

New Methods for the Interpolation and Interpretation of Lineaments in Aeromagnetic Data

by

Tomas Naprstek

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

© Tomas Naprstek, 2020

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian University/Université Laurentienne
Faculty of Graduate Studies/Faculté des études supérieures

Title of Thesis Titre de la thèse	New Methods for the Interpolation and Interpretation of Lineaments in Aeromagnetic Data	
Name of Candidate Nom du candidat	Naprstek, Tomas	
Degree Diplôme	Doctor of Philosophy	
Department/Program Département/Programme	Mineral Deposits/Precambrian Geology	Date of Defence Date de la soutenance June 16, 2020

APPROVED/APPROUVÉ

Thesis Examiners/Examineurs de thèse:

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

Dr. Desmond Rainsford
(Committee member/Membre du comité)

Dr. Clarence Virtue
(Committee member/Membre du comité)

Dr. Erwan Gloaguen
(External Examiner/Examineur externe)

Dr. Julia Johnson
(Internal Examiner/Examineur interne)

Approved for the Faculty of Graduate Studies
Approuvé pour la Faculté des études supérieures
Dr. David Lesbarrères
Monsieur David Lesbarrères
Dean, Faculty of Graduate Studies
Doyen, Faculté des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

I, **Tomas Naprstek**, hereby grant to Laurentian University and/or its agents the non-exclusive license to archive and make accessible my thesis, dissertation, or project report in whole or in part in all forms of media, now or for the duration of my copyright ownership. I retain all other ownership rights to the copyright of the thesis, dissertation or project report. I also reserve the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that this copy is being made available in this form by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner.

Abstract

Aeromagnetic data is one of the most widely collected types of geophysical data. In mineral exploration it can assist in mapping geological features, as well as indicate potential locations of economic interest. Due to the method in which aeromagnetic surveys are flown, an interpolation process must be completed before any map-based interpretation can be accomplished. One artifact common to many interpolation methods is that of “beading”, which is a discontinuous sequence of circular magnetic features that are at acute angles to the traverses, often caused by thin, linear geologic features such as dykes. Developing an interpolation method that “trends” or images these beads as continuous features on magnetic images would allow automatic and reliable quantitative methods to be used for interpretation by geologists and geophysicists.

First, a new interpolation method was developed for aeromagnetic data. Utilizing a Taylor derivative expansion and structure tensors, it iteratively enhances trends evident across flight lines to manifest as linear features on the interpolated grid. When applied to both synthetic data and field data, the new method showed improvement over standard bidirectional gridding, minimum curvature, and kriging methods for interpolating thin, linear features at acute angles to the flight lines.

Following this, a machine-learning interpolation approach was developed for aeromagnetic data using support vector machines and random forests. By using multiple standard interpolation methods as input to the machine-learning models, a filter-like approach was developed. These models could produce aeromagnetic maps that were overall more accurate than any single interpolation method, but not as effective as the Taylor derivative expansion method on lineament features.

Finally, convolution neural networks were applied to estimate the source parameters characterizing lineament anomalies. A synthetic aeromagnetic data modeler was used to vary relevant physical parameters, and a representative dataset of approximately 1.4 million images was developed. These were then used for training convolution neural networks to estimate the strike and depth of sources. Applying the trained networks to a real-world dataset that was interpolated by the Taylor derivative expansion method, they located a dyke and estimated a depth consistent with a previous borehole investigation.

Keywords

Aeromagnetics, artificial intelligence, convolution neural network, geophysics, interpolation, machine learning, parameter estimation, random forest, support vector machine.

Co-Authorship Statement

Chapters 2, 3, and 4 are composed of research papers that are each co-authored by Dr. Richard S. Smith. I undertook the research for all papers with guidance from Dr. Smith, and wrote the initial draft of all papers, which were then edited by Dr. Smith. Although the work and opinions are mine, because the chapters in this thesis are in the style intended for ultimate publication, the term “we” is used throughout the thesis. Chapter 2 has already been published in the journal *Geophysics*. Chapter 3 has been submitted to the journal *Geophysics*, and is in the process of being revised for resubmission. Chapter 4 will be submitted to a similar peer-reviewed journal.

Acknowledgments

I would first like to thank my supervisor, Dr. Richard S. Smith. His mentoring and support over the course of this research has been invaluable. He always encouraged me to find my own solutions and approaches, while providing me with excellent guidance. Once I began working as a researcher full-time, he continued to support me by being flexible in our meeting times and locations. This is not something that many supervisors would have accommodated, and for that I am truly grateful to him. I would also like to thank my other committee members Desmond Rainsford and Dr. Clarence Virtue. I appreciate their time and effort in the committee process, as well as reviewing my thesis. I would like to thank NSERC for funding this PhD.

I would greatly like to thank my colleagues at the National Research Council of Canada: Dr. Madeline Lee, Dr. George Leblanc, Gabriela Ifimov, Dr. Pablo Arroya-Morro, and Raymond Soffer. Madeline was kind enough to give me a chance at an internship with NRC during my PhD, which eventually led me to joining the team as a full-time researcher. I will always be grateful to her for that opportunity, as well as her overall mentorship during my initial years as a professional researcher.

Finally I would like to thank my family, particularly my ever supportive parents, Donna and Uri Naprstek, and my incredible spouse Janine. They are the reason I am who I am today.

Table of Contents

Abstract	iii
Co-Authorship Statement.....	v
Acknowledgments.....	vi
Table of Contents	vii
List of Tables	x
List of Figures	xii
Chapter 1	1
1 Introduction	1
1.1 Research Motivation	1
1.2 References.....	5
Chapter 2.....	9
2 A New Method for Interpolating Linear Features in Aeromagnetic Data	9
2.1 Abstract.....	9
2.2 Introduction.....	10
2.3 Method	13
2.3.1 User-Defined Variables	19
2.4 Synthetic Data Test	22
2.5 Field Data Test	31
2.6 Conclusions.....	37
2.7 Method's Source Code.....	38
2.8 References.....	38
Chapter 3.....	42

3 Applications of Machine Learning to Aeromagnetic Spatial Interpolation	42
3.1 Abstract	42
3.2 Introduction.....	43
3.3 Method	47
3.3.1 Predictor Set 1	51
3.3.2 Predictor Set 2.....	56
3.3.3 Predictor Set 3.....	58
3.4 Results.....	60
3.5 Case Study	65
3.6 Conclusions.....	71
3.7 References.....	72
Chapter 4.....	79
4 Convolution Neural Networks Applied to the Interpretation of Lineaments in Aeromagnetic Data.....	79
4.1 Abstract	79
4.2 Introduction.....	80
4.3 Method	85
4.3.1 Synthetic Training Data Generation	85
4.3.2 CNN Architecture	88
4.4 Results.....	92
4.4.1 Synthetic Data.....	92
4.4.2 Case Study	105

4.5 Discussion.....	110
4.5.1 Grid Cell Size.....	110
4.5.2 Width Estimation	110
4.5.3 Weaknesses and Methods of Improvement	111
4.6 Conclusions.....	113
4.7 References.....	114
Chapter 5.....	122
5 Conclusion	122
5.1 Summary of Research.....	122
5.2 Future Research	125
5.2.1 Investigate Multi-Trend Gridding for Additional Uses	125
5.2.2 Investigate Gridding Cell Effects.....	125
5.2.3 Utilize Deep Learning Methods for Aeromagnetic Interpolation.....	127
5.2.4 General Improvements to a CNN-Based Parameter Estimator.....	127
5.2.5 Estimate Additional Parameters Using CNNs	128
5.2.6 Develop a Similar Parameter Estimator for Other Anomaly Types	128
5.2.7 Develop a Machine Learning Parameter Estimator/Inversion Using Only the Aeromagnetic Flight Data.....	129

List of Tables

Table 1. The minimum, maximum, mean, median, and standard deviation values for each of the residual plots from Figure 9.....	28
Table 2. The hyperparameter choices for the RF method with predictor set 1.....	55
Table 3. The hyperparameter choices for the RF method with predictor set 2.....	57
Table 4. The hyperparameter choices for the SVR method with predictor set 2.....	57
Table 5. Summary error statistics of the four standard interpolation methods and all predictor sets (along with their associated ML method) applied to the test dataset.....	63
Table 6. A list of the parameters varied within GRAV_MAG_PRISM.	88
Table 7. The value ranges for each bin in the two classification CNN models. Note that there is no connection between classification bins across the two models (i.e. class 2 in the strike model does not imply that the depth model will also be class 2).	89
Table 8. The architecture of our CNN model. Note that the final dense layer’s neuron number “Out” depended on the feature that was being extracted.	91
Table 9. Specifications of each anomaly in the synthetic model.	93
Table 10. The comparisons between true parameters and the predicted results. Note that all lineaments contained some mix of predictions, so the results here list only the most populous predicted bin.....	95

Table 11. Summary of the noise analysis. Listed for both CNN models are the number of correctly located cells, out of the 1782 total cells that contain lineaments, as well as the mean and standard deviation of the absolute binned error. 105

List of Figures

- Figure 1. An example of a profile along an interpolated data grid. As can be seen, the Taylor Interpolation result (before normalization) smooths out the data while maintaining directional information, but does not always properly honor the real data cells (which contain measured data). The Final Result rectifies this by “normalizing” the Taylor Interpolation result, which “pulls” nearby interpolated data as real data cells are replaced with their measured values. The data at the right has been normalized the most. 15
- Figure 2. An example of the anisotropic searching to solve equation 9, with gridlines representing grid cells. The “X” represents the current σ_i location being investigated, the dots indicate the real data cells, the dashed line shows the eigenvector “search” direction, and the dotted lines, d_1 and d_2 show the straight-path distances to the σ_r locations. The double-sided arrows show an example of the “closest” nearby σ_r values. 17
- Figure 3. Flow chart of the main steps that occur in the method. 19
- Figure 4. The full 3 km x 3 km synthetic dataset of 5 m cells with Gaussian noise of 1 nT standard deviation added. The solid lines represent the 250 m line spacing data used for interpolation. Dashed grey lines and boxes represent the locations of the sources. 23
- Figure 5. Interpolation of the 250 m line spacing synthetic data set using bidirectional gridding (Akima splines) at a cell size of 50 m. All parameters were set to default within the software for this interpolation. 24

Figure 6. Interpolation of the 250 m line spacing synthetic data set using minimum curvature at a cell size of 50 m. The convergence criteria in the software were set to 99.5% pass tolerance, and 0.05% error tolerance, which it achieved in fewer than 100 iterations..... 25

Figure 7. Interpolation of the 250 m line spacing synthetic data set using kriging at a cell size of 50 m. A spherical variogram model was used with the nugget set to zero, and the sill and range were found to be 309 and 1951, respectively (default parameters in the software used). 26

Figure 8. Interpolation of the 250 m line spacing synthetic data set using our new method at a cell size of 50 m (no subsampling was used), $\varphi = 125$ m, $\theta = 10^\circ$, and the trend strength set to 100%. The interpolation process stopped at 50 iterations. 27

Figure 9. Residual plots and associated histograms of the synthetic model interpolations. The x -axes of the histograms are the residual value bins, and the y -axes are the number of cells that fall within the bins. a) Bidirectional gridding. b) Minimum curvature. c) Kriging. d) The new method..... 30

Figure 10. Minimum curvature interpolation of the Overby-Duggan aeromagnetic data set at a cell size of 80 m. The convergence criteria in the software were set to 99.5% pass tolerance, and 0.05% error tolerance, which it achieved in fewer than 500 iterations..... 32

Figure 11. Our new method’s interpolation of the Overby-Duggan aeromagnetic data set at a cell size of 80 m (subsamped from 40 m during interpolation), $\varphi = 200$ m, $\theta = 5^\circ$, and the trend strength set to 100%. The interpolation loop stopped at 97 iterations. The boudinage artifacts in the map are greatly reduced, having been trended into linear features..... 33

Figure 12. A vertical derivative map of the minimum curvature interpolation of the Overby-Duggan dataset..... 35

Figure 13. A vertical derivative map of our method’s interpolation of the Overby-Duggan dataset. 36

Figure 14. An example showing the definition of each type of cell after splitting each synthetic dataset. The interpolated cells are initially empty, and to be filled through the ML interpolation process..... 50

Figure 15. The synthetic dataset developed in PyGMI (Cole, 2019) that was utilized for testing the methods developed here. The vertical black lines show the simulated flight lines that were used as the input data for interpolation. Note that all plots of this dataset have the same colour range, and therefore only this plot shows the colourbar. 51

Figure 16. A depiction of the training process (top row) and the prediction process (bottom row) for predictor set 1. The predictors are a series of datasets (indicated by the multiple grids) that are calculated using the flight data only. With this approach, all training is completed with a single dataset’s known results (flight data). 55

Figure 17. A depiction of the training process (top row) and the prediction process (bottom row) for predictor sets 2 and 3. In this approach, multiple full datasets were split into simulated flight line data and predicted data, depicted by the three different sets of predictors in the training process. These were used to train a model, which was then applied to the test dataset. 58

Figure 18. The results from the synthetic testing. a) Minimum curvature. b) Predictor set 1, RF result. c) Predictor set 2, RF result. d) Predictor set 2, SVR result. e) Predictor set 3, RF result. f)

Predictor set 3, SVR result. Note that all plots have the same colour range as the full dataset plot in Figure 15..... 65

Figure 19. A flowchart showing the overall steps of predictor set 3, and its application to real-world data. Note that the standard interpolation methods are shown in dashed lines to highlight their unique situation as predictors, since they require the raw data (as do the other predictors), yet are predictors themselves. 66

Figure 20. The results from the case study showing a mixture of features. Specific areas have been circled for discussion. Four interpolation approaches are shown here: a) Minimum curvature, b) Multi-trend gridding, c) RF result, and d) SVR result. 69

Figure 21. The vertical derivative results from the case study. Specific areas have been circled for discussion. a) Minimum curvature. b) Multi-trend gridding. c) RF result. d) SVR result..... 70

Figure 22. A simple model showing a generic NN, where I is the input layer, h_n are the hidden layers, w_n are the associated layer weights, and O is the output layer. Each circle represents a "node" of data values, and the lines connecting them represent the weights that are iteratively updated during training. 82

Figure 23. Examples of the synthetic data for training. a) A lineament passing through the center of the image (a "hit"). It was generated with a susceptibility of 0.01, a width of 14 m, x and y positions of 250 m, a depth to top of 120 m, and a strike of 120° . b) The same dataset as a), with each cell separated to show the positioning of the data. The blue lines represent the physical location of the lineament, and the black box shows where the check is made to determine if the dataset is a "hit". c) The same lineament shifted to a y position of 350 m such that it no longer

passes through any of the center nine cells of the image (“no lineament”). d) The same lineament as in a), but with a secondary identical lineament passing in the $\pm y$ direction horizontally offset from the primary lineament. 87

Figure 24. Accuracy and loss plots from the training process for both the strike and depth models. Note that Loss does not have a unit, and the value is dependent on the type of loss function. 91

Figure 25. a) The synthetic model used for testing. The six lineaments being analyzed have been labelled “L1” to “L6” and the two background block anomalies have been labelled “B1” and “B2”. b) The model with the lineaments (in black) and blocks (in red) shown to indicate their physical locations. 93

Figure 26. The prediction results for the two source-model parameters. a) The strike predictions, ranging between 0° and 180° , in steps of 20° . b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m. 95

Figure 27. A scatter plot showing the cells that contain lineaments (black) and cells that were predicted to contain lineaments by the models (red). 97

Figure 28. Error histograms for the two models. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide. 98

Figure 29. Synthetic dataset with varying noise levels. Note that Figure 25 was generated with 0.1 nT of Gaussian noise. a) No noise, for comparison. b) Noise set to 1 nT. c) Noise set to 3 nT. d) Noise set to 10 nT. 99

Figure 30. The prediction results for 1 nT of Gaussian noise. a) The strike predictions, ranging between 0° and 180°, in steps of 20°. b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m. 100

Figure 31. The prediction results for 3 nT of Gaussian noise. a) The strike predictions, ranging between 0° and 180°, in steps of 20°. b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m. 101

Figure 32. The prediction results for 10 nT of Gaussian noise. a) The strike predictions, ranging between 0° and 180°, in steps of 20°. b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m. 102

Figure 33. Error histograms for 1 nT of Gaussian noise. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide. 103

Figure 34. Error histograms for 3 nT of Gaussian noise. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide. 104

Figure 35. Error histograms for 10 nT of Gaussian noise. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide. 104

Figure 36. A subsection of the Matheson dataset, centered on the dyke that has been analyzed in other studies. The specific location of interest has been circled for clarity. 107

Figure 37. Predictions from the strike model on the Matheson dataset. The colourbar shows the strike estimate in steps of 20° 108

Figure 38. Predictions from the depth model on the Matheson dataset. The colourbar shows the depth estimate in steps of 25 m. The specific area of interest for comparison to other studies is shown at expanded scale in the insert section, with the expected value of the 100 m to 125 m bin being predicted..... 109

Chapter 1

1 Introduction

1.1 Research Motivation

Aeromagnetic surveys are used around the world in a vast array of applications, from mapping geology to locating buried archaeological structures (Burger et al., 2006; Reynolds, 2011). In mineral exploration, magnetic surveys are extensively utilized, as they can lead directly to deposits of interest, or assist in mapping geological features that can point to their location (Telford et al., 1990). Due to the relatively low cost and ease in which aeromagnetic data can be acquired (Reeves, 2005), many nation-wide surveys have been undertaken by governments in countries such as Canada, Australia, and Finland to encourage mineral exploration (Nabighian et al., 2005).

Aeromagnetic surveys are completed by attaching one or (usually) more magnetometers to an aircraft (fixed-wing or helicopter), either directly in a “stinger”, or towed behind/below as a “bird” (Reeves, 2005). They are then flown in a direction for some distance, usually tens to hundreds of kilometers, before reversing direction and flying a parallel line. Each of these parallel “profile lines” are generally spaced anywhere from less than a hundred meters to thousands of meters apart from each other, resulting in high-spatial-density data along the profiles, and no data between them. The specific parameters of the survey (flight height, line spacing, etc.) are determined by the purpose of the survey, and the size of features that are sought to be resolved (Reid, 1980).

These magnetic surveys measure the magnetic field along the flight path, and anomalies are found by observing the deviation between the expected magnetic field at a location, which can be estimated very accurately using geomagnetic field models (Thébault et al., 2015), and the actual measured value. The rock below and around the survey locations causes the deviation from the expected magnetic field, with the amplitude depending linearly on the magnetic susceptibilities of the subsurface materials (Blakely, 1996). An aeromagnetic anomaly map can be utilized by geophysicists and geologists to assist in interpreting the geologic structure of a region, and potentially lead to areas where there are economic mineral deposits. Dykes are one feature that appear strongly in aeromagnetic data and are also useful in mapping regional geology. Dykes are intrusions of igneous rock, whose thickness varies from a few centimeters to tens of meters (Grotzinger and Jordan, 2010), and are often found in large quantities in an area, referred to as dyke swarms. Most dykes are composed of tholeiitic material (Pilkington and Roest, 1998), which is mafic igneous rock containing magnesium and iron, so magnetite (an iron oxide with a large magnetic susceptibility) is frequently a constituent mineral. Despite often being small in size, it is generally found that mafic dykes that are wider than 5 m will be found by most standard aeromagnetic surveys (Pilkington and Roest, 1998), and focused small-scale surveys can potentially resolve smaller dykes and features.

There are a number of processing steps that must be completed before aeromagnetic profile data can be interpreted, with a key step being interpolation of the flight-line data to a 2D grid so that an image and map can be generated. This interpolation process attempts to utilize the high-spatial-density data collected along-line to fill in the gaps existing between lines. This process

must respect the data measured on the flight line, while also attempting to generate interpolated data between the flight lines that smoothly varies. Unfortunately, this can introduce various artifacts in the data, such as those that often appear when interpolating thin, linear features, like the responses of dykes and dyke swarms. When these linear features pass through the map at non-perpendicular angles with respect to the flight lines, the response often manifests as a series of “bulls-eyes”, a “string of beads”, or a “boudinage” artifact (e.g. Keating, 1997; Smith and O’Connell, 2005; Guo et al., 2012; Geng et al., 2014). Interpretation is facilitated if the interpolation of the data looks more like the geological structures causing the anomalies.

Interpretation is primarily a manual process undertaken by a skilled geophysical/geological interpreter but can be augmented by automatic methods, which use either the profile data or the interpolated map data to estimate source parameters, such as the depth to the top of the magnetic sources, the width of the dykes, and their magnetic susceptibility. The automatic methods rely heavily on low-noise data, as they often use high-order derivatives of the data. If the data is not well pre-processed, or of sufficient quality, noise and other artifacts can corrupt the results and confuse the resulting interpretations.

With the growing use of machine learning in many research fields, methods such as support vector machines and convolution neural networks are becoming straight-forward to implement and apply by non-experts of artificial intelligence. These have been proven to be valuable tools in automatic processing and analysis of many different types of data. Machine learning has already been applied using aeromagnetic data in some contexts with success, such as predicting lithology (e.g. Yu et al., 2012; Cracknell and Reading, 2013; Cracknell and Reading, 2014;

Rodriguez-Galiano et al., 2015) or interpreting some aspects of magnetic anomalies (Aghaee Rad, 2019; Nurindrawati and Sun, 2019). Therefore, if used correctly, machine learning can be advantageous in both the interpolation and interpretation of aeromagnetic data.

The goal of this research was twofold:

- 1) Develop a new interpolation method, effective at interpolating thin, linear features, including any dyke or dyke-like feature. This could use high-order spatial derivatives, but ideally would not.
- 2) Develop a new interpretation method that does not inherently rely on high-order derivatives.

Chapter 2 describes a new numerical, iterative-based interpolation method¹ whose goal is to interpolate aeromagnetic grids to be free of the classic lineament-associated artifacts. The method does use high-order spatial derivatives. Chapter 3 discusses three different approaches to utilizing machine learning methods (specifically random forest regression and support vector machines) for general aeromagnetic interpolation improvement, none of which use spatial derivatives explicitly. Chapter 4 investigates the usage of machine learning (specifically convolution neural networks) for the automatic detection of dykes in aeromagnetic grids, as well

¹ After the paper was published, it has generally been referred to as “multi-trend gridding”, or “multi-trend interpolation”.

as interpreting their strike and depth to top of source. Finally, Chapter 5 summarizes the results of these three chapters and suggests potential future work.

This PhD thesis is composed of 5 Chapters, where Chapters 2, 3, and 4 are “stand-alone” research papers. As such, each Chapter will begin with an abstract, and all references are contained within each individual chapter. Note that although the work and opinions are mine, because the chapters in this thesis are in the style intended for ultimate publication, the term “we” is used throughout the thesis, referring to my co-author (and supervisor), Dr. Richard S. Smith.

1.2 References

Aghaee Rad, M. A., 2019, Machine learning of lineaments from magnetic, gravity and elevation maps, University of British Columbia. Retrieved from <https://open.library.ubc.ca/collections/ubctheses/24/items/1.0376558>.

Blakely, R. J., 1996, Potential theory in gravity and magnetic applications, Cambridge university press.

Burger, H. R., A. F. Sheehan, and C. H. Jones, 2006, Introduction to applied geophysics: Exploring the shallow subsurface, WW Norton.

Cracknell, M. J., and A. M. Reading, 2013, The upside of uncertainty: Identification of lithology contact zones from airborne geophysics and satellite data using random forests and support vector machines: *Geophysics*, 78, no. 3, WB113-WB126.

Cracknell, M. J., and A. M. Reading, 2014, Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information: *Computers & Geosciences*, 63, 22-33.

Geng, M., Q. Yang, and Y. Yuan, 2014, Trend enhancement in aeromagnetic maps using constrained coherence-enhancing diffusion filtering: *Geophysics*, 79, no. 1, J1-J9.

Grotzinger, J., and T. Jordan, 2010, *Understanding Earth: Sixth Edition*, New York: W. H. Freeman and Company.

Guo, L., X. Meng, and L. Shi, 2012, Gridding aeromagnetic data using inverse interpolation: *Geophysical Journal International*, 189, 1353-1360.

Keating, P., 1997, Automated trend reinforcement of aeromagnetic data: *Geophysical Prospecting*, 45, 521–534.

Nabighian, M. N., V. J. S. Grauch, R. O. Hansen, T. R. LaFehr, Y. Li, J. W. Peirce, J. D. Phillips, and M. E. Ruder, 2005, The historical development of the magnetic method in exploration, *Geophysics*, 70, 33ND-61ND.

Nurindrawati, F., and J. Sun, 2019, Estimating total magnetization directions using convolutional neural networks, *SEG Technical Program Expanded Abstracts*, 2019, 2163-2167.

Pilkington, M., and W. R. Roest, 1998, Removing varying directional trends in aeromagnetic data: *Geophysics*, 63, 446-453.

Reeves, C., 2005, *Aeromagnetic Surveys Principles: Practice and Interpretation*, GeoSoft.

Reid, A. B., 1980, Aeromagnetic survey design, *Geophysics*, 45, 973-976.

Reynolds, J. M., 2011, *An introduction to applied and environmental geophysics*, John Wiley & Sons.

Rodriguez-Galiano, V., M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, 2015, Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines: *Ore Geology Reviews*, 71, 804-818.

Smith, R. S., and M.D. O'Connell, 2005, Interpolation and gridding of aliased geophysical data using constrained anisotropic diffusion to enhance trends: *Geophysics*, 70, no. 5, V121-V127.

Telford, W. M., L. P. Geldart, and R. E. Sheriff, 1990, *Applied Geophysics*, Cambridge University Press.

Thébault, E., C. C. Finlay, C. D. Beggan, P. Alken, J. Aubert, O. Barrois, ... and E. Canet, 2015, International geomagnetic reference field: the 12th generation, *Earth, Planets and Space*, 67, 79.

Yu, L., A. Porwal, E. J. Holden, and M. C. Dentith, 2012, Towards automatic lithological classification from remote sensing data using support vector machines: *Computers & Geosciences*, 45, 229-239.

Chapter 2

2 A New Method for Interpolating Linear Features in Aeromagnetic Data

2.1 Abstract

When aeromagnetic data are interpolated to make a gridded image, thin linear features can result in “boudinage” or “string of beads” artifacts if the anomalies are at acute angles to the traverse lines. These artifacts are due to the under-sampling of these types of features across the flight-lines, making it difficult for most interpolation methods to effectively maintain the linear nature of the features without user guidance. Dykes and dyke swarms are typical examples of the type of geological feature whose magnetic response can cause these artifacts, and thus these features are often difficult to interpret. Many interpretation methods utilize various enhancements of the gridded data, such as horizontal or vertical derivatives, and these artifacts are often exacerbated by the processing. Therefore, interpolation methods that are free of these artifacts are necessary for advanced interpretation and analysis of thin, linear features. We have developed a new interpolation method that iteratively enhances linear trends across flight lines, ensuring that linear features are evident on the interpolated grid. Utilizing a Taylor derivative expansion and structure tensors allows the method to continually analyze and interpolate data along anisotropic trends, while honoring the original flight line data. We applied this method to synthetic data and field data which both show improvement over standard bidirectional gridding, minimum curvature, and kriging methods for interpolating thin, linear features at acute angles to the flight

lines. These improved results are also apparent in the vertical derivative enhancement of field data. The source code for this method has been made publicly available.

2.2 Introduction

An iconic aspect of many geophysical surveys is that the acquired data is spatially dense along traverse lines, and entirely devoid of data between these lines. This poses a unique interpolation challenge to utilize the high-density data, but avoid introducing artifacts in an interpolated map, image, or grid, while at the same time respecting the measured data. At its core, it is an aliasing issue, where features that occur across lines need to be handled appropriately or artifacts may occur (Reid, 1980). One such type of artifact is the aeromagnetic response of thin, linear features, like those often produced by dykes and dyke swarms (Pilkington and Roest, 1998). If these linear features are trending at non-perpendicular angles with respect to the flight lines, the response often manifests as a “string of beads” or “boudinage” artifact on the interpolated map (Keating, 1997; Smith and O’Connell, 2005; Guo et al., 2012; Geng et al., 2014). This is particularly prone to occurring when data is interpolated using the most common methods in mining geophysics, such as bidirectional splines (Bhattacharyya, 1969; Akima, 1970), minimum curvature (Briggs, 1974; Swain 1976; Mallet, 1989; Smith and Wessel, 1990) and kriging (for example, Hansen, 1993). Bidirectional splines interpolate both along flight lines and across them, inherently developing a grid with some directional bias (Keating, 1997). This leads to effective interpolation of anomalies that are perpendicular to the flight lines; however, it can lead to these beading artifacts when a linear anomaly is at an acute angle to the line data. These artifacts are also inherent in minimum curvature interpolation, as a trend at an acute angle to the flight line

data is under-sampled (Geng et al., 2014), thus leading to minimal data for the interpolation to be constrained by. This lack of constraints is found to result in a circular anomaly. Kriging interpolation results are similar to those obtained by bidirectional methods, in that regional trends can be effectively handled (Hansen, 1993), but local trends are not accounted for (Keating, 1997; Guo et al., 2012). Other methods (Cordell, 1992; Mendonca and Silva, 1994, 1995; Billings et al., 2002) have similar issues due to under-sampled anomalies (Guo et al., 2012).

Therefore, what is required to solve this type of artifact is a filter or interpolation method that can account for multi-directional, overlapping, under-sampled trends that are trending at a variety of angles with respect to the flight lines. There have been a number of fairly effective treatments of the issue, both by applying post-interpolation filters and by developing new interpolation methods as a whole. Keating (1997) inserts trend lines as new data between flight lines when local maxima and minima are discovered and these trend lines include nearby real data maximas/minimas. This circumvents the beading issue by essentially trending features as separate entities from the rest of the interpolation process. Yunxuan (1993) and Sykes and Das (2000) utilize the Radon transform (also known as the slant stack in seismic applications) for a variety of trend-based processes, including the enhancement of linear trends. Guo et al. (2012) developed an inverse interpolation methodology, which showed reductions in beading artifacts of trends when compared to minimum curvature results, particularly in the vertical derivative enhancements. Smith and O'Connell (2005) apply an anisotropic diffusion enhancement, which analyzes the structure of the data using structure tensors and iteratively smooths it along linear

trends. This approach was later improved by Geng et al. (2014) by constraining the process to be only applied in those highly anisotropic locations that contain thin, linear trends.

The method proposed here is most similar to these last two methods, as part of the process utilizes structure tensors to analyze the interpolated data, and iteratively enforce trends. The eigenvalues and eigenvectors of structure tensors have been effectively used in seismic applications (Fehmers and Höcker, 2003; Hale, 2010; Wu, 2017) to describe the strength and direction of anisotropy, and therefore can provide useful information on trends once an interpolation process has been applied. However, unlike other methods, we base our interpolation around a discrete version of the Taylor series expansion of two variables. This was chosen as, similar to the spline methods, it provides a simple yet flexible mathematical basis for the formulation of the problem and inherently will enhance trends by extrapolating features across flight lines. A further advantage of using this method is that the data does not necessarily need to be acquired along straight-line traverses (as with bidirectional methods).

We begin by describing the Taylor series interpolation, as well as the method of utilizing structure tensors for trend analysis. We then describe how we implement these two features in combination with a process that we refer to as “normalizing” the data to develop an iterative interpolation methodology. Finally, we show the capability of our method by applying it to both synthetic and field data sets, and compare the results with the maps produced by the readily available and commonly-used techniques of bidirectional splines, minimum curvature, and kriging. For ease-of-use, the C# source code has been made available.

