INTEGRATING GEOLOGIC DATA FROM THE THOMPSON NICKEL BELT INTO A GIS

By

DIANNE MICHALAK

A Thesis Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

Master of Science

Department of Geological Sciences
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

INTEGRATING GEOLOGIC DATA FROM THE THOMPSON NICKEL BELT INTO A GIS

BY

DIANNE MICHALAK

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

of

MASTER OF SCIENCE

Acknowledgements

Many thanks are extended to all who contributed to this document, in one way or another...

First, thanks to Dr. Halden for taking on the system and adopting an Interdisciplinary student. The project was entertaining and the people were inspirational, especially the CAMIRO TNB project researchers. Their dedication to science was motivating. Thank you to the CAMIRO TNB project industry and government sponsors who provided the data.

Introduction to the concept of data fusion is attributed to Dr. Moon. Dr. Ferguson and Dr. Benbow are credited with the task of editing and providing invaluable feedback. As a result, my thesis has seen more of the world than I, having travelled to Australia and back. Thanks, mate!

A big thank you to the folks from the Oceans Program Division at DFO for the opportunities to practice GIS as a professional as well as the ongoing support and interest in my work.

Special friends made the last two and a half years endurable, but more importantly, enjoyable. Specifically, Marta and Frank, Mandy, Kim, Jarod and Ami, Lisa, Marilyn and Nosa have all helped to keep this system functioning.

And, of course, the Family. Thank you.

Abstract

Integrating Geologic Data from the Thompson Nickel Belt into a GIS

Dianne Michalak

As the geologic community moves towards a digital environment, many still have to be convinced that data management is as necessary as chemical analysis. In fact, assessing the data and software needs of a project from a systems perspective is ideally suited to a GIS expert. The expert is able to assemble very detailed datasets into a comprehensive overview of a region. The interaction between people and data has to be monitored to determine the success of a system. Identifying limitations in terms of processes and software and adapting solutions to meet the challenge is also part of the system approach.

GIS project management issues particular to geological applications will be presented in the context of the CAMIRO Thompson Nickel Belt (TNB) project. Data are considered to be a resource and must be appropriately managed to realize its full potential. The benefits of designing and implementing a structured system of data management can be realized through time saving measures, efficient data storage and retrieval and the extraction of valuable information. Components of the system – people, processes, data, software and hardware – are examined with respect to the needs of the project.

Application of spatial analysis techniques and data fusion were used to create mineral potential maps of nickel for the TNB. Measures of belief, probability, disbelief and uncertainty present a flexible way of quantifying data and are used to weight attributes in data layers. The data layers are then combined to produce mineral potential maps. The process of entering, manipulating, calculating and interpreting data influences potential results.

This project identified issues that should be addressed in multidisciplinary projects. A realistic assessment of participants technical resources and timelines for data acquisition can increase use of data in GIS, hopefully evolving from simple visualization to more integrated analysis.

Table of Contents

Index of Tables

Index of Figures

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Geographic Information Systems (GISs) have functioned as a tool in many scientific disciplines in support of the decision making process. Its multidisciplined appeal is a result of the ability to combine the spatial location and quantitative and qualitative attributes of features. It is an obvious choice of tool for earth scientists. In particular, geologists have recognized the potential of GISs and have begun to structure their datasets for display, interpretation, analysis and modeling in a digital environment. However, the large volumes and diversity of geologic datasets pose a particular challenge with respect to data integration and answering geologic questions. Data fusion is one approach towards investigating spatial relationships amoungst multiple datasets.

GIS can be defined as either a tool or a system and, as a result of the terms duplicity, a range of definitions is necessary. Burrough (1986) proposes that a GIS is “a powerful set of tools for storing and retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes”. This definition focuses on the functions of a GIS. Alternatively, Star and Estes (1990) focus on the system aspect of a GIS in their definition: “An information system that is designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially-referenced data, as well as a set of operations for working with the data”.

Data fusion is a collection of tools, algorithms and data and is therefore difficult to define without restricting the full scope of the concept. Attempts at balancing the definition result in descriptions that mix what data fusion *is* with what data fusion *does* (Wald, 1999). Mangolini (1997) considers what data fusion does as "a set of methods, tools and means using data coming from various sources of different nature, in order to increase the quality (in a broad sense) of the requested information". The United States Department of Defense (DOD; Klein, 1993) developed their definition to read that data fusion is a "multilevel, multifaceted process dealing with the automatic detection, association, correlation, estimation, and combination of data and information from single and multiple sources".

