical property has a constant value. Each cell is also called a voxel, and the multiplicity of cells in the volume is called a mesh. Mesh parameters, i.e. cell size and dimensions, are assigned depending on the survey parameters, depth of target, and computational cost. The cell size is strongly dependent on the resolution of the observed data and the size of the area of the dataset. One forth to one fifth of the average line spacing in an airborne survey is a reliable guide to the width of the cells. Given the similar aerial coverage and initial resolution of the observed gravity and magnetic data, inversion using the same mesh facilitates direct evaluation as each cell represents a co-located pair of density and magnetic susceptibility (Williams, 2008 and Spicer et al., 2011).

The local mesh designed for this work extends 3125 m east-west, 4175 m north-south and 1575 m vertically. The width and length of cells within the local mesh were determined to be 25 meter. Because the sensitivity decreases with depth, it is recommends that the vertical size of cells increases incrementally with depth (Williams, 2008 and Spicer et al., 2011). At the top, the height of cells is half the width of the cells (12.5 m) and the height increases by a factor of 1.05 for each successively deeper layer.

4.5 Data sets

The airborne AeroTEM survey collected over the Victoria area in 2002 along south-north lines 100 meter apart included a magnetic data set with one sample every 1.3 m. Data preprocessing including tie-line leveling, IGRF correction, and microleveling was implemented by KGHM International. Because the along-line data spacing is significantly smaller than the cell size, the very high frequency information of the data cannot be reproduced by a model with a large cell size. To tackle this issue the data were down-sampled along the lines (Williams, 2008) to one sample every 10 m. Down-sampling also enhances time and computational efficiency of the UBC-GIF software by reducing the data set size.

For gravity data inversion, ground gravity data collected along the south-north lines at Victoria in 2009 were used. The data set were preprocessed by KGHM International including free air correction, terrain correction, latitude correction, and Bouger correction. The ground gravity grid was sampled to obtain data points with the sampling distance small enough (20 m) to make sure there is at least one sample per cell. Topographic data were obtained from the ground gravity survey, and they were used for the gravity and magnetic inversion.

4.5.1 Reference model

A geological model constructed based on the information from approximately 965 drillholes was supplied by KGHM International (Figure 4-2) to use as the framework for assigning constraints: reference model, upper bound model, and lower bound model. The location of drillholes is shown in the transparent 3D model in Figure 4-2. Lithological bodies are created using interpolating the information between the drillholes and extrapolating the information outside of

the drillholes. Down-hole physical property measurements were collected from seven drillholes with 20 cm increments. The down-hole measurements were statistically analyzed to assign the bounds for constrained inversion. The reference model properties are taken as the mean of the magnetic susceptibility and density contrast for each lithological unit. The upper and lower property bounds are initially assigned to be mean ± one standard deviation of the physical property

In several early recovered models, the histograms of physical property for some geological units looked flat with spikes at either end where the values cannot go beyond the lower and upper limits, in particular for metabasalt and norite. This suggests that the bounds restrict the values unrealistically. Perhaps there is an imperfect understanding of the distribution of the physical property of these lithological units, or there is some heterogeneity not described by the reference model, or there is an inaccuracy in the geological model (Williams, 2008 and Spicer et al, 2011).

In our case, we believe that the physical property of the norite was poorly sampled since it was intersected by only one borehole. For the metabasalt we believe that there is a significant heterogeneity in the physical properties that is not described by the supplied geological model. Therefore, wider range of variation was used for those lithological units. Rather than using one standard deviation, we set the bounds close to the minimum and maximum down-hole measurements for norite and metabasalt. Modified bounds and initial values (means) are summarized in Table 4-1. The inversion algorithm uses the magnetic susceptibility specified in SI units, while the density is specified as a density contrast by subtracting an average density of 2.97 g/cm^3.

Table 4-1- Density and magnetic susceptibility values used in the reference model, upper bound model and lower bound model for each lithological unit.

	Density (g/cm^3)			Magnetic Susceptibility ($SI*10^{-3}$)		
	Initial	Lower bound	Upper bound	Initial	Lower bound	Upper bound
Sulfide	3.24	3.18	3.57	87.0	28.2	240.0
Metabasalt	3.00	2.59	3.54	5.5	0.0	231.1
Metasediment	2.83	2.57	3.08	3.3	1.1	5.6
Magnetic metasediment	2.93	2.76	3.07	14.0	0.9	120.5
Diabase	3.01	2.97	3.31	10.9	10.1	121.5
Quartz diorite	2.90	2.84	2.96	3.9	1.6	6.2
Norite	2.82	2.66	3.29	7.6	0.0	163.5

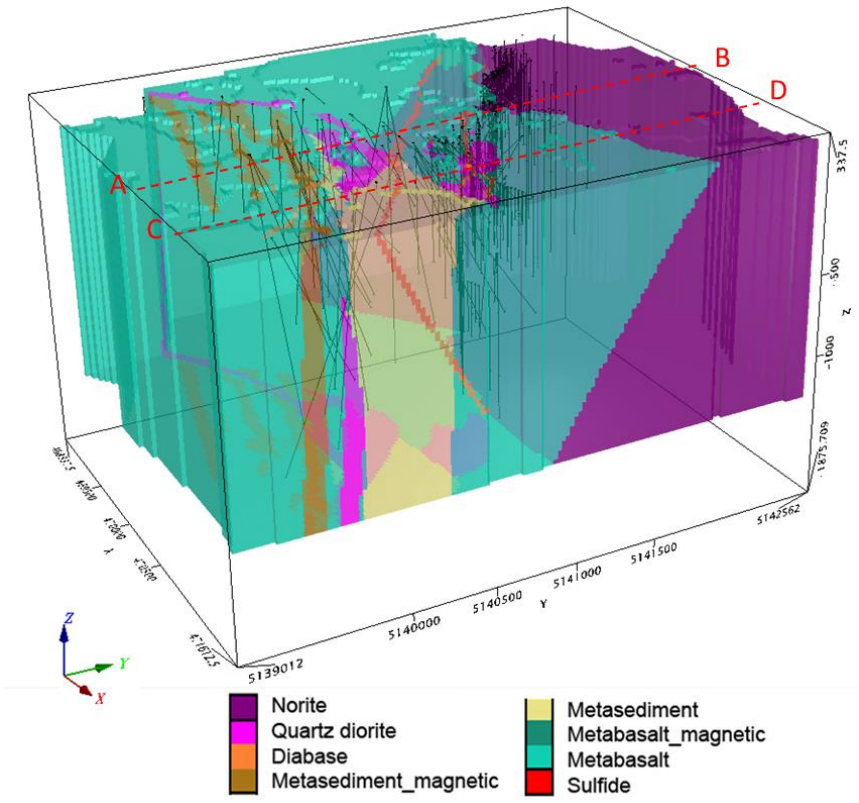


Figure 4-2- The initial geological model used to generate geological constraints for the inversion. Black lines represent traces of boreholes. Positive y-axis is north and positive x-axis is east. Two sections of the reference model along A-B and C-D are used to help interpret the inversion models.

4.5.2 Data preparation

Regional trend removal is performed to remove any field variations that cannot be explained by variations of the physical properties in the mesh (Li and Oldenburg, 1996). This is done by creating a regional mesh by adding padding cells to the local mesh. Regional cells have coarser size extending east-west, north-south, and vertically down. After the regional inversion, the mesh is modified by setting the physical property if the cells lying within the local mesh to zero. The calculated gravity and magnetic field of this modified mesh represents regional field at the location of the observed data. Subtraction of this regional field from the original data yields the residual data. The regional inversion is constrained to ensure the geophysical model is consistent with geological expectation (Williams, 2008). Upward continuation and grid subtraction can also be used to remove the regional-field contribution (Spicer et al., 2011). Finally, the residual data is leveled by subtracting the mean of the residual data, so that the mean of the residual data will be zero (Williams, 2008). The leveled residual data used for inversion are shown in Figure 4-3. The range of the residual magnetic data was 480 nT, which means that the uncertainty in the data is initially set to 18 nT (3nT+3% of 480 nT). The uncertainty attributed to gravity data was 0.037 mgal (1% of the residual data range).

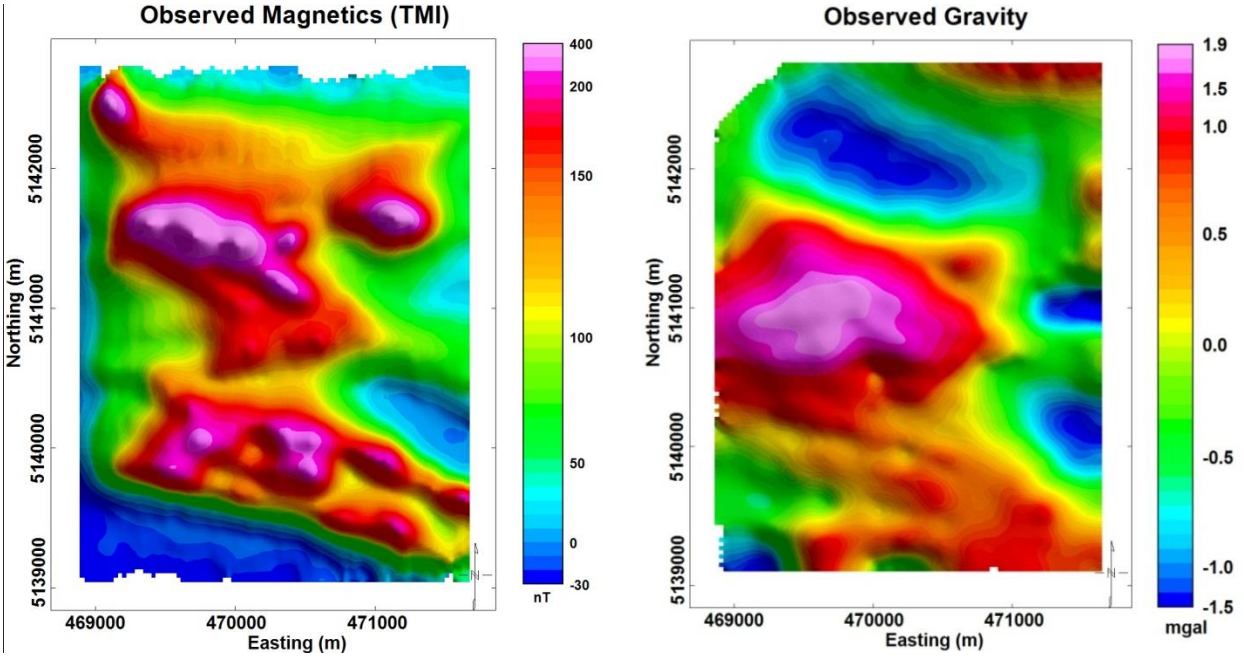


Figure 4-3- Leveled observed magnetic and gravity data input for inversion.

4.6 Results

To understand the degree of consistency between the initial geological model and the observed data, forward model of the reference model was calculated. Discrepancies between forward models of the reference model shown in Figure 4-4 and observed data are the result of a potential field signal in the measured data that is not present in the reference model. So, the initial model needs to be modified through constrained inversion to reproduce the observed data within some specified tolerance.

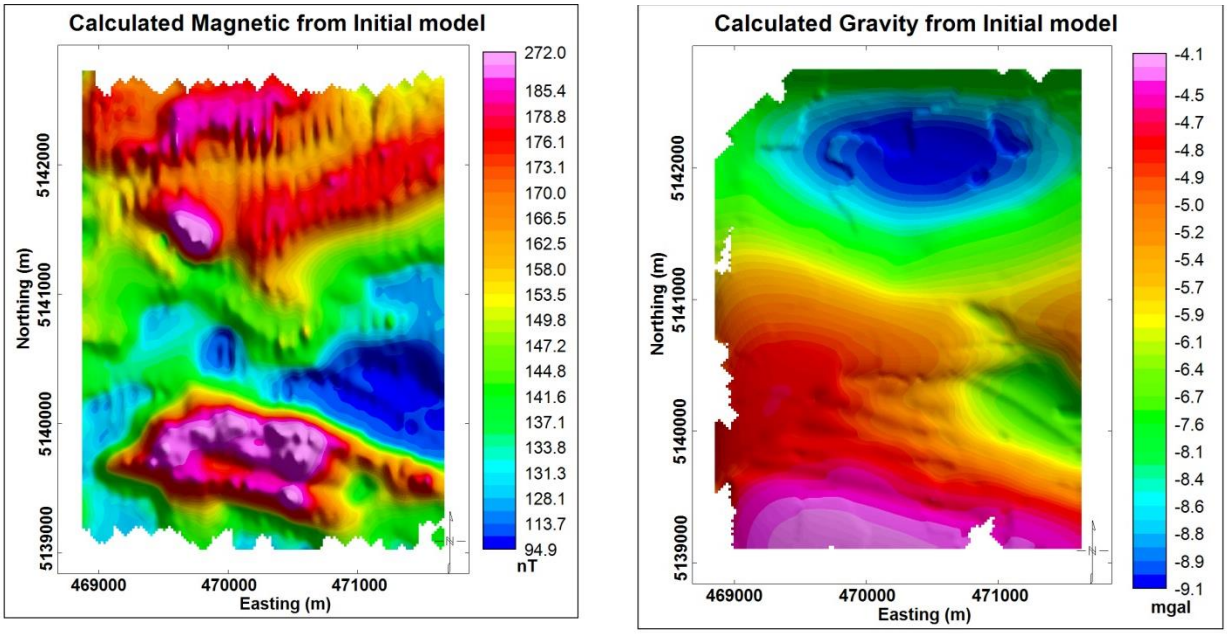


Figure 4-4- Forward models of initial model for density and magnetic susceptibility

For both gravity and magnetic constrained inversion, default values defined by the UBC-GIF facility for the alphas and weight functions are used. The predicted data calculated by forward modeling of the 3D inversion models are shown in Figure 4-5.