2.3 Method

The core of the interpolation process is based on a Taylor series expansion of two variables

(adapted from Abramowitz and Stegun (1970), page 880, equation 25.2.24):

$$f(i + m, j + n) \approx f(i, j) + m \frac{\partial f(i, j)}{\partial x} + n \frac{\partial f(i, j)}{\partial y} + \frac{1}{2} \left[m^2 \frac{\partial^2 f(i, j)}{\partial x^2} + 2mn \frac{\partial^2 f(i, j)}{\partial xy} + n^2 \frac{\partial^2 f(i, j)}{\partial y^2} \right], \quad (1)$$

where $f(i, j)$ is a data cell in a grid whose horizontal coordinates are represented by $x = i, y = j$,

and m and n are offsets to the x - and y -directions respectively. The derivatives are defined as

(Abramowitz and Stegun, 1970, pages 883-884)

$$\frac{\partial f(i, j)}{\partial x} = \frac{1}{2h} [f(i + 1, j) - f(i - 1, j)], \quad (2)$$

$$\frac{\partial f(i, j)}{\partial y} = \frac{1}{2h} [f(i, j + 1) - f(i, j - 1)], \quad (3)$$

$$\frac{\partial^2 f(i, j)}{\partial x^2} = \frac{1}{h^2} [f(i + 1, j) - 2f(i, j) + f(i - 1, j)], \quad (4)$$

$$\frac{\partial^2 f(i, j)}{\partial y^2} = \frac{1}{h^2} [f(i, j + 1) - 2f(i, j) + f(i, j - 1)], \quad (5)$$

$$\frac{\partial^2 f(i, j)}{\partial xy} = \frac{1}{4h^2} [f(i + 1, j + 1) - f(i - 1, j + 1) - f(i + 1, j - 1) + f(i - 1, j - 1)], \quad (6)$$

and h is defined as the absolute distance (in meters) between each point on the equispaced grid.

In this implementation, the spacing in the x and y directions is assumed to be the same. By

rearranging equation 1 for the eight combinations of (m, n) that are adjacent to location (i, j) , we

can solve for $f(i, j)$. As there are eight surrounding cells, this means there are eight separate

estimates for $f(i,j)$. A trimmed-mean filter (Hall, 2007) with $\alpha = 25\%$ is applied to these eight estimates (therefore the two smallest and two largest values are removed from the mean calculation), and the resulting value is recorded for location (i,j) in a new grid, $f_{NS}(x,y)$. This new cell will be a manipulation of the cell's previous value and the derivatives surrounding it (and hence the surrounding values). By applying this to every cell, a new grid with enhanced directional information is developed. However, to mitigate any potential discontinuities, cells that contain real data must also go through this process, thus changing them from their original (measured) value. To revert them back and properly honor the measured/real data, the cells surrounding them must also change, as a direct replacement would also cause discontinuities. Our method approaches this problem by applying a scaling factor to all cells, such that the real data cells are changed back to their original values, and the surrounding interpolated data is "pulled" along with it. We refer to this process as "normalizing" the data (Figure 1).

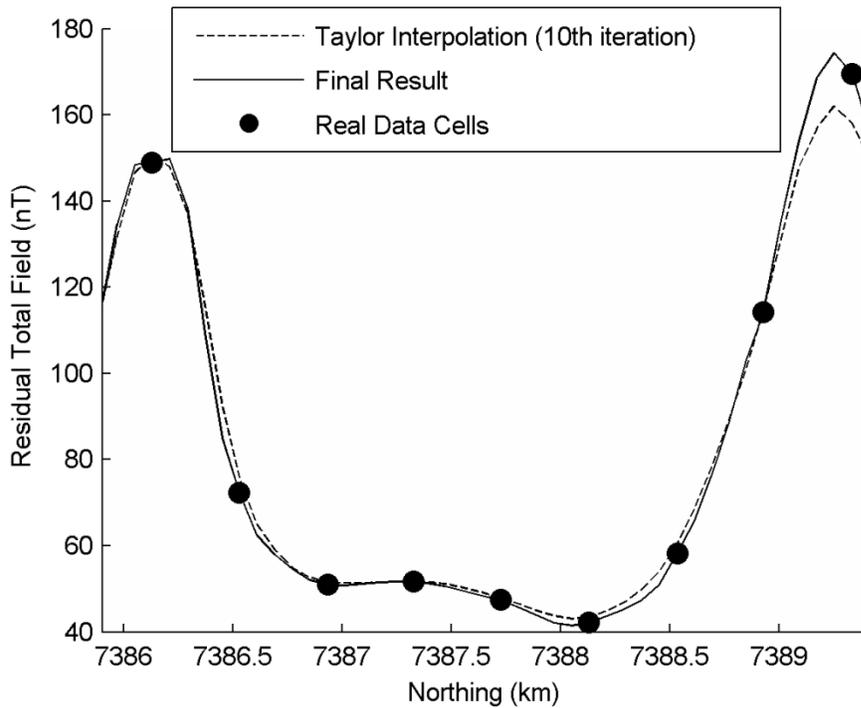


Figure 1. An example of a profile along an interpolated data grid. As can be seen, the Taylor Interpolation result (before normalization) smooths out the data while maintaining directional information, but does not always properly honor the real data cells (which contain measured data). The Final Result rectifies this by “normalizing” the Taylor Interpolation result, which “pulls” nearby interpolated data as real data cells are replaced with their measured values. The data at the right has been normalized the most.

To accomplish this normalization, we develop a grid of multipliers, $\sigma(x,y)$, comprised of real data multipliers, $\sigma_r(x,y)$, and interpolated data multipliers, $\sigma_i(x,y)$. The real data multipliers are defined as:

$$\sigma_r(i, j) = \left| \frac{f_r(i, j)}{f_{NS}(i, j)} \right| \quad (7)$$

which is the absolute value of the original real data cell, $f_r(i, j)$, divided by the cell's current value, $f_{NS}(i, j)$. The value of $\sigma_i(x, y)$ at each interpolated location comes from nearby $\sigma_r(x, y)$ values, which we find by completing a search along the path of greatest anisotropy. While a more simple approach, such as a mean of nearest neighbor $\sigma_r(x, y)$ values would be computationally quicker and less prone to noise, our goal of this method is to ensure anisotropic linear trends are enforced, and therefore following lines of anisotropy will help achieve this objective. To find the angle of greatest anisotropy, we calculate structure tensors, $S(i, j)$ at each interpolated data cell on the grid $f_{NS}(x, y)$. A 2 x 2 structure tensor is defined as (Smith and O'Connell, 2005):

$$S(i, j) = \nabla f(i, j) \nabla f(i, j)^T = \begin{bmatrix} \frac{\partial f(i, j)}{\partial x} \cdot \frac{\partial f(i, j)}{\partial x} & \frac{\partial f(i, j)}{\partial x} \cdot \frac{\partial f(i, j)}{\partial y} \\ \frac{\partial f(i, j)}{\partial y} \cdot \frac{\partial f(i, j)}{\partial x} & \frac{\partial f(i, j)}{\partial y} \cdot \frac{\partial f(i, j)}{\partial y} \end{bmatrix}, \quad (8)$$

where the dot symbol indicates scalar multiplication. We then calculate the eigenvalues and eigenvectors for the tensor at each grid point, as they describe both the strength and direction of any trend within that cell. Searching adjacent points along the trend in both positive and negative directions, we then calculate the interpolated cell's multiplier based on an inverse distance-weighted average of any $\sigma_r(x, y)$ values in those directions:

$$\sigma_i(i, j) = \frac{d_2 \left(\frac{\sigma_{r11} + \sigma_{r12}}{2} \right) + d_1 \left(\frac{\sigma_{r21} + \sigma_{r22}}{2} \right)}{d_1 + d_2}, \quad (9)$$

where σ_{r11} is the real data cell found in the positive direction, σ_{r12} is the real data cell “closest” to σ_{r11} and in a direction most perpendicular to the search path, σ_{r21} is the real data cell found in the negative direction, σ_{r22} is the real data cell “closest” to σ_{r21} in a similar way, and d_1 and d_2 are the straight-path distances to σ_{r11} and σ_{r21} respectively from the location of $\sigma_i(i,j)$. Figure 2 shows an example of this calculation.

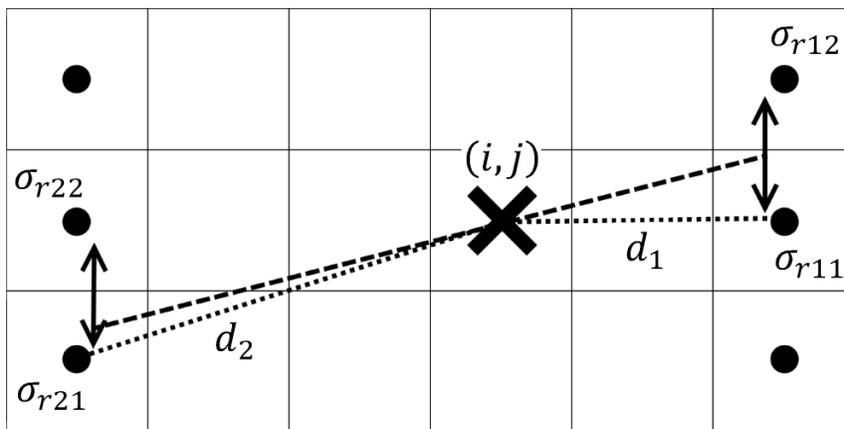


Figure 2. An example of the anisotropic searching to solve equation 9, with gridlines representing grid cells. The “X” represents the current σ_i location being investigated, the dots indicate the real data cells, the dashed line shows the eigenvector “search” direction, and the dotted lines, d_1 and d_2 show the straight-path distances to the σ_r locations. The double-sided arrows show an example of the “closest” nearby σ_r values.

Note that while averaging σ_{r12} and σ_{r22} may lessen the effect of normalization if a strong trend is found, they are implemented for a smoother normalization process, and to ensure that no erratic multiplier effect may occur. If no data is found in either direction (i.e. the edge of the grid is hit, or a maximum interpolation distance, ϕ , is reached before finding a real data cell), the

eigenvector search path is varied by the angle θ , a user-defined number of degrees, and the process is repeated until successful. A new grid, $f_{FS}(x,y)$, is then developed by applying all multipliers to their associated cells:

$$f_{FS}(x,y) = f_{NS}(x,y) \cdot (\tau \cdot \sigma(x,y)), \quad (10)$$

where τ is a user-defined “trend strength”. By repeatedly applying equation 10, and recalculating the Taylor interpolations and structure tensors every iteration, a final interpolated grid can be developed using the flowchart in Figure 3. This new grid will be comprised of real data cells that honor the flight-line data and interpolated data cells whose values are enhanced to enforce linear trends across the flight lines.

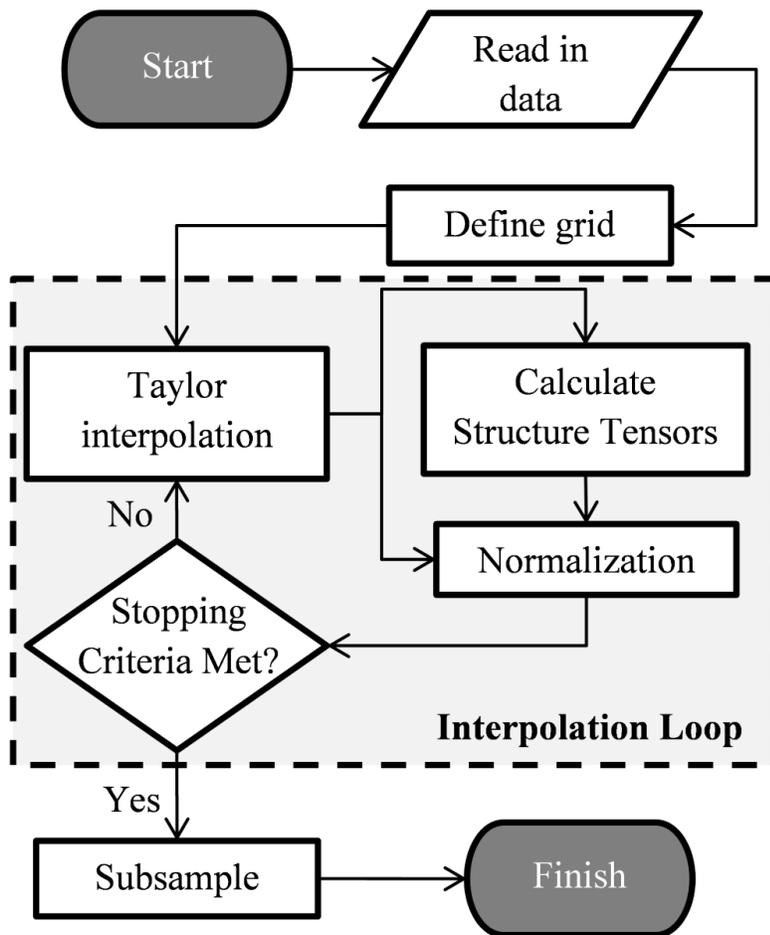


Figure 3. Flow chart of the main steps that occur in the method.

2.3.1 User-Defined Variables

As seen in the previous section, several parameters in this process must be defined by the user, as their effect can have significant consequences on the final resulting interpolation. Through extensive empirical testing, we have developed some basic guidelines to assist a user of this method. The maximum interpolation distance, φ , represents the maximum distance at which the method will search along the trend direction before stopping and varying the angle by θ , another

user defined variable. A maximum distance is required, as an interpolated data cell's eigenvector may be parallel to the flight lines, so that no real data cell will be along its search path. The authors have found that 50-100% of the average flight line spacing is generally an effective value for φ . However, this is left as a user-defined variable, as an astute user may wish to investigate trends that pass through the flight lines at very acute angles. These types of trends may require larger φ values to ensure that real data cells in the flight lines on both sides of an interpolated data cell are reached and utilized during the normalization process. The interpolation angle, θ , defines how much the trend direction will change if the maximum interpolation distance, φ , is reached before a real data cell is found during the normalization process. This angle should be small, generally 5-10° to ensure that relevant real data cells will not be skipped during the search process. However, it should be noted that a smaller value will increase the computation time of the method.

The trend strength, τ , represents another user-defined variable. In equation 10, τ can range from 0 to 1, depending on both the value entered by the user and the overall dataset's anisotropy. The user can enter a value ranging from 0 to 100, which represents how much data will be trended to the full effect. For example, at a trend factor of 75, the 75% most anisotropic data (as measured by a statistical analysis of the structure tensor's eigenvalues) is trended to the full extent, while the other 25% of the data will have a sliding reduction factor applied. The goal of this trend factor is to allow the user flexibility in what is trended, such that more or less strongly anisotropic features can be trended at various strengths.

Two user-defined steps not described in the previous section, but shown in the final flow chart (Figure 3) are the automatic stopping criteria and subsampling process. The interpolation loop may be set by the user to run a specific number of iterations; however an automatic stopping method is also available. After the current iteration, n , is completed, the resulting “corrected grid”, $f_{FS,n}(x,y)$ can be analyzed, calculating a “difference grid”:

$$f_{D,n}(x, y) = |f_{FS,n}(x, y) - f_{FS,n-1}(x, y)|. \quad (11)$$

This grid, $f_{D,n}(x,y)$, is compared to the previous difference grid, $f_{D,n-1}(x,y)$, and an average change is checked to determine if these differences are converging. If they are, the pass is recorded, and once this occurs three times (in total), the interpolation loop will end. These three checks are done to ensure that the method has properly converged. The other step is the option for subsampling the final grid. Following standard convention (Reid, 1980), the output cell size should be one-fourth or one-fifth the line spacing distance. However, since this method is essentially “smearing” data in trend directions and not trying to develop a line-of-best-fit across flight lines, it has been found through extensive testing that in this method it can often assist the interpolation and trending process to set the interpolation cell size to half that of the conventional output cell size (i.e. one-eighth to one-tenth the flight line spacing). Once the interpolation is complete, the method can subsample the entire grid up to the larger, more conventional, cell size. For strong linear features, the smaller cell size can assist in making the features better trended due to the effect of subsampling. However a small cell size must be used with caution, as it can

often remove weaker linear features, as they will now have “farther” to trend when connecting trends found in the real data cells.

2.4 Synthetic Data Test

To test the new method, a 3 km x 3 km synthetic aeromagnetic data set (Figure 4) was built using PyGMI (Cole, 2015). The data set, similar to the one used by Geng et al. (2014), consists of four trends along the southern extent at angles of 0°, 15°, 30°, and 45° with respect to north, another trend along the northern extent running east-west, and three isolated blocks in the north-east. All sources are at a depth of 100 m, and are 50 m in their depth extent. The isolated blocks are 35 m x 35 m in lateral extent, and the dykes are 5 m wide. The data is represented as though it was measured at a flying height of 100 m, the magnetic inclination and declination are 72.10° and -10.12° respectively, the Earth’s magnetic field intensity was set to 55000 nT, and the anomalies have a magnetic susceptibility of 1 SI. After generating the synthetic dataset with 5 m cell size, it was corrupted with Gaussian noise at a standard deviation of 1 nT, a value suggested by Smith and Salem (2005) as being typical for a large value of aeromagnetic noise. The data set was then subsampled as though it was flown at 250 m line spacing, indicated by the solid black lines. The data set was interpolated onto a 50 m grid using bidirectional gridding (Figure 5), minimum curvature (Figure 6), kriging (Figure 7), and our new method (Figure 8). All maps were interpolated in Geosoft’s Oasis Montaj software (Geosoft, 2018).

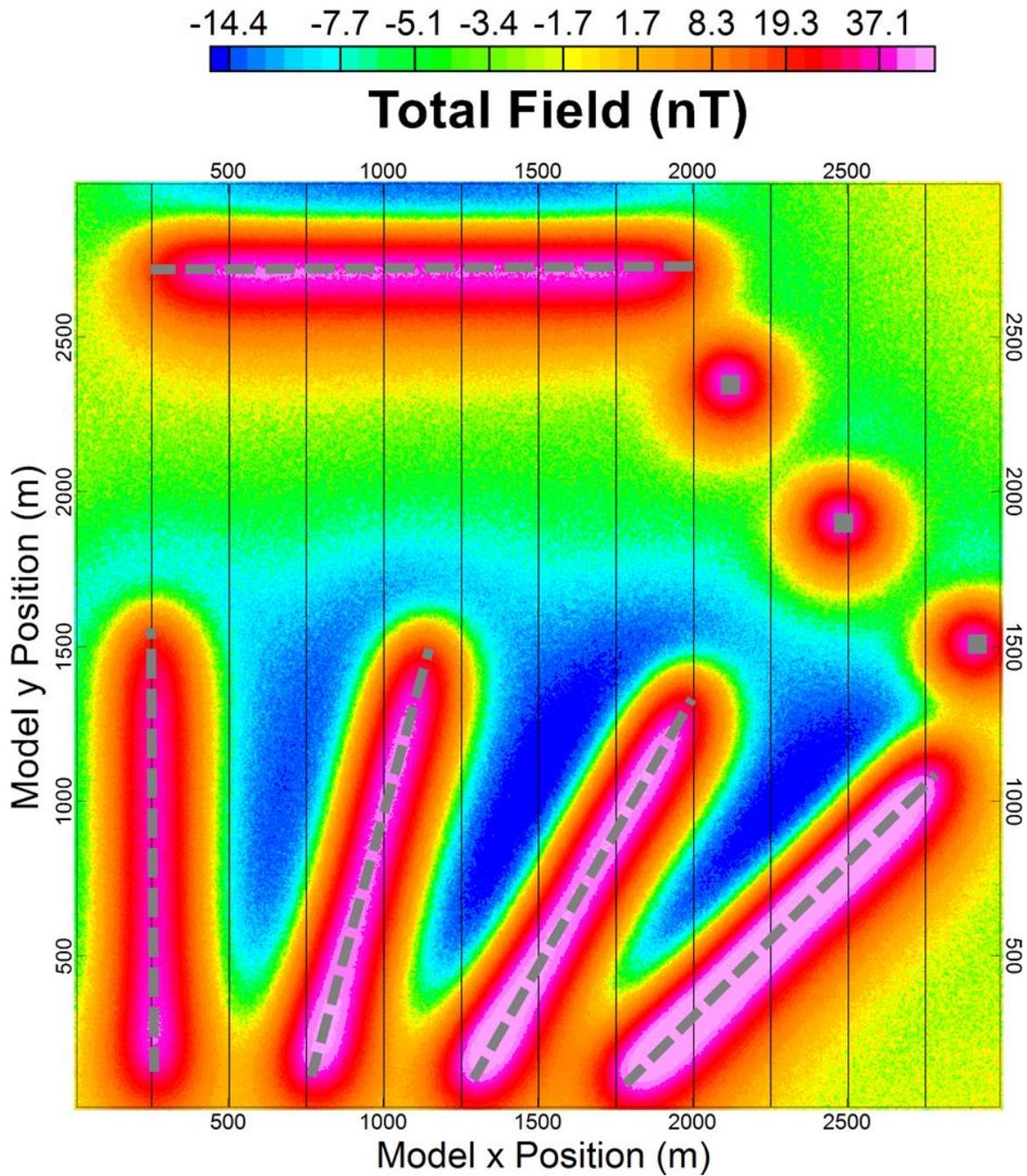


Figure 4. The full 3 km x 3 km synthetic dataset of 5 m cells with Gaussian noise of 1 nT standard deviation added. The solid lines represent the 250 m line spacing data used for interpolation. Dashed grey lines and boxes represent the locations of the sources.

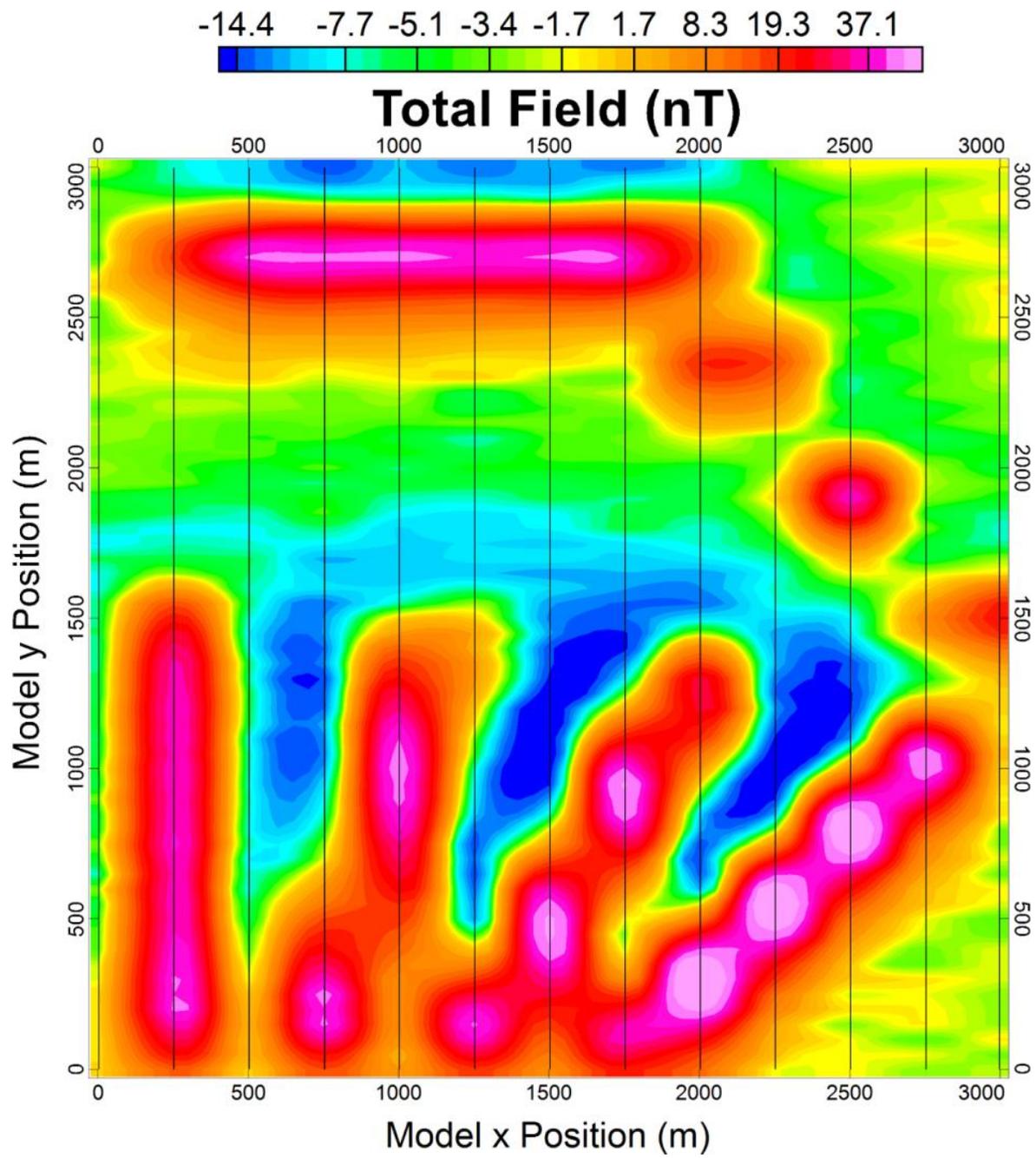


Figure 5. Interpolation of the 250 m line spacing synthetic data set using bidirectional gridding (Akima splines) at a cell size of 50 m. All parameters were set to default within the software for this interpolation.

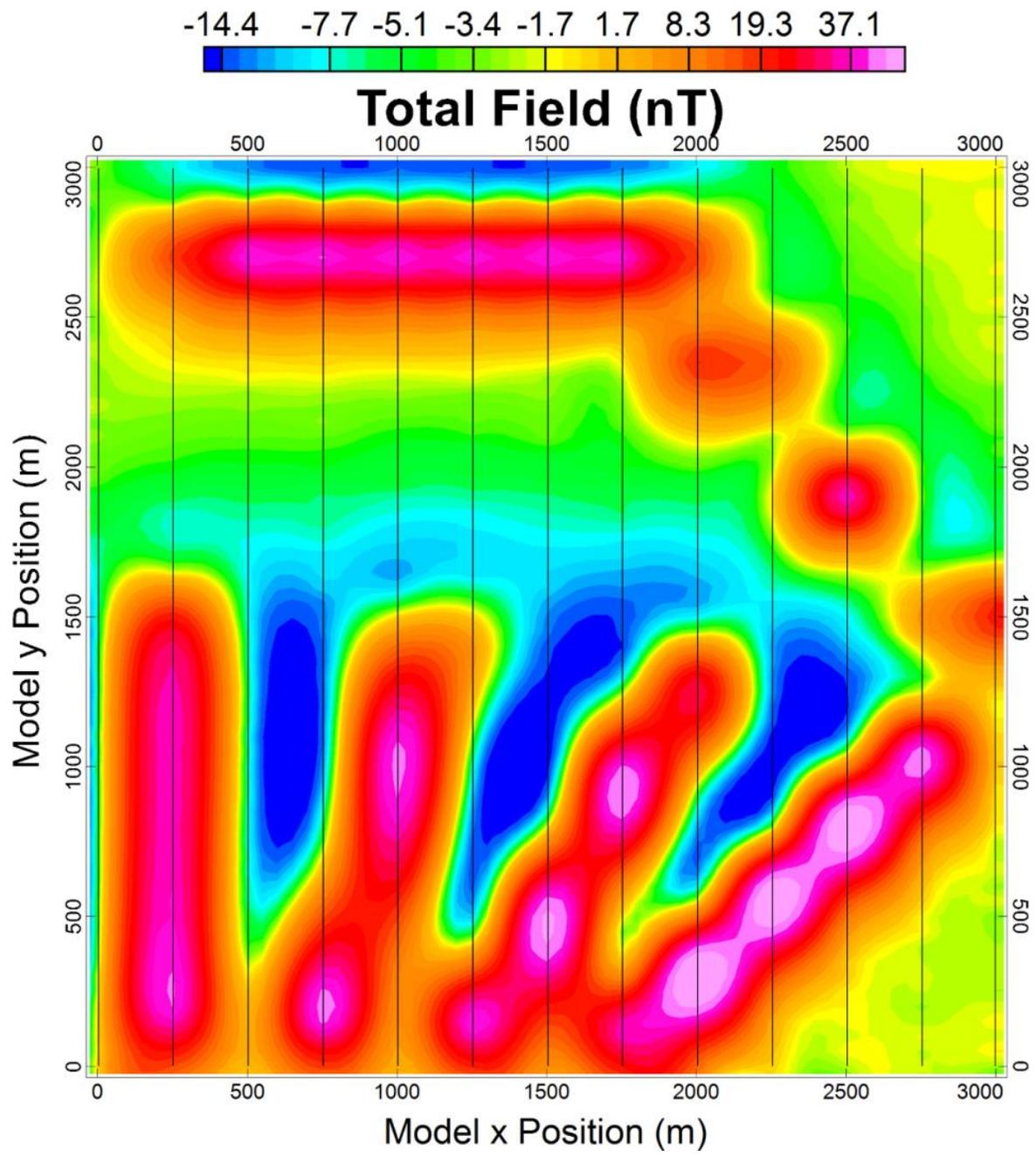


Figure 6. Interpolation of the 250 m line spacing synthetic data set using minimum curvature at a cell size of 50 m. The convergence criteria in the software were set to 99.5% pass tolerance, and 0.05% error tolerance, which it achieved in fewer than 100 iterations.

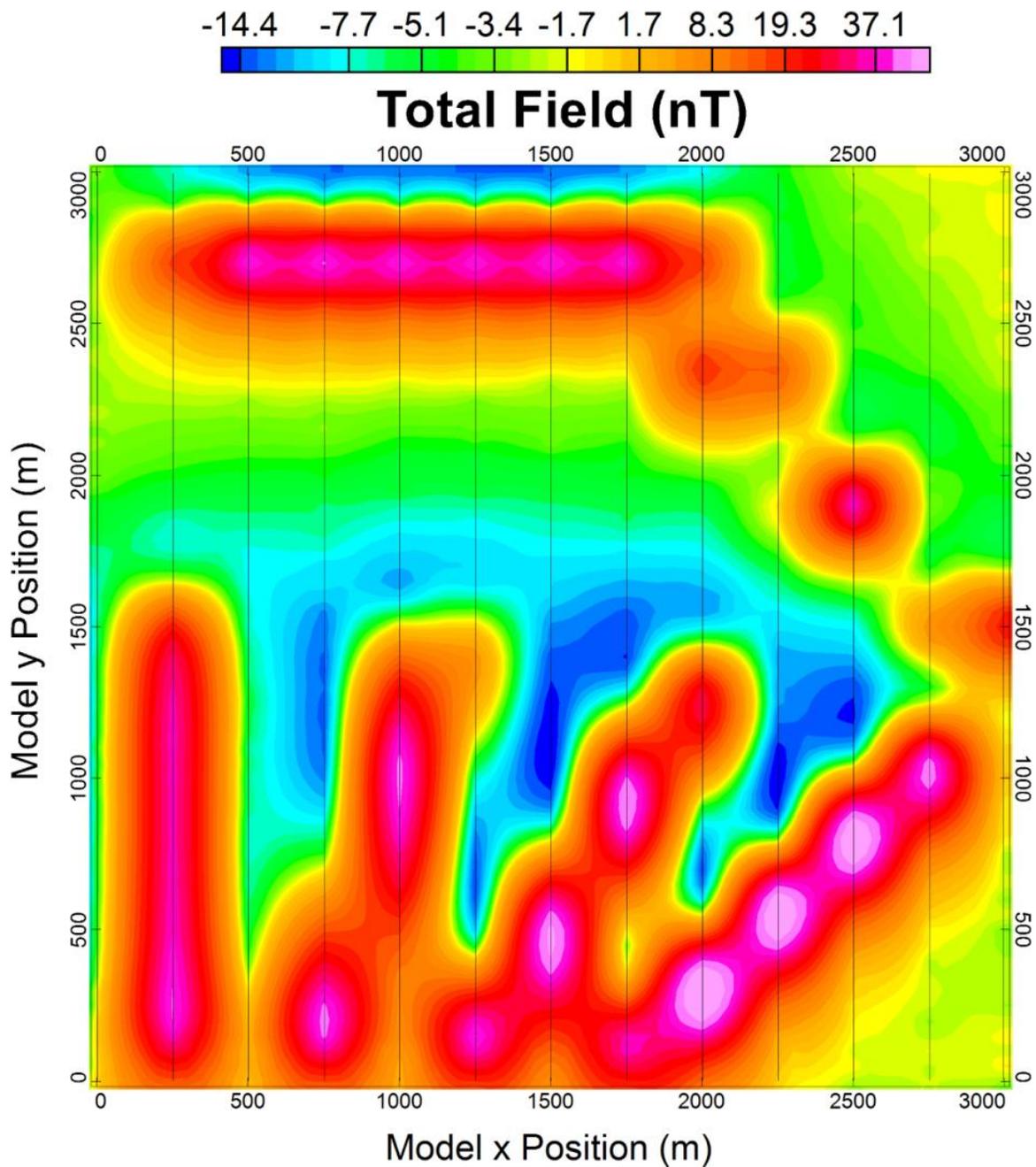


Figure 7. Interpolation of the 250 m line spacing synthetic data set using kriging at a cell size of 50 m. A spherical variogram model was used with the nugget set to zero, and the sill and range were found to be 309 and 1951, respectively (default parameters in the software used).