Conceptually, there are three different levels of fusion encompassing the pixel, attribute and decision (Pohl and van Genderen, 1998; Mangolini, 1994; Brandstatter and Sharov, 1998; Csatho and Schenk, 1998). Individually, these definitions do not reflect all the levels, nor other issues like quality or a framework concept. To fill the data fusion definition void, Buchroithner (1998) and Wald (1998) asserted that "data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater quality' will depend upon the application".

1.2 PROJECT SETTING

The thesis was developed as part of a geological project connected with CAMIRO (Canadian Mining Industry Research Organization) Project 97E-02. The area of interest extends along the Thompson Nickel Belt (TNB) in central Manitoba. The CAMIRO project was established as a joint venture between industry, government and academia to promote research and aid mineral exploration in a nickel producing region. The four year interdisciplinary project incorporates geological, geochronological, geophysical and geochemical studies that will:

1) Define the geology, stratigraphy, structure, geochronology and petrogenesis of the TNB with a particular emphasis on metallogenesis;
2) Refine geological, geochemical, and geophysical exploration tools relevant to the TNB and other komatiitic terranes, and
3) Aid in the identification of new exploration targets. The overall objective is to create a belt-wide synthesis which will allow us to compare and contrast the properties of the ore throughout the TNB as a guide to exploration.

The project was managed as two subprojects: 1) regional mapping, geochronology, tectonics, geophysics, and GIS data management and analysis; and 2) regional mapping, geochemistry and ore genesis. Data acquisition, organization, visualization, analysis and presentation used state-of-the-art GIS methods.

There are six objectives of the CAMIRO project:

1) Produce updated stratigraphic columns for the northern, central, and southern parts of the TNB, particularly the various lithofacies of the Ospwagan Group and the occurrences of ultramafic rocks within that sequence;

2) Define the detrital zircon age characteristics of the metasedimentary units and link them with precise U-Pb ages for mafic and ultramafic magmatism and for the adjacent Superior Province basement and Reindeer Zone;

3) Undertake detailed structural analysis of the TNB in the exposed shield and beneath the Phanerozoic cover using field, drillcore, seismic, magnetotelluric (MT) and potential field data;

4) Design, implement and undertake analysis of an integrated digital geoscience knowledge base system (GIS) for the overall CAMIRO project;

5) Generate a 3D model of subsurface structure from available geological, gravity, magnetic, seismic, rock property and bore hole data using GIS and 3D software tools; and

6) Refine regional structure, tectonic and metallogenetic models.

Each participant in the project contributes support through financial funding, release of data, field work, data analysis, theory development or new mapping. New information is combined with existing data in an effort to determine what knowledge exists as well as what gaps exist in the available data.

The data are compiled in a GIS for analysis and manipulation. The GIS component of the project extends to data management, creation of metadata and implementation of data fusion and other spatial techniques.

The results presented in this thesis reflect the experience of working with a variety of professionals including geochemists, geochronologists, geophysicists, structural geologists, company scientists, field geologists and technicians. Each contributed a wealth of knowledge about their datasets, the methods used to collect, analyze and interpret the information and suggestions as to how the information could be incorporated in a regional model.

Introducing individual contributions into a comprehensive collection required that the data met the needs of the project at a variety of scales. Perspectives ranged from microscale crystal analysis of pentlandites to crustal-scale structural features found along the TNB. Data and information obtained from localized studies are used to develop a regional model of the TNB and emplacement of nickel deposits.

The project builds on recent and ongoing investigations in the TNB: 1) CAMIRO Project 94-04 (Keays, Lesher, and Hulbert - Application of PGE Geochemistry to Exploration for Ni Sulphide Deposits); 2) the Manitoba Geological Services Branch's TNB project; and 3) Lithoprobe's Trans-Hudson Orogen Transect. Integration of results from these investigations provides a knowledge base from which the CAMIRO program can build.

The approach of building on previous work and integrating datasets is becoming popular as is the pooling of resources between industry, government and academia. Current multidiscipline geologic projects at the provincial level include Operation Superior: Northern Superior Province (2000, Manitoba Industry Trades and Mines). At the national level, the Geological Survey of Canada (GSC) is involved in projects such as EXTECH II (Exploration Technology Program II), NATMAP (Canada's National Geoscience Mapping Program), Lithoprobe, Ocean Drilling Program, NATGAM (National Gamma Ray Spectrometry Program) and Coronation Gulf Mineral Development Area (Nunavut) (2000, NRCAN).