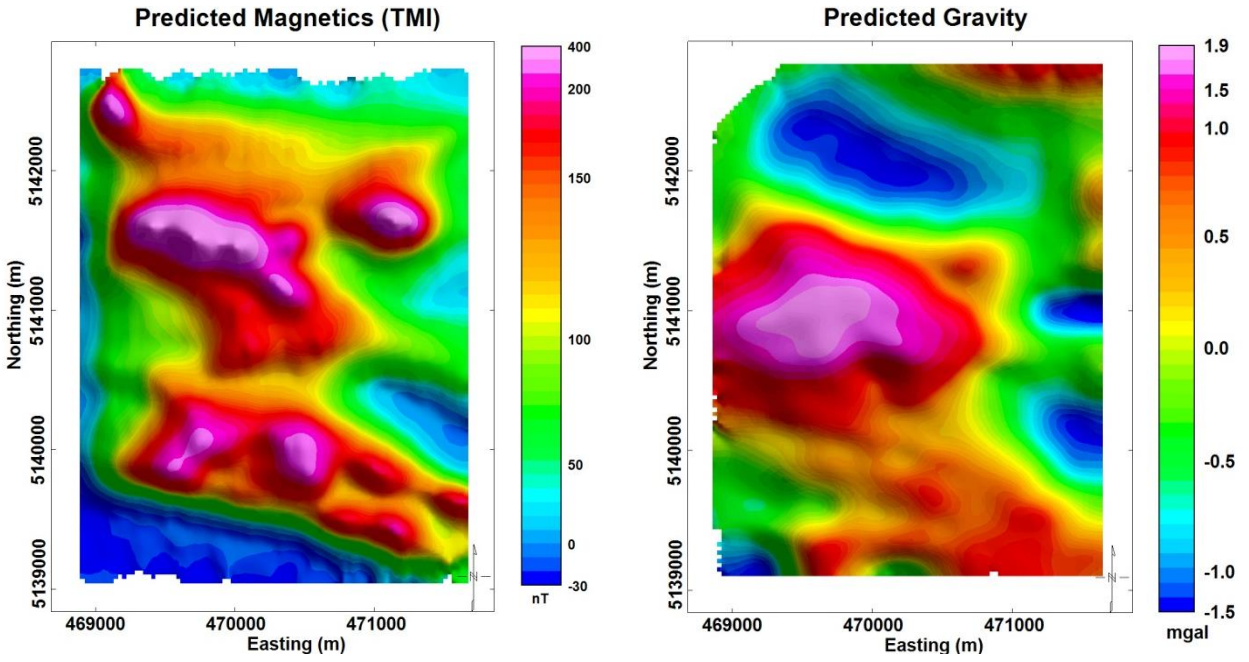


Figure 4-5- The calculated data from the magnetic susceptibility and density inversion models.

4.6.1 Magnetic susceptibility distribution

A 3D magnetic susceptibility model is shown in Figure 4-6. For detailed evaluation, an A-B section across the inversion model and the geological reference model is shown in Figure 4-7. (To locate the section across the reference model refer to Figure 4-2). The reference model and constraining bounds guide the inversion to enhance the magnetic susceptibility contrast of different lithological units; otherwise, the resultant model from the unconstrained inversion lacks structures and details of the geology. A comparison between vertical and horizontal sections of the unconstrained and the constrained models, shown in Figure 4-8, indicates the significance of using constraints to recover details and structures in the inverted model. The general geology and geometry of the reference model lithological units are preserved in the magnetic susceptibility

model. If the susceptibility model deviates from the initial geological model, then this indicates a subsurface lithological feature predicted by the magnetic data.

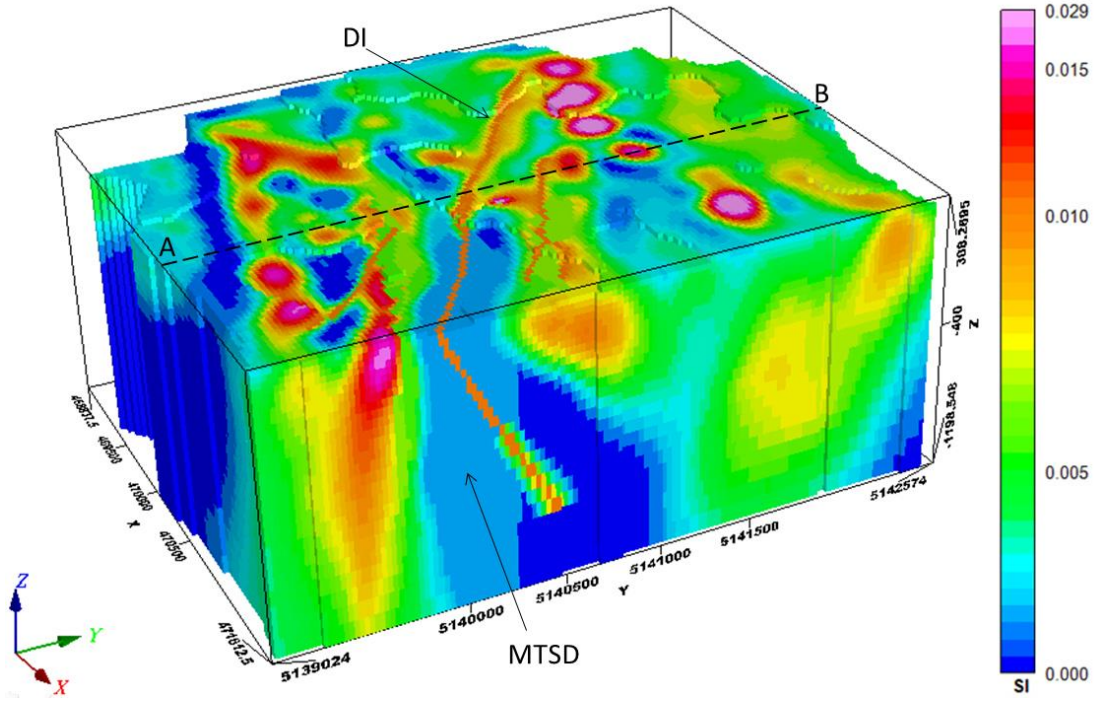


Figure 4-6- A perspective view of magnetic susceptibility model. Positive y-axis is north and positive x-axis is east. Dashed line (A-B) represents a south-north section line across the inversion model. DI shows the recovered diabase dyke, and MTSD points to the low magnetic susceptibility zone corresponding to non-magnetic metasediment.

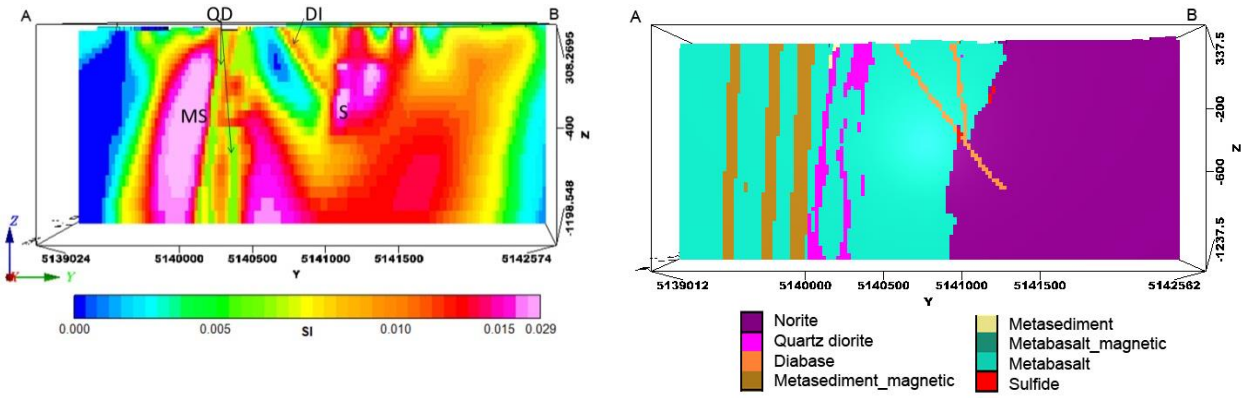


Figure 4-7- The A-B section across magnetic susceptibility model (left) and initial geological model (right). Marks on the magnetic susceptibility section describe: medium magnetic susceptibility of quartz diorite (QD); high magnetic susceptibility of sulfide (S), and recovered diabase dyke (DI).

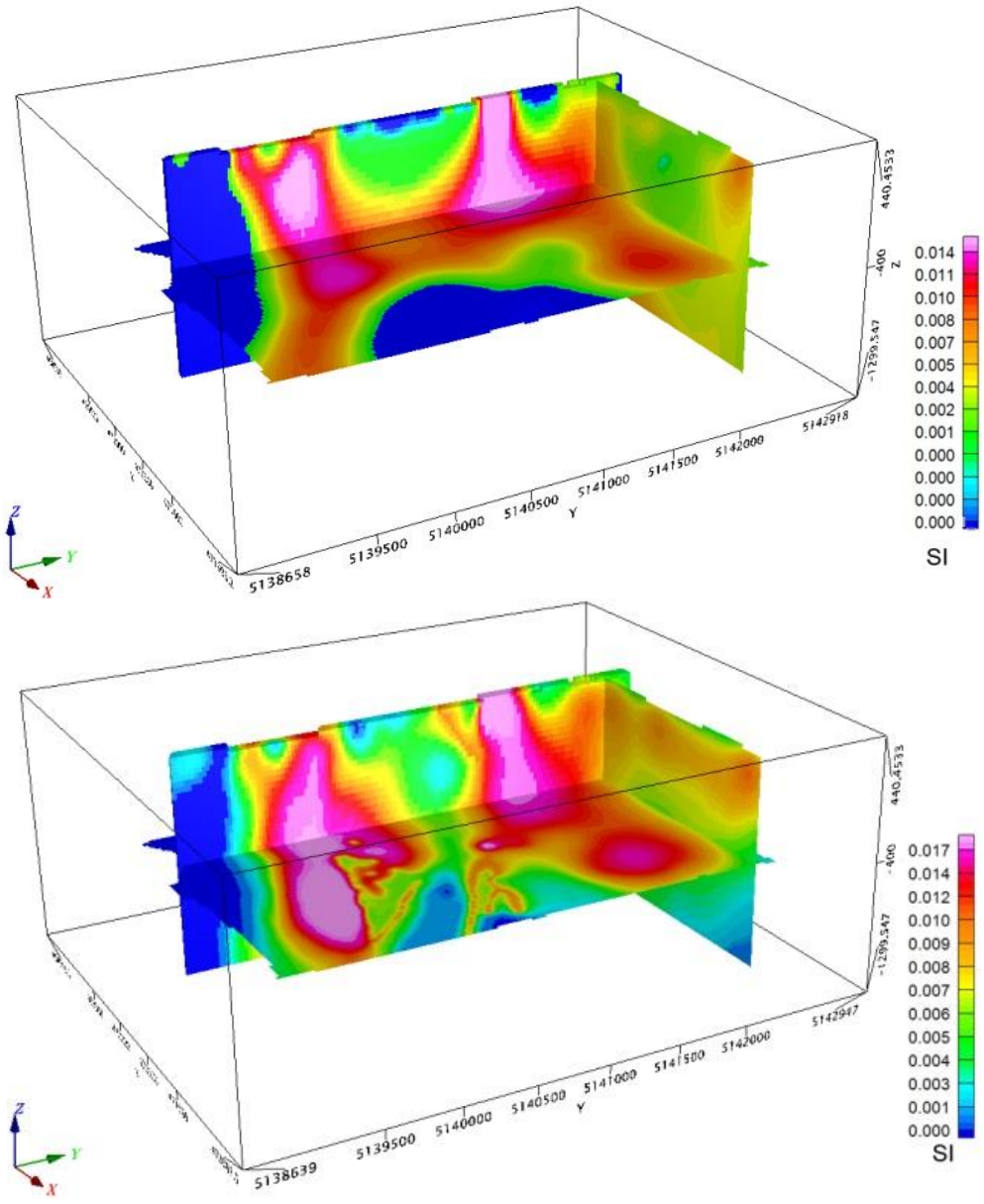


Figure 4-8- Vertical and horizontal sections of unconstrained (top) and constrained (bottom) inversion of magnetic data.

The narrow diabase dyke with high magnetic susceptibility in the top center of the section dipping towards B is clearly recovered in the inversion model marked by DI in Figures 4-6 and 4-7. However, this unit is a remnant of the initial and reference model and its presence probably indicates there is no information in the data to suggest that anything different is present in the model. This is an aspect of constrained geological inversion: the recovery of a feature consistent with the initial model does not imply that that feature is suggested by the data. This aspect can be a strength or a weakness of constrained inversion. Another feature with strong magnetic susceptibilities is the sulfide zones denoted by S in the section (Figure 4-7) near where the diabase dike intersects the Norite boundary. The section illustrates they are clearly recovered at depth, but larger in size compared with the sulfide zones in the reference model. As shown in the 3D model (Figure 4-2) and initial geological model section (Figure 4-7), the center and south of the model are dominated by metabasalt, but the inversion model shows a wide range of magnetic susceptibilities in these regions indicating a strong magnetic heterogeneity of metabasalt. Such a magnetic contrast in metabasalt is a good argument for introducing a geophysical sub-category: magnetic metabasalt which was also suggested by Perron et al. (2011). The north of the model is dominated by norite. It shows a variation in the inversion model, although its variation is not as strong and as broad as is evident with metabasalt. As expected, quartz diorite which is the most homogenous rock type at Victoria shows medium magnetic susceptibility with minor variation (marked by QD in the section). Magnetic metasediment is recovered in the inversion model as a high magnetic zone denoted by MS in the section (Figure 4-7). In the inversion model (Figure 4-6) metasediment corresponds to the low magnetic zone marked by MTSD at the east side of the model.

4.6.2 Density distribution

The gravity inversion model is represented as a 3D density distribution in Figure 4-9. A C-D section across the inversion model and the geological reference model is shown in Figure 4-10. The density model is less variable than the magnetic susceptibility model. Because of the application of constraints, the general features of the geological model are preserved in the inversion model. Vertical and horizontal sections of the unconstrained and the constrained models shown in Figure 4-11 indicate the significance of using constrains to recover details and structures in the inverted model.