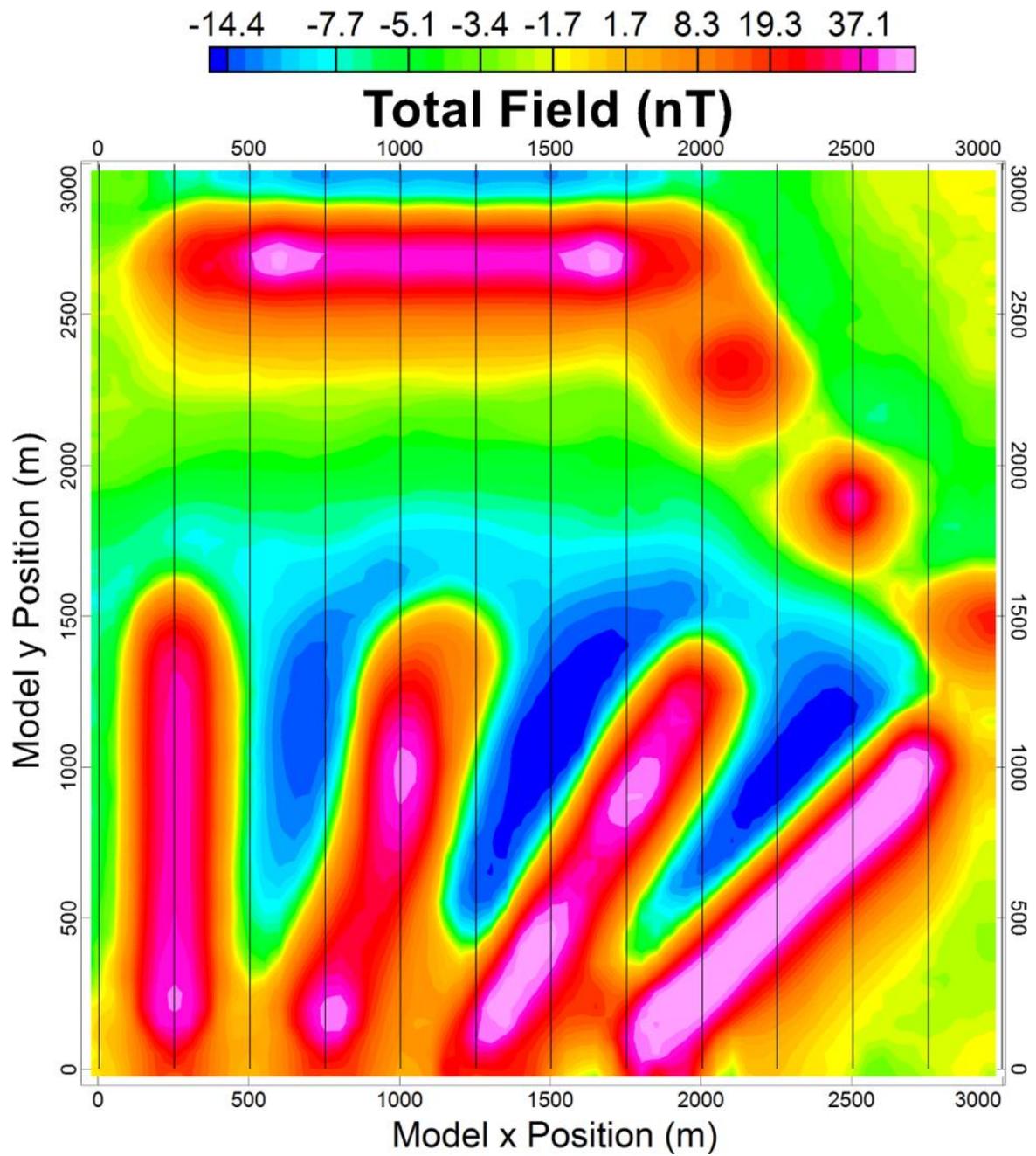


Figure 8. Interpolation of the 250 m line spacing synthetic data set using our new method at a cell size of 50 m (no subsampling was used), $\varphi = 125$ m, $\theta = 10^\circ$, and the trend strength set to 100%. The interpolation process stopped at 50 iterations.

All algorithms trended the east-west feature at the top of the image well. The three isolated features in the top right were not conjoined by any algorithm. Note that the bidirectional, minimum curvature, and kriging algorithms recover the amplitude of the middle isolated feature that was crossed by a flight line, but the other two features they shift to be on a flight line. The new algorithm does a better job at recovering the amplitude of the left-most circular feature, placing it between the flight lines. However the new method does not recover the center circular feature as well as the other methods. The 45° and 30° linear features at the bottom of the image show boudinage artifacts on the bidirectional, minimum curvature, and kriging grids, however the new method has effectively removed the artifacts. The 15° linear feature is essentially gridded as two separate anomalies in both the bidirectional gridding and the kriging results, whereas the minimum curvature and the new method have joined them. Figure 9 shows the residual plots of the methods, and Table 1 shows the minimums, maximums, means, medians, and standard deviations of these residuals.

Table 1. The minimum, maximum, mean, median, and standard deviation values for each of the residual plots from Figure 9.

Method	Minimum	Maximum	Mean	Median	Standard Deviation
Bidirectional gridding	-27.78	19.82	0.16	0.30	5.55
Minimum curvature	-23.60	16.69	0.32	0.17	4.98
Kriging	-27.64	17.80	0.07	0.47	5.81
New method	-22.63	18.50	-0.27	0.16	3.80

Unsurprisingly, bidirectional gridding has the most accurate result along the linear feature that runs perpendicular to the flight path, while the new method has a higher residual along the top edge the feature. This is likely due to the influencing effect of the low values along the edge of

the model, causing the new method to develop two thinner, linear features of polarizing values, rather than a single large feature with a drop off in value. The three standard methods have fairly similar residual grids, with minimum curvature being the most accurate. Overall however, the new method's residual values along the linear features are much smaller as compared to the other three. Additionally, the standard deviation of the residual is smallest in the new method. To be noted however is that there are larger residuals along several areas of real data in the new method. This is due to the new method sampling real data in a different, likely simpler way than the other three methods, and is accentuated by the large amount of noise added to the dataset.

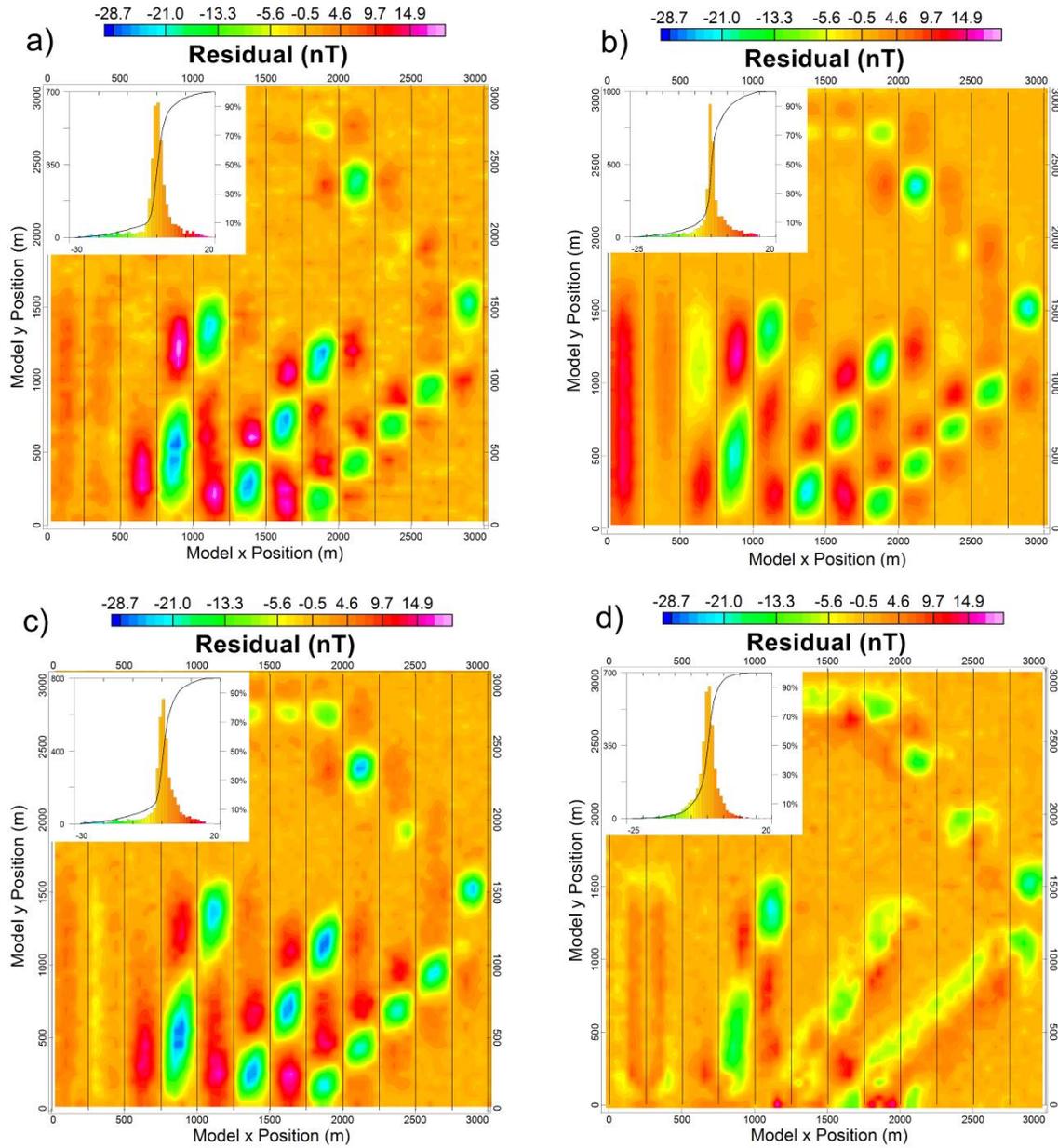


Figure 9. Residual plots and associated histograms of the synthetic model interpolations. The x -axes of the histograms are the residual value bins, and the y -axes are the number of cells that fall within the bins. a) Bidirectional gridding. b) Minimum curvature. c) Kriging. d) The new method.

2.5 Field Data Test

We then applied the method to a real-world example. An aeromagnetic data set from Nunavut, Canada (Overby-Duggan), was downloaded from the Natural Resources Canada geophysical data repository (Geological Survey of Canada, 2018), and a small section of it with linear features was extracted. The survey was flown at a line spacing of 400 m, and therefore it was interpolated with a cell size of 80 m using minimum curvature (Figure 10) and our new method (Figure 11). With some experimentation, we found that our method's results were better trended after completing the interpolation with a 40 m cell size, and subsampling up to the final 80 m cell size. Additionally, φ was set to half the line spacing at 200 m, θ was set to 5° , and τ was set to 100% to ensure the full trending effect would be applied. The automatic stopping option was used, completing after 97 iterations. Note that all coordinates are Eastings and Northings in UTM zone 13N.

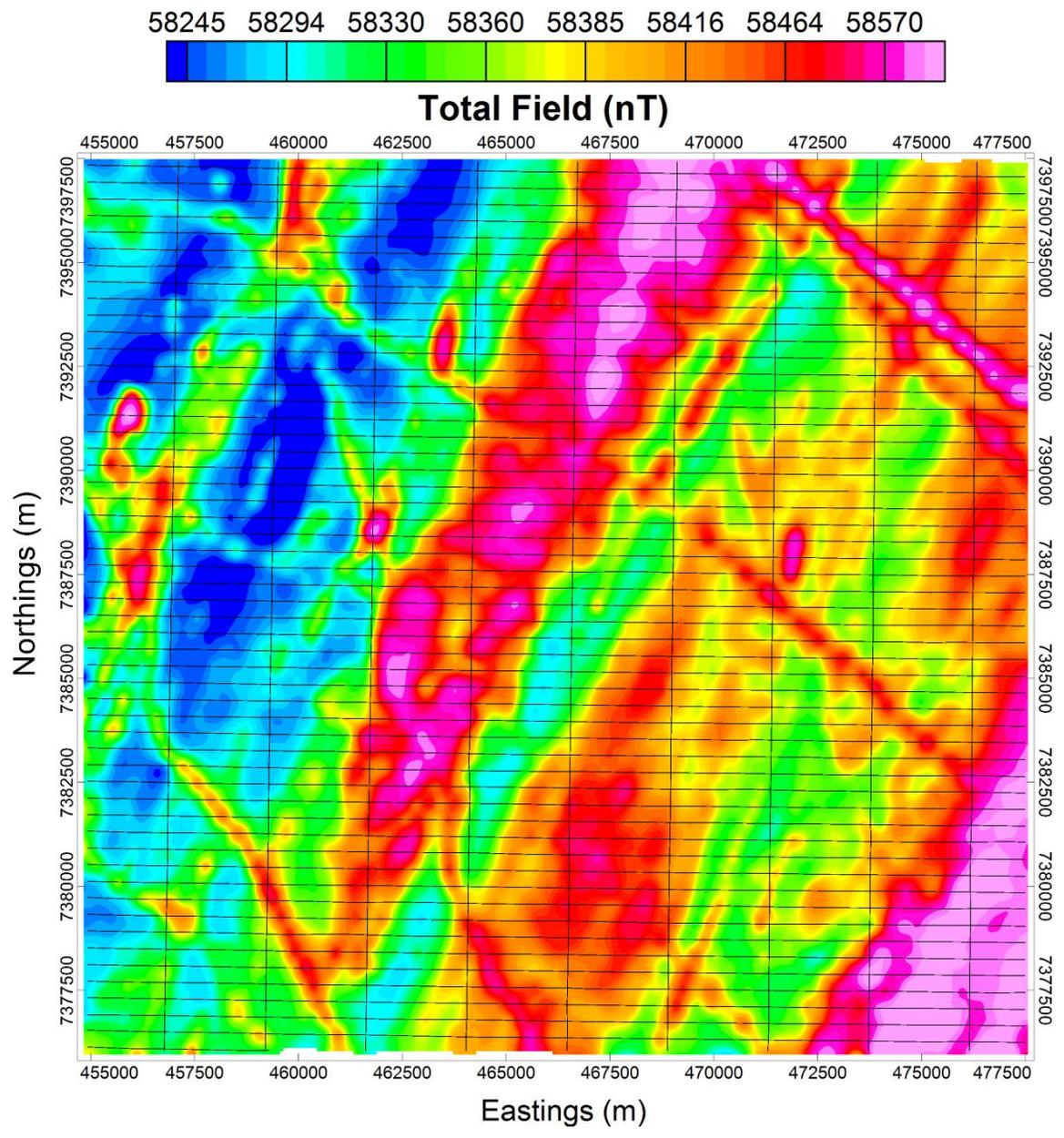


Figure 10. Minimum curvature interpolation of the Overby-Duggan aeromagnetic data set at a cell size of 80 m. The convergence criteria in the software were set to 99.5% pass tolerance, and 0.05% error tolerance, which it achieved in fewer than 500 iterations.

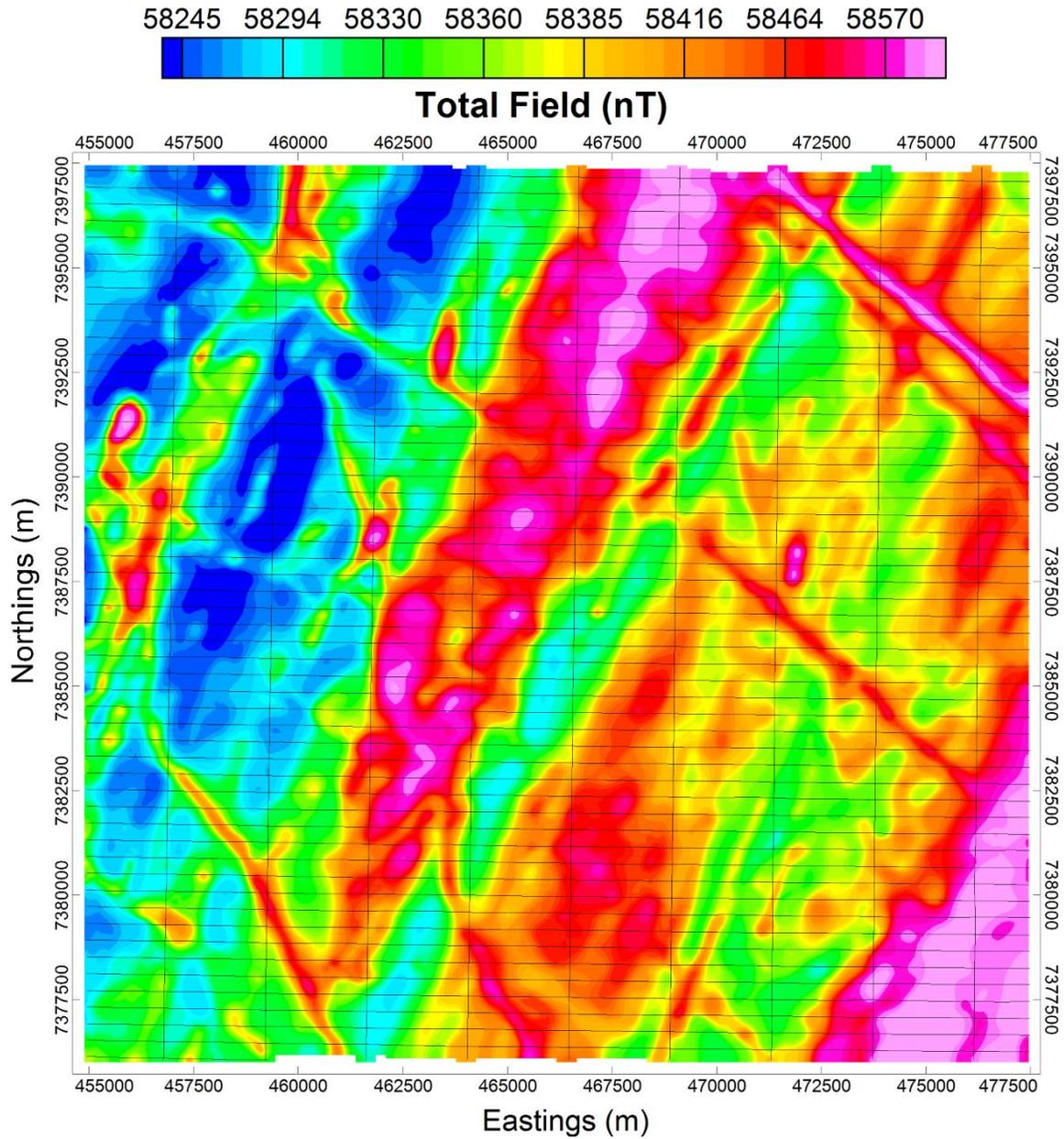


Figure 11. Our new method's interpolation of the Overby-Duggan aeromagnetic data set at a cell size of 80 m (subsamped from 40 m during interpolation), $\varphi = 200$ m, $\theta = 5^\circ$, and the trend strength set to 100%. The interpolation loop stopped at 97 iterations. The boudinage artifacts in the map are greatly reduced, having been trended into linear features.

Comparing the minimum curvature result with our new method, much of the beading artifacts have been removed, particularly the strong trends in the south-western corner, central-eastern side, and the north-eastern corner. Further, the linear trends appear as thinner, sharper features using the new method. The rest of the overall image has been kept similar to the minimum curvature results. We then applied a vertical derivative enhancement to both maps. Figure 12 is the vertical derivative from the minimum curvature method, and Figure 13 is calculated from the new grid. There is improvement to the result, with linear features less “beaded” and breaks in trends now connected. There are also some areas where no clear trend is occurring (e.g. compare the minimum curvature results with the new method’s results in the area at 7380000N, 467500E), and yet small linear features exist. This is most noticeable in the vertical derivative enhancement. If this was an area of interest, a lower trend strength may assist in interpretation, however since it is an area of low linear structure, minimum curvature is likely a more appropriate interpolation method to utilize.

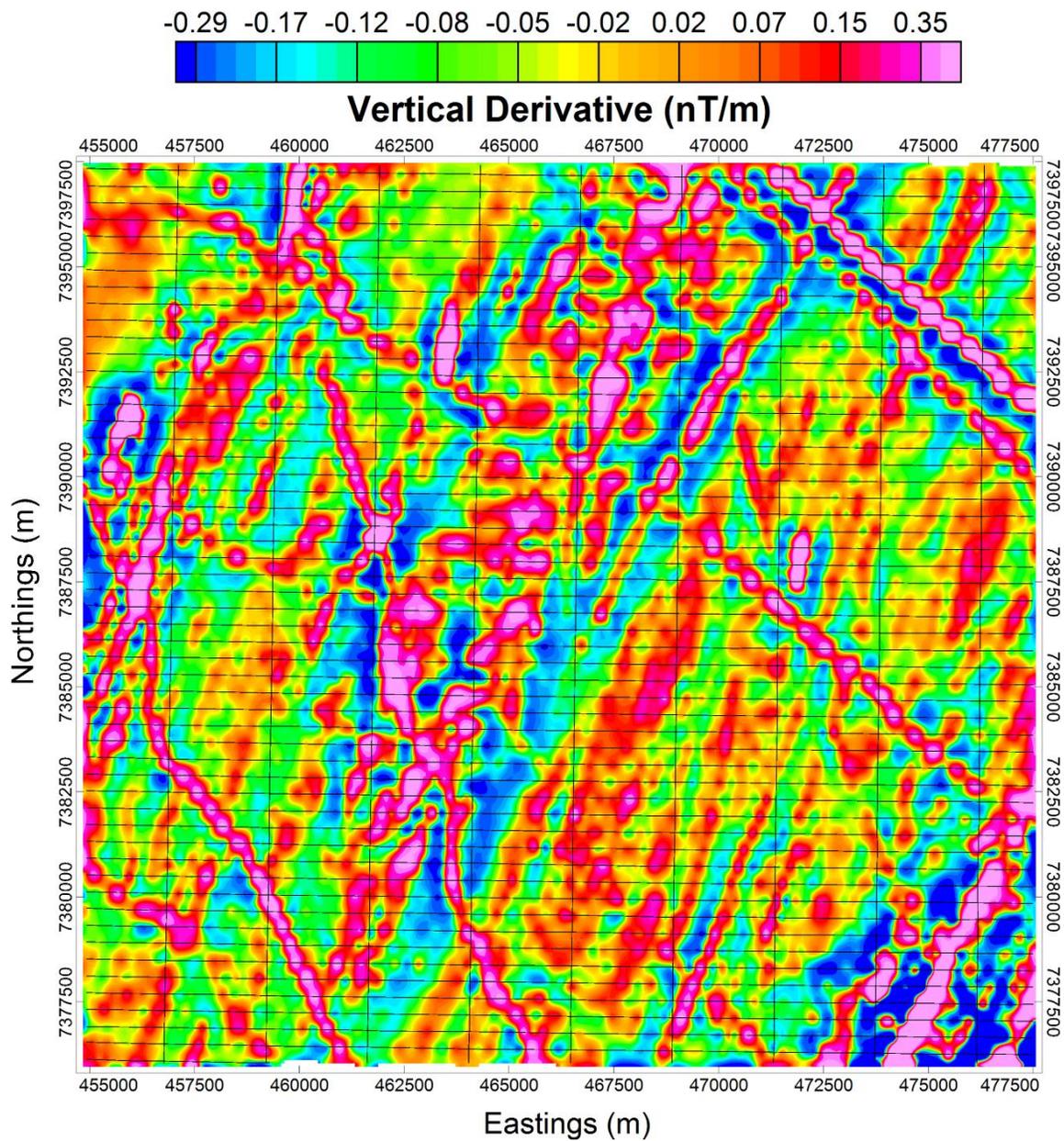


Figure 12. A vertical derivative map of the minimum curvature interpolation of the Overby-Duggan dataset.

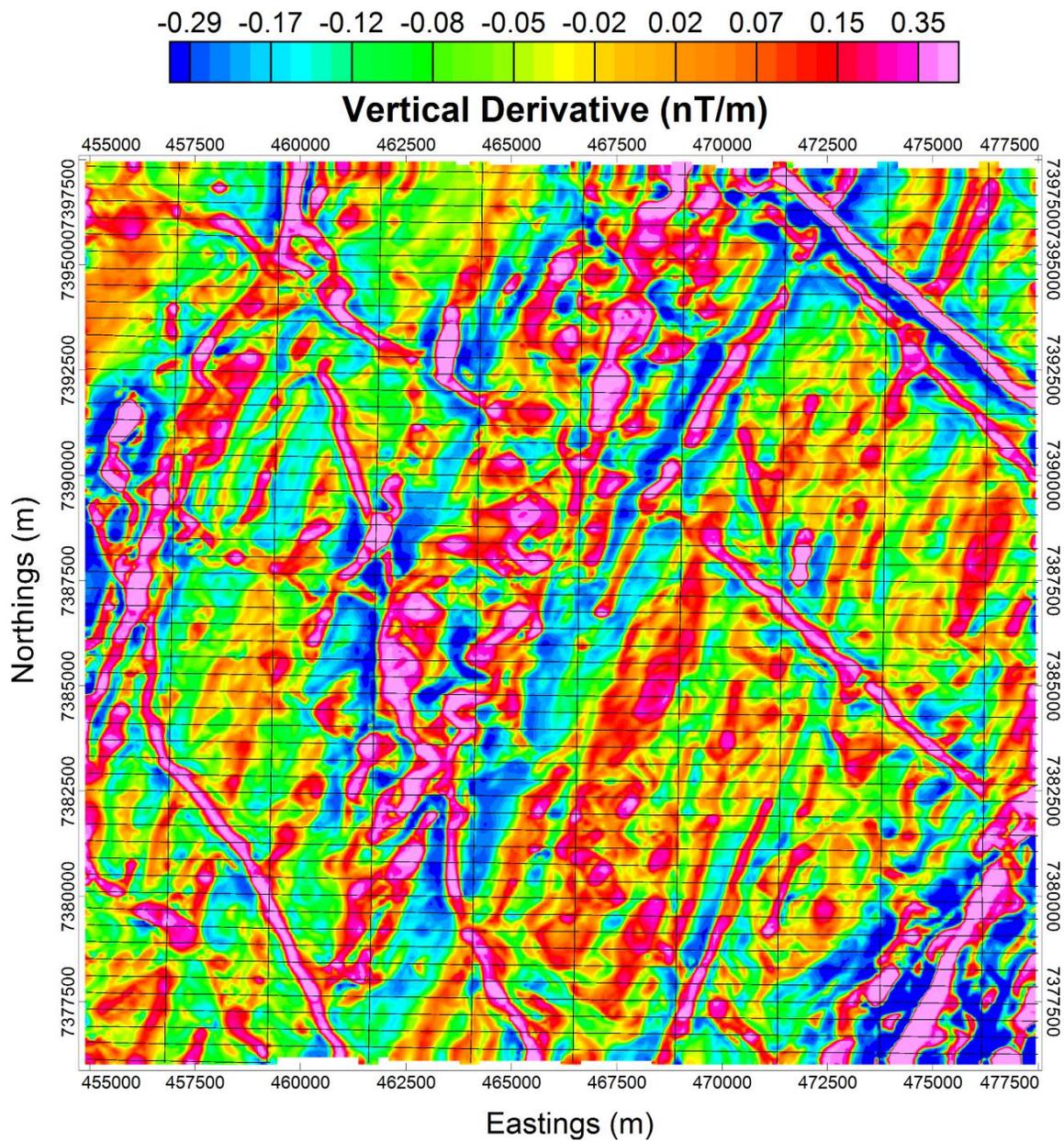


Figure 13. A vertical derivative map of our method's interpolation of the Overby-Duggan dataset.

2.6 Conclusions

When processing aeromagnetic data, many standard gridding methods have difficulty interpolating thin, linear features that lie at non-perpendicular angles to the flight lines. The resulting interpolation artifacts, often referred to as “beading” or “boudinage”, can make interpreting the data difficult, particularly when using analysis methods that involve derivatives, such as the vertical derivative enhancement. The goal of this research has been to develop a new interpolation method specifically for improving the results of datasets that contain thin linear features, such that they do not contain these artifacts. This iterative method utilizes a Taylor expansion of derivatives to interpolate data across the flight lines while maintaining linear features. However, to mitigate any discontinuities, this methodology must be applied to real data cells as well as the interpolated data cells. To honor the flight data, we then apply a “normalization” process, which returns the real data to its original values, while pulling the interpolated data along with it. To further enhance trends across flight lines, we apply this normalization along paths of highest anisotropy, as calculated using structure tensors.

After testing the method on both synthetic and field data, it can be concluded that this new method improves the resulting interpolation of this type of feature when compared to the widely-used interpolation methods of bidirectional gridding, minimum curvature, and kriging. Linear features that cross the flight lines at acute angles are maintained without the usual beading artifacts in both the total field and vertical derivative enhancement results.

This method involves several user-defined variables, and as such, will generally require some user experimentation. However, most data sets will result in fairly effective interpolations if the guidelines described in section 2.3.1 are followed. Additionally, since this method has been developed explicitly to solve the issues that can often occur to linear features, areas of data with little linear structure may result in weak trends that are noticeable in enhancements such as the vertical derivative. It is possible to reduce this trending by decreasing the trending parameter, τ .

2.7 Method's Source Code

Our method's C# source code is available at github.com/TomasNaprstek/Naprstek-Smith-Interpolation. The code can be run as a stand-alone method, or implemented into another program (for example, we have included a custom .dll for Geosoft's Oasis Montaj).

2.8 References

Abramowitz, M., and I. A. Stegun, 1970, Handbook of mathematical functions with Formulas, Graphs, and Mathematical Tables: Dover Publications.

Akima, H., 1970, A new method of interpolation and smooth curve fitting based on local procedures: Journal of the ACM, 17, 589-602.

Bhattacharyya, B. K., 1969, Bicubic spline interpolation as a method for treatment of potential field data: Geophysics, 34, 402-423.

Billings, S. D., R. K. Beatson, and G. N. Newsam, 2002, Interpolation of geophysical data using continuous global surfaces: Geophysics, 67, 1810–1822.

Briggs, I. C., 1974, Machine contouring using minimum curvature: *Geophysics*, 39, 39-48.

Cole, P., 2015, PyGMI, <https://github.com/Patrick-Cole/pygmi>, accessed 20 October 2015.

Cordell, L., 1992, A scattered equivalent-source method for interpolation and gridding of potential-field data in three dimensions: *Geophysics*, 57, 629-636.

Fehmers, G. C., and C. F. Höcker, 2003, Fast structural interpretation with structure-oriented filtering: *Geophysics*, 68, 1286-1293.

Geological Survey of Canada, 2018, Canadian Aeromagnetic Data Base, Airborne Geophysics Section, GSC - Central Canada Division, Geological Survey of Canada, Earth Sciences Sector, Natural Resources Canada. Data set accessed 30 August 2018 at <http://gdr.agg.nrcan.gc.ca/gdrdap/dap/search-eng.php>.