1.3 BACKGROUND

This research will investigate the process of integration geologic data into a GIS for the purpose of spatial analysis, including application of two methods of data fusion, the Fuzzy Logic (FL) method and the Dempster-Shafer Belief Theory (DSBT). Mineral exploration provides an opportunity to incorporate digital data management techniques and GISs with geological data fusion; an important link connecting the disciplines lies in data management. Geologists have historically worked with hard copy data. Current trends in the field reveal an emphasis towards digital data recording, manipulation and interpretation. Moreover, when data collected in the field is combined with geochemical and petrographic information, large amounts of data have to be managed and geographically referenced. New perspectives and techniques, particularly computer based skills and effective methods of data dissemination, are needed to manage and integrate the various data sets used by geologists. GISs is an environment that proves to be an effective and powerful tool for storage, display and analysis. Once in digital format, the data can be manipulated in many ways. Data fusion shows great potential for decision making purposes, particularly mineral exploration and in a broader context, land use.

Four themes will run through the thesis:

1) The transition of mapping techniques from traditional pen and paper field sketches to data collection via digital methods;
2) Incorporation of a variety of data formats into one system;
3) Information management; and
4) Presentation of experimental work.

From a systems perspective, there are five components to a GIS: People, Processes, Data, Software and Hardware (PPDSH). An information systems project is based on a well-designed system plan and will be explained with regards to PPDSH (Donaldson Dewitz, 1996).

1.4 OBJECTIVES

The chapters are organized to introduce data management issues particular to a geological project. The main goal is to demonstrate the importance data plays in the decision making process. Data are considered to be a resource and must be appropriately managed to realize its full potential. The benefits of designing and implementing a structured approach to data management can be realized through time saving measures, efficient data storage and retrieval and the extraction of valuable information. If the data needs of a project are properly assessed early on, time and money can be saved by eliminating the costs of acquiring inappropriate data. A range of criteria detailing specific data characteristics will be presented to help data managers assemble a robust collection. The management of data is a means to provide quality material to the end user.

The objectives of the thesis are:

1. To identify sources of data and parameters such as accuracy, method of data collection and propagation of error.
2. To outline the process of incorporating data into a multidiscipline project - how is it accessed, who can use it, how can it be used and how it can be integrated with other data.
3. To document issues concerning data distribution including first access to data, processing and redistribution of results and data compilations.
4. To determine the data needs of the geologic community, notably academic, industrial and governmental agencies.

5. To develop an exploration tool specific to the TNB using mineral potential maps created by data fusion techniques.

1.5 THESIS STRUCTURE

Chapter 1 introduces the CAMIRO project and defines GISs and data fusion. The components of a geological project will be identified in chapter 2, Systems Approach. Issues particular to Data Management will be addressed in Chapter 3. The geological context will be introduced in chapter 4, followed by application of Spatial Analysis in Chapter 5. Data Fusion will be addressed in Chapter 6. Discussion and concluding comments will end the thesis.

CHAPTER 2

SYSTEMS APPROACH TO A GIS PROJECT

2.1 INTRODUCTION

A Geographic Information System can be incorporated into a project of any size, ranging from a simple display of events on a map to an enterprise-wide business solution. The wide range of applications is possible because many components can be included in a GIS (Obermeyer and Pinto, 1994). In order to appreciate the scope of GISs, the term must be defined and put in context. Each word in the term has significant meaning. "Geographic" implies a spatial context, indicating that the location of a feature on earth is important. "Information" about that feature is the next element. Many databases store numerous observations about events, however information in a GIS is spatial in nature in that every piece of information, or attribute, can be linked to a location. The "System" component of a GIS suggests the need for a set of processes. The processes can exist within a specific software package, as a set of steps to bring data into software or as a more complicated network facilitating input, output and feedback between nodes in an organization.