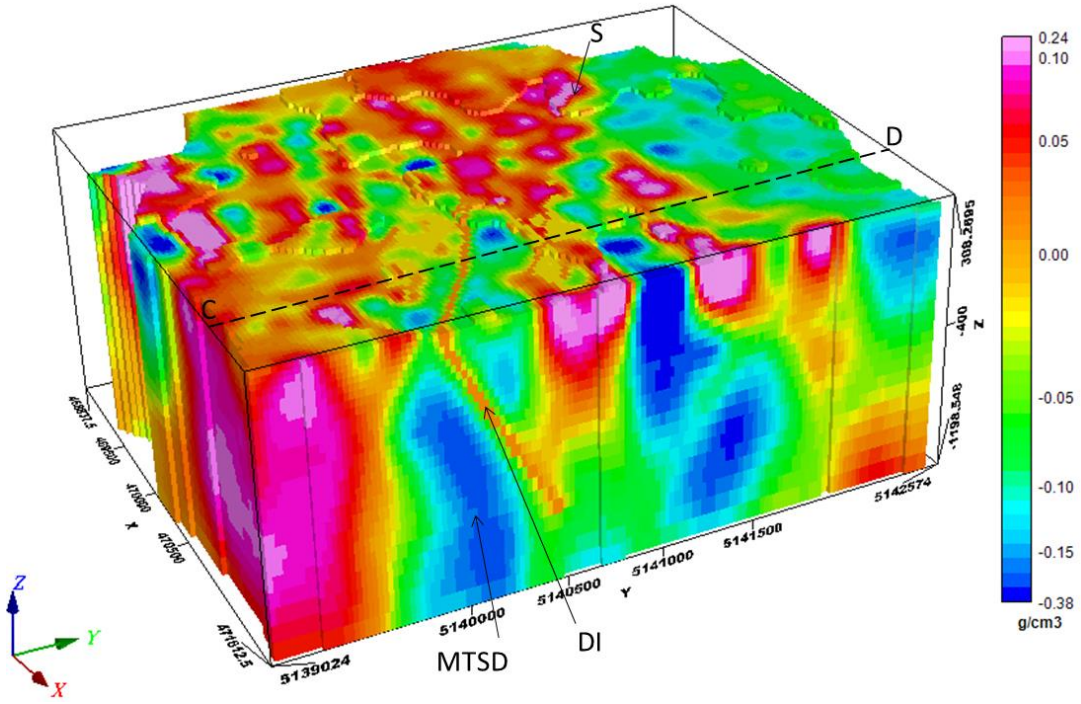


Figure 4-9- A perspective view of the density model. Positive y-axis is north and positive x-axis is east. Dashed line (C-D) represents a south-north section line across the inversion model. DI shows the recovered diabase dyke, MTSD points to the low density zone corresponding to non-magnetic metasediment, and S indicates the high density of sulfide in the inversion model.

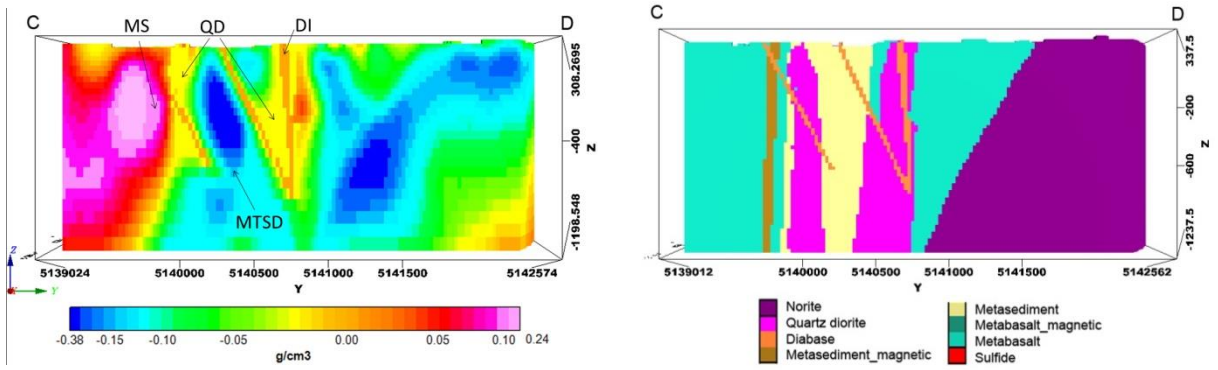
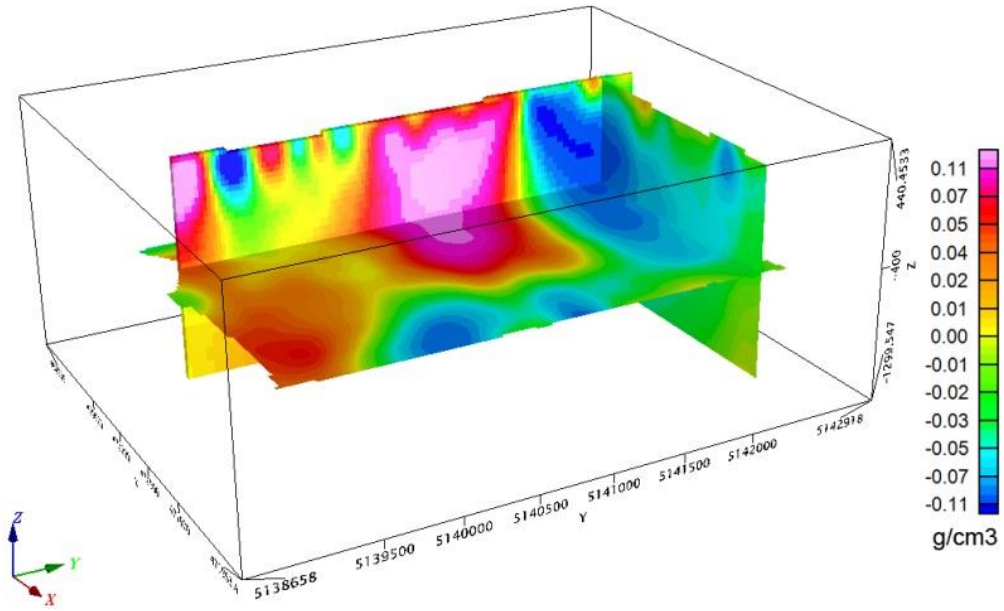


Figure 4-10- The C-D section across the density contrast model (left) and initial geological model (right). Marks on the density section indicate: medium density of quartz diorite (QD); high density of magnetic metasediment (MS), recovered diabase dyke (DI), and low density corresponding to metasediment (MTSD).



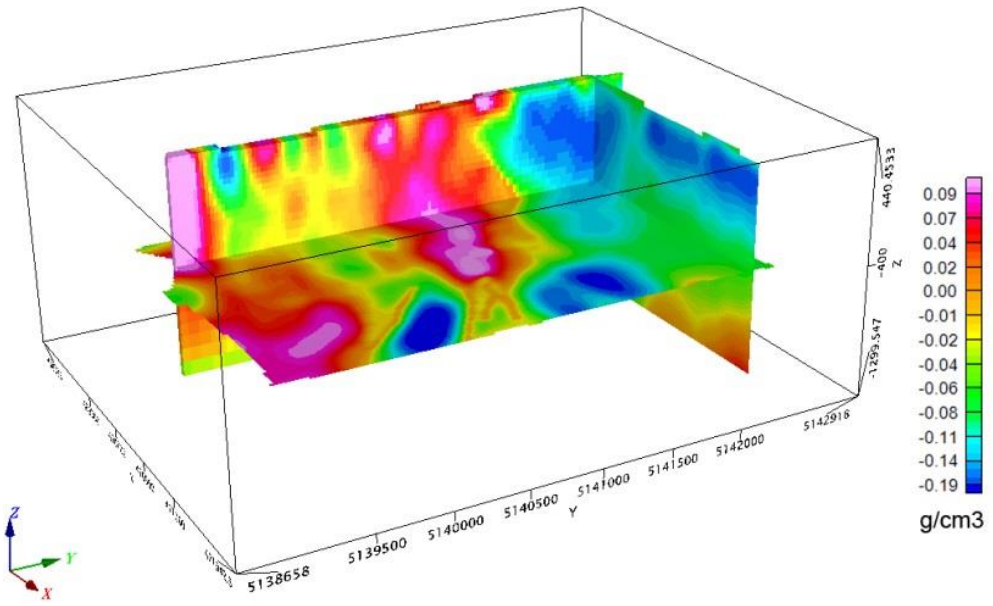


Figure 4-11- Vertical and horizontal sections of unconstrained (top) and constrained (bottom) inversion of gravity data.

The C-D sections across the inversion model and initial geological model are presented in Figure 4-10. In this discussion we focus on section C-D rather than A-B, as the former section contains metasediment which shows variability in density from low for non-magnetic metasediment to high for magnetic metasediment. The sections indicate that constraints helped to retain the general geology and geometry of the lithological units, although there are a number of differences. Features retained are the high density diabase dykes clearly observed in the inversion model (arrow labelled DI in Figures 4-9 and 4-10). As with the magnetic data, these features are unlikely to be indicated by the data, but are a remnant of the geological constraint. A high density zone to the north of the diabase dyke denoted by S in the inversion model (Figure 4-9) is interpreted to be a sulfide zone. High density in the metasediment zone is evident in the section which is marked by MS in Figure 4-10 and this is interpreted to be magnetic

metasediment. The metabasalt in the southern part of the section shows a heterogeneity, with a purple higher density zone being evident within the metabasalt. Quartz diorite corresponds to a medium density zone in the inversion model marked by QD in the section (Figure 4-10). The density of norite shows a smooth variation, from low density in the south part of the norite, to higher density to the north (at depth). Metasediment is recovered by low density zones marked by MTSD in Figures 4-9 and 4-10.

4.6.3 Classification (neural network)

The models derived from the inversion are numerical representations of the subsurface physical property distribution, consistent with the expected geology of the Victoria property. In general, these physical property distributions can be analyzed to derive meaningful patterns of variation related to lithology, mineralization, or alteration. In this work, our intent is to obtain a litho-prediction model using a neural network trained from down-hole density and magnetic susceptibility measurements. Classifying the inverted physical property distribution from the training of down-hole measurements assumes that the distributions would be the same. Although, the inversion tends to smooth the model, Figure 4-12 illustrates the distribution of histograms of magnetic susceptibility and density values measured in the drillholes and derived from the physical models along drillholes.

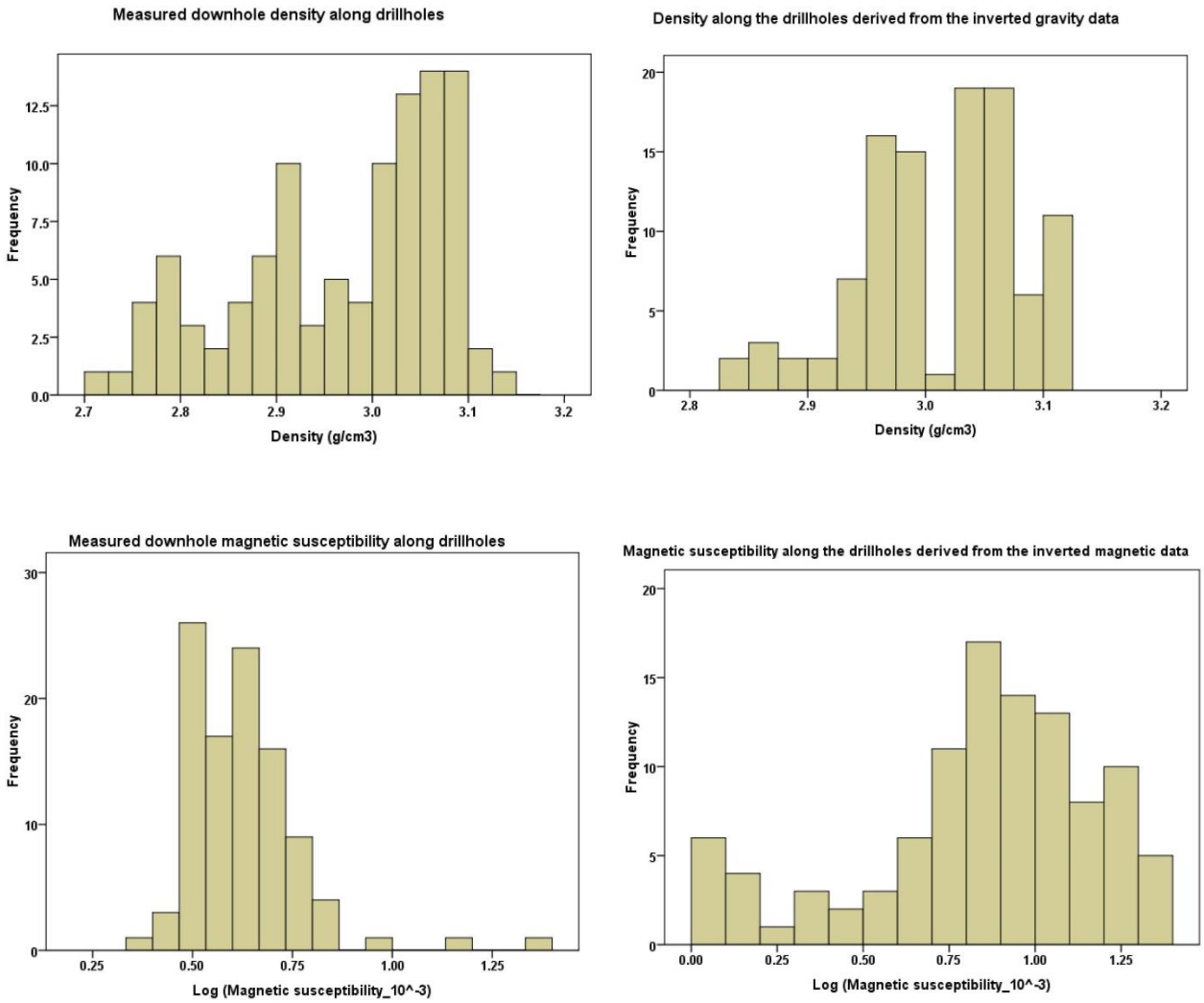


Figure 4-12- Histograms of density and magnetic susceptibility measured in the drillholes and derived from the inverted models along drillholes

We used the same neural network procedure outlined by Mahmoodi and Smith (2015b), except we only used the density and magnetic susceptibility. During the training process, a 2-11-8 network was selected as an optimal structure with a predictive accuracy of 72%. In the 2-11-8 network, the 2 represents the number of neurons in the input layer (number of variables), the 11

is the number of hidden layers, and the 8 is the output layer (number of rock types). The trained network takes two variables; i.e. magnetic susceptibility and density of each cell, and provides the rock type probabilities for that cell. Then the highest probability is picked as the predicted lithological unit. This part of the work was performed using the MATLAB neural network toolbox.

4.6.4 Lithological unit model

One crucial concern which must be considered in litho-type prediction is the physical property overlap between two or more rock types. Our previous study (Mahmoodi and Smith, 2015b) revealed the possibility of misclassification of rock types with similar physical property. For example, at Victoria, diabase and sulfide have similar physical properties, and are mutually misclassified as each other. In the case of misclassification, spatial attributes and geometry of the rocks in the initial geological model helps to determine the correct classified rock. The primary output of the neural network classified the main part of the diabase as sulfide. When plotted in 3D space, the geometry of the sulfide body and its correlation with diabase in the initial model meant that this zone could be confidently reclassified as diabase. Likewise, similar physical properties of quartz diorite and norite cause misclassification. Once again, the spatial separation of these two lithologies (norite is observed only at the north part of the model, and quartz diorite is present at the center and the south), meant that the classifications could be changed with confidence. Therefore, the norite predicted by the neural network at the center and south part of the model was represented as quartz diorite in the litho-prediction model. The resultant litho-prediction model, shown in Figure 4-13, was in a good correlation (89%) with the expected geology of the Victoria property, but also shows a number of discrepancies due to a) the

inversion modifying the initial geological model, b) limitations in the resolving power of the inversion, or c) classification errors or ambiguities.