Geng, M., Q. Yang, and Y. Yuan, 2014, Trend enhancement in aeromagnetic maps using constrained coherence-enhancing diffusion filtering: *Geophysics*, 79, no. 1, J1-J9.

Geosoft Inc., 2018, Oasis Montaj, Toronto, Canada.

Guo, L., X. Meng, and L. Shi, 2012, Gridding aeromagnetic data using inverse interpolation: *Geophysical Journal International*, 189, 1353-1360.

Hale, D., 2010, Image-guided 3D interpolation of borehole data: CWP report, 656.

Hall, M., 2007, Smooth operator: Smoothing seismic interpretations and attributes: *The Leading Edge*, 26, 16-20.

Hansen, R. O., 1993, Interpretive gridding by anisotropic kriging: *Geophysics*, 58, 1491-1497.

Keating, P., 1997, Automated trend reinforcement of aeromagnetic data: *Geophysical Prospecting*, 45, 521–534.

Mallet, J. L., 1989, Discrete smooth interpolation, *ACM Transactions on Graphics (TOG)*, 8, 121-144.

Mendonca, C. A., and J. B. C. Silva, 1994, Equivalent data concept applied to the interpolation of potential-field data: *Geophysics*, 59, 722–732.

Mendonca, C. A., and J. B. C. Silva, 1995, Interpolation of potential-field data by equivalent layer and minimum curvature: A comparative analysis: *Geophysics*, 60, 399–407.

Pilkington, M., and W. R. Roest, 1998, Removing varying directional trends in aeromagnetic data: *Geophysics*, 63, 446-453.

Reid, A. B., 1980, Aeromagnetic survey design, *Geophysics*, 45, 973-976.

Smith, W. H. F., and P. Wessel, 1990, Gridding with continuous curvature splines in tension: *Geophysics*, 55, 293-305.

Smith, R. S., and M.D. O'Connell, 2005, Interpolation and gridding of aliased geophysical data using constrained anisotropic diffusion to enhance trends: *Geophysics*, 70, no. 5, V121-V127.

Smith, R. S., and A. Salem, 2005, Imaging depth, structure, and susceptibility from magnetic data: The advanced source-parameter imaging method: *Geophysics*, 70, no. 4, L31-L38.

Swain, C. J., 1976, A FORTRAN IV program for interpolating irregularly spaced data using the difference equations for minimum curvature: *Computers and Geosciences*, 1, 231-240.

Sykes, M. P., and U.C. Das, 2000, Directional filtering for linear feature enhancement in geophysical maps: *Geophysics*, 65, 1758-1768.

Wu, X., 2017, Building 3D subsurface models conforming to seismic structural and stratigraphic features: *Geophysics*, 82, no. 3, IM21-IM30.

Yunxuan, Z., 1993, Radon Transform Application to the Improved Gridding of Airborne Geophysical Survey Data: *Geophysical Prospecting*, 41, 459-494.

Chapter 3

3 Applications of Machine Learning to Aeromagnetic Spatial Interpolation

3.1 Abstract

Advances in machine learning (ML) theory, as well as the availability of both the computation power required and the programming libraries for languages such as Python and R have enabled easy implementation and usage of these algorithms by non-experts. These methods are new tools to solve both novel and established problems in a wide array of fields, but must be investigated thoroughly, as an effective solution for one problem may not be appropriate in a related context, despite similarity. Following approaches in seismic imaging and remote-sensing, we have applied two ML methods, random forest (RF) regression and support vector regression (SVR), to the problem of interpolating aeromagnetic data. Developing three sets of spatial predictors, trained RF and SVR models were created and applied to a test dataset. Results were mixed, with one predictor set showing poorer results than standard interpolation methods such as minimum curvature, and the other two predictor sets resulting in interpolations that were as good as or marginally better than the standard methods. The most effective predictor set was a “filter”-based approach, where several standard interpolation methods were fed into the ML models as predictors, along with other relevant neighbourhood information. This allowed the ML models to produce interpolations that were incrementally better than any one of the standard interpolations. Using a real-world dataset, it is found that this approach is moderately effective, resulting in maps that can be seen as a mix of the standard interpolations, utilizing the strengths of each

method where appropriate. We conclude that ML can offer improvement over standard interpolation methods for aeromagnetic data if utilized properly. However, we also note that the effort required for the small improvement with our approach may not be worthwhile unless the data is particularly difficult to interpolate.

3.2 Introduction

Machine learning (ML) is the general term describing the application of algorithms developed to discover relationships and patterns within datasets. Many ML methods can be used to develop advanced regression models to predict outputs, based on various inputs of data. They have been extensively applied in a wide variety of contexts, and across many disciplines as the ease of using ML methods has increased and become more widely available in commonly used programming languages. In this study, we investigate the potential of ML as applied to aeromagnetic interpolation, and therefore focus on the realm of spatial problems. This is an area that a large number of ML studies have investigated in a variety of different contexts. Many have focused on three of the most common ML methods: support vector regression (SVR), random forest (RF), and neural networks (NNs). Most methods originated in image processing and computer vision, where the primary issues addressed include enhancing low resolution images and reconstructing noisy or corrupted images (e.g. Ni and Nguyen, 2007; Tang et al., 2011; Ohashi and Torgo, 2012; Huang et al., 2015). In these cases, the interpolated cells of data are often surrounded by a large number of spatially close real data cells, allowing for very accurate interpolations, even when a high percentage (>50%) of the original image is corrupted.

ML has been applied to aeromagnetic data in other contexts. For example, aeromagnetic data has been one of several datasets used for predicting lithology or mineral occurrence in a number of studies (e.g. Yu et al., 2012; Cracknell and Reading, 2013; Cracknell and Reading, 2014; Rodriguez-Galiano et al., 2015). However, specifically in aeromagnetic data interpolation, there is limited literature available. Liu et al. (2018) present a methodology for the reconstruction/interpolation of magnetic data, primarily for missing base station data. They utilize SVR, RF regression, gradient boosting regression, and long short-term memory recurrent NNs, concluding that the final of these is the most effective on both line and gridded data. However, the grid data is reconstructed with only a single line of data missing between lines of collected data. In the context of aeromagnetic data, we would generally expect somewhere in the range of 3 to 5 lines of data to be interpolated.

If we expand the field of application of ML to other geophysical contexts, spatial interpolation with ML has been used extensively. For instance, seismic data has had several developments in ML-based interpolation. Jia and Ma (2017) describe an effective interpolation method that utilizes SVR with window-based predictors. It results in comparative or better results than the $f-x$ method on both synthetic and field data examples. However, their method requires a pre-interpolation step, thus resulting in their method essentially being a filter, albeit an effective one. They later improve the efficiency of their method (Jia et al., 2018) using “intelligent” training data choices. Instead of utilizing all data when training the regression model, they only choose to use data that has a high “variance”, and is therefore more likely to provide useful information.

They show that this approach is extremely effective, substantially reducing the time required for their method, while keeping the results approximately the same as the full-dataset approach.

A series of studies in remote sensing have proven that the inclusion of spatial variables can greatly enhance the results of the prediction of geophysical characteristics with ML. Since most ML algorithms do not inherently involve spatial relations, including spatial variables can lead them to be more accurate and similar to more traditional interpolation methods. For instance, Li et al. (2011) compare a large number of interpolation methods, including several novel ML approaches, on the spatial interpolation of mud content in southwestern Australia. They conclude that RF, when using spatial predictor variables such as the latitude and longitude of each predicted cell, was among the most effective of all interpolation methods. They also conclude that their approach of combining RF with ordinary kriging gave better results as compared to using either approach individually. Behrens et al. (2018) also studied the usage of spatial predictors in ML, applying it to the interpolation of soil maps. They show that utilizing specific spatial relations, referred to as Euclidean Distance Fields (EDFs), improve ML interpolation results. Similarly, Hengl et al. (2018) show that Euclidean buffer distances can likewise improve results, by associating all locations with each other, based on their relative distance from one another. Overall, these studies show the potential of ML for spatial interpolation.

In this research, we have applied ML to the spatial interpolation of aeromagnetic data.

Aeromagnetic interpolation is a well-established field of research (e.g. Bhattacharyya, 1969; Briggs, 1974; Swain, 1976; Smith and Wessel, 1990; Cordell, 1992; Hansen, 1993; Mendonca

and Silva, 1994, 1995; Keating, 1997; Billings et al., 2002; Smith and O'Connell 2005; Geng et al., 2014; Naprstek and Smith, 2019), as it is a critical step in the processing of the data, and the quality of interpolation can always be improved. Aeromagnetic data is collected along flight lines with high-density sampling, between which there is a complete absence of data. As such, interpolating to fill in these gaps with the most reasonable data is a problematic task. There are many effective methods currently available for approaching this problem; however ML offers a potential new avenue of solving this problem. We approach this by applying two of the most widely used and effective ML methods found in the above literature, RF and SVR, to the spatial interpolation of aeromagnetic data, focusing on the different predictors that can guide the ML. These results are compared to standard aeromagnetic interpolation methods to better understand the future potential in this field. We begin by describing the two ML methods, and the supporting theory required to understand our investigation. We then describe three separate approaches to our goal of aeromagnetic interpolation with ML. The first approach is an attempt to follow similar studies by both training and applying a ML model on the same dataset, in an effort to make a highly specialized model. The second approach instead focuses on developing a more general model, by training the ML methods on a series of datasets, before applying them to an independent, but similar aeromagnetic dataset. Finally, we extend the second approach into a "filter"-based method, by utilizing standard aeromagnetic interpolation methods to assist in guiding the ML algorithms. To evaluate each method, we utilize a synthetic dataset that we can determine the exact errors on. The best approach is then applied to a real data set and compared to standard aeromagnetic interpolation methods in the case study section.

3.3 Method

There are a wide variety of ML methods; however in the context of spatial interpolation, RF and SVR regression have been the most widely explored. In the above referenced literature, the results overwhelmingly show that RF regression is the most accurate, as well as providing smoother results with fewer outliers. The reasoning for this is generally attributed to the core feature of RFs: they average over a large number of de-correlated regression trees (Were et al., 2015; Huang et al., 2015), which prevents over-fitting and smooths out the final predictions (Liu et al., 2018). However, other ML methods are shown to produce encouraging results, such as Jia and Ma (2017) where they applied SVR to the interpolation of seismic data. We approached this research by focusing on both RF and SVR methods, as implemented in the scikit-learn Python library (Pedregosa et al., 2011).

RF regression is a method that takes the ML method of decision trees (Breiman et al., 1984), and develops a large “forest” of these trees each with a result which can be averaged to give the final result (Breiman, 2001). Each tree selects (or “bags”) a subset of the total predictor dataset to analyze (Breiman, 1996). No data is removed however, ensuring that while each tree has its own subset of the total data, there is overlap between bagged data. This ensures that the trees are not cross-correlated and reduces over-fitting (Huang et al., 2015; Liu et al., 2018), which is a common problem in regression methods. The out-of-bag data for each tree can be used to evaluate its accuracy, and can also be used to determine each predictor’s importance to the overall model (Rodriguez-Galiano et al., 2015; Were et al., 2015). Due to the averaging approach

of RF, it is generally more robust than other ML methods, and able to handle noise and outliers well (Breiman, 2001; Li et al., 2011).

The core principle of SVR (Cortes and Vapnik, 1995; Drucker et al., 1997) is to map the predictor variables to a high-dimensional space in an effort to simplify non-linear relations (Were et al., 2015; Jia and Ma, 2017; Behrens et al., 2018; Liu et al., 2018). This mapping is completed through the use of a user-defined kernel function (Gunn, 1998). Once in this new space, a hyperplane (dependent on the kernel chosen) is developed by minimizing a loss function based on user-defined cost and slack variables (Rodriguez-Galiano et al., 2015). These variables allow for more or less flexibility in the hyperplane and any errors it may encounter. Once developed, new data can be analyzed using this hyperplane's solution.

As an end-user, we must understand and utilize two primary aspects of ML: the input data known as predictors, and the user-defined parameters that guide the ML model's development known as "hyperparameters". The predictors are variables that relate to the final output of the ML model, and are represented entirely throughout the dataset(s) being analyzed. It is necessary that any value that is to be predicted through a ML model has associated values for all predictors the model was trained with. Predictors that are relevant to the final data should be chosen carefully, as uncorrelated data has the possibility to influence the ML model development in negative ways. The focus of our study is the effect of various predictors on the resulting aeromagnetic interpolation, and which predictors will lead us to the most accurate interpolated grid. The hyperparameters vary depending on the ML method, but include aspects such as what

kernel is used in SVR or how many trees should be developed in RF regression. They can be thought of as similar to variables such as “tension” in minimum curvature interpolation, guiding the final result while leaving the core methodology the same. In each of our approaches below, we completed an extensive series of tests to determine effective hyperparameters. Note that if not explicitly mentioned in the methodology, we left all other hyperparameters as their default value in the scikit-learn functions. For further information on these two ML methods, we refer the reader to the sources and associated literature we have cited. Additionally, we recommend exploring the extensive online resources and tutorials for these (and other ML) methods that can greatly assist in both understanding the theory behind the methodologies, as well as implementing them for applications such as the one we present here. For specific information on the RF and SVR methods we used and their associated hyperparameters, see the scikit-learn functions *sklearn.ensemble.RandomForestRegressor* and *sklearn.svm.SVR*.

To test our research, a series of synthetic datasets were forward modelled using PyGMI (Cole, 2019) at various cell sizes ranging from 20 m to 50 m. They were then subsampled as though they were flown in aeromagnetic surveys, with flight-line spacings equal to approximately five times the cell sizes. This split each dataset into two, as shown in Figure 14: flight data (grey cells) with cells to be interpolated (white cells), and original data without flight data. The simulated flight dataset is used as raw data input for the ML algorithms, and the original synthetic dataset is used for both training and comparison purposes.

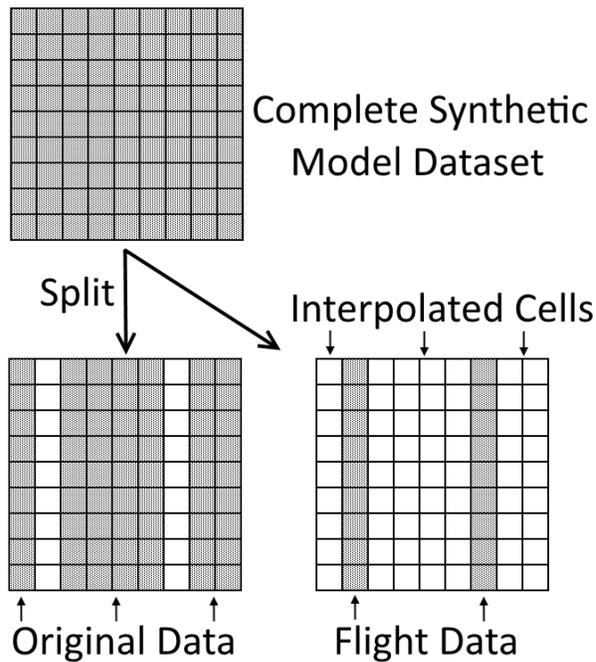


Figure 14. An example showing the definition of each type of cell after splitting each synthetic dataset. The interpolated cells are initially empty, and to be filled through the ML interpolation process.

Specifically due to a requirement of ML methods, all datasets were individually normalized to range from zero to one, rather than their original total field value in nanoteslas. This is a recommended approach for most ML methods, as simple linear discrepancies in data ranges between datasets can potentially cause issues when training models. One synthetic dataset was chosen as the “test” data to which all methods would be applied to enable a comparison of results, and therefore was not utilized in any training. The full dataset can be seen in Figure 15. For each predictor set and ML method, a series of statistics were calculated, and the summary of all of these are shown in the Results section of this paper.

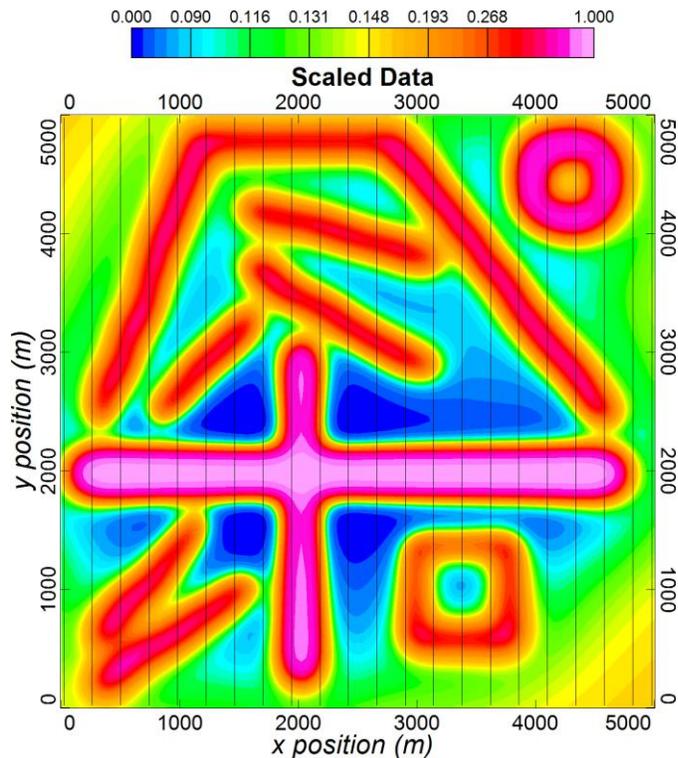


Figure 15. The synthetic dataset developed in PyGMI (Cole, 2019) that was utilized for testing the methods developed here. The vertical black lines show the simulated flight lines that were used as the input data for interpolation. Note that all plots of this dataset have the same colour range, and therefore only this plot shows the colourbar.

3.3.1 Predictor Set 1

The first technique we investigated involved both training and applying a ML method on the same dataset. This followed other approaches in the literature, such as Cracknell and Reading (2014). In it they attempt to predict lithology in a specific geographic location, by utilizing data including a digital elevation model (DEM), the latitude and longitude coordinates of each location, and Landsat spectral imagery. Each of these were used as a predictor dataset that covers

the entire area, and were combined with an incomplete dataset of ground truth data for the lithology. By training a ML method on the predictor datasets with the associated ground truth data, all other locations without a lithology ground truth can be predicted. Due to the uniqueness of the ML model, it therefore would not give reliable results in another area, as it has been trained to produce the best result for that specific geographical area.

Unfortunately, aeromagnetic data interpolation will generally not have the advantage of accessing multiple other datasets to utilize as predictors. In fact, aeromagnetic data is often one of the initial datasets used to investigate a location for mineral exploration, due to the ease and cost-effectiveness of the data collection process. For instance, in the Cracknell and Reading (2014) example described, aeromagnetic data was used as one of the predictors. Therefore, we utilized this ML approach with aeromagnetic interpolation using only data that we assumed all aeromagnetic surveys would contain. This is limited to only the aeromagnetic data itself and the associated location information. We therefore followed several studies from the literature, and developed a set of predictors based on the physical location of aeromagnetic data within a specific dataset:

- 1) Neighbourhoods (Ohashi and Torgo, 2012):
 - a. Mean of 3 cell radius (neighbourhood 1)
 - b. Mean of 6 cell radius (neighbourhood 2)
 - c. Mean of 12 cell radius (neighbourhood 3)
 - d. Distance weighted mean (DWM) of 3 cell radius (DWM neighbourhood 1)

- e. DWM of 6 cell radius (DWM neighbourhood 2)
 - f. DWM of 12 cell radius (DWM neighbourhood 3)
 - g. Standard deviation of:
 - i. Neighbourhood 1
 - ii. Neighbourhood 2
 - iii. Neighbourhood 3
 - h. Ratios of the standard means and the distance weighted means:
 - i. Neighbourhood 1 / Neighbourhood 2
 - ii. Neighbourhood 1 / Neighbourhood 3
 - iii. Neighbourhood 2 / Neighbourhood 3
 - iv. DWM neighbourhood 1 / DWM neighbourhood 2
 - v. DWM neighbourhood 1 / DWM neighbourhood 3
 - vi. DWM neighbourhood 2 / DWM neighbourhood 3
- 2) Euclidean distance fields (EDFs) (Behrens et al., 2018):
- a. Distance from minimum x value (essentially an x position)
 - b. Distance from minimum y value (essentially a y position)
 - c. Distance from the North-Western corner of the area
 - d. Distance from the North-Eastern corner of the area
 - e. Distance from the South-Western corner of the area
 - f. Distance from the South-Eastern corner of the area
 - g. Distance from the center of the area

3) Buffer distances (Hengl et al., 2018):

a. Distance from data point to every other data point in the area

Therefore, there were 22 pre-defined predictors (neighbourhoods and EDFs), and a dataset-specific number of buffer distance predictors. The neighbourhood predictors were expected to generate basic information regarding the surroundings of each cell, as well as the effective change that occurred as the neighbourhood size was increased. The three radii chosen (three cells, six cells, and twelve cells) were set to represent the local values of: the nearest flight line data, the flight line data on either side (of the cell being analyzed), and the four nearest flight lines respectively. This assumes that the data has been gridded following standard aeromagnetic convention of a grid cell size equaling one-fourth to one-fifth the flight line spacing. The EDFs were expected to assist the interpolation by developing linear or radial trends from different reference points. Thus, if regional gradients were present, they could be accounted for simply by the interpolated cell's position in the area. Finally, the buffer distances were expected to develop high-dimensional (i.e. many-variable) relationships, such that, for instance, if a specific cell was found to be influential, another cell's relative proximity to it could be an indicator of its interpolated value. These predictors were calculated for all simulated flight data cells, and the aeromagnetic response for each flight data cell was associated with its set of predictors. These were used to train a ML model using both the RF and SVR methods. The predictors were then calculated for all interpolated data cells, and using the trained model, predicted data for the interpolated data cells were developed (Figure 16). The hyperparameters used for the RF method

can be seen in Table 2. No appropriate solution could be developed using the SVR method with this predictor set.

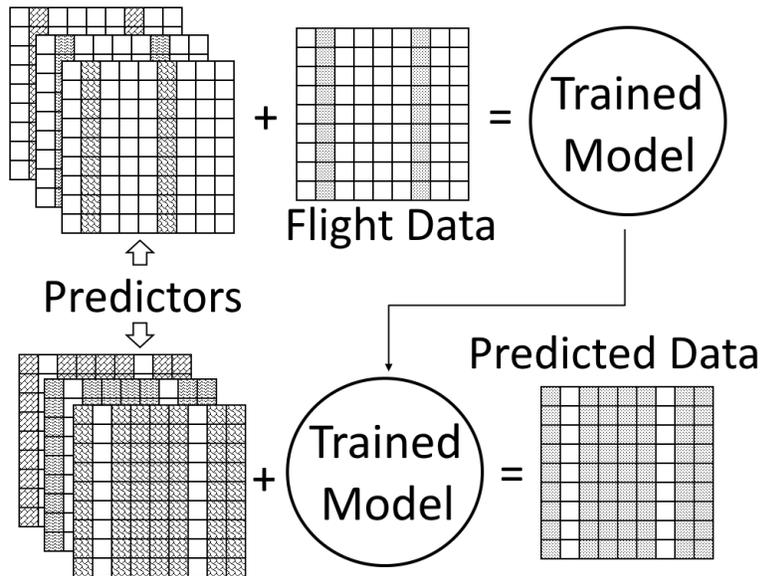


Figure 16. A depiction of the training process (top row) and the prediction process (bottom row) for predictor set 1. The predictors are a series of datasets (indicated by the multiple grids) that are calculated using the flight data only. With this approach, all training is completed with a single dataset's known results (flight data).

Table 2. The hyperparameter choices for the RF method with predictor set 1.

Hyperparameter	Property Set To:
<i>n_estimators</i>	100
<i>min_samples_leaf</i>	5
<i>min_samples_split</i>	10
<i>max_features</i>	log2
<i>max_depth</i>	50

3.3.2 Predictor Set 2

Our second predictor set utilized some of the same predictors as in predictor set 1; however, the overall training approach was adjusted. By training a ML method on a series of synthetic datasets where the full ground truth is known, we developed a model that could then be applied to other datasets for evaluation. This meant that the area-specific predictors of EDFs and buffer distances could not be used in this approach, as these predictors required locational information unique to each dataset. To replace them, the 20 closest real data points to each interpolated cell, analogous to simple inverse distance weighting, were included. As these predictors involve only relative straight-path separation between an interpolated cell and real data cells, these could be calculated for any dataset. We once again utilized the neighbourhood predictors; however, through extensive empirical testing, it was found that in this approach, many of them offered no benefit to the final result. Therefore, they were removed. This gave 26 predictors:

- 1) DWM of all real data values in a radius of:
 - a. 3 cells (neighbourhood 1)
 - b. 6 cells (neighbourhood 2)
 - c. 12 cells (neighbourhood 3)
- 2) Ratio of:
 - a. Neighbourhood 1 / Neighbourhood 2
 - b. Neighbourhood 1 / Neighbourhood 3
 - c. Neighbourhood 2 / Neighbourhood 3
- 3) 20 closest real data values

Predictor datasets were developed from six synthetic aeromagnetic models comparable to Figure 15, and the RF and SVR methods were trained using the hyperparameters seen in Table 3 and Table 4. Once trained, each method was then given the test dataset (the flight line data from Figure 15), and plotted (see Figure 17 for the outline of the process).

Table 3. The hyperparameter choices for the RF method with predictor set 2.

Hyperparameter	Property Set To:
<i>n_estimators</i>	500
<i>min_samples_leaf</i>	50
<i>min_samples_split</i>	50
<i>max_features</i>	log2
<i>max_depth</i>	500

Table 4. The hyperparameter choices for the SVR method with predictor set 2.

Hyperparameter	Property Set To:
<i>kernel</i>	rbf
<i>C</i>	1000
<i>gamma</i>	0.01
<i>epsilon</i>	0.001

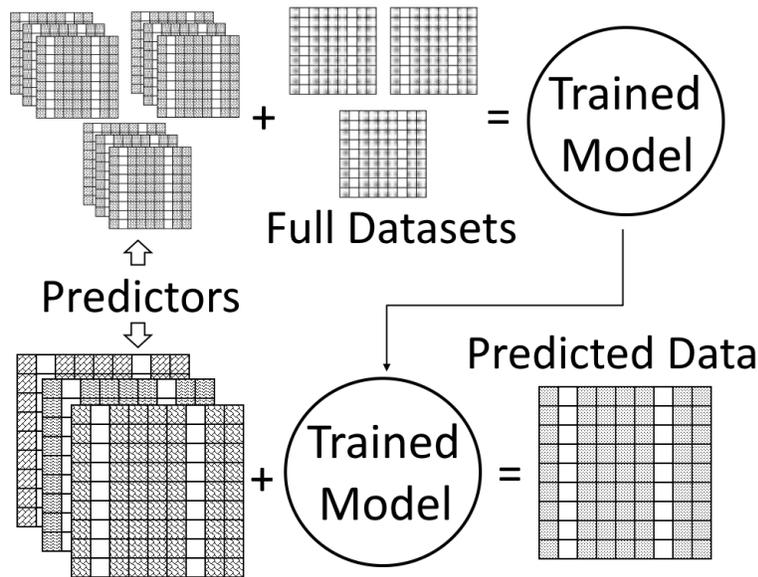


Figure 17. A depiction of the training process (top row) and the prediction process (bottom row) for predictor sets 2 and 3. In this approach, multiple full datasets were split into simulated flight line data and predicted data, depicted by the three different sets of predictors in the training process. These were used to train a model, which was then applied to the test dataset.

3.3.3 Predictor Set 3

The final predictor set investigated a modified approach to predictor set 2. By feeding in the end result of standard aeromagnetic interpolation methods as additional predictors, this predictor set was developed to act as more of a “filter” approach than an original interpolation. The expectation was that the ML models would utilize the varying strengths of each standard interpolation method to develop a better final interpolation grid than any single one of them. Therefore, this third set of predictors consisted of:

- 1) Minimum curvature result (e.g. Briggs, 1974; Swain 1976; Smith and Wessel, 1990)

- 2) Kriging result (e.g. Hansen, 1993)
- 3) Bidirectional spline result (e.g. Bhattacharyya, 1969; Akima, 1970)
- 4) Multi-trend gridding result (Naprstek and Smith, 2019)
- 5) Distance weighted mean of all real data values in a radius of:
 - a. 3 cells (neighbourhood 1)
 - b. 6 cells (neighbourhood 2)
 - c. 12 cells (neighbourhood 3)
- 6) Ratio of:
 - a. Neighbourhood 1 / Neighbourhood 2
 - b. Neighbourhood 1 / Neighbourhood 3
 - c. Neighbourhood 2 / Neighbourhood 3

Note that we removed the 20 closest real data values as predictors, due to the fact that they would offer little unique information when compared to the results from each of the standard interpolation techniques. However, the neighbourhood predictors were kept, as it was expected that they would continue to provide unique information regarding a cell's surroundings, and the effective change (and rate of range) that was occurring nearby. The standard aeromagnetic gridding results were generated using Geosoft's Oasis Montaj methods (Geosoft, 2019); except for multi-trend gridding which was developed using the custom Oasis Montaj dll provided by Naprstek and Smith (2019). As with predictor set 2, predictor datasets were developed from six synthetic models, the two ML methods were trained, and each method was then given the test dataset (Figure 17). Due to the similarity in predictors, it was found that the same

hyperparameters used in predictor set 2 (Table 3 and Table 4) were effective with this predictor set

3.4 Results

Applying each predictor set and ML method, we developed our series of trained models and resulting interpolations on the test dataset. A set of statistics were then calculated to measure how well all methods compared. The absolute value of the error between each interpolation and the original data was calculated, and then the mean, standard deviation, and signal-to-noise ratio (SNR) were found (Table 5). The SNR was calculated using the definition in Jia and Ma (2017), who complete a similar process for their investigation:

$$SNR = 10 \log_{10} \left(\frac{\|I\|^2}{\|I_n - I\|^2} \right)$$

where I and I_n are the original data and the ML result respectively. Note again that all total field values in the synthetic datasets were normalized to range between zero and one for the purposes of training and predicting.

All interpolated results of the synthetic testing are shown in Figure 18. For comparison, the dataset's interpolation with minimum curvature is shown in Figure 18a. The minimum curvature result is accurate, with only some “beading” issues (Keating, 1997; Smith and O’Connell, 2005; Geng et al., 2014; Naprstek and Smith, 2019) along the angled lineaments, as well as a curving of the square feature in the south-east corner.

As noted previously, the SVR result with predictor set 1 was very poor, and therefore is not shown here. However, the RF result with predictor set 1 is shown in Figure 18b. It can be seen that the overall interpolation is also quite flawed when compared to all other results, however still presents a reasonable interpolation when the broad features are examined. In particular, this method is actually fairly effective at reconciling the angled lineaments on the northern half of the dataset. The central values of the lineaments are arguably less affected by the beading issues than that of the minimum curvature result. Overall however, the result is still inaccurate, with effects such as “bulging” between flight lines, particularly noticeable along the edges of the central horizontal lineament. The predictor set 1 statistics clearly indicate that this method, at least in its current implementation, is not a viable approach to aeromagnetic interpolation with ML. Despite these results, we felt that this approach was important to show, as it exemplifies how the context of the data, and any related data that is available for predictors, is important when utilizing ML. Despite this style of approach (training and predicting on the same dataset) working well in the literature referenced, particularly environmental prediction applications, it is clear that it does not work as effectively in the specific context of aeromagnetic interpolation.