The most familiar context where the term "GIS" is used is in reference to a piece of software. ArcInfo, MapInfo, ArcView, Maptitude and Idrisi are examples of well-known products commonly referred to as GISs. In this context, GISs are applied as tools to perform specific tasks such as data display or analysis. As a system, GIS can describe the operational infrastructure and processes involved with data. The infrastructure includes hardware such as computers, plotters and digitizers; examples of processes include quality control checks, metadata collection and maintenance of datasets. Some parties prefer to use the term Geographic Information Science. Goodchild (1992) defined Geographic Information Science as "the generic issues that surround the use of GIS technology, impede its successful implementation, or emerge from an understanding of its potential capabilities." By replacing "system" with "science", some believe the term relays the impression that science is a large part of applying a GIS to a problem (Schuurman, 1999). The facility to model complex earth systems quantitatively and make predictions about how the system will respond is based in scientific method. Those who use a GIS in this capacity are still using it as a tool in a larger scientific study dealing with spatial and temporal data. Whether as a science or a system, a GIS integrates many facets of a complex spatial problem and offers great potential for exploring data with quantitative and qualitative methods (Clarke, 1997).

As the application of GISs in numerous business and research environments continues to evolve, the stratification of GISs from a system to a science may become increasingly important, particularly from a staffing perspective. Anyone can learn how to use a piece of software but it is more challenging to implement a multifaceted research program based on scientific principle.

A GIS expert has a unique set of skills that pull the Geographic, Information and System elements in a GIS together. The skill set includes software-specific functionality in addition to cross platform integration of hardware and software to meet the needs of the project and to make data convey required information.

2.2 DATA, INFORMATION, KNOWLEDGE AND GIS

GIS is a tool that supports the decision making process, based on the concept that data leads to information and information leads to knowledge. Decision makers rely on their knowledge of an issue. The knowledge they have results from available information. The information is a product of data and how the data are summarized. Figure 2.1 contains a cartoon of the Knowledge Pyramid. It explains how a GIS supports the evolution of data from information to knowledge.

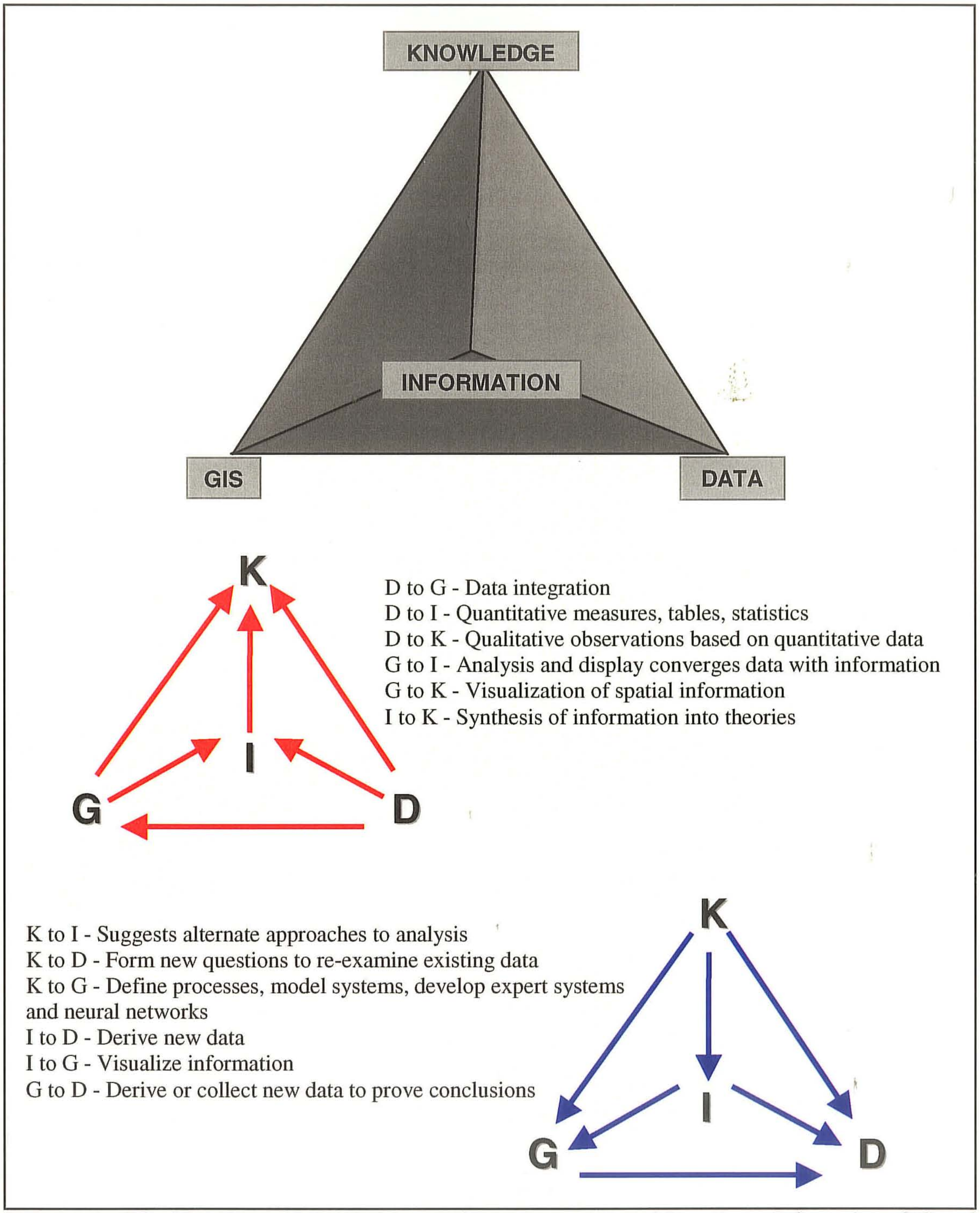
Raw data in tabular form is difficult to summarize. Graphs and charts are accepted visual aids in many disciplines. Often, the data in graphs and charts are spatial in nature. Displaying the data in its spatial context often reveals new information not seen in traditional media. Spatial trends can be visualized in a GIS by creating thematic maps of attributes in a dataset. Colours, sizes and symbols can be customized to highlight results. By classifying data into categories, the data are summarized as information.