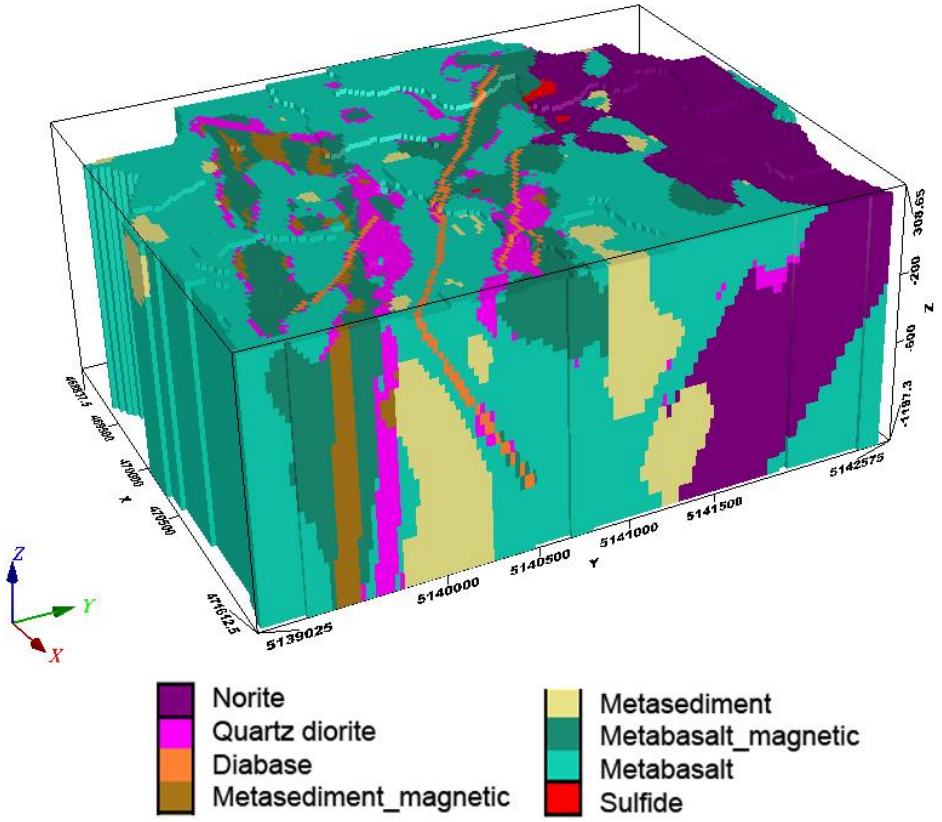


Figure 4-13- A perspective view of the litho-prediction model. Positive y-axis is north and positive x-axis is east.

To better visualize and compare the predicted rocks with the initial geological model, they are subdivided into regions in Figures 4-14 and 4-15 to show their spatial correlation. Table 4-2 summarizes the spatial correlation between prediction model and initial geological model.

Table 4-2- Confusion matrix summarizing the correlation of litho-prediction model and initial geological model. The number is the total number of cells in the mesh as classified in the litho-prediction model.

		Litho-prediction model							
		Meta-basalt	Magnetic meta-sediment	Diabase	Quartz diorite	Meta -sediment	Norite	Sulfide	Magnetic meta-basalt
Initial geological model	Meta-basalt	85915 (46.6%)	-	-	3159 (1.7%)	5107 (2.8%)	-	-	90352 (49.0%)
	Magnetic meta-sediment	4119 (20.4%)	14276 (70.8%)	-	1645 (8.2%)	138 (0.7%)	-	-	-
	Diabase	-	-	5113 (100%)	-	-	-	-	-
	Quartz diorite	2284 (12.7%)	839 (4.7%)	-	14428 (80.5%)	382 (2.1%)	-	-	-
	Meta-sediment	3621 (25.1%)	619 (4.3%)	-	3107 (21.5%)	7083 (49.1%)	-	-	-
	Norite	7487 (6.5%)	-	-	240 (0.2%)	3753 (3.2%)	103579 (89.6%)	557 (0.5%)	-
	Sulfide	-	-	-	-	-	-	122 (100%)	-

Figure 4-14 includes metasediment, magnetic metasediment, diabase, and sulfide zones. Diabase is recovered in the prediction model due to the geological constraint. The geometry of these dykes helped to differentiate diabase from the sulfide zone. The sulfide zone has distinguishable physical properties which makes it easily differentiable from most of the surrounding rocks. The prediction result confirms the existence of pre-known sulfide zones (albeit slightly larger) and detects two new zones (marked in Figure 4-14) that have been interpreted as sulfide. However, these could also be diabase, so further follow-up is required to confirm these zones. Nearly half of the metasediment in the predicted model correlates with what is classified as metasediment in the initial geological model. However, the predicted model identifies a new body of metasediment (marked in Figure 4-14) which corresponds to low density and low magnetic susceptibility in the inversion models. In this location, the initial geological model is not

consistent with the two geophysical models. As Table 4-2 indicates magnetic metasediment is mainly (70%) correlated with this unit in the initial model. This correlation is observable in Figure 4-14 where they have similar spatial distribution.

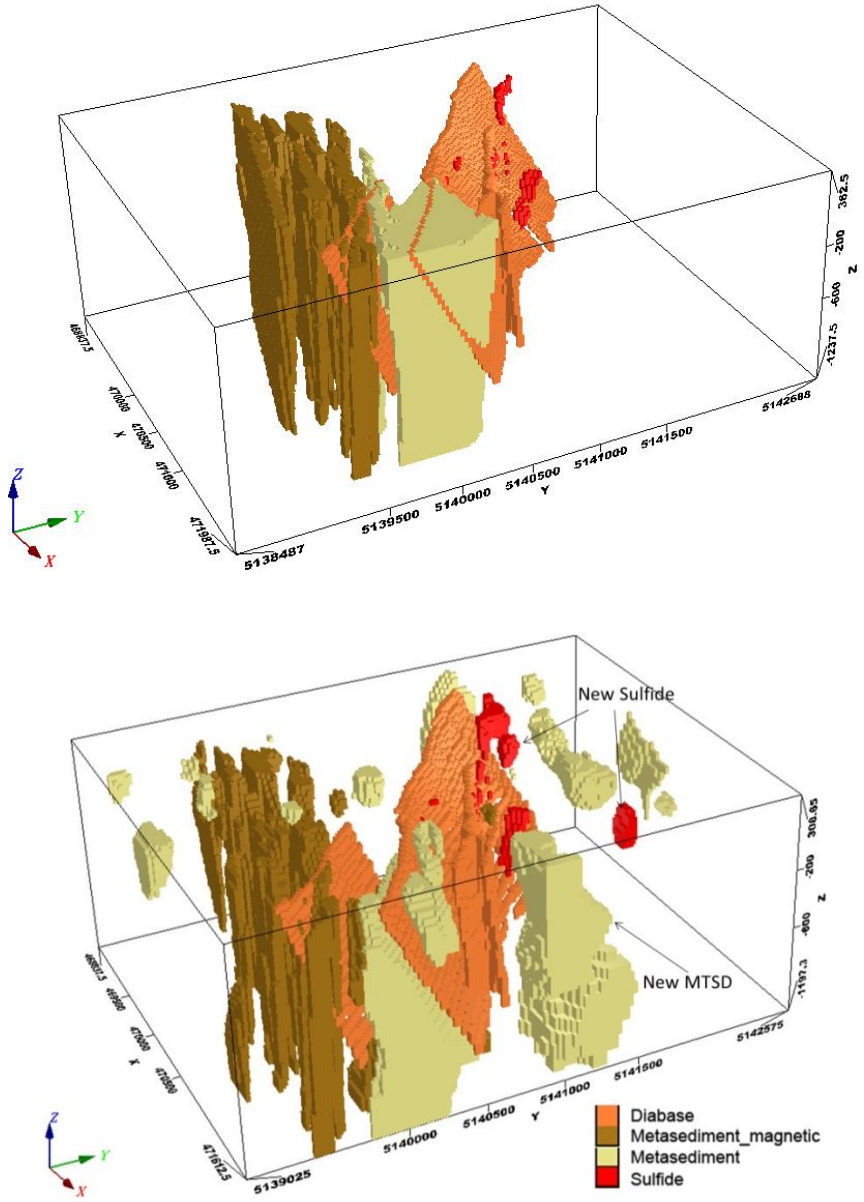
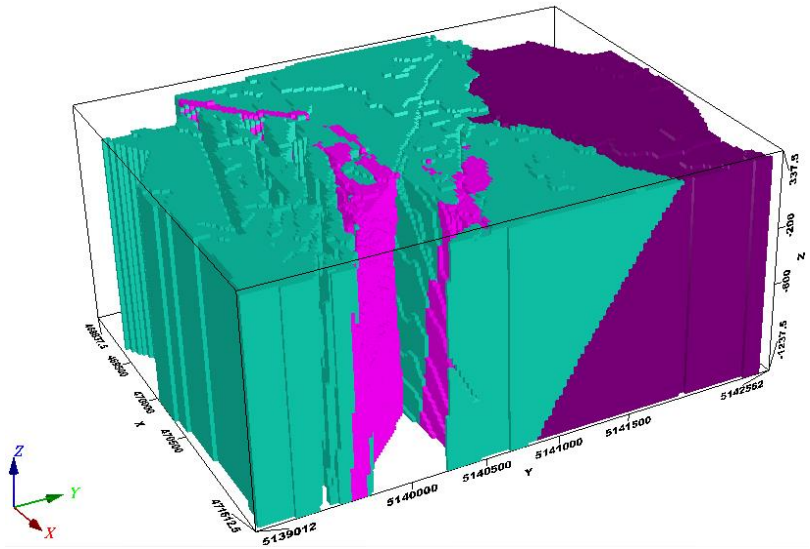


Figure 4-14- Metasediment, magnetic metasediment, diabase, and sulfide zones in the litho-prediction model (bottom) and initial geological model (top). Positive y-axis is north and positive x-axis is east. New predicted sulfide zones and a new metasediment zone are marked on the litho-prediction model.

Table 4-2 reveals that all cells predicted as norite were also defined as norite in the initial geological model, so the general geometry of norite is preserved (Figure 4-15). However, some peripheral parts of the norite are predicted as metabasalt, quartz diorite and metasediment which results in updating the location of the boundary of the geological model to be consistent with the physical property data. Quartz diorite in the predicted model occupies larger space than the initial model (Figure 4-15). About 80% of the predicted quartz diorite was quartz diorite in the initial model. The remaining 20% was classified as medium density and low magnetic portions of metabasalt, metasediment, and magnetic metasediment. Metabasalt is the most dominant rock type at the Victoria which we have subdivided based on variations in the magnetic susceptibility. Table 4-2 indicated that the metabasalt in the initial model is equally divided into metabasalt and magnetic metabasalt. The magnetic part of metabasalt is represented by dark green in Figure 4-15.



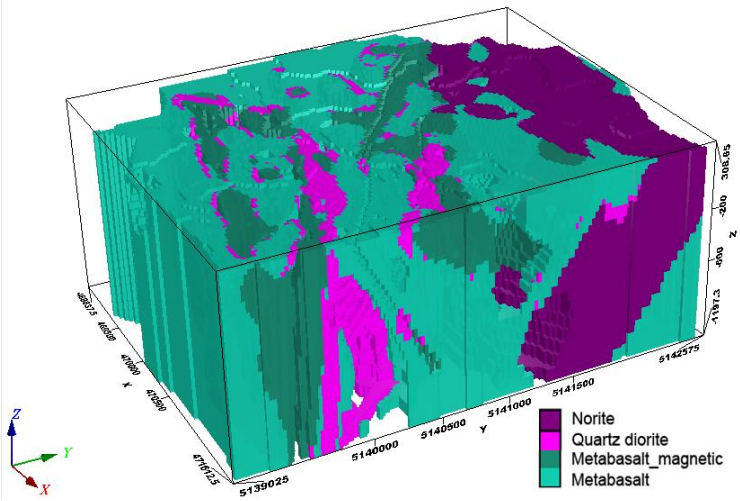


Figure 4-15- Metabasalt, magnetic metabasalt, norite and quartz diorite in initial geological model (top) and litho-prediction model (bottom). Positive Y-axis is north and positive x-axis is east.

4.7 Discussion

Many factors are involved in an inversion problem and control the output. Data uncertainty which determines the level of misfit should be specified wisely to balance between an excessively smooth model and an excessively noisy model. The range of observed data can suggest the initial uncertainty value, and further evaluation of each inversion model can guide the user towards more appropriate values. The uncertainty level suggested by Williams (2008) for gravity data (1% of the range of the residual data) was well suited to our data set and we found no further adjustment was necessary. However, for magnetic data the suggested value (5% of the range of the residual data plus 5nT) resulted in an unreasonably smooth inversion model and high data misfit. So, the level of uncertainty was decreased to 3% of the range of data

plus 3 nT. In geophysical inversion, preprocessing such as regional trend removal is necessary to yield an acceptable result otherwise the physical properties might be distributed unrealistically to reproduce the observed data. The existence of zones of extreme values in the corners and bottom of the mesh indicates inaccurate regional trend removal or improper mesh size (Williams, 2008).

Having a clear understanding of the physical characteristics of the rock type is critical to establish a reliable link between geology and geophysics. While using bounds to constrain the inversion, statistical characterization of the lithological units should be well understood. The upper and lower bounds are critical parameters as they determine the range to which physical properties assigned to the cells can deviate from the value in the reference model if the reference model is not consistent with the geophysical data. Physical heterogeneity of the geological units should be well understood and incorporated into the constraints. If heterogeneity is not taken into account, the standard deviations extracted directly from statistical analyses might not deliver realistic estimates of the upper and lower bounds, therefore further modification and adjustment of the bounds may be necessary. For example, at Victoria norite was intersected by only one drillhole, and insufficient measurements of the physical properties of norite meant that its physical characteristics were poorly defined. This impacted the inversion generating a distribution of physical properties for norite that were unrealistic, with most of the samples assigned the upper or the lower bound. Physical heterogeniety of metabasalt which was not present in the initial model also resulted in difficulties when assigning proper bounds. We found that reassigning the upper and lower bounds to be the maximum and minimum values measured for the physical property of norite and metabasalt, it was possible to obtain a distribution which looked more normal.