Looking at the results of predictor set 2, it can be seen that both the RF method (Figure 18c) and the SVR method (Figure 18d) have resulted in effective interpolations. Overall the RF method appears to produce a smoother result, however has more beading issues than the SVR method. The SVR method has very well-developed lineaments, and minimizes the beading effects such that all angled lineaments appear better interpolated than the minimum curvature result. Additionally, the central portion of the round feature in the north-east corner is also better

resolved in the SVR result as compared to the RF result. Unfortunately, the SVR result suffers from a bulging issue between flight lines, similar to the RF result of predictor set 1, again most noticeably along the edges of the central horizontal linear feature. Looking at the statistics, both results from using predictor set 2 appear to be comparable to the standard interpolation methods. In spite of the bulging issues in the SVR result, the rest of the interpolation is effective, creating less errors on average than the RF result, such that it is essentially on par with the minimum curvature interpolation. However, even if some areas are as accurate or better, the appearance is not what is expected of potential field data, as it is not varying smoothly. So despite the statistics indicating that overall the results are approximately equivalent to minimum curvature, we argue that the results are in some ways worse. Therefore, this approach clearly has potential; however it requires some guidance (likely in the form of alternate or additional predictors) that would further constrain the data to look more like potential-field data.

Finally, we can look at predictor set 3. The RF result (Figure 18e) and SVR result (Figure 18f) appear to be both very accurate and smoothly varying. When we compare the visual results of predictor set 3 with the minimum curvature result, we can see that there is less beading along lineaments, as well as a tighter structure in both the north-east circular feature, and the south-east square feature. Similar to the predictor set 2 results, the SVR interpolation has less beading effects than the RF method (noticeable on effectively all lineaments), but introduces some very small noise effects along the bottom edge of the central horizontal feature between flight lines that are not present in the RF result. However, as a whole it can be seen that predictor set 3 combines better elements from each of the standard interpolation methods. This approach can be

almost seen as an “intelligent” averaging of each interpolation, combined with other statistics (neighbourhood information) about the data. Multi-trend gridding and bidirectional splines are effective at handling lineaments, however can introduce noise (in the form of imposed structure) in areas with little linear coherency. Conversely, minimum curvature and kriging will, in general, produce smoother results across all areas, but can be subject to artifacts such as beading along lineaments. By combining the methods in this way, we can arrive at an incrementally better interpolation, as evidenced by the results in Table 5 that show that when using predictor set 3, both the RF and SVR results are (in this dataset) better than all the individual interpolation methods.

Table 5. Summary error statistics of the four standard interpolation methods and all predictor sets (along with their associated ML method) applied to the test dataset.

Method	Mean of absolute error	Std. Dev.	SNR
Minimum Curvature	0.0145	0.0215	18.03
Kriging	0.0174	0.0262	16.09
Bidirectional Spline	0.0121	0.0218	17.70
Multi-trend gridding	0.0100	0.0208	18.64
Predictor set 1 – RF	0.0251	0.0354	14.63
Predictor set 1 – SVR	0.1064	0.1038	3.90
Predictor set 2 – RF	0.0149	0.0232	17.33
Predictor set 2 – SVR	0.0137	0.0220	18.06
Predictor set 3 – RF	0.0103	0.0174	19.78
Predictor set 3 – SVR	0.0097	0.0173	19.59

a)

b)

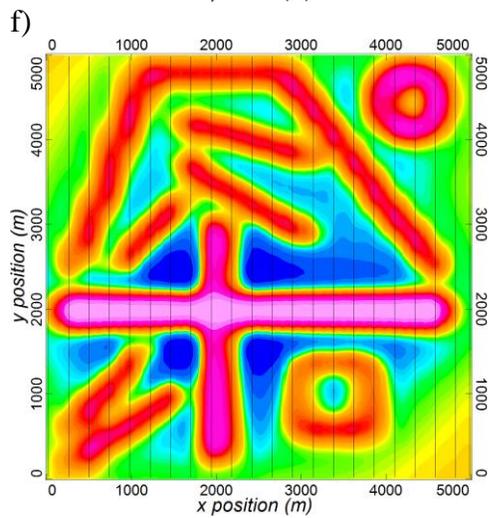
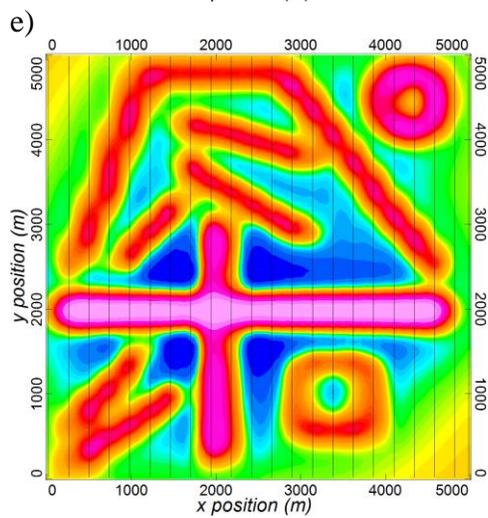
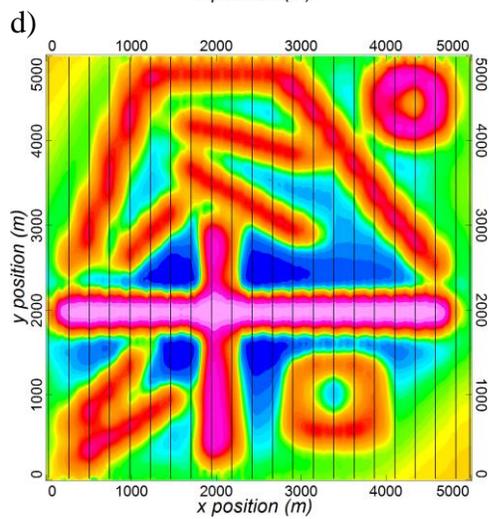
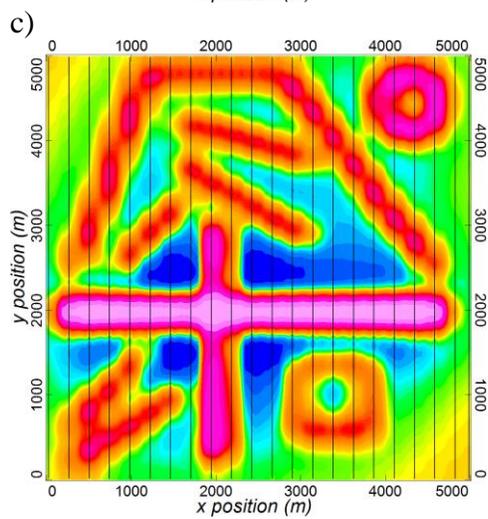
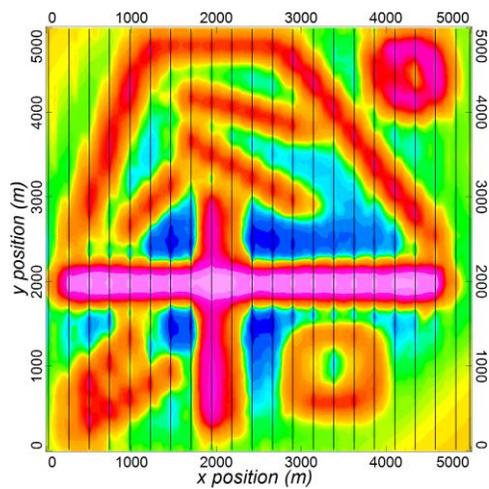
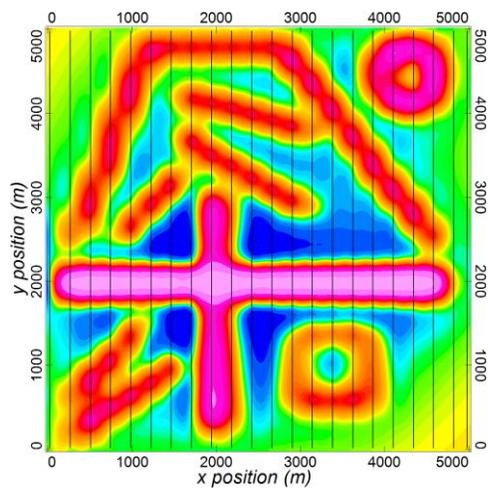


Figure 18. The results from the synthetic testing. a) Minimum curvature. b) Predictor set 1, RF result. c) Predictor set 2, RF result. d) Predictor set 2, SVR result. e) Predictor set 3, RF result. f) Predictor set 3, SVR result. Note that all plots have the same colour range as the full dataset plot in Figure 15.

3.5 Case Study

To better understand the applicability of this method on real-world data, we took both predictor set 3 trained models and applied them to a dataset from northern Canada (Somerset Island). This dataset was downloaded from Natural Resources Canada's online geophysical data repository (Geological Survey of Canada, 2019), and a subsection of it was extracted for use. The overall process for this method using predictor set 3 from the methodology section can be seen in Figure 19.

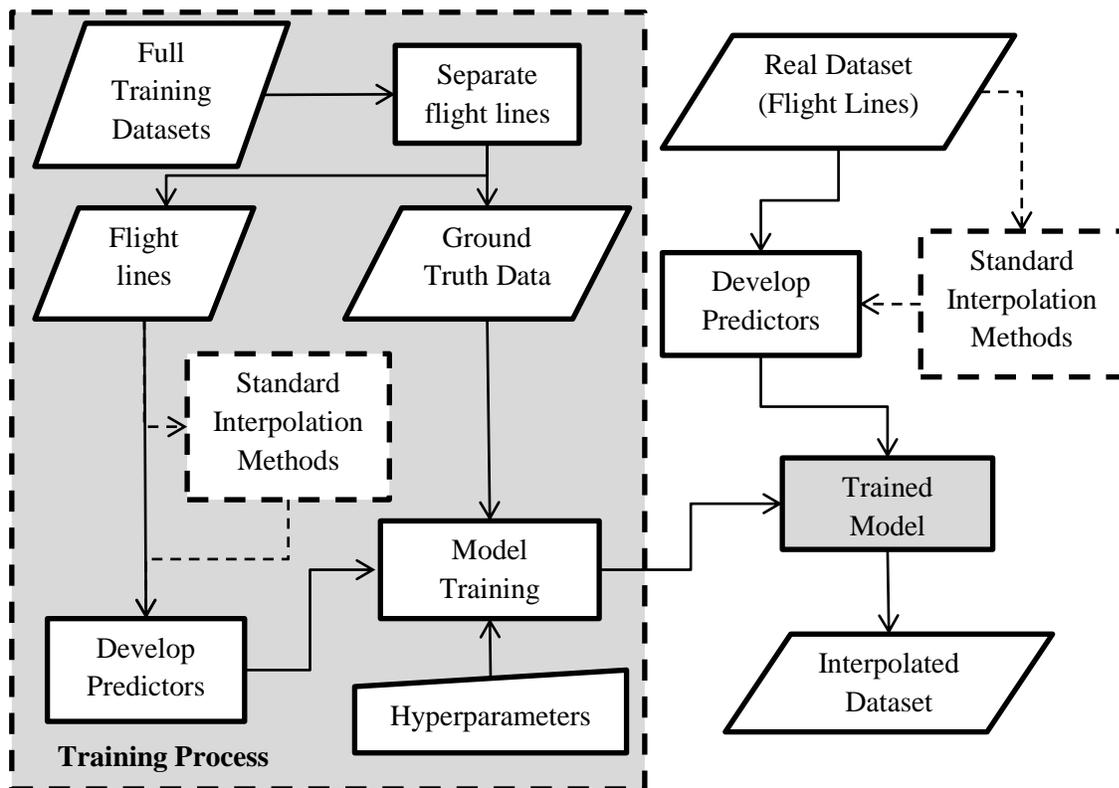


Figure 19. A flowchart showing the overall steps of predictor set 3, and its application to real-world data. Note that the standard interpolation methods are shown in dashed lines to highlight their unique situation as predictors, since they require the raw data (as do the other predictors), yet are predictors themselves.

The results of minimum curvature and multi-trend gridding are compared against the RF and SVR results in Figure 20. We can see that all four results are fairly similar, with two of the more noticeable differences highlighted by circled areas. The eastern circle shows a linear feature that is best interpolated in the multi-trend gridding result, while the minimum curvature interpolation results in a beading effect. Both ML approaches result in much less beading than the minimum

curvature interpolation, with SVR more strongly trending the lineament, similar to the multi-trend gridding result. Conversely, the western highlighted area shows a series of circular features with little linear structure. The multi-trend gridding method here introduces some small artifacts, while the minimum curvature result does not. Here the two ML approaches follow the minimum curvature response, keeping the features unconnected and circular, which is likely more appropriate given the surrounding data.

Beyond that, there are not many differences that can be seen without vertical derivative enhancement results which are presented in Figure 21. In this, the advantages that the ML methods offer can be seen quite well. The central linear feature that is circled shows the differences between the minimum curvature and multi-trend gridding, and how the thin linear feature is better trended by the multi-trend gridding result. Additionally, it can be seen in both ML approaches that they too better trend the feature. The north-east circled area shows the opposite, where a feature that is less well-defined in terms of linear structure in the total field is better rectified in the vertical derivative by the minimum curvature result as compared to the multi-trend gridding result. Comparing these to the ML interpolations, it can be seen that both RF and SVR approaches interpolated the data better than either individual standard interpolation method, resulting in well-defined features. If we compare just the ML methods for this section, it is also clear that the SVR result is more distinct in its separation of individual features than the RF result. Further comparing just the two ML approaches we can see differences in the south-east circled area. The feature here is fairly nebulous, and unfortunately the RF result appears to be affected by noise, leading to out-of-place values near to the flight lines that are not present in

any of the other results, including the SVR result. Finally, if we compare the ML methods to the standard methods in the north-west corner, we can see some small horizontal noise effects near the flight lines in both ML methods that are not present in the standard methods.

Overall, the results are as expected: both ML methods offer maps that are, in effect, a mix of the standard gridding results, attempting to utilize the strengths of each. In much of the data shown, particularly the total field maps, there is little difference between any of the four interpolation methods. However, in some cases the ML results are as good as the more appropriate standard gridding method (for the given type of feature), and in specific situations it can be seen that the ML approaches actually appear to be more effective than either of the two standard gridding methods shown. It must be noted though, that in some particular areas both ML methods result in interpolations containing noise in the vertical derivative enhancement. Therefore we believe that as a whole, this approach has a possibility of improving interpolation results, even when standard interpolation methods already offer effective maps. It could, however, be argued that the benefits of the ML approach are minimal, at least in this specific dataset's case given the effort that was required. It is possible that in a more difficult situation such as large flight line separation or noisy data that this approach may offer a more substantial advantage.

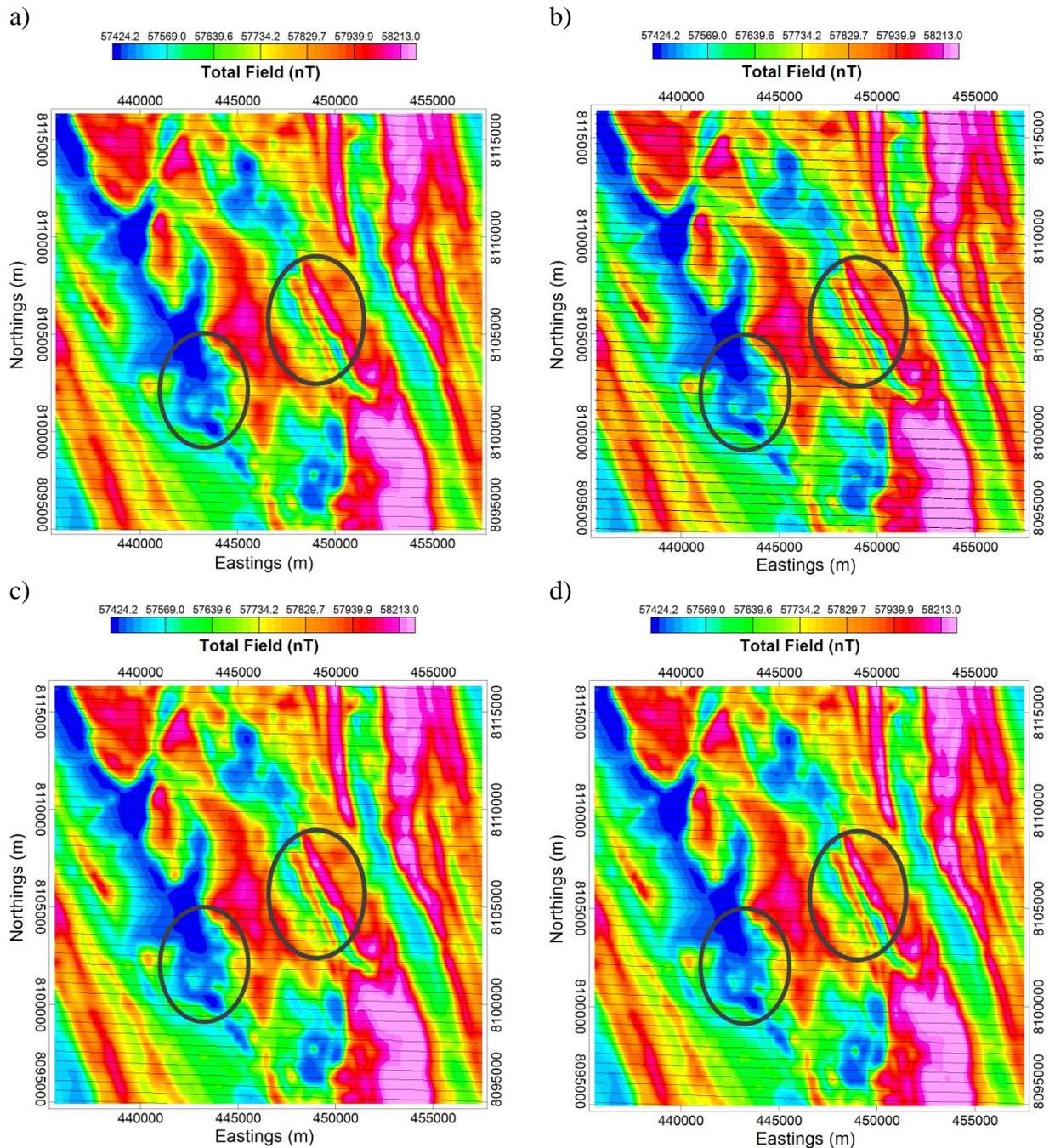


Figure 20. The results from the case study showing a mixture of features. Specific areas have been circled for discussion. Four interpolation approaches are shown here: a) Minimum curvature, b) Multi-trend gridding, c) RF result, and d) SVR result.

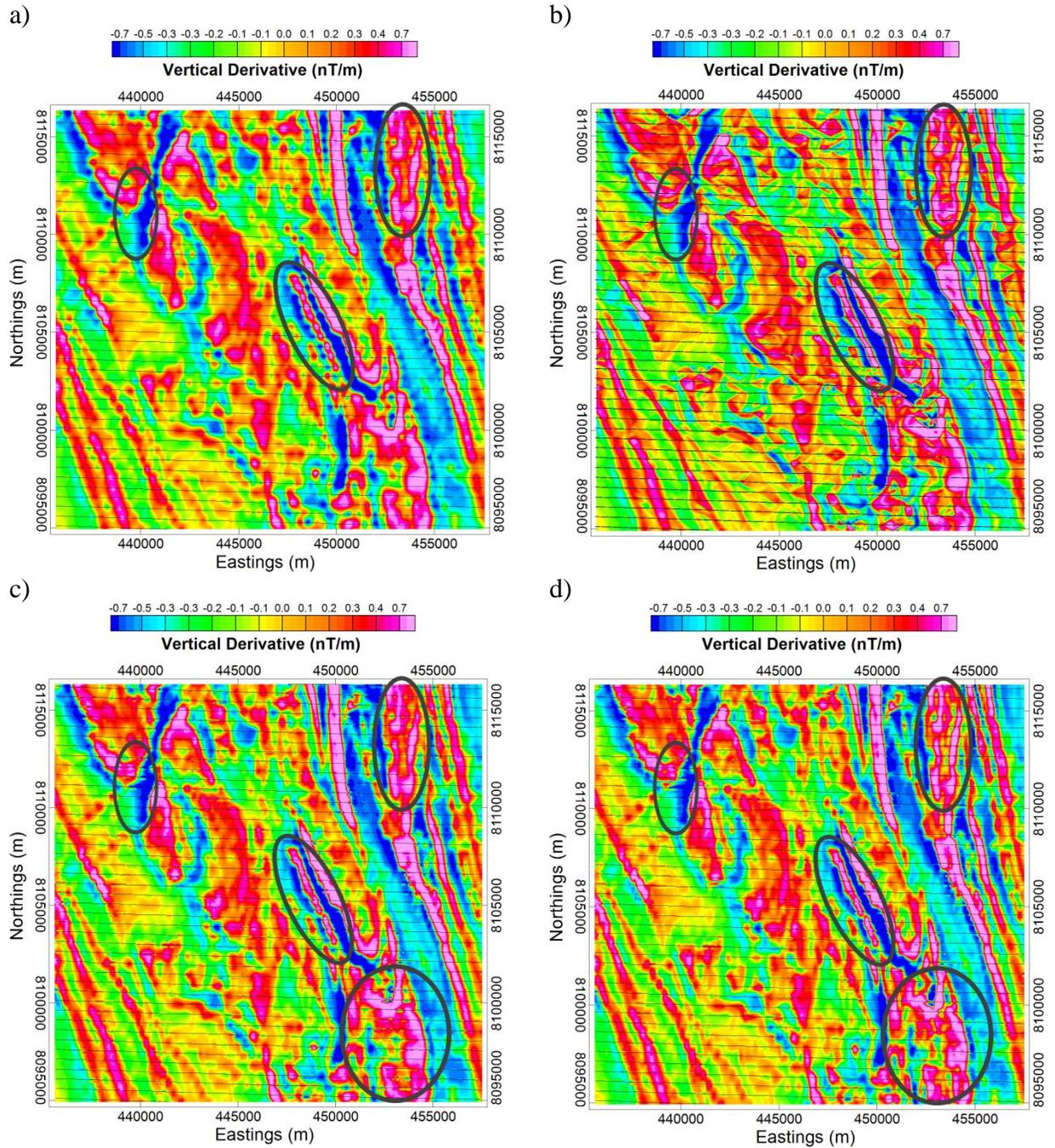


Figure 21. The vertical derivative results from the case study. Specific areas have been circled for discussion. a) Minimum curvature. b) Multi-trend gridding. c) RF result. d) SVR result.

3.6 Conclusions

We have shown that ML methods have potential for improving aeromagnetic data interpolation. While applying ML spatial interpolation using predictors similar to methods previously presented in the literature was not effective (predictor set 1), we have found that by training an RF model and an SVR model on multiple synthetic datasets, a new aeromagnetic dataset can then be effectively interpolated. Various predictors for the ML model can be used, and we have shown one example (predictor set 2) that is comparable to standard interpolation methods such as minimum curvature, but does not offer any competent improvement. However, if the results of several standard interpolation methods are also fed in as predictors, the resulting filter-like ML model can actually offer a theoretically better result, as the strengths of multiple interpolation methods are “intelligently” used where appropriate. When this process is applied to a real-world dataset, we can see that the results are similar, with areas of linear structure being more effective than the minimum curvature result, and areas with little linear structure being more effective than the multi-trend gridding result. In some cases, the ML interpolations arrive at results that are more effective than any single standard interpolation method. While both SVR and RF are effective with controlled datasets such as the synthetic ones generated in this investigation, the case study shows that SVR is overall the more effective method, at least in this approach. It interpolated thin linear features more effectively than RF, and shows less artifacts in the vertical derivative images.

However, it can be argued that due to the inherent properties of ML algorithms, it is possible that they are not best suited for the “raw” interpolation of aeromagnetic data, especially when

compared to standard interpolation methods that are designed to simulate the physical nature of potential fields. Despite the effective results in the case study, there were areas of poorer interpolation in the ML results, showing noise not present in the standard interpolation methods.

Therefore, we believe that if utilized properly, the filter-like ML approach shown here offers unique opportunities in improving aeromagnetic interpolation by training models to benefit from the most effective aeromagnetic interpolation methods. However, we also note that the effort required for the small improvements may not be worthwhile unless the dataset is particularly difficult to interpolate. It is possible that by incorporating other datasets, or completing this step as part of a whole ML-based inversion, this process could be utilized to truly develop novel information.

An area for future investigation would be to utilize neural networks, similar to how we have used SVR and RF regression. A number of the studies we referenced in the introduction have explored similar experiments in their respective contexts with neural networks, showing positive results.

3.7 References

Akima, H., 1970, A new method of interpolation and smooth curve fitting based on local procedures: *Journal of the ACM*, 17, 589–602, doi: 10.1145/321607.321609.

Behrens, T., K. Schmidt, R. A. Viscarra Rossel, P. Gries, T. Scholten, and R. A. MacMillan, 2018, Spatial modelling with Euclidean distance fields and machine learning: *European journal of soil science*, 69, no. 5, 757-770.

- Bhattacharyya, B. K., 1969, Bicubic spline interpolation as a method for treatment of potential field data: *Geophysics*, 34, 402–423, doi: 10.1190/1.1440019.
- Billings, S. D., R. K. Beatson, and G. N. Newsam, 2002, Interpolation of geophysical data using continuous global surfaces: *Geophysics*, 67, 1810–1822, doi: 10.1190/1.1527081.
- Breiman, L., 1996, Bagging predictors: *Machine learning*, 24, no. 2, 123-140.
- Breiman, L., 2001, Random forests: *Machine learning*, 45, no. 1, 5-32.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1984, *Classification and regression trees*: Wadsworth & Brooks, Cole Statistics/Probability Series.
- Briggs, I. C., 1974, Machine contouring using minimum curvature: *Geophysics*, 39, 39–48, doi: 10.1190/1.1440410.
- Cole, P., 2019, *PyGMI: Python Geophysical Modelling and Interpretation*, South Africa Council of Geosciences, <http://patrick-cole.github.io/pygmi/>, accessed December 2018.
- Cordell, L., 1992, A scattered equivalent-source method for interpolation and gridding of potential-field data in three dimensions: *Geophysics*, 57, 629–636, doi: 10.1190/1.1443275.
- Cortes, C., and V. Vapnik, 1995, Support-vector networks: *Machine learning*, 20, no. 3, 273-297.

Cracknell, M. J., and A. M. Reading, 2013, The upside of uncertainty: Identification of lithology contact zones from airborne geophysics and satellite data using random forests and support vector machines: *Geophysics*, 78, no. 3, WB113-WB126.

Cracknell, M. J., and A. M. Reading, 2014, Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information: *Computers & Geosciences*, 63, 22-33.

Drucker, H., C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, 1997, Support vector regression machines: *Advances in Neural Information Processing Systems*, 9, 155-161.

Geng, M., Q. Yang, and Y. Yuan, 2014, Trend enhancement in aeromagnetic maps using constrained coherence-enhancing diffusion filtering: *Geophysics*, 79, no. 1, J1–J9, doi: 10.1190/geo2013-0132.1.

Geological Survey of Canada, 2019, Canadian Aeromagnetic Data Base, Airborne Geophysics Section, GSC - Central Canada Division, Geological Survey of Canada, Earth Sciences Sector, Natural Resources Canada. Data set accessed 6 April 2019 at <http://gdr.agg.nrcan.gc.ca/gdrdap/dap/search-eng.php>.

Geosoft Inc., 2019, Oasis Montaj.

Gunn, S. R., 1998, Support vector machines for classification and regression: Image Speech and Intelligent Systems Research Group, University of Southampton.

Hansen, R. O., 1993, Interpretive gridding by anisotropic kriging: *Geophysics*, 58, 1491–1497, doi: 10.1190/1.1443363.

Hengl, T., M. Nussbaum, M. N. Wright, G. B. Heuvelink, and B. Gräler, 2018, Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables: *PeerJ*, 6, e5518.

Huang, J. J., W. C. Siu, and T. R. Liu, 2015, Fast image interpolation via random forests: *IEEE Transactions on Image Processing*, 24, no. 10, 3232-3245.

Keating, P., 1997, Automated trend reinforcement of aeromagnetic data: *Geophysical Prospecting*, 45, 521–534, doi: 10.1046/j.1365-2478.1997.3450272.x.

Jia, Y., and J. Ma, 2017, What can machine learning do for seismic data processing? An interpolation application: *Geophysics*, 82, no. 3, V163-V177.

Jia, Y., S. Yu, and J. Ma, 2018, Intelligent interpolation by Monte Carlo machine learning: *Geophysics*, 83, no. 2, V83-V97.

Li, J., A. D. Heap, A. Potter, and J. J. Daniell, 2011, Application of machine learning methods to spatial interpolation of environmental variables: *Environmental Modelling & Software*, 26, no. 12, 1647-1659.

Liu, H., Z. Liu, S. Liu, Y. Liu, J. Bin, F. Shi, and H. Dong, 2018, A Nonlinear Regression Application via Machine Learning Techniques for Geomagnetic Data Reconstruction Processing: *IEEE Transactions on Geoscience and Remote Sensing*, 99, 1-13.

Mendonca, C. A., and J. B. C. Silva, 1994, Equivalent data concept applied to the interpolation of potential-field data: *Geophysics*, 59, 722–732, doi: 10.1190/1.1443630.

Mendonca, C. A., and J. B. C. Silva, 1995, Interpolation of potential-field data by equivalent layer and minimum curvature: A comparative analysis: *Geophysics*, 60, 399–407, doi: 10.1190/1.1443776.

Naprstek, T., and R. S. Smith, 2019, A new method for interpolating linear features in aeromagnetic data: *Geophysics*, 84, no. 3, 1-35, doi: 10.1190/geo2018-0156.1. [Chapter 2 of this thesis]

Ni, K. S., T. Q. Nguyen, 2007, Image superresolution using support vector regression: *IEEE Transactions on Image Processing*, 16, no. 6, 1596-1610.

Ohashi, O., and L. Torgo, 2012, Spatial interpolation using multiple regression: 2012 IEEE 12th International Conference on Data Mining, 1044-1049.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, 2011, Scikit-learn: Machine learning in Python: *Journal of machine learning research*, 12, 2825-2830.

Rodriguez-Galiano, V., M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, 2015, Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines: *Ore Geology Reviews*, 71, 804-818.

Smith, R. S., and M. D. O'Connell, 2005, Interpolation and gridding of aliased geophysical data using constrained anisotropic diffusion to enhance trends: *Geophysics*, 70, no. 5, V121–V127, doi: 10.1190/1.2080759.

Smith, W. H. F., and P. Wessel, 1990, Gridding with continuous curvature splines in tension: *Geophysics*, 55, 293–305, doi: 10.1190/1.1442837.

Swain, C. J., 1976, A FORTRAN IV program for interpolating irregularly spaced data using the difference equations for minimum curvature: *Computers and Geosciences*, 1, 231–240, doi: 10.1016/0098-3004(76)90071-6.

Tang, Y., P. Yan, Y. Yuan, and X. Li, 2011, Single-image super-resolution via local learning: *International Journal of Machine Learning and Cybernetics*, 2, no. 1, 15-23.

Were, K., D. T. Bui, Ø. B. Dick, and B. R. Singh, 2015, A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape: *Ecological Indicators*, 52, 394-403.

Yu, L., A. Porwal, E. J. Holden, and M. C. Dentith, 2012, Towards automatic lithological classification from remote sensing data using support vector machines: *Computers & Geosciences*, 45, 229-239.

Chapter 4

4 Convolution Neural Networks Applied to the Interpretation of Lineaments in Aeromagnetic Data

4.1 Abstract

Parameter estimation in aeromagnetism is an important tool in geological interpretation. Due to aeromagnetic data being so prevalent, it can often be used to assist in understanding the geology of an area as a whole, or for locating potential areas of further investigation for mineral exploration. Methods that automatically provide information such as the location and depth to the source of anomalies are useful to the interpretation, particularly in areas where a large number of anomalies exist. Unfortunately, many of the current methods rely on high-order derivatives, and are therefore susceptible to noise in the data. Convolution neural networks (CNNs) are a subset of machine learning methods that are well-suited to image processing tasks, and do not inherently require these higher-order derivatives. They have been shown to be effective at interpreting other geophysical data, such as seismic sections, and so we have developed a neural network architecture for estimating the location, strike, and depth of lineaments in aeromagnetic maps. One of the drawbacks of deep learning methods, however, is that they require an extensive library of labelled information to be well-trained and robust. To approach this, we utilized a synthetic aeromagnetic data modeler to vary relevant physical parameters, and developed a representative dataset of approximately 1.4 million images. These were then used for training classification CNNs, with each class representing a small range of results. Applying this method to a dataset from north-eastern Ontario, Canada, a dyke whose

depth has been previously located with borehole data was correctly estimated within the error of the classification bin size of 25 m. One of the advantages of the method is that the results for all dyke-like features in the area are simple to display on a map and interpret.