Information results from data. The process from data to information is simple because facts are placed in context with other facts. Observations are based on how the data are displayed and spatial and temporal trends can be detected by comparing datasets captured at different points in space and time. As more information accumulates, observations can change from noting what occurs at a snapshot in time and space to noting the changes occurring over time and space. The evolution marks the entrance of knowledge.

According to the Merriam-Webster Online Collegiate Dictionary (Merriam-Webster Online 2001), knowledge is defined as “the sum of what is known : the body of truth, information, and principles acquired by mankind”. The truthfulness of knowledge depends on the robustness and accuracy of data and information that lead to its being.

The evolution of information from data to knowledge is dependent on observation. Improving data and information display affects the quality of knowledge. GISs provide visualization and classification tools that enhance data display. The process does not stop at knowledge. Instead, knowledge can lead to redesigning sampling surveys for data collection. Knowledge can in turn become data as the sum of what is known increases. Further study is thereby supported with the same system. GISs also provide analysis and reclassification tools to help summarize information. Additionally, data and information are combined with knowledge when systems are modeled in a GIS.

A GIS functions as the link between data, information and knowledge. For the link to work, all the components of the GIS must compliment each other. The components of a typical GIS include devices for data input, display, storage, analysis and output.



D to G - Data integration
D to I - Quantitative measures, tables, statistics
D to K - Qualitative observations based on quantitative data
G to I - Analysis and display converges data with information
G to K - Visualization of spatial information
I to K - Synthesis of information into theories

K to I - Suggests alternate approaches to analysis
K to D - Form new questions to re-examine existing data
K to G - Define processes, model systems, develop expert systems and neural networks
I to D - Derive new data
I to G - Visualize information
G to D - Derive or collect new data to prove conclusions

Figure 2.1: The Knowledge Pyramid is a graphic representation of how Data, Information, GIS and Knowledge relate to each other. GIS is a tool used to manipulate data, extract information and present knowledge.

2.3 GIS COMPONENTS

A GIS needs data to function. Getting data into a GIS is the most time consuming and costly aspect of any project (Korte, 1997). In addition, the data must be organized and managed with focus placed on meeting project goals. How data will be entered into the database is as important as what data to enter. Input devices such as scanners, digitizers, graphics tablets, Global Positioning Systems (GPS) and digital cameras combined with keyboard entry can accommodate most methods of transferring data to a digital environment. The time, costs and skills required for each option must be determined and then matched to available resources. It is worth repeating that implementing data input is typically the most expensive phase of a GIS project.

Once data are entered into a GIS, it must be visualized. Display options include dynamic interaction between a presenter and data with a graphical user interface (GUI) on a screen or through a data projector. Static displays such as posters or reports can also be used. Effective display of spatial data optimizes inherent properties of the data structure. Spatial data structure includes raster images and vectors (points, lines and polygons). The display medium should communicate information derived from the dataset.

Data in a GIS format are a valuable resource. Access to the files should be efficient and appropriate for the project. Often, large volumes of data build up over the life of a project. To effectively use the data collection, it must be stored in a medium that users can access. For example, field crews in remote locations may not have access to offsite network connections and in this situation, saving data to compact disks may provide the optimum storage solution. Alternatively, a busy office environment involving many users could benefit from network distributed data sets so that the data manager would not have to burn endless copies of data to compact disk. Resources also must be specified for data backup.