Unconstrained inversion commonly tends to produce smooth, structureless models in which only large scale geological features are represented. To recover structures and small scale features, prior information in the form of geological constraints is required. We found that sharp boundaries of petrophysically distinct units such as the diabase dyke were recovered only because they are constrained by the reference model. If they were not constrained, the physical properties could be adjusted to the point that they characterize another lithological unit, which is different from the initial geological model.

Using supervised classification techniques to predict rock type is a powerful tool, but care should be taken while interpreting the results. Rock types with similar physical property are more likely to be misclassified as each other. We were aware of this ambiguity from our comprehensive statistical analyses of down-hole physical property measurement for each rock type (Mahmoodi and Smith, 2015a, b). The impact of overlap of physical properties creating ambiguity can be resolved by involving geological expectation for the specific area. The expected location and geometry of the geological body can help to differentiate these rocks in the final predicted model. For example, diabase dykes with unique geometry at the center of the model was initially misclassified as sulfide due to their similar physical properties. Our prior experience and the geological model were invoked to justify modifying the derived lithological model and hence get a better understanding of the subsurface geology.

4.8 Conclusion

Physical characteristics of lithological units were incorporated into the 3D geological model to constrain the inversion of ground gravity and airborne magnetic data collected at Victoria. The

UBC-GIF software allows for assigning an initial model, a reference model, lower bound model, and upper bound model. Using constraints enables the inversion to recover magnetic susceptibility and density models that are consistent with the geophysical data and expected geology.

The standard use of inversion results is to look for anomalous features in these physical properties model. An alternate approach is to use these magnetic and density distributions to classify the data based on proper criteria to derive lithological, alteration or mineralization model.

At Victoria, a large database of drillholes was used to train a supervised classification of the physical properties data with the intent of updating the present geological model. The resultant litho-predicted model satisfied the expected geology, and also added some features to the geological model to be consistent with the geophysical data. The main differences from the initial geological model are: the prediction model suggests a more extensive body of quartz diorite which is more likely related to the offset dykes and considered as a zone of interest for exploration, as offset dykes contain sulfides. Also and new larger sulfide zone was identified, which might indicate a mineralization zone. Due to an ambiguity in the physical properties, this sulfide could also be a diabase; however, diabase normally occurs in sheets and the new zone is compact like a sulfide deposit. Other updates made by the litho-prediction model are: a new body of metasediment at the center of the volume of investigation; a new magnetic subcategory of metabaslt; and new locations for boundaries that are consistent with the geophysical model. Heterogeneity of some rock types became apparent, specifically metabasalt, suggested by Perron et al. (2011) but not initially understood in the reference model.

4.9 References

Ames, D., and Farrow, C. (2007). Metallogeny of the Sudbury mining camp, Ontario, in Goodfellow W.D., ed., Mineral deposits of Canada: A synthesis of major deposit-types, district metallogeny, the evolution of geological provinces, and exploration methods. Geological Association of Canada, Mineral Deposits Division, Special Publication no. 5, 329–350.

Baldwin, J., Bateman, A., and Wheatley, C. (1990). Application of neural network to the problem of mineral identification from well logs. The Log Analyst, 31, 279–293.

Benaouda, D., Wadge, G., Whitmarsh, R. B., Rothwell, R., and MacLeod, C. (1999). Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the Ocean Drilling Program. Geophysical Journal International , 136, 477-491.

Bosch, M. M. (2001). Joint inversion of gravity and magnetic data under lithologic constraints. The Leading Edge, 877–881.

Boszczuk, P., Cheng, L. Z., Hammouche, H., Roy, P., Lacroix, S., and Cheilletz, A. (2011). A 3D gravity data interpretation of the Matagami mining camp, Abitibi Subprovince, Superior Province, Québec, Canada. Application to VMS deposit exploration. Journal of Applied Geophysics, 75, 77–86.

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

Carlos, D. U., Uieda, L., and Barbosa, V. C. (2014). Imaging iron ore from the Quadrilátero Ferrífero (Brazil) using geophysical inversion and drillhole data. Ore Geology Reviews, 61, 268–285.

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

Dias, F., Barbosa, V., and Silva, J. (2009). 3D gravity inversion through an adaptive-learning procedure. Geophysics, 74(3), I9-I21.

Farmer, R., and Adams, S. J. (1998). Facies recognition using neural networks. New Zealand Petroleum Conference. Queenstown.

Fergoso, E., and Gallardo, L. (2009). Cross-gradients joint 3D inversion with applications to gravity and magnetic data. Geophysics, 74, L31–L42.

Fullagar, P. K., and Pears, G. A. (2007). Towards geologically realistic inversion. In B. Milkereit (Ed.), In "Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration", (pp. 444-460).

Granek, J. (2011). Computing geologically consistent models from geophysical data. BC, Canada: University of British Coloumbia.

Killeen, P. (1997). Borehole geophysics: Exploring the third dimension. In A. Gubins (Ed.), Exploration 97: Fourth Decennial International Conference on Mineral Exploration, (pp. 31–42).

Konaté, A. A., Pan, H., Fang, S., Asim, S., Ziggah, Y. Y., Deng, C., et al. (2015). Capability of self-organizing map neural network in geophysical log data classification: Case study from the CCSD-MH. Journal of Applied Geophysics, 118, 37-46.

Li, Y., and Oldenburg, D. (1996). 3-D inversion of magnetic data. Geophysics, 61, 394–408.

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

Louro, V.H.A. and Mantovani, M.S.M. (2012) 3D Inversion and modeling of magnetic and gravimetric data characterizing the geophysical anomaly source in Pratinha I in the southeast of Brazil. Journal of Applied Geophysics, 80, 110-120. http://dx.doi.org/10.1016/j.jappgeo.2012.01.013

Mahmoodi, O., and Smith, R. (2015a). Clustering of downhole physical property measurements at the Victoria property, Sudbury for the purpose of extracting lithological information. Journal of Applied Geophysics, 118, 145-154.

Mahmoodi, O., and Smith, R. (2015b). Supervised classification of down-hole physical properties measurements using neural network to predict the lithology. Submitted to the Journal of Applied Geophysics

Maiti, S., and Tiwari, R. (2009). A hybrid Monte Carlo method based artificial neural networks approach for rock boundaries identification: a case study from KTB Borehole. Pure and Applied Geophysics, 166, 2059–2090.

Maiti, S., and Tiwari, R. (2010). Neural network modelling and an uncertainty analysis in Bayesian framework: a case study from the KTB borehole site. Journal of Geophysical Research: Solid Earth.

Maiti, S., Tiwari, R. K., and Kumpel, H. J. (2007). Neural network modelling and classification of lithofacies using well log data: a case study from KTB borehole site. Geophysics Journal International, 169, 733–746.

McDowell, G. M., King, A., Lewis, R. E., Clayton, E. A., and Grau, J. A. (1988). In-situ nickel assay by prompt gamma neutron activation wireline logging. SEG Annual Meeting. New Orleans, Louisiana.

McDowell, G., Fenlon, K., and King, A. (2004). Conductivity-based nickel grade estimation for grade control at Inco‟s Sudbury mines. SEG Annual Meeting. Denver, Colorado.

Mwenifumbo, C., and Mwenifumbo, A. (2012). Borehole geophysical logging in the Flin Flon Mining Camp. Geological survey of Canada. Geological survey of Canada, 75.

Ojha, M., and Maiti, S. (2013). Sediment classification using neural networks: An example from the site-U1344A of IODP Expedition 323 in the Bering Sea. Deep-Sea Research II.

Perron, G., Fullagar, P., Pears, G., Phillips, N., Williston, C., Gerrie, V., and Everest, J. (2011). 3D litho-prediction model from in-situ physical rock property logging and constrained potential fields data inversion. In “Proceedings of 31st Gocad Meeting, Nancy”.

Portniaguine, O., and Zhdanov, M. (2002). 3-D magnetic inversion with data compression and image focusing. Geophysics, 67, 785–796.

Qi, L., and Carr, T. R. (2006). Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas. Computers and Geosciences, 32, 947–964.

Ribeiro, V. B., Louro, V. H., and Mantovani, M. S. (2013). 3D Inversion of magnetic data of grouped anomalies—Study applied to São José intrusions in Mato Grosso, Brazil. Journal of Applied Geophysics, 93, 67-76.

Shamsipour, P., Schetselaar, E., Bellefleur, G., and Marcotte, D. (2014). 3D stochastic inversion of potential field data using structural geologic constraints. Journal of Applied Geophysics, 111, 173–182.

Spicer, B., Morris, B., and Ugalde, H. (2011). Structure of the Rambler Rhyolite, Baie Verte Peninsula, Newfoundland:Inversions using UBC-GIF Grav3D and Mag3D. Journal of Applied Geophysics, 75, 9–18.

UBC Geophysical Inversion Facility. (2013). A program library for forward modelling and inversion of magnetic data over 3D structures Version 5.0.

Vella, L., and Emerson, D. (2009). Carrapateena: Physical properties of a new Iron-oxide copper-gold deposit. Journal Article ASEG Extended Abstracts, 1 – 13.

Wang, G., Zhu, Y., Zhang, S., Yan, C., Song, Y., Ma, Z., et al. (2012). 3D geological modeling based on gravitational and magnetic data inversion in the Luanchuan ore region, Henan Province, China. Journal of Applied Geophysics, 80, 1-11.

Williams, N. (2008). Geologically constrained UBC-GIF gravity and magnetic inversions with examples from the Agnew-Wiluna greenstone belt, Western Australia. Vancouver, BC, University of British Columbia.

Williams, N., and Dipple, G. (2007). Mapping subsurface alteration using gravity and magnetic inversion models. In B. Milkereit (Ed.), Exploration 07: Fifth Decennial International Conference on Mineral Exploration, (pp. 461-472).

Wong, P., Jian, F., and Taggart, L. (1995). A critical comparison of neural networks and discriminant analysis in lithofacies, porosity and permeability predictions. Journal of Petroleum Geology, 18(2), 191-206.

Zeyen, H., and Pous, J. (1993). 3-D joint inversion of magnetic and gravimetric data with a priori information. Geophysical Journal International , 112, 244–256.

Zhou, J., Zhang, X., and Xiu, C. (2015). Lithological characterization and its application based on three-dimensional structure-coupled joint inversion of gravity and magnetic data. International Journal of Geosciences, 6, 230-237.

Chapter 5

5 Rock boundary selection from more than one geophysical log using principal component and derivative analyses

5.1 Abstract

Down-hole geophysical measurements provide a continuous high resolution log trace which reflects the variation in the physical properties of subsurface lithological units. Inflection points of the geophysical log trace typically corresponds to rock boundaries. Thus, the zero values in the second derivative of the trace have been proposed as an indicator of rock boundaries. Density and gamma ray response logs measured at the Victoria property were used in this research. As the derivative method is a univariate technique, principal component analysis was used to extract one variable that is indicative of the variation of gamma ray response and density logs. Many zero crossings are evident on the extracted variable, but some are a consequence of noise. Selecting the significant zero crossings on the basis of the operator width is subjective, but we used two less subjective approaches: interpreting the importance curve or the layer thickness data. In our case, the layer thicknesses of the 67 most important layers seemed to display reasonable character. The lithological units selected by geologists had excellent correlation with the boundaries detected by the wavelet method. The results from the method we describe also showed additional information describing variation of physical properties within and between layers not identified by the geologist.

5.2 Introduction

Down-hole geophysical log measurements can provide continuous, high-resolution data that represents the physical properties of rocks surrounding the borehole. Changes in the lithology are distinguishable in the physical logs based on their physical property contrast (Mwenifumbo and Mwenifumbo, 2012). High noise levels, local heterogeneities, and aliasing of high frequency variations can adversely impact the detection of lithological units from the physical logs. Ideally, we would like a method to produce a smooth and unchanged log through homogenous units, while keeping the edges between individual units sharp. This process is called blocking or zoning (Cowan and Cooper, 2003 and Cooper and Cowan, 2009).

Blocking conventionally is performed using a median filter with adjustable window size. As the window size increases, the geophysical log becomes smooth at the expense of blurring the edges. An alternate method of picking layer boundaries is to look for inflection points in the measured properties log, which can be done by computing the second derivative of the log and then identify zero values. A simple calculation of the second derivative can be used, but the result might be confusing due to noise in data. Application of the 'Mexican hat' wavelet was introduced by Cooper and Cowan (2009) as a capable tool of computing the second derivative of logs. Using the discrete wavelet transform (DWT) to block the log traces delivers sharp boundaries with flat segments (completely denoised) between layer boundaries. The only drawback of DWT is that the location of boundaries was limited to the powers of 2. To tackle this issue, the continuous wavelet transform (CWT) has been applied to block the geophysical logs. It analyzes the log at different scales to evaluate the variation of frequency content with distance. The resultant denoised log produced by the CWT has completely flat segments with sharp boundaries (Cooper and Cowan, 2009). Davis and Christensen (2013) suggested using the

piecewise linear approximation of the wavelet to compute the second derivative of logs. This second-derivative operator is simple and can be implemented quickly.

This derivative analyses have been successfully applied on geophysical logs measured in sedimentary environments (Cowan and Cooper, 2003; Cooper and Cowan, 2009; Davis and Christensen, 2013). Down-hole density and gamma ray response measured within hole FNX1182 at Victoria property (located in Sudbury, Canada) were used to evaluate how effectively this technique can be applied to a complex igneous/metamorphic environment. As the techniques described above analyze only one physical log each time, we are proposing the application of a Principal Component Analysis (PCA) to derive one variable containing the maximum variation in two (or more) physical logs. The boundaries detected on the principal component scores log include significant changes in both density and gamma ray response logs. Matlab code supplied by Davis and Christensen (2013) was then used to analyze the PCA data.

5.3 Physical properties data set

Previous studies at the Victoria property (Mahmoodi and Smith, 2015) introduced down-hole gamma-ray response and density measurements as two physical properties which can differentiate lithological units. Therefore, they are used in this study to detect boundaries corresponding to change in lithological units down the hole. The physical logs collected from hole FNX1182 with a vertical resolution of 20 cm were used to detect lithological boundaries down the hole. Data acquisition was undertaken in 2011 by DGI Geoscience on behalf of KGHM International. To lower a multi-parameter probe to measure the data a winch and steel cable

system was used. Although the hole was NQ sized, caliper log was also acquired and this was used to correct for the borehole size.

The collected data were preprocessed to remove erroneous data. Due to the different sampling intervals for the density and gamma-ray response data, the weighted average of adjacent measurements were computed every 20 cm to resample all the logs to a uniform sample interval. The density measurements showed a normal distribution, but gamma-ray response data appeared to be more like a log-normal distribution. Since normal distributions are better for PCA (Krzanowski, 2000), the data were normalized by taking the logarithm of the gamma-ray counts. The physical logs we analyzed are represented in the two left most panels of Figure 5-1.