4.2 Introduction

Assisted interpretation of aeromagnetic data is a well-researched field, with many methods being developed since aeromagnetism was introduced as a geophysical tool. The reasoning initially was to assist geophysicists and geologists in interpreting data; however, in more recent years this has expanded to include the fact that there is now more aeromagnetic data collected than is reasonable to be interpreted manually. As such, automated and semi-automated methods are increasingly useful, either to focus interpreters onto specific locations for detailed analysis, or for advanced interpretation that may not be feasible when observing aeromagnetic data over a large area. There are a number of studies (e.g. Phillips, 2000; Fairhead and Williams, 2006; Pilkington and Keating, 2009; Pilkington and Tschirhart, 2017) that summarize and compare many of the most popular analysis techniques used in aeromagnetism, covering methods such as the vertical derivative (Hood, 1965), analytic signal (Nabighian, 1972; Roest et al., 1992), Euler deconvolution (Thompson, 1982; Reid et al. 1990), tilt angle (Miller and Singh, 1994; Verduzco et al., 2004; Salem et al., 2007), and local wavenumber (Thurston and Smith, 1997). These methods are used for a variety of purposes, ranging from the simple enhancement of various structural features such as contact mapping, to determining source parameters such as the depth to the top of a source body, the thickness of the body, and the magnetic susceptibility of the rock unit. Many of these methods are still being built upon, or entirely new approaches are being

developed for both edge detection/structure mapping (e.g. some examples from the last ten years include: Cascone et al., 2012; Holden et al., 2012; Lee et al., 2012; Ferreira et al., 2013; Hidalgo-Gato and Barbosa, 2015; Tschirhart and Morris, 2015; Foks and Li, 2016; Sun et al., 2016; Oliveira et al., 2017; Nathan et al., 2020) and source parameter estimation (e.g. some examples from the last ten years include: Ulla et al., 2010; Cooper, 2014; Cooper, 2015; Curto et al., 2016; Tontini et al., 2018). Many of these involve high-order derivatives, which rely on quality aeromagnetic data that are collected with high-precision GPS and precise magnetometers with effective airborne compensation.

Due to the continually lowering barrier for entry of usage, machine learning methods are being applied increasingly in a wide array of applications. Online resources and coding libraries in high-level languages such as Python and R are allowing non-experts to train machine learning models in their own specialized fields. One of the most popular machine learning methods has been Neural Networks (NNs), which have been shown to be highly effective at a large number of various data processing and analysis tasks, including many within the geophysics field (for example, see summaries such as: Poulton, 2002; Peters et al., 2019; Saikia et al., 2020). As supervised methods, they take data with a pre-determined solution (known as labelled data) into their input layer, and iteratively adjust weightings in “hidden” layers between it and the output layer (Figure 22). This iterative adjustment, known as training, allows the weights to be “learned”, such that when the labelled data is input, the output is close to the expected result. Once trained, the model containing the set of learned weights may be applied to unlabelled data, resulting in predictions for this new data.

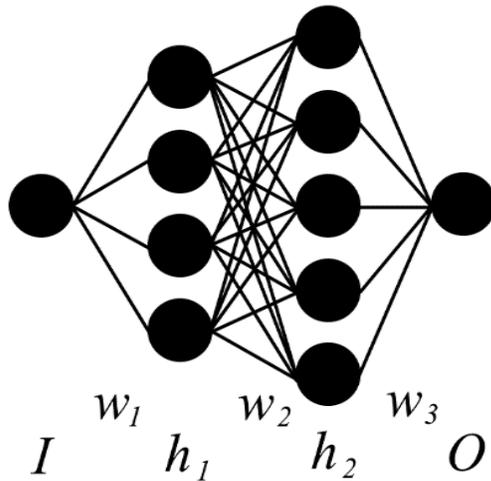


Figure 22. A simple model showing a generic NN, where I is the input layer, h_n are the hidden layers, w_n are the associated layer weights, and O is the output layer. Each circle represents a "node" of data values, and the lines connecting them represent the weights that are iteratively updated during training.

Convolution Neural Networks (CNNs) are a further subset of machine learning methods, whose primary usage is for image and video data input. CNN's nodes contain images (or matrices) rather than singular values, and the weights are convolution filters. In the context of geophysics, CNNs have been effective with seismic data, used for applications such as fault detection (e.g. Pochet et al., 2018; Cunha et al., 2020), salt classification (e.g. Waldeland and Solberg, 2017; Shi et al., 2018), and horizon tracking (Yang and Sun, 2020). They have also been used in aeromagnetics, such as Nurindrawati and Sun (2019) who utilized CNNs to estimate the total magnetization direction of anomalies, and Aghae Rad (2019) who applied CNNs to aeromagnetic, gravity, and elevation data to determine geological lineament locations.

The CNN model architecture is comprised of a number of specific components, including convolution layers, pooling layers, dropout layers, dense layers, activation functions, and the optimization algorithm. Each of these must be experimented with, as no single architecture is appropriate across all applications. Convolution layers are the primary method in which the CNN functions, developing the weights that will iteratively improve the final output. Of key importance to them is the user-defined kernel size, which determines how large of a convolution filter to apply to the input matrix. Pooling layers (e.g. Riesenhuber and Poggio, 1999; Scherer et al., 2010; Krizhevsky et al., 2012) apply a small moving window to the output from a previous convolution layer in an effort to reduce the spatial size (and therefore processing time) while still retaining pertinent information. Max pooling, which keeps the maximum value in each window, and average pooling, which returns the mean of the window, are the two most common pooling filters. Dropout layers (Srivastava et al., 2014) are a standard approach to reducing overfitting the CNN model to the training data, which is a common problem in deep learning. By removing, or “dropping”, a set number of random neurons in a layer, the CNN model is forced to become more robust and not rely on a select few features to predict an entire aeromagnetic dataset/image. Activation functions are essentially logistic gates that aim to assist in normalizing neurons and introduce non-linearity to the CNN. They are applied after every layer in the CNN, including at the final output; however each layer’s activation function does not need to be the same and can vary depending on the circumstances of the model and application. Dense/fully-connected layers in the context of CNNs refer to the standard approach of converting the multi-dimensional image to a vector (termed “flattening”) before the final stages of the CNN model. In most CNN models,

several convolution layers will be part of the CNN model's analysis, before one or more dense layers are added to complete the analysis and output the final result of the model. The optimization algorithm controls the overall "learning" that the neural network accomplishes. It determines how the error is minimized during training, based on the loss calculated between the CNN model's predicted values and the training data's true values.

In this research we applied CNNs to the problem of aeromagnetic data interpretation in the context of dykes, which appear as thin, linear features. Dykes often appear very strongly in aeromagnetic maps, and can greatly assist in the geologic interpretation of a region. Of primary interest to us was determining the location, strike, and depth to the top of the source of the lineaments. Our approach was to use a synthetic aeromagnetic modeller to generate an extensive dataset that varied relevant physical parameters, and use it to train CNN models to predict these parameters in new data, whether synthetic or real. One of the key advantages of this approach is that unlike many other parameter estimation methods, it does not explicitly rely on high-order derivatives that can be greatly affected by issues such as noise in the data. Additionally, it does not require any user input once trained, using only the total-field grid. We begin by describing the parameters of the synthetic data generated for training, followed by the architecture for our CNN models. We then apply the trained models to a difficult synthetic dataset to determine the capabilities of the method. Finally, we test the effectiveness of the method on a real-world aeromagnetic dataset that includes a previously-researched lineament in northern Ontario, Canada, whose location and depth is known.

4.3 Method

4.3.1 Synthetic Training Data Generation

As CNNs require significant amount of labelled training data, the first step of this research was to develop a streamlined approach of generating useful aeromagnetic data. To accomplish this, we used GRAV_MAG_PRISM, a Matlab program developed by de Barros et al. (2013). The primary reasoning for using this was that it could be incorporated into a Matlab script, allowing for the automatic generation of a significant number of combinations of the relevant physical parameters: strike, depth, width, and susceptibility. While not evaluated as parameters in this study, width and susceptibility were also varied to ensure a diverse range of resulting realistic datasets for training. A generic suite of models was developed (Figure 23a), and all relevant parameters were assigned a range of values (Table 6). Note that the strike is measured clockwise from the positive x -axis, such that 0° is along the x -axis and 45° is halfway between the positive x -axis and the negative y -axis. This approach also makes two key assumptions to limit the number of variables for training. First is that all data was generated with an inclination of 90° and declination of 0° . This means that the trained CNN model will only be appropriate for data that has been reduced-to-pole. However, as this is a common process, this should not be an issue for most aeromagnetic datasets, unless they are at low-latitudes (Blakely, 1996). In those cases, it should be possible to simply invert this approach, such that it is trained to use reduced-to-equator data. The second assumption was that no dip variation in the dykes was introduced, and therefore all are assumed to be vertically dipping. Allowing changes in the dip angle would introduce another major variable that would need to be controlled, and ideally estimated. Many dykes in

Canada do not significantly deviate from a vertical dip (Bates and Halls, 1990; Ruffman and Greenough, 1990; Tschirhart and Morris, 2012; Ernst, 2014), and so we believed this to be a valid assumption for this approach (R. E. Ernst, personal communication, October 10, 2019).

All datasets were sampled at a cell size of 25 m, with x and y extents of 500 m, resulting in 21 by 21 cell “windows”. After generating the dataset, a check was completed to determine if the lineament’s physical size fell within any of the nine center cells of the image (Figure 23b). If it did, then the dataset was determined to be a “hit”, while any that do not meet this criteria were labelled as “no lineament” (Figure 23c). Two additional factors were added in to the source datasets to ensure the neural network structure and weights (CNN model) were robust: a secondary “off-center” lineament (Figure 23d) with the same properties as the primary lineament, and a background rectangular anomaly with a susceptibility of $1/100^{\text{th}}$ that of the primary lineament. As such, every combination of physical parameter values had four resulting datasets: the single primary lineament, two lineaments, the primary lineament with a background anomaly, and the combination of the two lineaments with the background anomaly. This led to 365904 individual synthetic aeromagnetic datasets/images to be used for training. Following standard approaches for machine learning training data augmentation (e.g. Simard et al., 2003; Krizhevsky et al., 2012) these images were then rotated by 90° , 180° , and 270° to garner an additional 1097712 images. Finally, 1350 images of generic block-like anomalies without any lineaments present were also generated and added to the “no lineament” dataset. The final total dataset for training therefore comprised of 300608 images that were labelled as a “hit”, and 1164358 images that were labelled as “no lineament”, totalling 1464966 training images. This

total amount was randomly split, resulting in 67% of the data used for training, and the remaining 33% used for validation.

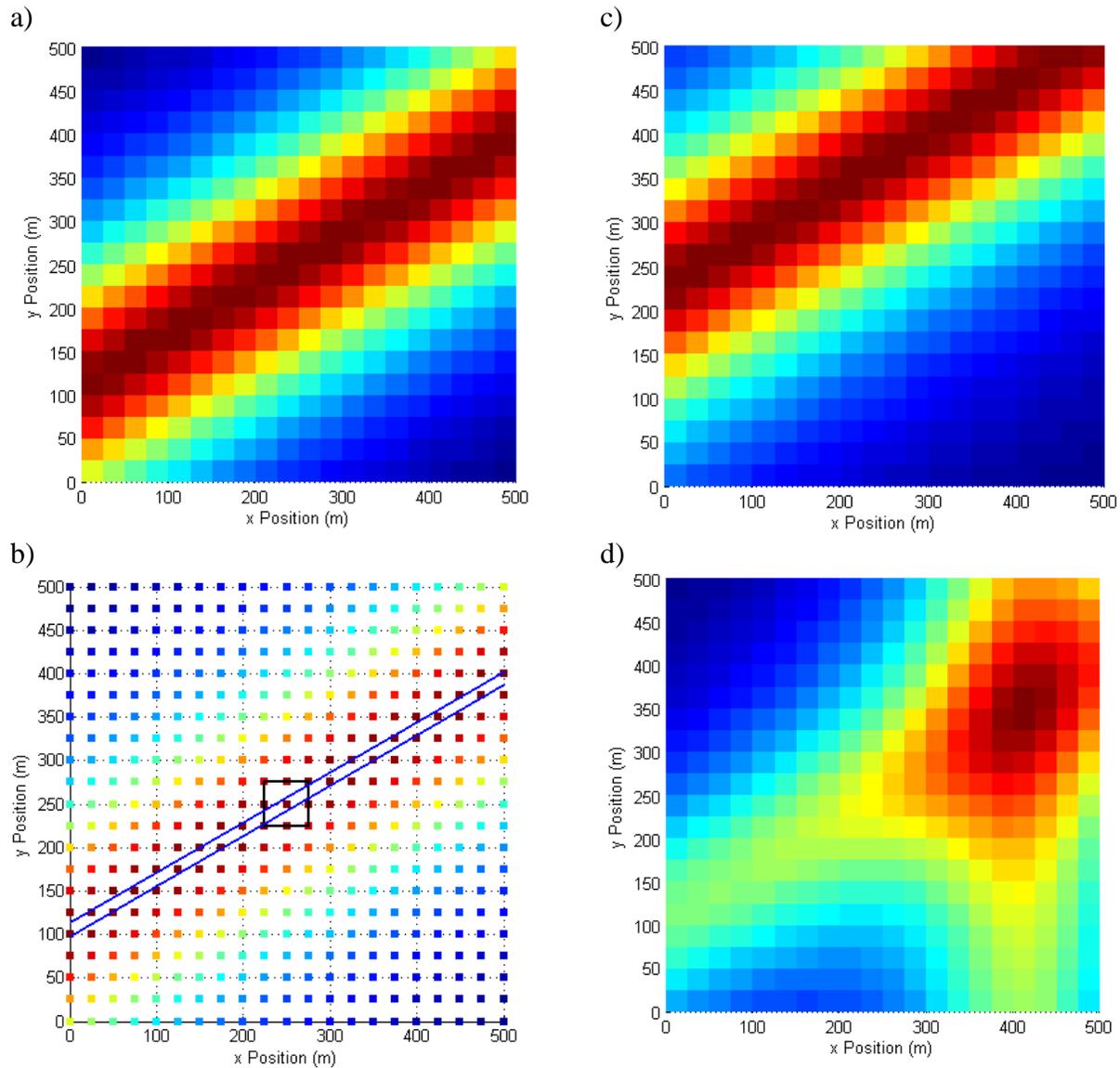


Figure 23. Examples of the synthetic data for training. a) A lineament passing through the center of the image (a "hit"). It was generated with a susceptibility of 0.01, a width of 14 m, x and y positions of 250 m, a depth to top of 120 m, and a strike of 120° . b) The same dataset as a), with

each cell separated to show the positioning of the data. The blue lines represent the physical location of the lineament, and the black box shows where the check is made to determine if the dataset is a hit”. c) The same lineament shifted to a y position of 350 m such that it no longer passes through any of the center nine cells of the image (“no lineament”). d) The same lineament as in a), but with a secondary identical lineament passing in the $\pm y$ direction horizontally offset from the primary lineament.

Table 6. A list of the parameters varied within GRAV_MAG_PRISM.

Parameter (units)	Values
Susceptibility (SI units)	0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1
Width (m)	5, 9.5, 14, 18.5, 23, 27.5, 32, 36.5, 41, 45.5, 50
x Position (m)	50, 150, 250
y Position (m)	50, 150, 250
Depth to Top (m)	30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330
Strike (°)	0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165

4.3.2 CNN Architecture

The two parameters being determined (strike and depth to top of source) were trained as independent CNN models, but both utilized the same architecture and were implemented in Python using TensorFlow (Abadi et al., 2015). The models were treated as multi-classification problems, where “bins” of source parameter values were set as separate classes, similar to the approach by Nurindrawati and Sun (2019). While regression models would be more appropriate for predicting these two continuous parameters, classification models require less training data than an equivalently robust regression model, and therefore we determined it to be the better approach for these initial experiments. The bin ranges for each model can be seen in Table 7.

Note that since most depth estimates take into account the airborne sensor altitude, classes 1 and 2 in the depth model will not be applicable to most aeromagnetic datasets.

Table 7. The value ranges for each bin in the two classification CNN models. Note that there is no connection between classification bins across the two models (i.e. class 2 in the strike model does not imply that the depth model will also be class 2).

Class	Strike Range (°)	Depth Range (m)
1	0 – 20	0 – 25
2	20 – 40	25 – 50
3	40 – 60	50 – 75
4	60 – 80	75 – 100
5	80 – 100	100 – 125
6	100 – 120	125 – 150
7	120 – 140	150 – 175
8	140 – 160	175 – 200
9	160 – 180	200 – 225
10	No lineament present	225+
11	N/A	No lineament present

The process of determining the architecture of any neural network is highly exploratory, and depends greatly on the specific data and context it is being applied to. We tested a number of different architectures, before concluding on the one seen in Table 8, which is similar to that of the one described in Cunha et al. (2020). In this architecture, we used stochastic gradient descent as our optimization algorithm, with the learning rate set to 0.001, and our loss as “sparse categorical crossentropy”. The rectified linear unit activation function (e.g. Nair and Hinton, 2010) was used as our activation function for all convolution and dense layers, as it has been shown to accelerate the training process (e.g. Krizhevsky et al., 2012). The exception to this was the final output layer where a softmax activation was used, which produces a normalized

probability that suits our treatment of this problem as a multi-classification model. Thus, the final output after the softmax function was a probability for each possible class, the sum of probabilities being equal to 1. This also meant that the number of neurons at the output depended on the model. As seen in Table 7, the strike model had 10 output classes, whereas the depth model had 11. The TensorFlow implementation of CNNs also contains an option for class weighting in imbalanced datasets. Through experimentation, it was found to be beneficial to set the “no hit” class to 10%, thus introducing a bias of the training process towards “hit” data.

A key aspect of this method was to normalize each cell of the images before inputting them to the CNN training:

$$d_n(x, y) = \frac{d(x, y) - \mu}{\sigma}$$

where $d(x, y)$ is a single cell’s value, μ is the mean of the entire image, and σ is the standard deviation of the entire image. This was done to ensure that simple shifts in the background total field would not affect the final result, and instead all training would focus on the change in shape and size due to the physical parameters varying.

Each model was trained for 75 epochs, with each epoch taking approximately 270 seconds to complete using TensorFlow’s GPU implementation on an NVIDIA GTX 1050Ti graphics card with 4 GBs of RAM. The resulting accuracy and loss plots can be seen in Figure 24. At the end of training, both models had at least 95% accuracy on both the training data and the validation data.

Table 8. The architecture of our CNN model. Note that the final dense layer's neuron number "Out" depended on the feature that was being extracted.

Layers	Conv2D	Dropout (%)	Conv2D	Dropout (%)	Max Pooling	Conv2D	Dropout (%)	Conv2D	Dropout (%)	Max Pooling	Dense	Dropout (%)	Dense	Dense
Neurons	20	5	20	5		50	10	50	10		16	20	32	Out
Kernel Size	3x3		3x3		2x2	3x3		3x3		2x2				

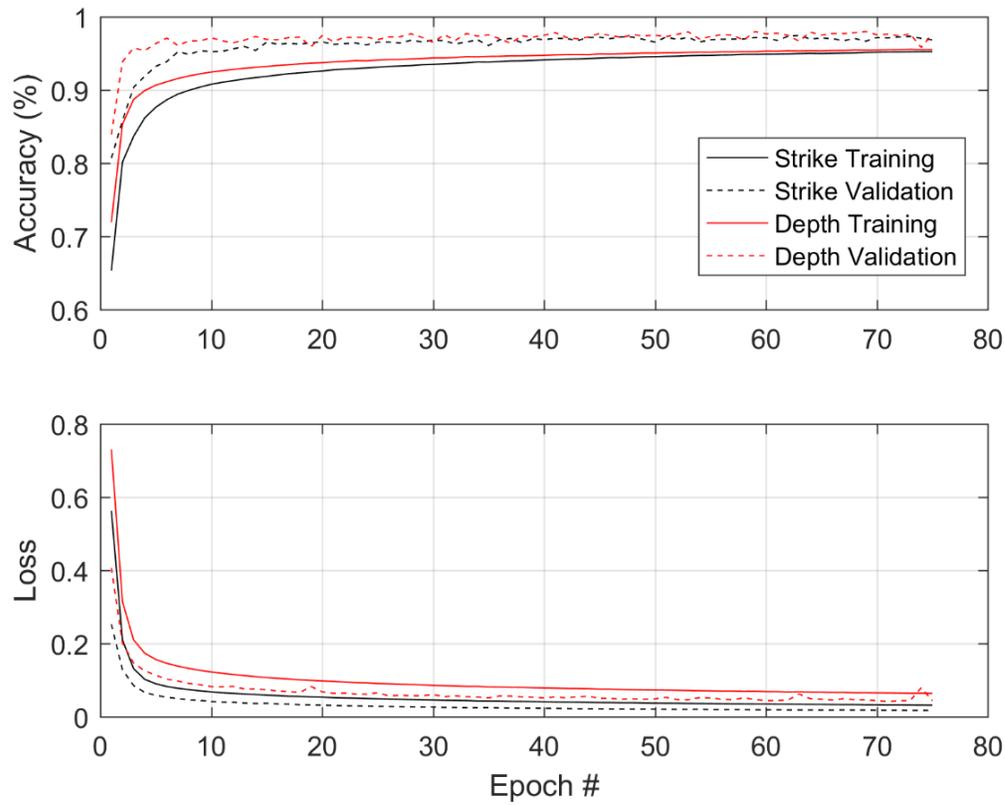


Figure 24. Accuracy and loss plots from the training process for both the strike and depth models. Note that Loss does not have a unit, and the value is dependent on the type of loss function.

4.4 Results

As previously discussed, all training data were 21 by 21 pixel images. This window size was chosen as it would, in general, be large enough to contain enough of any realistically sized lineament (i.e. under 50 m in width, up to several km in depth to top) to be interpretable when placed in the center of the window. This meant that any image that was to be classified by the models had to be examined using this window size. To achieve this, a sliding window approach was developed, such that central pixels of each window would be analyzed and assigned their respective classification. Note that this meant that any input image would be reduced in size around all edges by 10 cells (i.e. no padding was applied). The image for each sliding window was normalized using the same method as the training data. Normalizing each window separately greatly helps the method in isolating local features, as a normalization of the total dataset at once would lead to only the strongest of anomalies being detected.

4.4.1 Synthetic Data

Both models were first applied to a difficult synthetic dataset generated using GRAV_MAG_PRISM (Figure 25). The dataset had a cell size of 25 m by 25 m, and a total extent of 5000 m by 5000 m. The model contained six lineaments and two background block anomalies whose physical specifications can be found in Table 9. The values in the table were chosen as they contain multiple instances that are at or near the split point between classes, therefore making this a particularly difficult dataset. The model was developed with an inclination of 90° , and declination of 0° , a background field of 50,000 nT, and 0.1 nT of Gaussian noise.

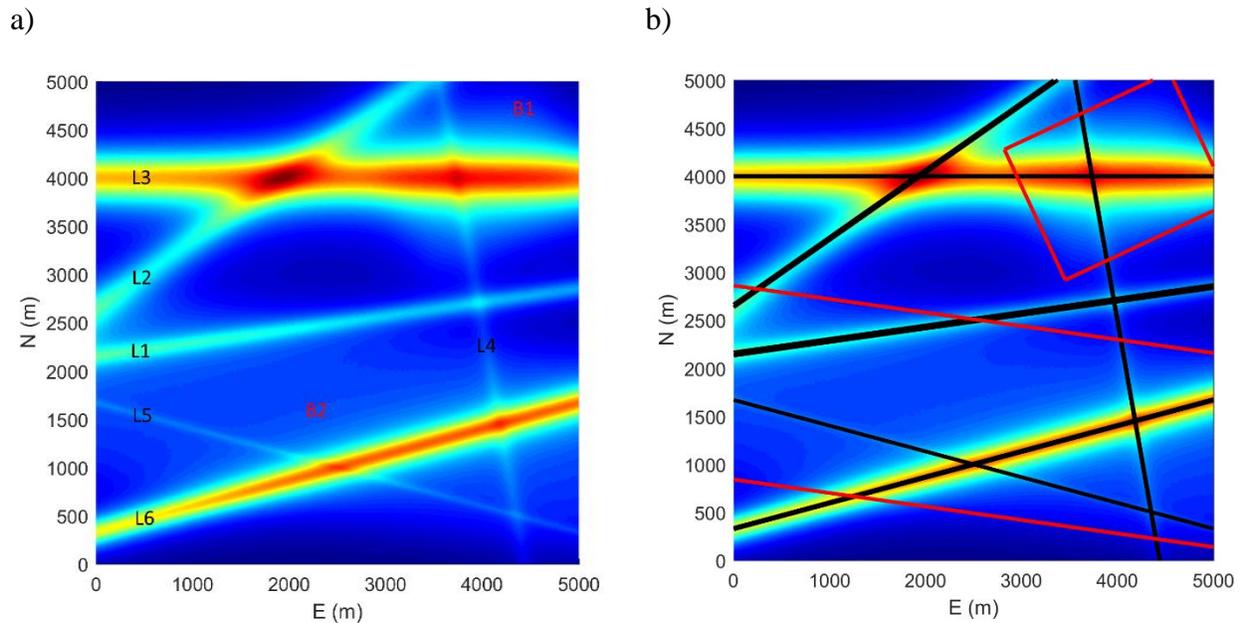


Figure 25. a) The synthetic model used for testing. The six lineaments being analyzed have been labelled “L1” to “L6” and the two background block anomalies have been labelled “B1” and “B2”. b) The model with the lineaments (in black) and blocks (in red) shown to indicate their physical locations.

Table 9. Specifications of each anomaly in the synthetic model.

Anomaly	Susceptibility	Width (m)	Depth to Top (m)	Strike (°)
L1	0.01	33	100	172
L2	0.03	25	185	145
L3	0.46	5	248	0
L4	0.008	8	80	80
L5	0.014	2	35	15
L6	0.07	15	130	165
B1	0.001	N/A	300	65
B2	0.001	N/A	300	8

Each model took approximately 7 seconds to process the dataset, and the results are shown in Figure 26. As can be seen, both CNN models correctly identified all six individual lineaments,

detecting L1 and L6 in their entirety. However L4 was only occasionally identified, partially as it has the weakest magnetic response of the source models, but primarily because it had the greatest interference with the other stronger lineaments (it interacts with all other lineaments in the model). The training data did not include extensive models on multi-lineament interactions, and as such, the L4 case is particularly difficult for these CNN models. The background anomalies B1 and B2 appear to have no effect on the resulting lineament predictions. Table 10 lists the results of strike and depth predictions for each lineament in comparison with their true values. Considering every lineament contained some variation in predicted results, we listed on the table the most prominent bin predictions. An obvious error can be seen in the strike estimator where L1 is predicted to be in the 0° to 20° bin, rather than the 160° to 180° bin. However, due to the symmetry of the angles near 180° this is actually a fairly minor error, being only a single bin away from the actual value, similar to the L2 strike result. The other four lineaments were predicted in the correct bins. The depth estimations were similar, with three lineaments being correctly predicted, and two lineaments, whose values were near a bin transition, were predicted in an adjacent bin. Unfortunately, the L4 depth wasn't as consistently predicted as its strike, and so characterizing this dyke is more difficult. If we examine the predicted values on the image in Figure 26b, there is approximately an even split between the 50 m to 75 m bin and the 75 m to 100 m bin. This could be interpreted as an indication that the lineament is both difficult to interpret and is a value near the split of the two bins, which in this case is 80 m.

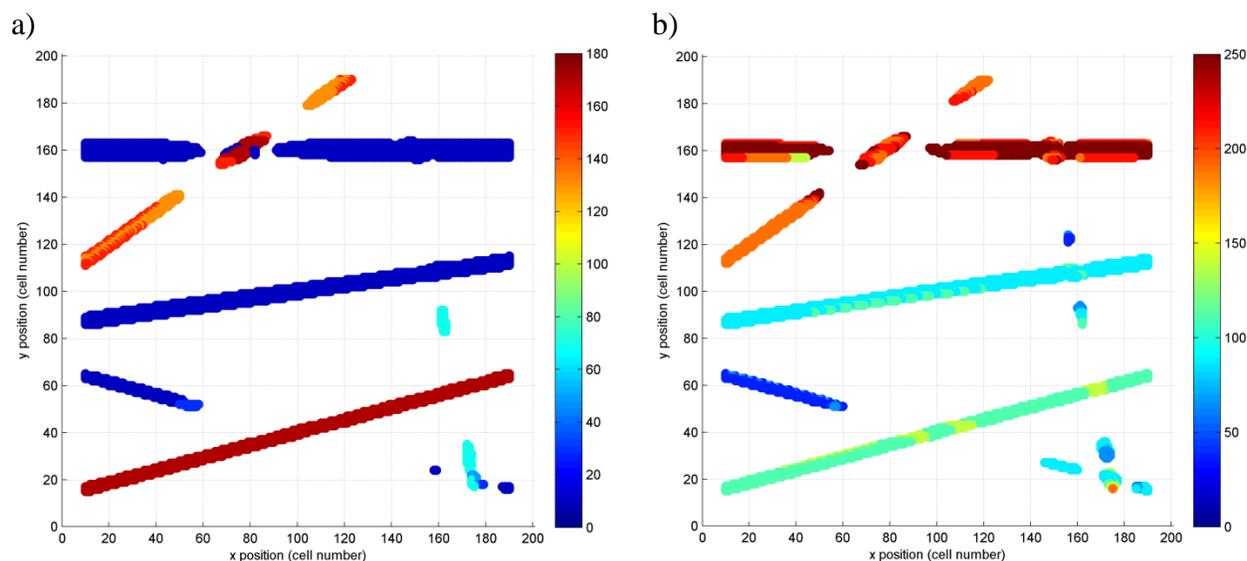


Figure 26. The prediction results for the two source-model parameters. a) The strike predictions, ranging between 0° and 180° , in steps of 20° . b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m.

Table 10. The comparisons between true parameters and the predicted results. Note that all lineaments contained some mix of predictions, so the results here list only the most populous predicted bin.

Anomaly	Strike ($^\circ$)	Predicted Strike	Depth to Top (m)	Predicted Depth to Top
L1	172	0° to 20° bin	100	75 m to 100 m bin
L2	145	120° to 140° bin	185	175 m to 200 m bin
L3	0	0° to 20° bin	248	250 m and higher bin
L4	80	60° to 80° bin	80	No clear prominent bin
L5	15	0° to 20° bin	35	25 m to 50 m bin
L6	165	160° to 180° bin	130	100 m to 125 m bin

If we focus on the cells predicted to be lineaments and compare them with the true source model, we can better see the capabilities of the lineament detection. Figure 27 shows the predicted location as red dots and the actual locations as black dots. Comparing only the correctly

predicted locations (the red dots that coincide with black dots), the true data were first binned into the same classes as the CNN model (following Table 7), followed by an error calculation of the true values minus the model predictions. Error histograms for the two models can be seen in Figure 28. Overall it can be seen that both models were very effective, with low overall error. The most noticeable outlier in the strike predictor is due to the previously noted issue with L1 and the near symmetry between the 160° to 180° bin and the 0° to 20° bin.