Internet map servers (IMSs) have emerged as a viable data dissemination solution for some organizations. Data can be accessed at any time from any office or location with internet connections, therefore any individual can view or download data. Most IMSs can be customized to different levels of security thereby creating a flexible tool that can meet demands on any organization. Alternatively, tape drives, zip disks, compact disks, RAIDs and data servers offer different levels of flexibility; in combination, they can work to meet the logistical needs of the project.

Successful data analysis is a result of planning and organizing the data such that questions can be answered with the chosen software. Topology, single versus double precision, editing capabilities and field formats are only a few issues that, if left unconsidered, can considerably compromise a project (Burrough and McDonnell, 1998). Although most GIS programs implement the same functions, each vendor provides functions specific to niches in the market. To choose optimal software, a list of needs should be compared to product specifications before a purchase is made. Alternatively, an application can be developed specifically for a project. Even more planning is necessary with custom-made software since all functions must be programmed. Long-term projects have additional software issues. Upgrades to new products should be anticipated. Keeping abreast of market trends and being able to adapt to new technologies and standards may allow incorporation of more effective tools.

Data output devices allow data to be disseminated. Printers, plotters, storage media and download sites permit data to be easily disseminated. Data mobility can be enhanced by software that converts a dataset into other formats. Therefore, data converters can also be considered as output devices. It is often necessary to use more than one software package in the course of a project and that will require data conversion.

2.4 GIS AS A TOOL

Software specific functions make each GIS product unique. However, common features are found in every GIS package. These common features provide the functionality expected of a typical GIS and can be categorized into display, database, editing and analysis tools (Aronoff, 1989).

2.4.1 DISPLAY

Data display is the most apparent feature of a GIS. A dataset appears as a layer that can be viewed with other datasets, also in the form of layers. In a vector-based GIS, a layer may be defined by a theme or a feature type as determined by the software. A theme is a collection of information with a common element. For example, a field crew may map an outcrop and include hand sample locations, transect lines and local geology in one layer. The theme is the outcrop and all information pertaining to the outcrop is included in the layer. Alternatively, a layer based on feature type would distinguish between point, line and area features. Using the outcrop example, three layers would be created. The point feature layer would include hand sample locations. The transect would be in the line feature layer and geology in the area feature layer. Areas are also referred to as polygons.

Raster-based GIS software incorporates layers of data organized into a uniformly spaced grid. The smallest element of the grid is called a picture element, or pixel (Lillesand and Kiefer, 1994). Each pixel displays a colour that represents the data value for that location. Geophysical data, commonly collected as a grid network of points, is interpolated and modeled to produce a surface which represents the sampling area. The amount of detail retained in the grid output is a function of pixel size which in turn is a function of the sampling density and model algorithm.

Figure 2.2 shows an example of a geological map in vector and raster format. In the vector layer, panel A, geological units are displayed as polygons. Notice the sharp boundary between polygons. This layer was converted to a raster format shown in panel B. The boundary is no longer clearly defined. Depending on the scale of the vector map and the resolution of the raster data, small features may not be rendered during the conversion.

Remotely sensed images, such as satellite data or orthophotos, are another type of raster data. The data stored in a raster image is an impression of the surface as viewed through the remote sensing platform. The platform collects information with sensors tuned to specific bands of the electromagnetic spectrum. Figure 2.3 shows the electromagnetic spectrum frequency and wavelengths. The resolution of the image depends on the distance of the viewing platform from the viewing area and the instrumentation. Sensors that collect data in the visible portion of the spectrum render images similar to photographs. The quality of the image is dependent on light, weather and atmospheric conditions at the time of the survey. Detection of wavelengths that range outside of the visible portion is not dependent on time of day. Moisture content in the air, specifically in the form of clouds, affects sensors differently. Radar sensors can detect the surface through cloud cover whereas optical sensors can not (Lillesand and Kiefer, 1994).