This derivative technique is a univariate analysis, so the number of variables should be reduced to a single variabe capable of adequately representing the variation in the data. PCA is a statistical tool which is used to reduce the dimensionality of the data (Krzanowski, 2000). The PCA was performed on the log of the gamma-ray response and the density logs, and the first component represented more than 95% of the variance shared among the density and the gamma-ray response (see the trace on the right of Figure 5-1). There was not considerable information related to lithological changes described by PC2. We also experimented with including the magnetic susceptibility log in the analysis, but found that inclusion of this log did not improve the number or quality of the lithological boundaries described by PC1. A further advantage of using the PC1 is that low-amplitude high-frequency variations, generally attributed to noise, are suppressed in the PC1.

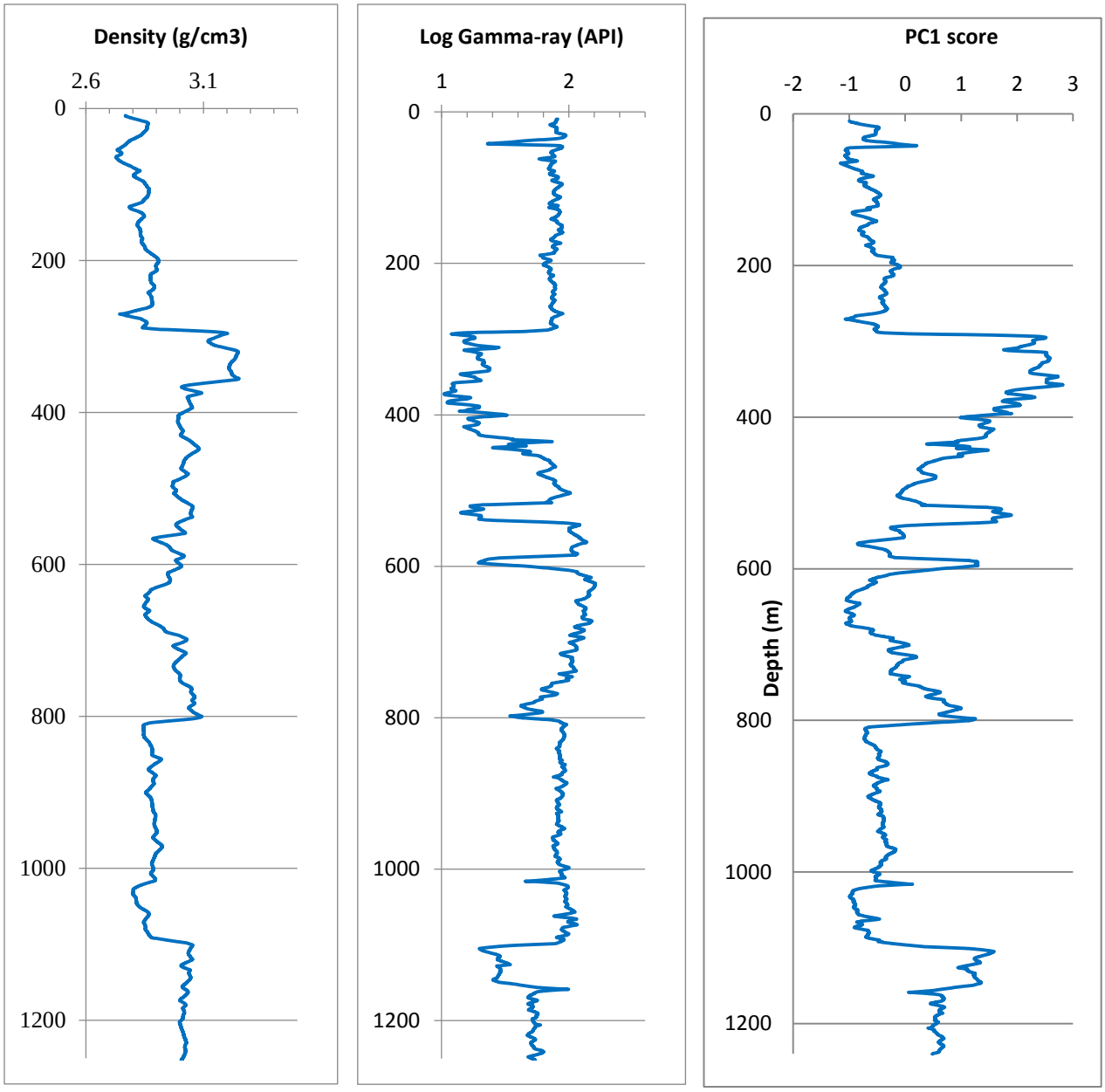


Figure 5-1- Density, gamma-ray response logs of FNX1182 and the PC1 scores plot which represents 95% of the variation of the two physical logs.

5.4 Theory and/or Method

The continuous wavelet transformation Wf of a geophysical log f(t) is represented as a function of scale (s) and position (u):

$$Wf(u,s) = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{s}}\Psi^*\left(\frac{t-u}{s}\right)dt$$

where Ψ indicates the wavelet used and * represents the complex conjugate and t is the distance down the hole. Using different scales of the wavelet permits analyzing the log trace to detect either low or high frequency variation. Figure 5-2 shows an example of a signal being analyzed with low and high scale wavelets.

Figure 5-2- Two examples, showing a low-scale and a high-scale wavelet being using to analyze a signal (http://www.mathworks.com/help/wavelet/gs/continuous-wavelet-transform.html).

The low-scale wavelet is a compressed wavelet which detects high-frequency variations related to rapidly changing signal. The high scale wavelet looks like a streched form of the low scale wavelet which detects low frequency variation related to slowly changing signal.

The procedure begins with the lowest scale of the wavelet at the t =0 location (step 1 in Figure 5-3). Then, the scale remains unchanged and the wavelet shifts to a different t value to cover the whole log trace (step 2 in Figure 5-3). The shift distance for the down-hole measurement data equals the distance between two adjacent data points. Next, the amount of distance between two adjacent data points is added to the scale, and the wavelt covers the trace with a larger scale (step 3 in Figure 5-3). This process repreats unitil the trace is analyzed with the largest scale. The lowest scale is assigned to be the distance between two adjacent data points, and two-third of the length of log trace is used for the largest scale to make sure variations with the longest possible wavelength in the trace are captured by the wavelet.

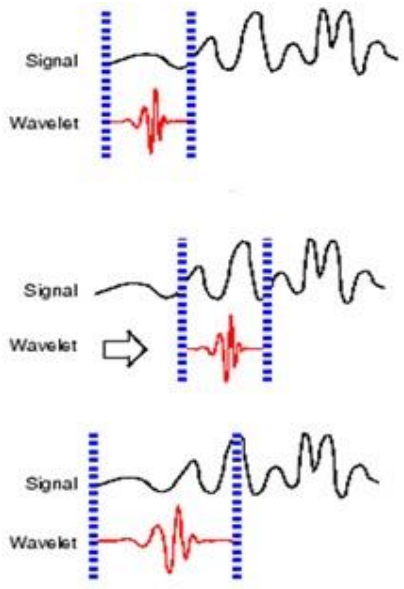


Figure 5-3- Step 1: the wavelet with smallest scale starts at the zero location; step 2: the wavelet moves by the shift distance to cover the whole trace; step 3: the scale increases, and the trace is analyzed by a larger scale wavelet (http://www.mathworks.com/help/wavelet/gs/continuous-wavelet-transform.html).

Cooper and Cowan (2009) proposed using the 'Mexican hat' wavelet, which is the second derivative of a Gaussian function:

$$\Psi(t) = \frac{2}{\pi^{0.25}\sqrt{3\sigma}}\left(\frac{t^2}{\sigma^2} - 1\right)e^{-\frac{t^2}{2\sigma^2}}$$

In this work, instead of the Mexican hat its piece-wise linear approximation (Davis and Christensen, 2013) is used as an operator to take the second derivative of the log. In order to convolve the log with the differential operator in the space-domain, the Fourier transform of the log is multiplied by the Fourier transform of the operator, and then the convolved result is transformed back to the space domain.

The derivative analysis was performed on PC scores data using a Matlab code written by Davis and Christensen (2013). The range of operator width was assigned between 1 and 800 m as the maximum width should be two-thirds of the length of the trace to make sure the longest wavelength variation is analyzed (Davis and Christensen, 2013). This process results in a matrix containing the transformed data at each depth for different operator widths. If the transformed data are plotted against the operator width for each depth, zero contours represent inflection or edge points detected at different scales, *u*. To locate the positions of lithological boundaries, a scale needs to be selected, and only the zero contours intersecting the scale should be traced back to an operator width of zero (Davis and Christensen, 2013).

5.5 Results

The color image of the transformed data for each operator width is plotted on Figure 5-4.a. Positive and negative deflection zones are indicated by red and blue colors, and black contours

correspond to zero values which indicate the inflection point in the log data. This plot can be used to locate the boundaries detected by a specific operator width. For example, the operator width of 700 m intersects the black contours at three points. If these contours are traced back, three boundaries are located at depths of 291 m, 452 m, and 1099 m.

There are many contours at smaller scales because edge points related to high frequency noise events are detected as boundaries. On larger scales, the log is analyzed to detect more significant edge points. The proper operator width selection is a critical step in this work as a scale that is too small tends to detects inflection points related to noise content of the data, and a value too large only detects very significant (large scale) boundaries. The blue line in Figure 5-4.b shows the lithological boundaries detected when using an operator width of 30m. This width is subjectively determined to be away from boundaries related to noise content of the log trace. The number of detected boundaries is determined by selecting an operator width, which is a subjective decision.

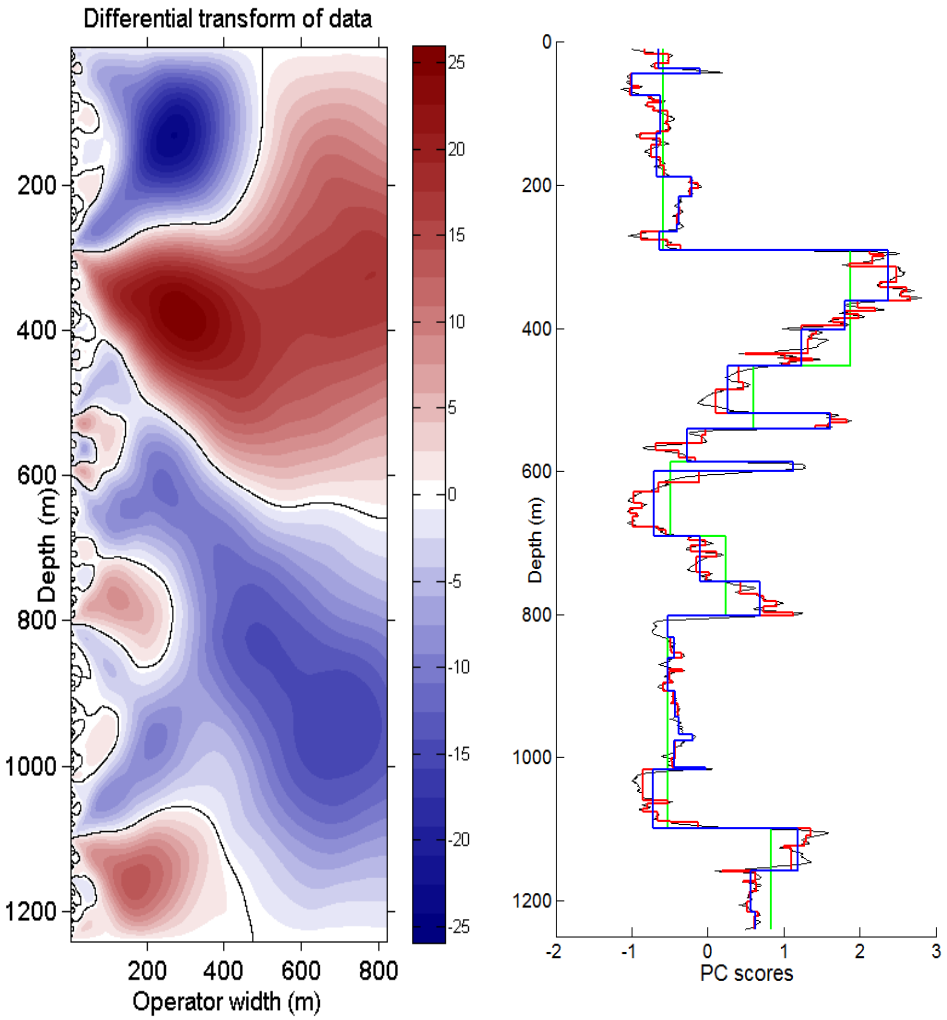


Figure 5-4- a) Second derivative of the transform of the PC scores along the hole. Negative and positive deflections of the trace are represented by blue and red regions, respectively. Black contours indicates where a zero crossing and hence an inflection point occurs; b) Black line: the original trace of PC scores; other curves are different blockings. Green line: seven most important boundaries; red line: 67 most important detected boundaries; blue line: layer boundaries detected by an operator width of 30m.

To avoid a subjective result, proper criteria are required to pick the layers based on their importance. High amplitude deflections of a log trace relative to neighboring values of the trace indicates important layers. We follow Davis and Christensen (2013) in defining the layer importance as the average value of the transformed data within each region shown in Figure 5-

4.a. The normalized layer importance obtained by dividing the mean values for each region by the maximum mean region value is shown in Figure 5-5, and the numbers indicate the seven most important layers.

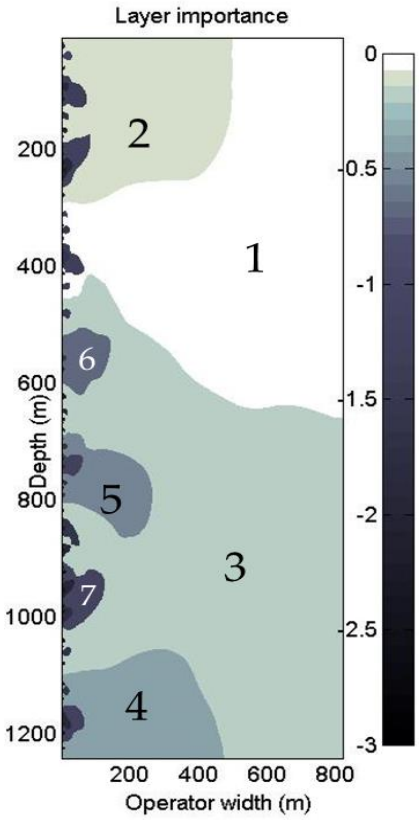


Figure 5-5- The importance layer graph. The numbers indicate the seven most important layers detected by derivative analyses.