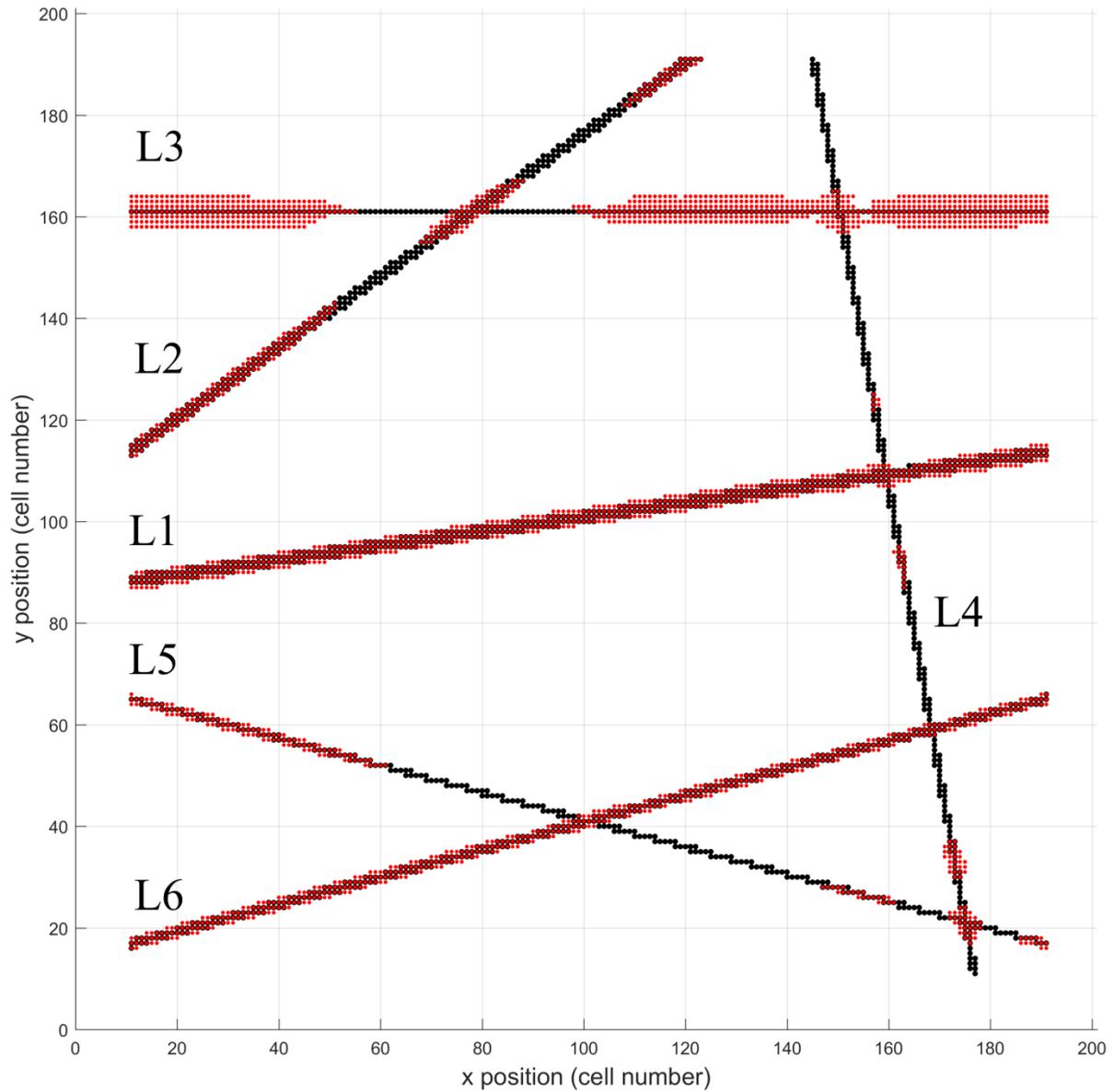


Figure 27. A scatter plot showing the cells that contain lineaments (black) and cells that were predicted to contain lineaments by the models (red).

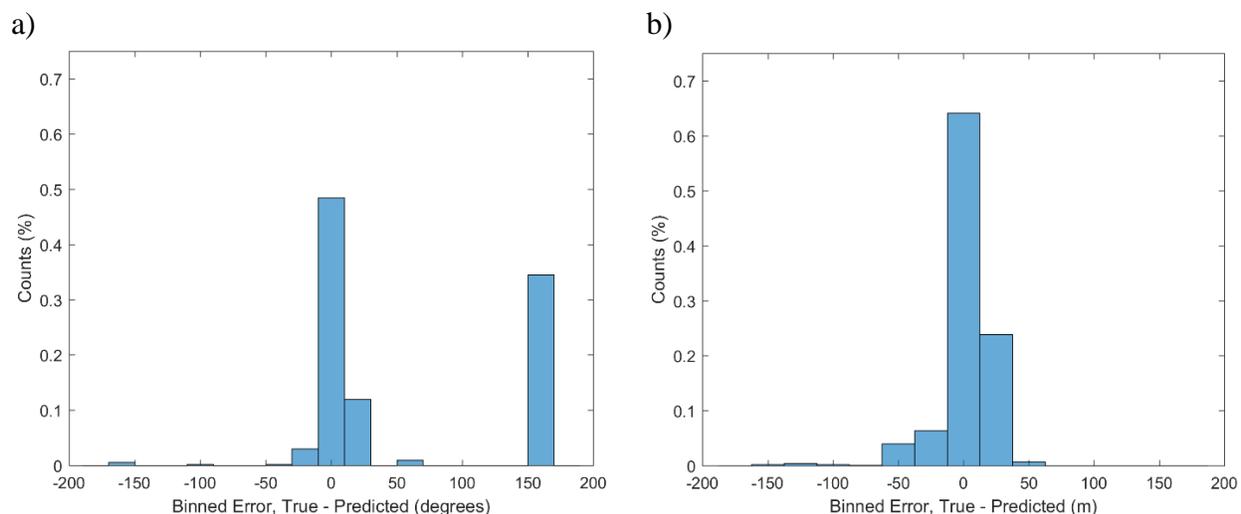


Figure 28. Error histograms for the two models. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide.

4.4.1.1 Noise Effects

A study was undertaken to investigate the impact that increasing the noise level in the test images has on the network prediction. It is important to note that the training imagery was generated with no noise.

The same synthetic dataset that was generated with 0.1 nT of Gaussian noise was re-generated using 1 nT, 3 nT, and 10 nT of Gaussian noise (Figure 29). These data were then input into the trained CNN models, and the results can be seen in Figure 30, Figure 31, and Figure 32. The strike model is clearly less affected by the noise, however both models cannot identify L4 by the 3 nT level, followed by not identifying L5 at the 10 nT level. At the 10 nT level the depth model's predictions become inconsistent, despite clearly detecting the lineaments. The depth

model also begins to result in several spots of false positives in the center of the image, which worsens as the noise level increases.

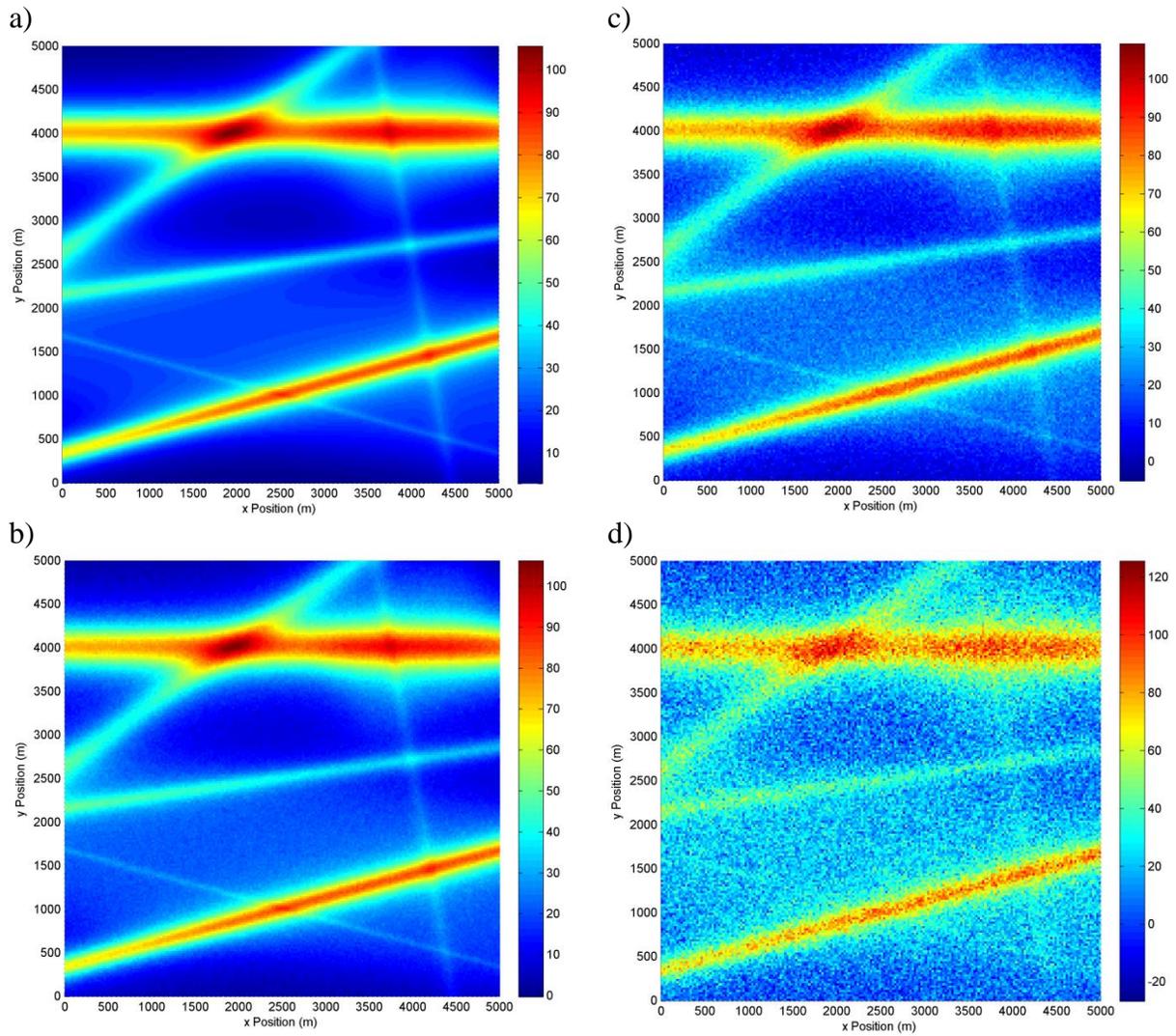


Figure 29. Synthetic dataset with varying noise levels. Note that Figure 25 was generated with 0.1 nT of Gaussian noise. a) No noise, for comparison. b) Noise set to 1 nT. c) Noise set to 3 nT. d) Noise set to 10 nT.

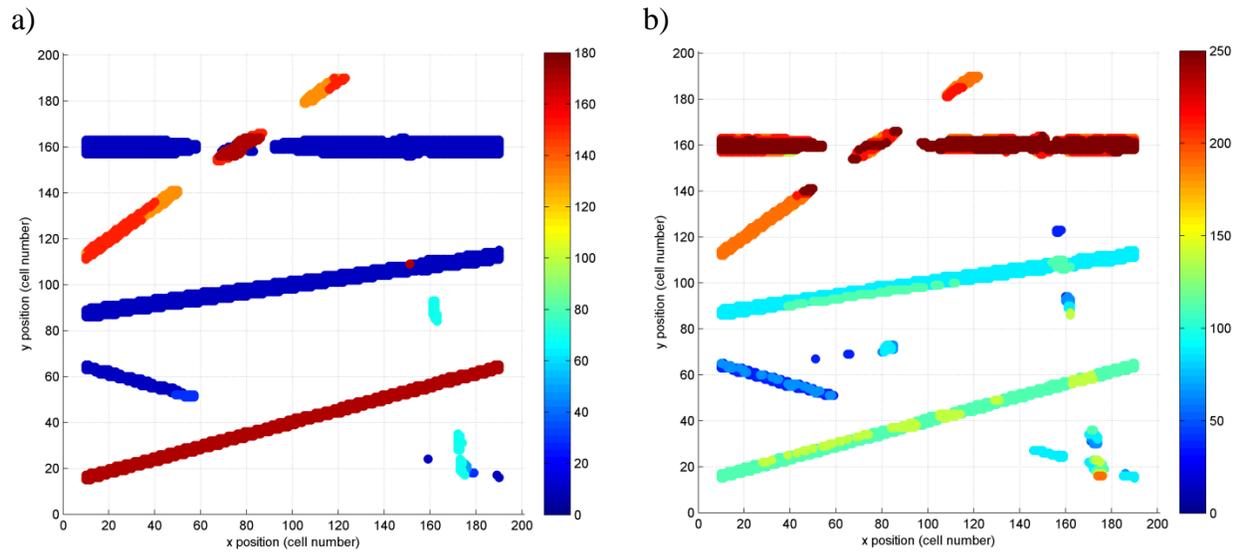


Figure 30. The prediction results for 1 nT of Gaussian noise. a) The strike predictions, ranging between 0° and 180° , in steps of 20° . b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m.

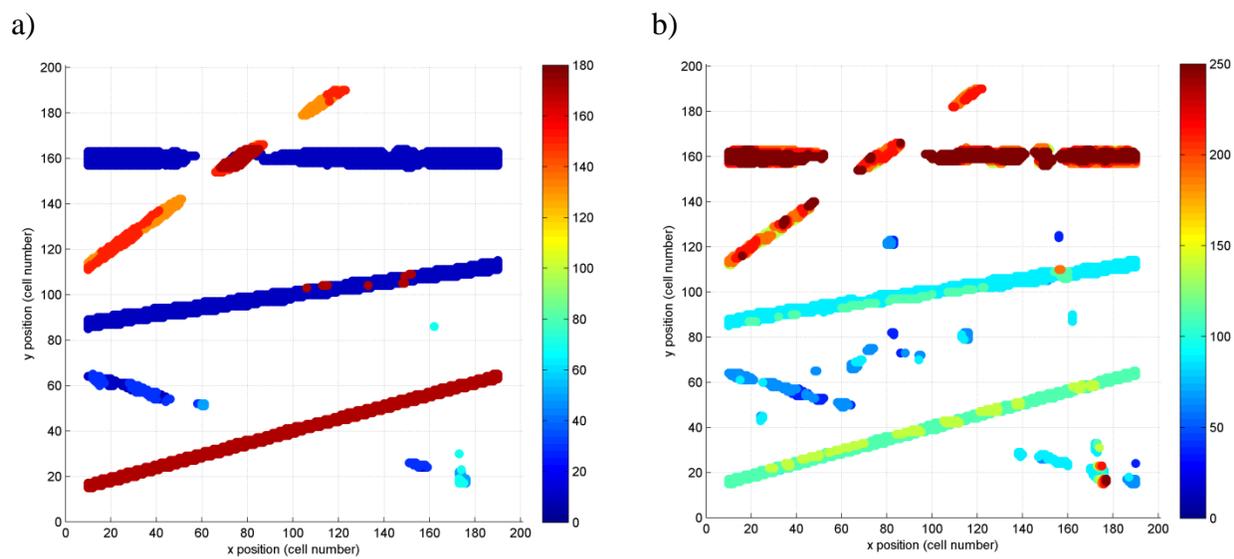


Figure 31. The prediction results for 3 nT of Gaussian noise. a) The strike predictions, ranging between 0° and 180° , in steps of 20° . b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m.

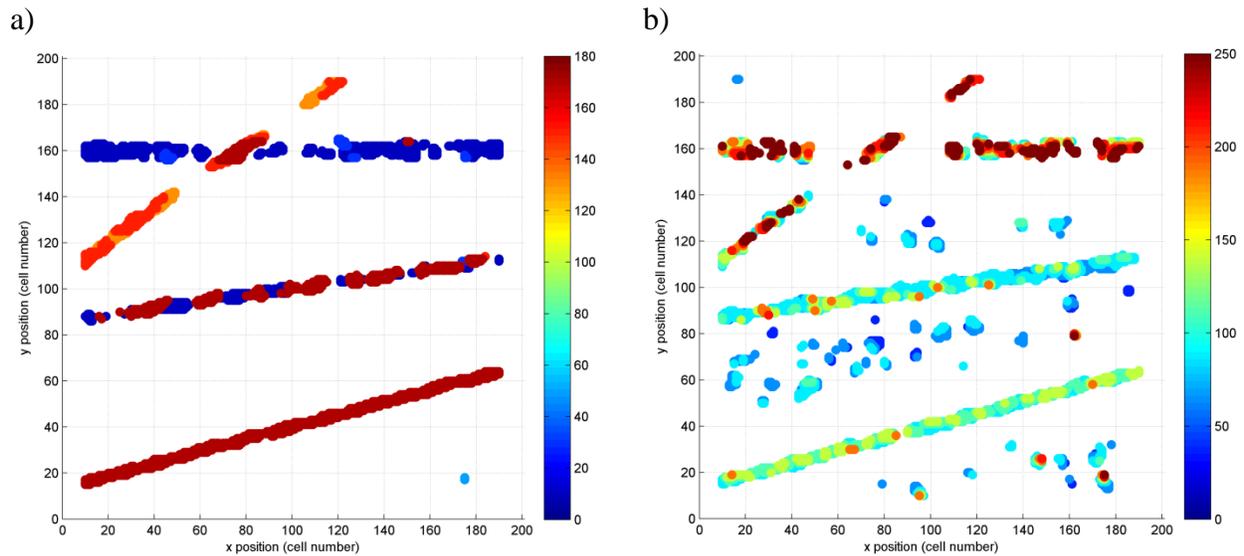


Figure 32. The prediction results for 10 nT of Gaussian noise. a) The strike predictions, ranging between 0° and 180° , in steps of 20° . b) The depth predictions ranging between 0 m and greater than 250 m, in steps of 25 m.

Following the same approach as before, the true strike and depths were binned according to the CNN model classes, and the error between these parameters and those of the predicted models' at the correctly located cells were plotted in histograms. These can be seen for each noise level in Figure 33, Figure 34, and Figure 35. Clearly there is very little effect from the 1 nT and 3 nT noise levels, however the 10 nT noise causes the depth estimates to become much poorer, while technically enhancing the strike predictions due to the difficulty in predicting a consistent angle for L1 (see Figure 32a). It should be noted that the histograms are plotted as a percentage of the total number of correctly predicted cells. To see more complete statistics, Table 11 lists the number of correctly predicted cells (out of a total of 1782 cells within the true model that contain lineaments), the mean, and the standard deviation for both the strike and depth models at each of

the analyzed noise levels. These confirm that there is very little difference in the results for noise levels up to 3 nT of Gaussian noise. Between the 3 nT and 10 nT level the models begin to have difficulty, the depth model more so than the strike model, though both are still able to clearly locate four of the six lineaments.

Noise in aeromagnetic data is primarily due to incomplete aircraft compensation, however as per the Geological Survey of Canada standards for aeromagnetic surveys (Coyle et al., 2014), the noise must be well-below the values used here. Hence, our method should be robust to noise in most circumstances. However, if noise is an issue in a dataset, the training data could be corrupted with noise to make the CNN models more robust to these effects.

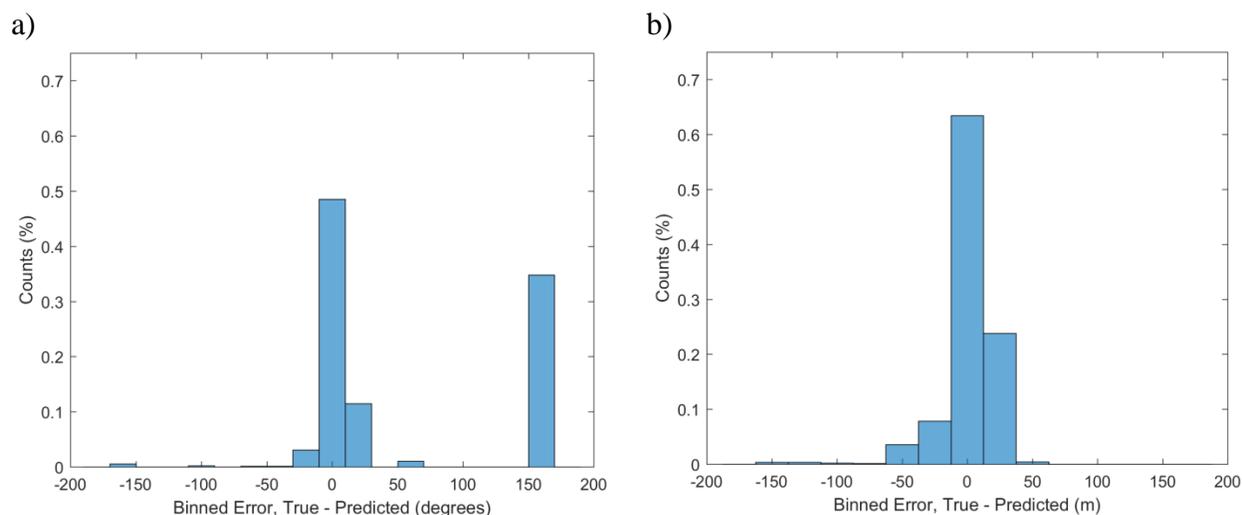


Figure 33. Error histograms for 1 nT of Gaussian noise. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide.

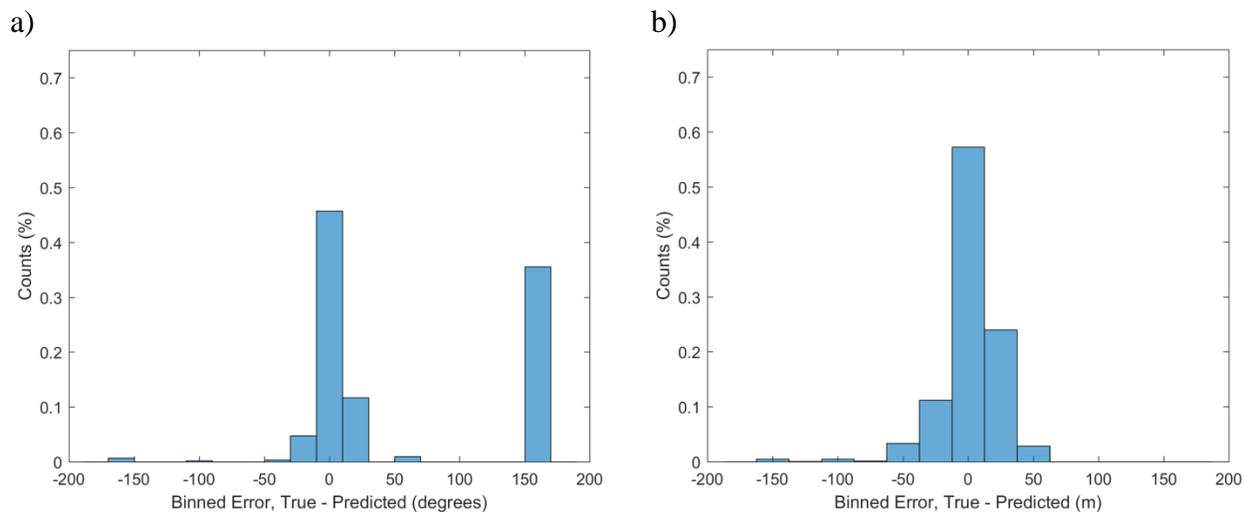


Figure 34. Error histograms for 3 nT of Gaussian noise. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide.

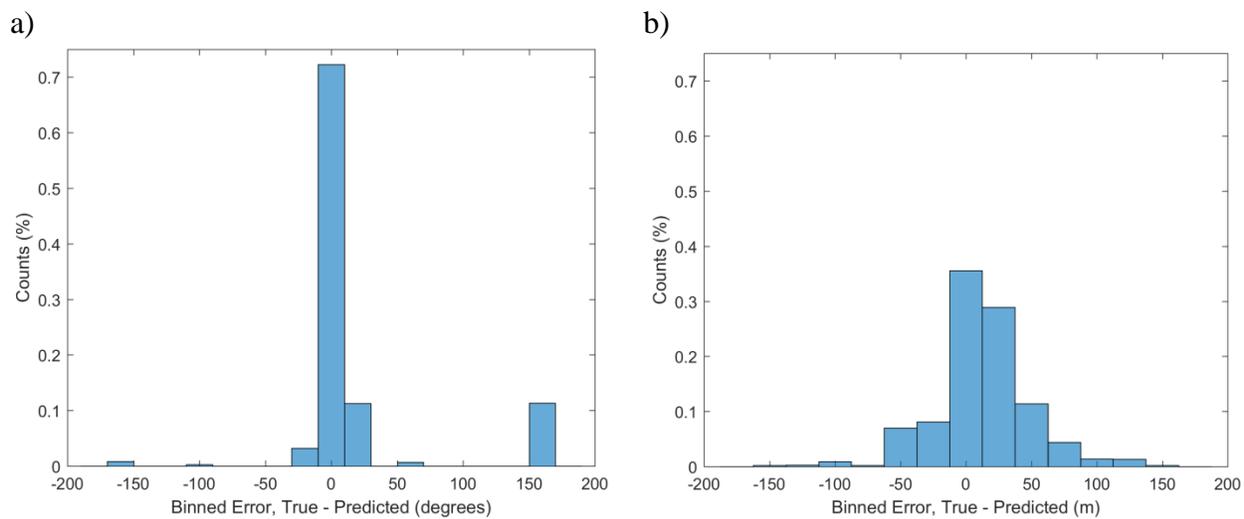


Figure 35. Error histograms for 10 nT of Gaussian noise. a) The result of the strike estimator. Bins are 20° wide. b) The result of the depth to top estimator. Bins are 25 m wide.

Table 11. Summary of the noise analysis. Listed for both CNN models are the number of correctly located cells, out of the 1782 total cells that contain lineaments, as well as the mean and standard deviation of the absolute binned error.

Noise (nT)	Strike Model Correctly Located Cells	Strike Absolute Binned Error (°)		Depth Model Correctly Located Cells	Depth Absolute Binned Error (m)	
		Mean	Standard Deviation		Mean	Standard Deviation
0.1	1288	60.1	74.2	1293	11.1	18.5
1	1278	60.5	74.3	1266	11.3	18.7
3	1245	62.3	74.5	1220	13.5	20.1
10	959	23.2	52.0	996	26.6	27.6

4.4.2 Case Study

To confirm the effectiveness of this method on real data, it was applied to the Matheson dataset from northern Ontario, Canada (Ontario Geological Survey, 2003). This data was flown N30°W, approximately perpendicular to the local geological strike, at a line spacing of 200 m, with the magnetic sensor in a bird approximately 70 m above ground. This dataset was chosen as a specific location in the survey area has been analyzed by several other studies (Vallée et al., 2004; Salem et al., 2005; Thurston and Smith, 2007) for dyke depth, and as such, those results can be used as comparison. Additionally, the dataset includes borehole data that was collected close to the analyzed location, indicating a depth of 41 m to the top of the magnetic body.

Therefore it is expected that the depth should be approximately 111 m below sensor, however all three studies estimated a larger depth, ranging from 139 m to 145 m. Also note that the other studies calculated the depth estimate on profile data, which is not possible for our image-based

approach. As such, a subset of the dataset centered on the analyzed dyke location (Figure 36) was extracted, interpolated at a 40 m cell size using multi-trend gridding (Naprstek and Smith, 2019), and reduced-to-pole using Oasis Montaj (Geosoft, 2020). The specific location of interest (E519470 m, N5377100) is circled. Note that the 40 m gridded data was subsampled to 25 m in Oasis Montaj before being input to the CNN models, to keep the input data the same size as the data the models were trained on.

The two CNN models each took approximately 28 seconds to predict, and Figure 37 and Figure 38 show the results. Overall, the strike estimator appears to give reasonable results along all major lineaments. The depth estimator shows the zoomed portion that we are primarily interested in, with a circle showing the cells that are closest to the previously analyzed profile data. The values in the circled portion include both the 75 m to 100 m bin (light blue) and the 100 m to 125 m bin (light green); however the majority of the predicted values are in the 100 m to 125 m bin. Focussing on our specific location of interest at the center of the circle (inset of Figure 38), the cell that encapsulates this position has a predicted depth is in the correct bin, as do almost all other cells surrounding it.

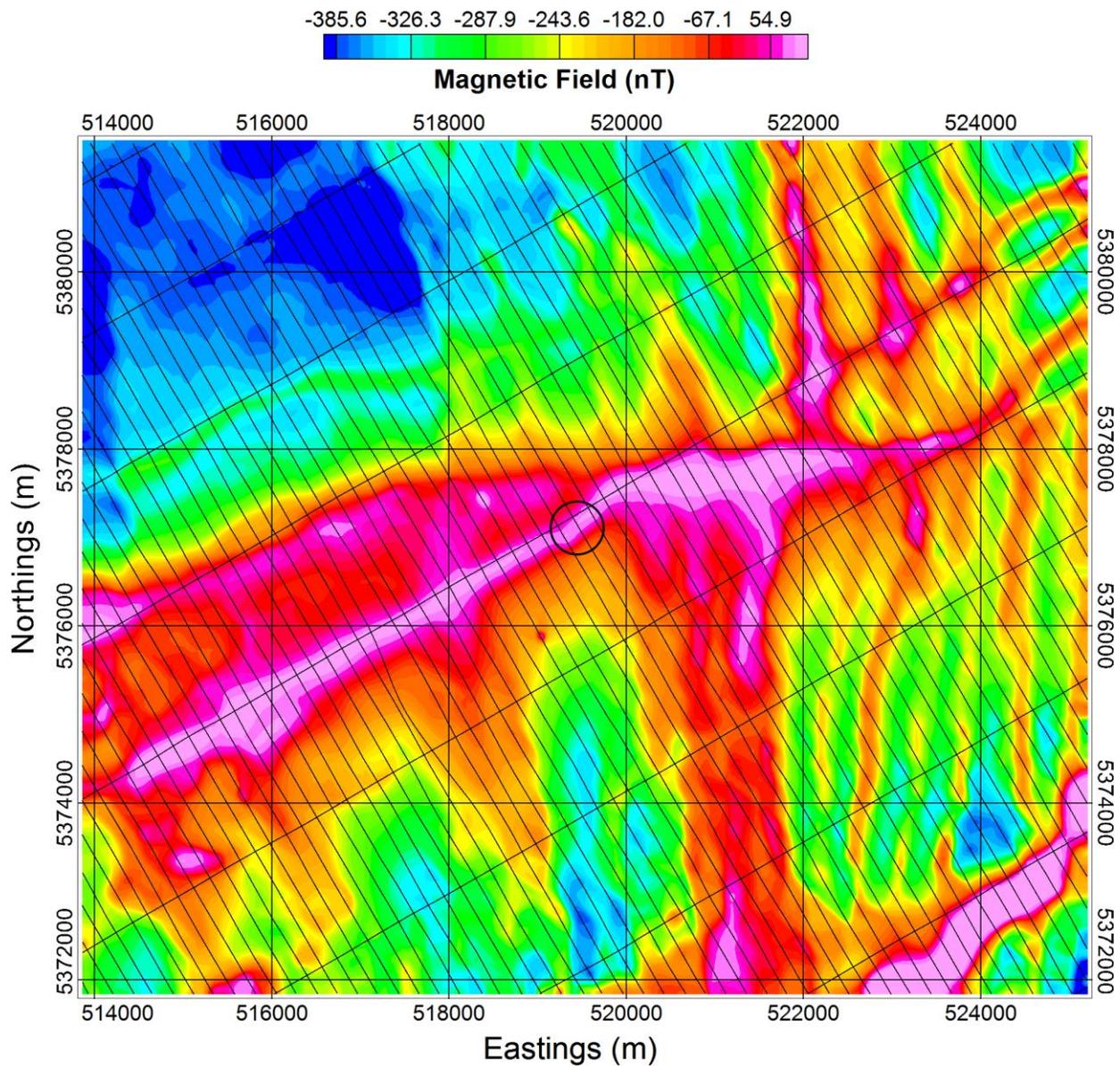


Figure 36. A subsection of the Matheson dataset, centered on the dyke that has been analyzed in other studies. The specific location of interest has been circled for clarity.