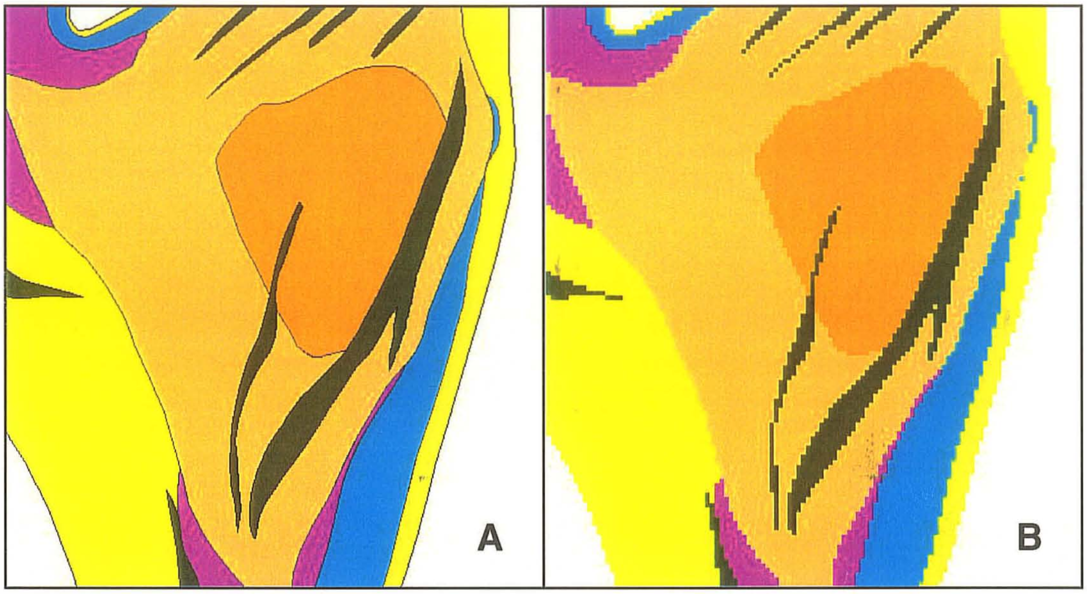


Figure 2.2: Examples of vector and raster data. Panel A shows the original vector file of a geology map. The layer was converted to a grid, a type of raster data specific to ArcView, shown in Panel B. Notice the difference along the boundary of features. The pixelation seen in Panel B is the result of the size of the pixels.

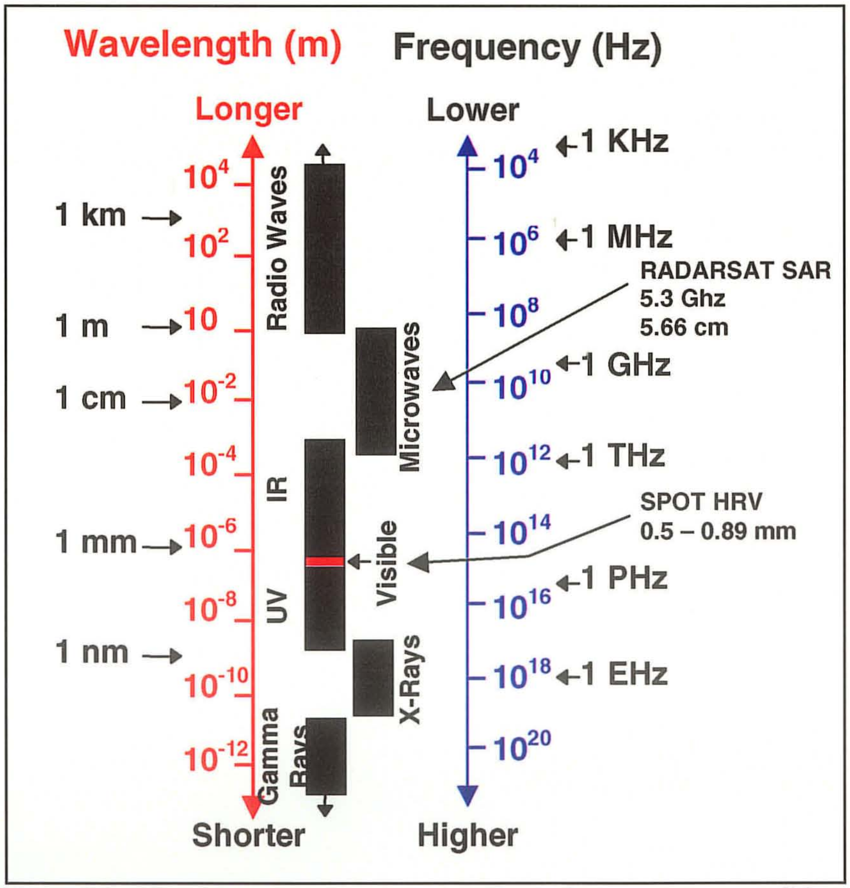


Figure 2.3: The electromagnetic spectrum. Raster images capture conditions at the Earth's surface using information from different parts of the electromagnetic spectrum. Adapted from CCRS website glossary.