All detected layers are ordered in Figure 5-6 based on their normalized layer importance (Davis and Christensen, 2013). The thickness of all these detected layers is also plotted in Figure 5-6 (in the same order). As illustrated, the importance of layers shows a significant drop after the 7th layer. The seven most important lithological boundaries are plotted with the green line in Figure 5-4.b. One approach would be to use the seven most important layers as a cut off, but there does

seem to be additional valid geological information in the data, so we would prefer to use more layers. There does not seem to be other significant drops in the importance curve that would suggest another threshold.

Another approach is to look at the character of the layer thickness plot in the bottom panel of Figure 5-6. The thicknesses vary rapidly, with the thicknesses varying by large amounts from one layer to the next layer of lesser importance. However beyond layer 67, the character of the variation of layer thicknesses changes, with the thicknesses not varying as much from layer to layer. We interpret this as the point at which the noise starts to have an impact on the blocking of the data.

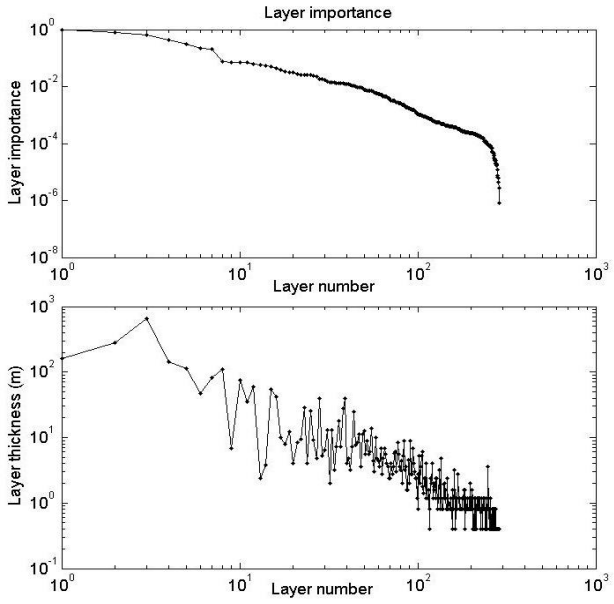


Figure 5-6- Normalized layer importance and layer thickness for all detected layers by derivative analysis in decreasing order of importance. A significant drop appears after the 7th layer, and the pattern in thickness variation changes after 67th layer.

The 67 most important layers are shown in Figure 5-4.b and repeated in Figure 5-7 for comparison with the lithological boundaries logged by geologists (black horizontal lines). All

geological boundaries between rock types are detected with the PCA wavelet technique, although they are to varying extents misplaced compared with the geologists' results. When there is a 20 cm interval of measurement, the results of wavelet analysis can enhance the accuracy of the boundaries selected. Aside from the major boundaries, the boundaries plotted within a rock unit add more details about the heterogeneities within units. One example is the sub-layers seen within the quartz diorite (QD) units, likely due to subtle changes in density; or the significant changes in physical properties that generated boundaries within the metagabbro (MTGB) at 450m and 750m.

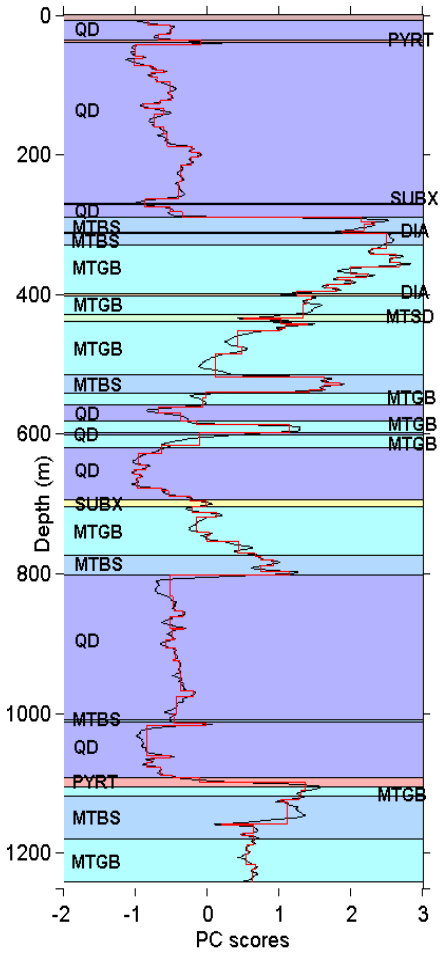


Figure 5-7- Lithological boundaries logged by geologists and the 67 most important layers (red line) detected by derivative analysis.

5.6 Discussion

Derivative analysis is a reliable and fast technique to detect layer boundaries on physical logs. In addition to significant changes on the physical log, minor within-layer changes related to heterogeneities can also be detected by the derivative analysis. Although the location of boundaries is objectively determined by picking the inflection point, the number of detected layers depends on subjective selection criteria. Wider operators detect significant edges on physical logs and they are less impacted by noise, but finer details are ignored. On the other hand, a smaller operator width results in finer details but there is a greater tendency to be adversely impacted by noise. The balance between signal and noise should be adjusted by careful selection of operator width. To eliminate noisy results (usually very thin layers) a lower threshold can be assigned for the thickness. The threshold filters the boundaries to generally pick layers above the specified minimum thickness. This approach is more useful in areas with simple stratigraphy and well-layered sedimentary rocks. In more complex metamorphic/igneous environment where rocks can occur with highly variable thicknesses, selecting a proper threshold is practically difficult. For different data sets, different approaches may yield better results. The layer importance and thickness graphs display data in a way that can lead to a means for selecting the number of layers or the minimum thickness. Selecting the number of layers which corresponds to significant drop on the layer importance graph is one valid approach. A second approach is to look at the thicknesses of the layers ordered by importance. We argue that when layers with low importance and thicknesses vary randomly around small values (seen at the layer thickness graph tail), they are more likely to be related to the noise. In the case of the data analyzed here, experimentation led us to conclude that this approach seemed to be the best method.

5.7 Conclusions

Down-hole geophysical logs provide valuable information about the physical properties of rock surrounding a drillhole. Inflection points on these logs can correspond to layer boundaries which separate two layers or lithological units with distinct physical properties. A piece-wise linear approximation of the Mexican hat wavelet is an operator to compute the second derivative of the log and zeros in the second derivative can indicate lithological boundaries. We extend the method (previously applied to one physical log) by applying it to principal component scores obtained from multiple logs (in our case density and gamma-ray response logs). Applying the operator with a range of operator widths to transform the data can be used to analyze variations in the PC1 scores at different scales. Wider operators detect significant changes in the log, while smaller operators look for higher frequency variation. The proper operator width is the one which represents more detailed information, while disregarding the boundaries related to high frequency noise content in the data. The operator width selection is a subjective way to pick the layers. Less subjective ways are to look at the layer importance, or the variation in the thicknesses as the importance decreases. We interpreted a change in the character of the thickness variation that occurred at 67 layers. Using this threshold, we found that the picked interfaces agreed well with lithological contacts identified by a geologist, demonstrating that this technique can detect boundaries even within an igneous/metamorphic environment. Additional boundaries were also identified and slight variations in depth could be used as a means of checking the rock boundaries previously logged by geologists.

5.8 References

Cooper, G. R. J., Cowan, D., (2009). Blocking geophysical borehole log data using the continuous wavelet transform. Exploration Geophysics 40(2), 233–236.

Cowan, D., Cooper, G. R. J., (2003). Wavelet analysis of detailed drillhole magnetic susceptibility data, Brockman Iron Formation, Hamersley Basin, Western Australia. ASEG Extended Abstracts 2003 (January (2)), 1–4.

Davis, A., and Christensen, N., (2013). Derivative analysis for layer selection of geophysical borehole logs. Computer and Geosciences 60, 34-40.

Krzanowski, W.J., (2000). Principles of multivariate analysis – A user's perspective. Oxford University Press, New York.

Mwenifumbo, C., and Mwenifumbo, A. (2012). Borehole geophysical logging in the Flin Flon Mining Camp. Geological survey of Canada. Geological survey of Canada, 75.

Chapter 6

6 Conclusion

6.1 Summary of research findings

Three geophysical logs were analyzed using different methods to: 1) find a meaningful pattern of variation in the physical properties data; 2) relate them to the lithological units; and 3) utilitize this link further to predict lithological units in a drillhole and in 3D space in a complex igneous/metamorphic environment. During the analyses, we found that the reliability of the results was enhanced by using knowledge we had obtained from the physical unit analysis. It suggested when the physical properties data are used to define distinct homogenous physical units, there is less overlap in the physical properties so there is also less ambiguity in classification; however, lithological units are sometimes preferred when more specific results are required. In terms of the methodology used, the capability of pattern recognition techniques was assessed for analyzing down-hole physical properties data. The Victoria property was an interesting case for this study since methods for systematically interpreting down-hole physical properties have not been well developed in hard rock environment for mineral exploration compared to hydrocarbon exploration. The results obtained from physical property characterization of rock types and lithological unit prediction using geophysical logs were promising which guaranteed the reliability and applicability of this study in hard rock environment. Due to the geological characteristics of an igneous/metamorphic environment and the overlap between physical properties of rock types, analysis and interpretation are more

complex compared with sedimentary environment. This thesis determined the critical factors, such as overlap between physical properties of rock types and heterogeneity of physical properties of rock types, to be considered while analyzing and interpreting the physical data. Although the results are applicable to the Victoria property, the same procedures can be a starting point for work at other sites.

In Chapter 2, the down-hole data from five holes were separately analyzed using the fuzzy k-means clustering to find patterns of variation in the data. We have outlined a workflow that allows us to use the multiple outputs of the fuzzy k-means algorithm, which describes the fuzziness of the results, to help in the interpretation of the data. The workflow begins with the data preprocessing required to deliver a suitable dataset. Then, the number of clusters should be determined to undertake the fuzzy k-means algorithm. An analysis of the fuzzy membership values, confusion index, and variation in physical properties can be used to modify the raw clustering result delivered by the algorithm.

The fuzzy k-means clustering was used successfully to identify four homogenous physical log units with unique physical properties at Victoria. The four physical log units is fewer in number than the 10 identified by the geologist. Lithological units with similar physical properties were logically grouped together into a physical unit. For example, metagabbro and metabasalt were represented by cluster1. The heterogeneity of metagabbro and metabasalt, which was not represented in the lithological classification done by the geologist, was apparent in the clustering as a relatively high confusion index and changes in the fuzzy membership values within the clusters. Quartz diorite and norite were dominant in cluster2. Metasediment and quartzite were grouped in cluster3. A new physical unit, “HiMag”, was manually introduced to represent extremely high magnetic rock types, primarily corresponding to diabase and sulfide zones.

Transition zones were identified to explain gradual changes in fuzzy membership values and the confusion index between two adjacent physical units. Sharp boundaries between the physical units were also verified by peaks in the confusion index and modifications to interpreted boundaries were made if necessary. The results could detect locations where despite the anomalous variation in physical properties, the rock type remains unchanged. These locations may draw attention to the core to recheck the log and possibly modify the rock-type classification.

Overall, the results proved that the physical properties vary in a meaningful pattern responding to the changes in lithological characteristics. I found that the gamma-ray log and the density log were very useful at subdividing the data space. If additional information similar to this could also be collected, then perhaps more lithological units could be identified. This information could provide a further means of identifying more physical units.

Given the identified patterns of physical properties discussed in the previous section, a classifier could be designed to quantitatively utilize the patterns for prediction. In Chapter 3, the down-hole physical properties were used to predict a) the lithological units and b) the physical log units in the hole. In lithological prediction the results contained seven different rock types, but the predictions were unreliable or ambiguous. The physical unit prediction provided more reliable, less ambiguous results at the expense of eliminating detailed lithological variation, as only four physical log units were identified. The training process for lithological prediction had 83% accuracy, while the accuracy for physical unit prediction reached 95%. The overlap of physical properties of rock types reduces the network's ability to distinguish similar rock types. Ideally, the prediction process should be able to predict all the main lithologies and possibly identify variations within those lithologies. In my results, seven lithological or four physical log units

could be predicted to varying certainties. One way of increasing the number of units identified and the accuracy of the predictions is to incorporate more logging data that relates to lithology. For example, four channel gammy-ray data and/or spectral gamma data.

A careful interpretation of the prediction results proved that the network performance is impacted by the number of data belonging to each class involved in the training process. Because the rock types with a large population contribute more data in training, the network shows a preference for the dominant rock type rather than less populous rock types when they have similar physical properties. In this work, the data were randomly divided into training, testing, and validation data. Using a different method to divide the training dataset may reduce the influence of the size of each class on the training process.

The prediction results were verified by observing the minerals and rock types in the core samples. Similarity of the physical properties between metagabbro and metabasalt is due to their identical to similar composition and mineralogy. The heterogeneity of their physical properties is partially caused by patches of inconsistent grain size and a variable percentage of clinopyroxene. Metasediment and quartzite are produced by metamorphism of sedimentary rocks. These metamorphic rocks have anomalously high gamma-ray responses because it is interpreted that their protolith contains higher concentrations of radioactive minerals. The concentration of denser mafic mineral and/or sulfide minerals increase the density and also the magnetic susceptibility of rocks. High magnetic susceptibility of the diabase is produced by a high concentration of magnetite.

In chapters 2 and 3, the magnetic data was included in the analysis. It was felt that this did not have a strong determination on the number of physical units I could define or the ability to

identify a different lithology. This was primarily because the rocks were primarily non-magnetic, but a small number of samples were magnetic. These samples were statistically insignificant on defining clusters or training neural networks. I noted that when there was two rock types with overlapping physical properties (e.g. mafic rocks and sulfides), they would fall in one physical-log-unit cluster and the network would classify them on the basis of the more common rock type (mafic rocks). Hence parameters like magnetic susceptibility are most useful for subdividing a well-defined physical log unit or lithology, for example non-magnetic or magnetic metasediments. Thus magnetic susceptibility might not be too useful for lithological classification, but it might be useful for mineral exploration as the non-magnetic and magnetic parts could potentially be used to identify alteration that has destroyed or created magnetite or pyrrhotite. A similar argument applies to the resistivity log. This log was not included in the lithological analysis, but low resistivity zones could be reflecting areas were the rock has been altered to create sulfides.