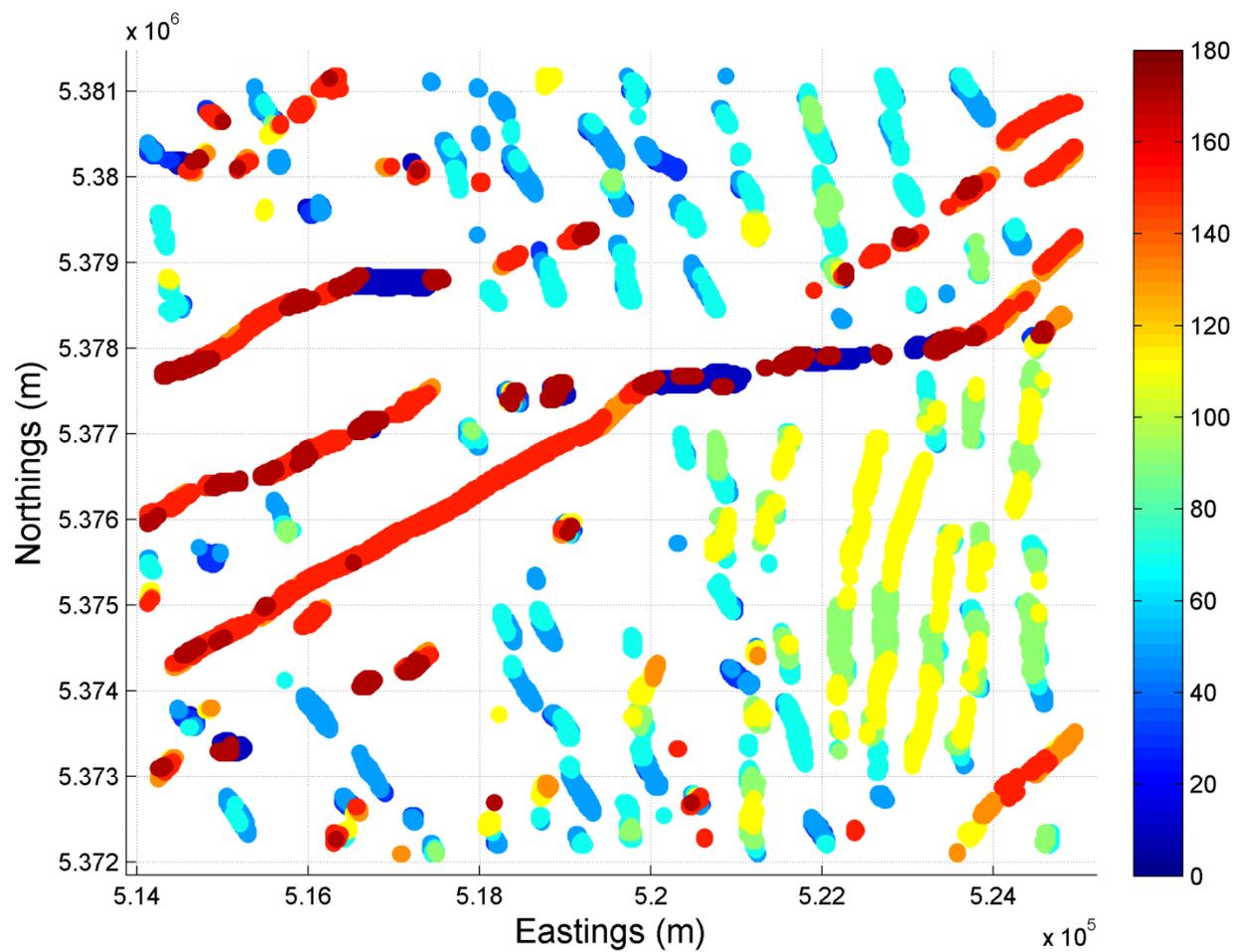


Figure 37. Predictions from the strike model on the Matheson dataset. The colourbar shows the strike estimate in steps of 20° .

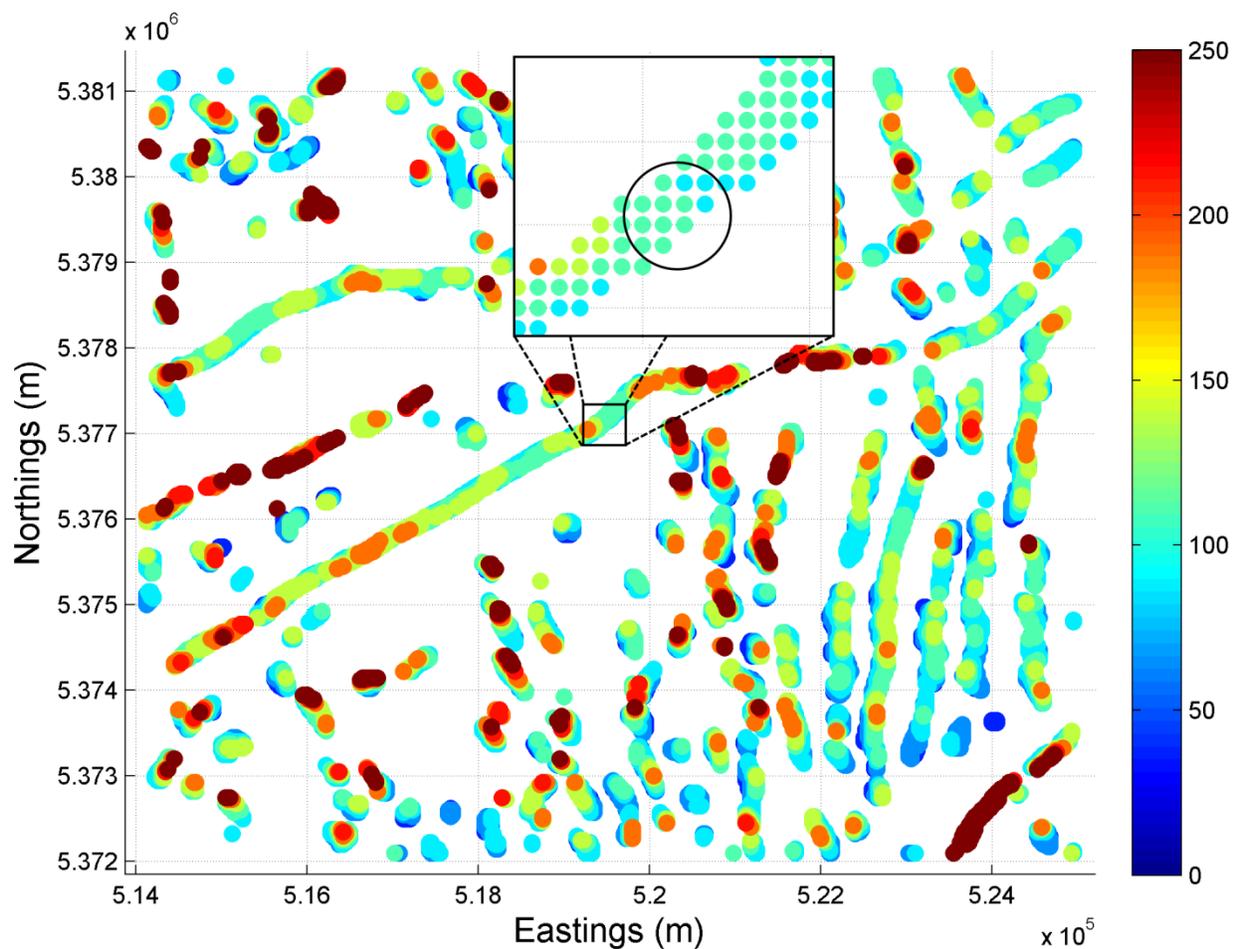


Figure 38. Predictions from the depth model on the Matheson dataset. The colourbar shows the depth estimate in steps of 25 m. The specific area of interest for comparison to other studies is shown at expanded scale in the insert section, with the expected value of the 100 m to 125 m bin being predicted.

4.5 Discussion

4.5.1 Grid Cell Size

An important factor in this process is the choice of cell size. As this methodology uses gridded aeromagnetic data, the profile data must be interpolated to create the total field grids, which requires the selection of a grid cell size. All training in this approach was completed with 25 m cells. However, in aeromagnetic data, the grid cell size is normally 20% to 25% of the flight-line spacing, which in the Matheson dataset is 40 m or 50 m. Our initial analysis found that reasonable results were obtained when a 40 m cell size was used, except there was much more variation in the values of the predicted strike and depth. This was particularly noticeable at the area of interest due to its angle with respect to the grid itself, i.e. the grid cells are oriented along the cardinal directions, but the dyke of interest strikes South-West to North-East. The interpolated data was then subsampled to 25 m (the cell size of all training data), and the result was much more consistent, as seen in Figure 38. Procedures that may ensure successful results on future datasets include training at the required cell size for a specific dataset, developing more robust CNN models by using multiple cell sizes when training, or operate on grids rotated to appropriate angles to suit the dataset.

4.5.2 Width Estimation

As observed on Figure 27, the width of the zone with red dots corresponds approximately with the width of the lineaments. Therefore, by estimating this width, we can get an upper bound on the width of the lineament. Since the method is checking for the central nine cells for classification as a “hit” (Figure 23b), it can generally be assumed that the number of cells

predicted will be overestimated, which is confirmed by the red zones in Figure 27 being wider than the black zones. For instance, L1 has a width of 33 m, and the CNN model predicts three to four cells (red dots) wide. We can subtract two cells due to the overestimation (one from either side), and assume that the lineament is no wider than two cells (50 m). Additionally, since it estimates that in some cases the zone would only be a single cell wide, it may be on the lower end of the width range (i.e. a width somewhere between a single cell, 25 m, and two cells, 50 m, likely closer to 25 m).

4.5.3 Weaknesses and Methods of Improvement

One of the obvious drawbacks of utilizing a classification CNN is that continuous values cannot be predicted. This approach allowed for an easier training step of the process, but it results in less fidelity when analyzing data. The bins we chose are small enough to be useful for interpretation; however it would be understandable for the next step of this research to attempt to develop better resolution for its results, likely using a regression CNN, although increasing the number of bins in a classification CNN may be a viable approach as well.

Another drawback of this method is that it cannot handle overlapping lineaments. With relatively minimal effort, a solution that would indicate “multiple lineaments present” within a cell could be made. However, without a fairly significant change to the methodology it is unlikely more detail than this would be possible, as it would become difficult to discern the difference between overlapping lineaments and block-like anomalies. However, if made the focus of future work,

large datasets that focus on this aspect could be developed in an effort to solve this specific use-case.

A primary area for future investigation is to advance this method to be able to interpret the dip, width, and susceptibility of lineaments. Predicting the dip and width will likely be possible using same general approach as for the depth. Given enough data with varying dip angles and widths within the range required, the network may be predictable to a similar accuracy as the strike and depth. The primary difficulty may stem from the fact that both the dip and width, in general, have a smaller range of realistic values, i.e. a realistic dip of a dyke may be only -20° to $+20^\circ$ away from vertical, and the width of dykes is generally no larger than 100 m. It is possible that rather than using the total magnetic field as an input grid, it might be easier to extract the width and dip from the first or second vertical derivative, the tilt-angle derivative (Salem et al., 2007), or all three quantities. Due to the fact that we are normalizing windows before evaluation, no information regarding susceptibility can be discerned in this approach. This is because susceptibility only affects the magnitude of the response, not the shape or size of the anomaly. It is possible that by using the methodology shown here as a baseline, and training a large amount of data without normalizing it that the susceptibility can also be estimated. However, processing steps such as IGRF correction, base station subtraction, and others would have to be carefully controlled, as each of them will affect the amplitude of the final magnetic field that is analysed.

An effective approach to the overall improvement of this method may be to use a number of large synthetic datasets, similar to the synthetic model in the Results section, and apply the same

sliding window technique to develop the training data imagery. Each dataset could be generated at multiple noise levels, and with various cells sizes, to ensure the CNN models are robust to these effects. This approach would enable a large amount of complex and more realistic windowed datasets to be developed for training.

4.6 Conclusions

We have presented a new approach for the automatic interpretation of lineaments in aeromagnetic data using deep learning. Specifically, we have developed a classification convolution neural network approach that is accurate in its predictions of the location, strike, and depth of lineaments. In the synthetic dataset it detected all six lineaments, correctly predicting the strike in four of the six, with the other two lineaments being predicted in adjacent bins. It also correctly predicted three of the depths, with another two whose values were near to the edge of the classification bins being predicted as the adjacent bin. The final lineament's depth was not as clearly predicted, however approximately 50% of the predicted cells did indicate the correct bin. Applying this method to a real world dataset, a dyke's location, strike, and depth were correctly predicted, as verified by associated borehole data. Clear avenues for improvement and future research are available, however the current approach can be used to predict the location, strike, and depth to top of lineaments in aeromagnetic maps.

4.7 References

- Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, ... and X. Zheng, 2015, TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.
- Aghaee Rad, M. A., 2019, Machine learning of lineaments from magnetic, gravity and elevation maps: M.Sc. thesis, University of British Columbia.
- Bates, M. P., and H. C. Halls, 1990, Regional variation in paleomagnetic polarity of the Matachewan dyke swarm related to the Kapuskasing Structural Zone, Ontario, Canadian Journal of Earth Sciences, 27, 200-211.
- Blakely, R. J., 1996, Potential theory in gravity and magnetic applications, Cambridge university press.
- Cascone, L., S. Campbell, C. Green, A. Salem, and D. Fairhead, 2012, ACLAS: A new automatic method of defining potential field lineaments using coherency analysis, SEG Technical Program Expanded Abstracts 2012, 1-6.
- Cooper, G. R., 2014, The automatic determination of the location, depth, and dip of contacts from aeromagnetic data, Geophysics, 79, J35-J41.
- Cooper, G. R., 2015, Using the analytic signal amplitude to determine the location and depth of thin dikes from magnetic data, Geophysics, 80, J1-J6.

Coyle, M., R. Dumont, P. Keating, F. Kiss, and W. Miles, 2014, Geological Survey of Canada aeromagnetic surveys: Design, quality assurance, and data dissemination.

Cunha, A., A. Pochet, H. Lopes, and M. Gattass, 2020, Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data, *Computers & Geosciences*, 135, *preprint article 104344*.

Curto, J. B., T. Diniz, R. M. Vidotti, R. J. Blakely, and R. A. Fuck, 2015, Optimizing depth estimates from magnetic anomalies using spatial analysis tools, *Computers & Geosciences*, 84, 1-9.

de Barros, A., S. Bongiolo, J. de Souza, F. J. F. Ferreira, and L. G. de Castro, 2013, Grav mag prism: a matlab/octave program to generate gravity and magnetic anomalies due to rectangular prismatic bodies, *Brazilian Journal of Geophysics*, 31, 347-363.

Ernst, R. E., 2014, *Large Igneous Provinces*, Cambridge University Press.

Fairhead, J. D., and S. E. Williams, 2006, Evaluating normalized magnetic derivatives for structural mapping, *SEG Technical Program Expanded Abstracts 2006*, 845-849.

Ferreira, F. J. F., J. de Souza, A. de B. e S. Bongiolo, and L. G. de Castro, 2013, Enhancement of the total horizontal gradient of magnetic anomalies using the tilt angle. *Geophysics*, 78, J33-J41.

Foks, N. L., and Y. Li, 2016, Automatic boundary extraction from magnetic field data using triangular meshes, *Geophysics*, 81, J47-J60.

Geosoft Inc., 2020, Oasis Montaj.

Hidalgo-Gato, M. C., and V. C. Barbosa, 2015, Edge detection of potential-field sources using scale-space monogenic signal: Fundamental principles, *Geophysics*, 80, J27-J36.

Holden, E. J., J. C. Wong, P. Kovesi, D. Wedge, M. Dentith, and L. Bagas, 2012, Identifying structural complexity in aeromagnetic data: An image analysis approach to greenfields gold exploration, *Ore Geology Reviews*, 46, 47-59.

Hood, P., 1965, Gradient measurements in aeromagnetic surveying, *Geophysics*, 30, 891-902.

Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012, Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 1097-1105.

Lee, M., W. Morris, J. Harris, and G. Leblanc, 2012, An automatic network-extraction algorithm applied to magnetic survey data for the identification and extraction of geologic lineaments, *The Leading Edge*, 31, 26-31.

Miller, H. G., and V. Singh, 1994, Potential field tilt—a new concept for location of potential field sources, *Journal of Applied Geophysics*, 32, 213-217.

Nabighian, M. N., 1972, The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: its properties and use for automated anomaly interpretation, *Geophysics*, 37, 507-517.

Nair, V., and G. E. Hinton, 2010, Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on Machine Learning, 807-814.

Naprstek, T., and R. S. Smith, 2019, A new method for interpolating linear features in aeromagnetic data: *Geophysics*, 84, no. 3, 1-35, doi: 10.1190/geo2018-0156.1. [Chapter 2 of this thesis]

Nathan, D., A. Aitken, E. J. Holden, and J. Wong, 2020, Imaging sedimentary basins from high-resolution aeromagnetism and texture analysis, *Computers & Geosciences*, 136, 104396.

Nurindrawati, F., and J. Sun, 2019, Estimating total magnetization directions using convolutional neural networks, *SEG Technical Program Expanded Abstracts*, 2019, 2163-2167.

Ontario Geological Survey, 2003, Ontario airborne geophysical surveys, magnetic and electromagnetic data, Matheson area, Geophysical data set 1101 - Revised.

Oliveira, S. P., F. J. Ferreira, and J. de Souza, 2017, EdgeDetectPFI: An algorithm for automatic edge detection in potential field anomaly images—application to dike-like magnetic structures, *Computers & Geosciences*, 103, 80-91.

Peters, B., E. Haber, and J. Granek, 2019, Neural networks for geophysicists and their application to seismic data interpretation, *The Leading Edge*, 38, 534-540.

Phillips, J. D., 2000, Locating magnetic contacts: a comparison of the horizontal gradient, analytic signal, and local wavenumber methods, SEG Technical Program Expanded Abstracts 2000, 402-405.

Pilkington, M., and P. B. Keating, 2009, The utility of potential field enhancements for remote predictive mapping, Canadian Journal of Remote Sensing, 35, S1-S11.

Pilkington, M., and V. Tschirhart, 2017, Practical considerations in the use of edge detectors for geologic mapping using magnetic data. Geophysics, 82, J1-J8.

Pochet, A., P. H. Diniz, H. Lopes, and M. Gattass, 2018, Seismic Fault Detection Using Convolutional Neural Networks Trained on Synthetic Poststacked Amplitude Maps, IEEE Geoscience and Remote Sensing Letters, 16, 352-356.

Poulton, M. M., 2002, Neural networks as an intelligence amplification tool: A review of applications, Geophysics, 67, 979-993.

Reid, A. B., J. M. Allsop, H. Granser, A. T. Millett, and I. W. Somerton, 1990, Magnetic interpretation in three dimensions using Euler deconvolution. Geophysics, 55, 80-91.

Riesenhuber, M., and T. Poggio, 1999, Hierarchical models of object recognition in cortex, Nature Neuroscience, 2, 1019-1025.

Roest, W. R., J. Verhoef, and M. Pilkington, 1992, Magnetic interpretation using the 3-D analytic signal, Geophysics, 57, 116-125.

Ruffman, A., and J. D. Greenough, 1990, The Weekend dykes, a newly recognized mafic dyke swarm on the eastern shore of Nova Scotia, Canada, *Canadian Journal of Earth Sciences*, 27, 644-648.

Tschirhart, P., and B. Morris, 2012, Grenville age deformation of the Sudbury impact structure: evidence from magnetic modelling of the Sudbury diabase dyke swarm, *Terra Nova*, 24, 213-220.

Saikia, P., R. D. Baruah, S. K. Singh, and P. K. Chaudhuri, 2020, Artificial Neural Networks in the domain of reservoir characterization: A review from shallow to deep models, *Computers & Geosciences*, 135, 104357.

Salem, A., D. Ravat, R. Smith, and K. Ushijima, 2005, Interpretation of magnetic data using an enhanced local wavenumber (ELW) method, *Geophysics*, 70, L7-L12.

Salem, A., S. Williams, D. Fairhead, R. Smith, and D. Ravat, 2007, Interpretation of magnetic data using tilt-angle derivatives, *Geophysics*, 73, L1-L10.

Scherer, D., A. Müller, and S. Behnke, 2010, Evaluation of pooling operations in convolutional architectures for object recognition, *International Conference on Artificial Neural Networks*, 92-101.

Shi, Y., X. Wu, and S. Fomel, 2018, Automatic salt-body classification using a deep convolutional neural network, *SEG Technical Program Expanded Abstracts 2018*, 1971-1975.

Simard, P. Y., D. Steinkraus, and J. C. Platt, 2003, Best practices for convolutional neural networks applied to visual document analysis, *International Conference on Document Analysis and Recognition 2003*, 958-963.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, 2014, Dropout: a simple way to prevent neural networks from overfitting, *The Journal of Machine Learning Research*, 15, 1929-1958.

Sun, Y., W. Yang, X. Zeng, and Z. Zhang, 2016, Edge enhancement of potential field data using spectral moments, *Geophysics*, 81, G1-G11.

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

Thurston, J. B., and R. S. Smith, 1997, Automatic conversion of magnetic data to depth, dip, and susceptibility contrast using the SPI (TM) method, *Geophysics*, 62, 807-813.

Thurston, J., and R. Smith, 2007, Source location using total-field homogeneity: Introducing the SLUTH method for depth estimation, *The Leading Edge*, 26, 1272-1277.

Tontini, F. C., R. J. Blakely, V. Stagpoole, and H. Seebeck, 2018, Semi-automatic determination of dips and depths of geologic contacts from magnetic data with application to the Turi Fault System, Taranaki Basin, New Zealand, *Journal of Applied Geophysics*, 150, 67-73.

Tschirhart, P., and B. Morris, 2015, Improved edge detection mapping through stacking and integration: a case study in the Bathurst Mining Camp, *Geophysical Prospecting*, 63, 283-295.

Ulla, J., R. Smith, C. Samson, and M. Vallée, 2010, Automation of the SLUTH method: a novel approach to airborne magnetic data interpretation, *Near Surface Geophysics*, 8, 519-528.

Vallée, M. A., P. Keating, R. S. Smith, and C. St-Hilaire, 2004, Estimating depth and model type using the continuous wavelet transform of magnetic data, *Geophysics*, 69, 191-199.

Verduzco, B., J.D. Fairhead, C. M. Green, and C. MacKenzie, 2004, New insights into magnetic derivatives for structural mapping, *The Leading Edge*, 23, 116-119.

Waldeland, A. U., and A. H. S. S., Solberg, 2017, Salt classification using deep learning, 79th EAGE Conference and Exhibition 2017, 1-5.

Yang, L., and S. Z. Sun, 2020, Seismic horizon tracking using a deep convolutional neural network, *Journal of Petroleum Science and Engineering*, 187, 106709.

Chapter 5

5 Conclusion

5.1 Summary of Research

This thesis has shown the development of two new aeromagnetic interpolation approaches and a new aeromagnetic interpretation method, with an overall focus on enhancing the results associated with lineament type anomalies.

In Chapter 2, a new interpolation method (now referred to as “multi-trend gridding”) was developed with the goal of interpolating thin, linear features without artifacts. The method is an iterative, numerical derivative approach, where strong trends are found within the data, and enhanced across the flight lines in an effort to preserve lineaments. A key step in the process is normalization, wherein the real data from the flight lines is preserved, and nearby data is increased or decreased depending on the local spatial anisotropy. It offers significant improvement over traditional methods when interpolating linear features that are at acute angles to the flights lines, allowing for a final grid that is much more appropriate for lineament analysis. However, due to its primary goal of interpolating lineaments, other, more nebulous features such as large block-like anomalies with no strong linear pattern are not as well interpolated when compared to traditional methods.

Following the research in Chapter 2, I was interested in overcoming the weakness of multi-trend gridding in dealing with non-lineament anomalies, without sacrificing the strength of

interpolating lineaments. Chapter 3 investigated this through the use of machine learning techniques applied to aeromagnetic interpolation, without the explicit use of derivatives, which can be noisy in some cases. The research showed that support vector machine (SVM) and random forest (RF) methods may be able to interpolate entirely on their own, but they are primarily effective at combining the best features from various interpolation methods into a single, final solution. The final aeromagnetic map was more accurate than any of the single interpolation results, although only marginally. These non-“deep learning” approaches to machine learning give the researcher more control, as they are able to input data that they believe to be important, such as including the varying neighbourhood information in Chapter 3. However, compared to deep-learning approaches, this means that less “hidden” connections will generally be found, and require a much more focused effort by the researcher to arrive at new and effective results. There is a wealth of information and studies on interpolation/image-enhancement using SVMs and RFs, as well as geoscience applications wherein multiple physical datasets are combined to predict a related feature. However, aeromagnetic interpolation falls somewhere between these, as it is similar to an image enhancement, but still has physical properties that can be used to constrain the results. As such, this area could be further investigated; however it is possible the wrong tools are being used to solve the problem.

Chapter 4 investigated the interpretation of linear features in aeromagnetic data using Convolution Neural Networks (CNNs). By developing an extensive amount of synthetic data, two CNN models were trained to be able to locate lineaments in aeromagnetic maps, as well as estimate their strike and depth to the top of the source. Applying two models to a real-world

dataset, it was found that the depth to the top of a dyke was estimated correctly when compared to borehole data. Location and interpretation techniques work better if the data is trended and linear like the underlying geology, and free from interpolation artifacts. Therefore the case study dataset was interpolated using the multi-trend gridding method from Chapter 2. I did not use the machine learning method developed in Chapter 3, as the results would require greater effort and only be marginally better on the linear feature of interest. This approach developed the CNNs as classification models, where each class was a small range of values, in steps of 20° for the strike and 25 m for the depth. Compared to a regression model, this approach allowed for a quicker training process, and more confidence in each predicted bin range.

Overall, these three Chapters form a new pipeline approach for interpolating and interpreting lineaments in aeromagnetic data. After data collection and standard processing (IGRF subtraction, levelling, etc.), the profile data can be interpolated using Chapter 2's multi-trend gridding method. If the dataset contains multiple types of features that the user wishes to investigate (e.g. both lineaments and rounder anomalies), Chapter 3's filter-like interpolation may optionally be used to bring the strengths of multiple interpolation methods together. Finally, Chapter 4's CNN models may be applied to the aeromagnetic map, resulting in estimations for the location, strike, and depth of any lineaments present. This new pipeline offers a valuable set of tools for anyone interested in understanding lineament-type features in aeromagnetic data.

5.2 Future Research

5.2.1 Investigate Multi-Trend Gridding for Additional Uses

Overall, the interpolation method developed in Chapter 2 is fairly complete. However, a potential area for future development would be to investigate any use for the trending values that are developed during the iterative process. The structure tensor and associated data calculated is used to focus the trending interpolation by understanding the local anisotropy. It is possible that this information could be used for additional interpolation purposes or for supplemental interpretation purposes since they are effectively calculating localized gradients. These small-scale changes could provide potentially useful information to a skilled interpreter, or be given to an automatic interpretation method. For example, the trend direction obtained could be fed into Chapter 3's interpolation methodology as a separate predictor set, offering valuable localized information.

5.2.2 Investigate Gridding Cell Effects

A fundamental issue appeared while I was comparing Chapter 2's new interpolation method to traditional methods and their implementation in geophysical software. Cell size, cell position, and the cell values can all be handled very differently, while on the surface appearing to be the exact same approach. Before interpolating, a decimation process must be applied on the original flight line data, to reduce the high-spatial-density data to a lower-spatial density appropriate to the grid. This decimation includes some form of averaging. The first part of this process is setting the cell positions and sizes. In my method, I set the cell position to be at the center, and measure out half a cell size in the x and y directions (both positive and negative) to define the

extent of each cell. However, another popular approach is to set the cell reference position to the minimum x and y position in the cell, and measure out a full cell in the positive x and y directions. This simple difference can have noticeable effects when comparing grids of aeromagnetic data, particularly if the cells include local minimas or maximas. Likewise, the approach for averaging the flight line data within the cells can differ. The simplest of these is to take the average of all values that fall within the cell. However, just as easily a distance-weighted average of the same values could be taken, as measured from the center of the cell (or the bottom-left if the cell reference point is at the minimum x and y position in the cell). Once again, this generally will not greatly affect the final results, but when applying advanced interpretations involving multiple derivatives, these factors are very important to consider, and anyone analyzing aeromagnetic maps should understand how the location of the grid and cells are defined.

Similar to Chapter 2, a primary point of interest during Chapter 4's research involved cell sizes. All training data was developed at a cell size of 25 m, and while they could be applied to datasets with other cell sizes, in the case study I found that the results were not as optimal when the cell size was left at 40 m. By simply subsampling the Matheson dataset to 25 m, a more consistent solution was found for the depth estimates. The cell size also affected the minor conclusion that these CNN models might have the potential to predict the width of the lineaments. Since each model first determined if a cell included a lineament, the resulting scatter plot of lineament detection could be simply counted to determine how many cells a lineament covered, perpendicular to its strike. However, the cell discretization limits the accuracy of this technique.

A study on varying cell positions and averaging methods during the gridding process, and the resulting effects on both the interpolation and interpretation would be beneficial. Based on the information from both Chapter 2 and Chapter 4, I believe there are a number of follow-on effects that these processes influence. A simple study where these are varied on a specific dataset, focussing on a specific anomaly may show how great of an effect these can have on the final data products.

5.2.3 Utilize Deep Learning Methods for Aeromagnetic Interpolation

A clear area for future improvements to the work described in Chapter 3 is to utilize deep learning methods in the three different approaches shown therein. For instance, using neural networks to interpolate may improve the results over approaches 1 and 2, and utilizing a CNN in approach 3 may garner an improvement over the SVM and RF regression methods.

5.2.4 General Improvements to a CNN-Based Parameter Estimator

A significant number of potential areas of research were identified in Chapter 4. While the methodology shown there is effective, a number of lessons learned and improvements to the approach were made clear through the investigation. I believe the best approach to improving the depth estimation would be to begin by developing a more robust synthetic dataset for training. A series of models (similar to the synthetic set that the method was tested on in the Results section) should be created and sampled at various cell sizes and with various amounts of noise. These could then be used to train a full regression CNN, which may improve the resolution of the results.

5.2.5 Estimate Additional Parameters Using CNNs

Following the methodology in Chapter 4, estimating the width and dip of a lineament may be possible as well. One approach might be to simply add them in as factors to the CNN model. The caveat to this is that these two physical parameters generally vary less than the depth (i.e. the dip in a lineament is generally only non-vertical by 10° to 20°), and therefore detecting the relatively smaller changes would require a focused effort from the training data. Another approach might be to modify the CNN to incorporate other data such as the vertical derivative and/or tilt derivative, which may better show the changes from these other physical parameters.

To use the approach of Chapter 4 to solve for the susceptibility, the normalization step would need to be removed. If this is done, then a very similar method could be used, however it would require much more data for training, given the large variability of realistic susceptibility values. The interaction between host rock and anomalies would also need to be extensively modelled, as the susceptibilities of the host and lineament would affect the total amplitude values.

5.2.6 Develop a Similar Parameter Estimator for Other Anomaly Types

The overall approach from Chapter 4 could also potentially be transferred to other interpretations for aeromagnetic data, such as parameter estimation for block-like anomalies, rather than just the linear features. This would obviously require an equal amount of training data with variations focused on the parameters that affect those features, however in theory the method would be equivalent.

5.2.7 Develop a Machine Learning Parameter Estimator/Inversion Using Only the Aeromagnetic Flight Data

A final potential area for research would be to essentially combine the interpolation and interpretation components of aeromagnetic data. If a machine learning model was trained using only the flight line data itself, but given the spatial context of all data, it is feasible that it could be made to estimate anomaly locations and source parameters, allowing for a full use of the high spatial-resolution flight line data. If properly set up, this approach could essentially act as an inversion, calculating the most likely location and size of various anomalies, and use that information to develop a final map for the user to view. This map would be a mix of the measured flight data and an interpolation using forward model data from the estimated source bodies.