When the neural network was reasonably reliable in the lithological unit prediction, the idea was extended to lithological prediction in 3D space. Inputs for the neural network were density and magnetic susceptibility models derived from geologically constrained inversion. In Chapter 4, the down-hole physical property measurements helped in two ways: firstly they made a significant contribution to the constrained inversion and secondly were used for converting the inversion models to the lithological model. In the first instance, accurate statistical parameters of physical properties for each lithological unit intersected in the drillholes allowed me to properly constrain the physical properties of each unit in the 3D geological model. During early stages of working on the constrained inversion of the gravity and magnetic data, it was obvious that the limited measurement of norite were not truly representative of the rock type elsewhere in the

area. This issue impacted the inversion by an unrealistic effect: pushing the recovered physical properties to upper and lower bounds that had been specified in the inversion. This same effect was not seen for the rocks that were well represented by down-hole measurements. The problem with the norite was addressed by widening the upper and lower bounds. Initially, I used the mean plus and minus the standard deviation for the upper and lower limits respectively; more reasonable results were obtained when I used the maximum and minimum values from the physical properties logs in norite as the upper and lower bounds. In the second instance, the down-hole measurements were used to train the neural network to classify the petrophysical models derived from the inversion and produce the litho-prediction model. The modifications delivered by the litho-prediction model updated the previous geological model to be consistent with the geophysical data. A more extensive body of quartz diorite was suggested by the litho-prediction model. This is of interest for further exploration activities as quartz diorite at the Victoria property can be related to the offset dykes where sulfide mineralization is more likely to occur. A larger sulfide zone was identified in the litho-prediction and a new sulide zone was identified, which might indicate a mineralized zone (or a diabase). In addition, a new body of metasediment at the center of the volume of investigation and; a new magnetic subcategory of metabasalt was introduced to describe the heterogeneity of metabasalt; however, both magnetic and non-magnetic metabasalt with significant differences in the physical properties were represented as metabasalt in the geological model. Locations for boundaries were modified that are consistent with the geophysical model. The ambiguity of classification for specific rock types due to the overlap of physical properties was resolved by using the geological model as a guide. The geometry of predicted bodies helped to differentiate between two rock types with similar physical properties; i.e diabase and sulfide zone. Based on prior knowledge, we know diabase

normally occurs in sheets and the sulfide is a more compact zone. Thus, the shape of the geological model can contribute to differentiate between diabase and sulfide zone where the classifier failed to do so.

During my project, I also looked at derivative analysis to de-noise the geophysical data and pick the layer boundaries. This is a robust and simple technique, but was only designed to work on one physical property log. Work in previous chapters suggests that both the gamma-ray and density logs are useful for defining lithological boundaries. In Chapter 5, I extended the method to work on at least two physical property logs, by using the first principal component of the log data. The derivative analysis method is helpful to provide a clear image of the physical property variation as a function of the depth. This technique applies a range of operator widths to find inflection points, which correspond to the layer boundaries on the geophysical log. Wider operators, which have a high degree of noise reduction, only detect major boundaries related to the strong inflections on the geophysical log. The smaller operator widths allow more detailed boundaries to become apparent. At very small operator widths, the degree of noise reduction decreases and boundaries are observed almost everywhere on the geophysical log, which contain a portion of boundaries related to noise. Different criteria are used to pick the desirable boundaries. Care should be taken while choosing the criteria to pick the boundaries since a subjective decision of the operator width is involved in this process. Plots of the thickness of all detected layers and their importance, which depends on the degree of deflection on the geophysical log, can be used to deliver a less subjective result. The choice involves a trade-off between the importance of the layers and their thicknesses. Generally, a significant drop at the layer importance plot or a change in the variation of layer thicknesses plotted as a function of importance can suggest a value for the operator width. For example, the point at which the

thickness of less important layers starts to fluctuate around low values indicates the portion of the boundaries most likely related to the noise content of the data. In this work, using this point as a threshold to pick the layers provided more detailed geological information while disregarding the boundaries related to high-frequency noise content. Although, variation of physical properties within an igneous/metamorphic environment is more complicated than in a sedimentary environment, this work demonstrated that using this technique is reliable even in a complex case like the Victoria property.

6.2 Future Work

6.2.1 Efficient re-logging of core samples

In Chapter 2 where the down-hole physical data were analyzed, a number of significant discrepancies between identified physical units and lithological units were observed. It is worthwhile rechecking the core-log at places such as these where the lithological unit is unchanged but the physical logs show significant changes. This checking might result in a modification of the lithological choice made by the geologists, or the recognition of a new lithological type.

6.2.2 Data analysis to detect altered or mineralized data

There are different factors which can control the physical properties of rocks. Other than lithology, mineralization and alteration are two phenomena which can be reflected in

geophysical logs (Keary, et al., 2002). In Chapters 2 and 3 the down-hole geophysical logs were analyzed only in relation to changes in lithological units. It might also be possible to obtain detailed information on alteration and mineralization within the holes and incorporate them into the analysis. The differences in alteration or mineralization might be evident from the magnetic and resistivity logs, or from the fact that they have different spatial frequencies or interfaces from the lithology. If the spatial frequencies are different, it might be useful to separate the variations related to mineralization and alteration from lithological changes using a Fourier transform in some way, and enhance the signal-to-noise ratio.

6.2.3 Getting better results by using a different method to assign the training data

As shown in Chapter 3 the number of data belonging to each class can affect the training process, so that the neural network favors a choice of the dominant rocks over less populous rocks with similar physical properties. Using a different method for data selection is recommended to see if the impact of the relative number of data on the training process is reduced. Assigning similar number of data for each class was suggested by Benaouda et al. (1999) but it under-represents the population of dominant rocks. A better method might be designed to keep the balance between the major and minor rock types in the prediction process. Block-adapted learning can also be tried to counteract the bias towards classes with a large sample set. In this method, weights are updated after a complete class of training samples has been fed to the network. This compensates for dissimilarities in sample size in contrast to data-adapted learning where the weights are updated after each sample (Bishop, 1995).

6.2.4 Analyzing core sample measurements

As the cost of down-hole data collection can be high, or the logistics difficult, physical property measurements on core samples can be considered as an alternative (Smith et al., 2012). Pecman (2014) collected physical properties using hand-held tools on two of the Victoria holes (FNX1182 and 700150). Further study could investigate whether the smaller sample set of physical properties measurements on these core samples is reliable for the purposes of characterizing the rock types using classification and clustering techniques. A correlation between the results obtained from down-hole measurements would demonstrate the efficiency of this data set for physical property data analysis.

6.2.5 Collect data in more holes and a greater variety of types of down-hole measurements

The small number of samples of some rock types in the area resulted in an inaccurate understanding of their physical characteristics. I also found that the range of physical properties of metagabbro and metabasalt from the limited down-hole samples did not represent the heterogeneity of these rocks in the 3D volume. This insufficient information likely adversely impacts the prediction results in Chapters 3 and 4. For better understanding of the physical characteristics of these rocks, more physical properties should be measured. For future down-hole surveys, the holes must be selected in such a way that the collected data can accurately represent physical characteristics of all rock types. Unfortunately this might require drilling holes in areas perceived as unprospective.

Beside the number of data, more variety of measurements to characterize lithological units can enhance the accuracy of the results. For example, in addition to total gamma-ray counts

(between 0.1 and 3.00 MeV), if four-channel spectrometry could be collected with energy windows for K, U and Th, than that would allow further physical units to be identified when there are for example low or high values of K, U and Th. The density is usually measured with a gamma-gamma tool, as the number of measured counts is a function of the density. There are also spectral gamma tools and these can be used to identify whether there are atoms in the rock with low or high atomic number.

6.2.6 Unbiased criteria for detection of layer boundaries

In Chapter 5 subjective decisions are involved in choosing the criteria to pick the layer boundaries. I found that the layer thickness and layer importance can possibly help to determine a threshold where the layer importance significantly drops and the thicknesses of layers begin to fluctuate around low values. This procedure provides more objective results compared with a specified operator width; hence the ideal case is to design a criterion which is capable of eliminating the impact of noise as much as possible while keeping the detailed variation in the geophysical logs related to geological phenomena. A low-pass filter can remove high-frequency noise from the data. However, it reduces the total number of detected layers, and only important layers related to significant variations remain in the result. Another suggestion is to define a more robust criterion combining the importance and the thickness of a layer together in some way (e.g. multiplication). If all detected layers are ordered based on this new criterion, those layers with higher importance and greater thickness are placed at the top. These layers can be easily picked by finding the place that the combination suddenly gets less important.

The derivative analysis method of Davis and Christensen (2013) is a uni-variable method which can only be applied to one geophysical log each time. We used principal component analysis to derive one variable from two geophysical logs. If it is preferred to use only geophysical logs (not PC scores), one possible approach that could be tested is to process each geophysical log separately and then to weight the importance of the layer by the importance of that layer, taking into account the relative importance of the same layers as determined from other logs.

6.3 References

Benaouda, D., Wadge, G., Whitmarsh, R. B., Rothwell, R.G. and MacLeod, C. (1999). Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the Ocean Drilling Program, Geophysical Journal International, 136, 477-491. DOI: 10.1046/j.1365-246X.1999.00746.x

Bishop, C. (1995). Neural networks for pattern recognition. New York: Oxford University.

Davis, A., and Christensen, N., (2013). Derivative analysis for layer selection of geophysical borehole logs. Computer and Geosciences 60, 34-40.

Keary, P.; Brooks, M.; and Hill, I., (2002). An introduction to geophysical exploration. 3rd Edition, Blackwell Scientific Publications, Oxford, 262 p.

Pecman, M.-J., (2014), The geological significance of rock property measurements taken on core samples at the KGHM Victoria property, Sudbury, Ontario. Honours thesis, Department of Earth Sciences, Laurentian University.

Smith, R., Shore, and Rainsford. (2012). How to make better use of physical properties in mineral exploration.

Theodoridis, S., and Koutroumbas, K. (2003). Pattern recognition, second edition. New York: Academic Press.

Appendix

In chapter 3 the goal was to predict physical units in a second borehole based on the information from an unsupervised classification in an initial borehole. However, the statistical distributions might be different in the second borehole, so an alternate approach is to use an unsupervised classification of all the physical properties data available for both boreholes; this broader understanding of the statistics can then be used to predict the physical units in the second borehole. The result of combining the datasets in boreholes FNX1168 and FNX1182 and then clustering of physical property measurements and using this to predict the physical logs in FNX1182 is shown in Figure A1 below. The results are different from those shown in Figure 3-4. Although the part between 700 and 800 m or the unit at 1400 m depth on Figure A1 is better predicted compared to Figure 3-4, in this example, we found that combining the datasets increases the overlap between physical units and confuses the classification. Such confusion impedes differentiation between physical units as they are not well separated in the physical properties space. In general, the best strategy would be to include as large a dataset as possible in the unsupervised classification, but for this dataset it is arguable that it provided a significant improvement. Although, the result is acceptable for cluster 2 representing quartz diorite and HiMag units repressing diabase and sulphide zone, other rock types are not well represented by specific clusters. As the result shows, the overlap between clusters prevents metasediment and quartz diorite from forming into two separate classes. Metasediment and quartzite are mainly classified with quartz diorite in cluster 2. Cluster 1 together with cluster 3 represents both metabasalt and metagabbro.

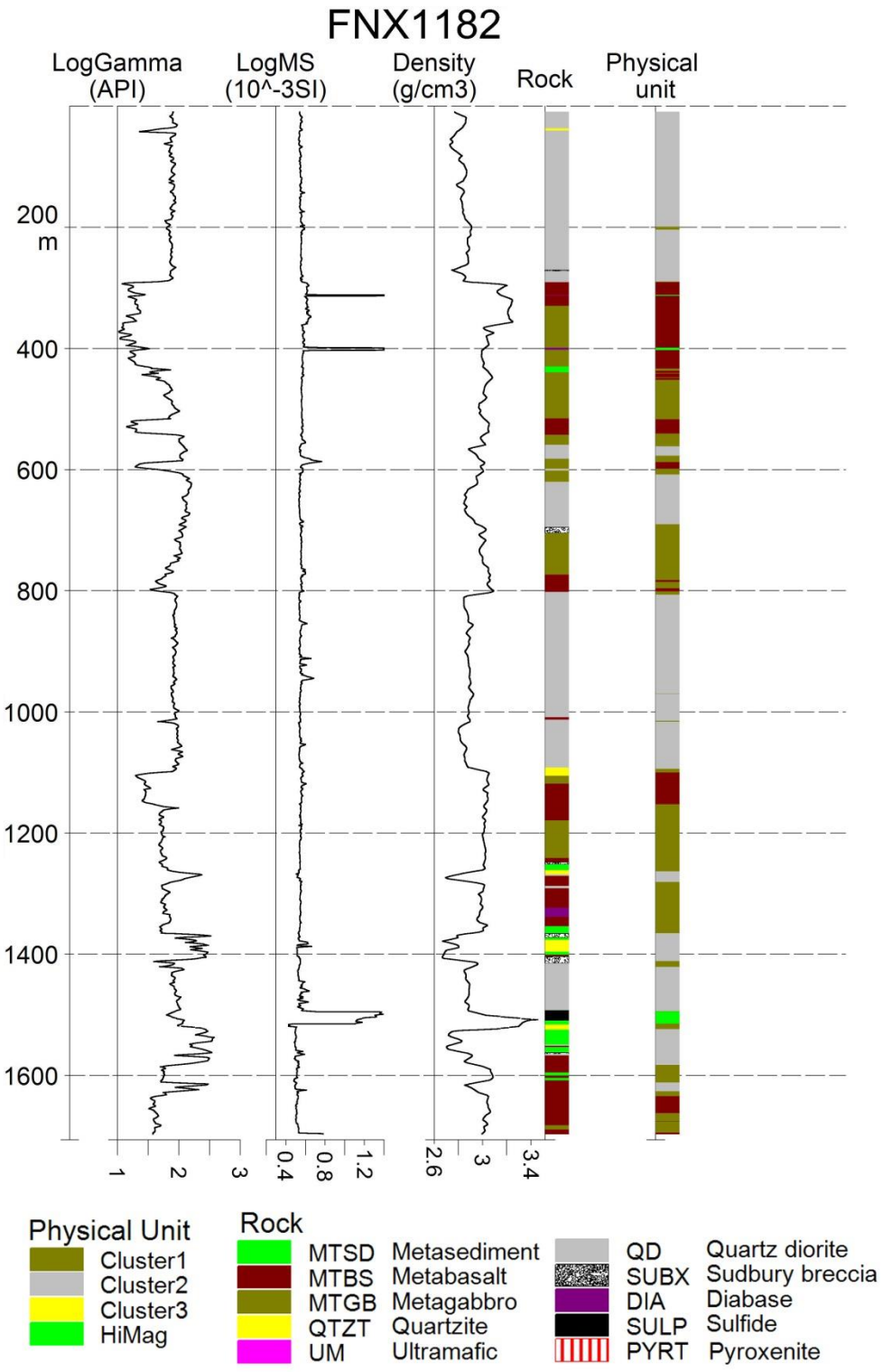


Figure A2 - Unsupervised classification of the combined datasets of FNX1168 and FNX1182.