2.4.2 DATABASE

A database is a collection of data that have been sorted and stored as files (Clarke, 1997). In a GIS, data can be stored as either raster or vector files. Two types of data are stored in the file – spatial data and attribute data. Each pixel in a digital image is identified by its location in the grid and its digital number. The location of a pixel can be measured on a Cartesian grid, producing a line and column location. If the image is geo-referenced then the pixel can also be identified by the coordinate system. The pixel is displayed according to the classification scheme (i.e. colour) assigned to the layer. The geologic maps shown in Figure 2.2 both use the same legend. For example, each pixel representing ultramafic rocks is coloured purple. In a digital image, each pixel has a value, or digital number, assigned to it. The digital number is inherent to the data structure.

Vector databases are structured differently and are more like spreadsheets (Burrough and McDonnell, 1998). A unique number identifies each feature in a vector layer. Unlike raster images, more than one attribute can be stored in a vector database. Attributes describe characteristics of each feature. The attributes can be used to label features or as data for a variety of graphs. Each attribute can be displayed spatially since it is linked to a feature with geographic coordinates. The spatial data are stored separately from the attribute data.

Data can be entered into a GIS by either digitizing from hard copy or scanned maps or creating features such as points, lines of polygons from a table of coordinates and attributes. Digitizing requires either a hard copy map and a digitizing tablet or a digital image and editing ability. The hard copy map is fastened to the digitizing tablet. Control points are chosen to set the grid coordinate system of the software to match the coordinate system of the map. A puck is used to trace features from the hard copy map. The trace is captured as series of nodes and lines. Each node assumes the coordinates as determined from the control points. The coordinates provide the spatial information needed to integrate the data into a GIS.

Heads up digitizing is similar. Instead of tracing a piece of paper with a puck, an image on the screen is traced with the mouse. This method is one way of converting raster data to vector data. There are automated processes that attempt to vectorize a raster image. This output is fairly coarse and lines appear jagged like a staircase. Attributes can be entered during the digitizing process.

Data that already exist in digital format with spatial attributes can be brought into a GIS. Point coordinates are the easiest to manage. Each point must have an X and Y coordinate that are used to place the point in geographic space. Latitude and longitude are often expressed as decimal degrees. Easting and northings, used in the Transverse Mercator projection, are other possibilities. The entry of lines and areas is more complicated. A point has no dimension, a line has 1 dimension and an area has 2 dimensions. Special routines are needed to convert a set of coordinates into a line or area. The routines are able to read start and end point information. The start and end points for an area should be identical to ensure the area is closed.

Converting data from other software is also an option. Computer automated design (CAD) packages store coordinate information for each feature. The database capacity in CAD systems is usually limited to parameters used to display the feature such as codes for colour, line thickness or symbols.

The structure of the database may limit the amount of data stored in any field. For example, in ArcView data are stored in tables. The fields in each table must be defined as number, string, boolean or date. The maximum number of characters in any field is limited to 255. Field names can not contain spaces and are limited to a maximum of 10 characters. If the field name starts with a number, it is changed to a "z".

An option for entering data into the GIS data table format is to establish a link to an existing database. Structured Query Language, or SQL, provides the functionality to define and manipulate data from a relational database. ArcView can retrieve a well-designed, complex database stored in Microsoft Access through the SQL protocol. Reducing the need for data entry and conversion also reduces the opportunity for error.

2.4.3 EDITING

A feature in a GIS layer is located in a plane defined by a coordinate system. A new feature can be added to the layer by selecting the appropriate tool and drawing its outline. For example, a polygon showing a new mining claim can be added to the existing polygons in the layer. Although ArcView provides for this simple process, it does not maintain topology.

Topology is the spatial relation among features (Burrough and McDonnell, 1998). In addition to identifying the coordinates of a feature, a topological dataset also stores information about neighbouring features. For example, Saskatchewan and Manitoba share a border. In ArcView, two polygons would be needed to show both provinces. In this case, the feature is identified by the attributes assigned to it. For example, the feature is a polygon and the polygon depicting Manitoba would contain a field with a text descriptor of "Manitoba". The shared boundary is drawn once for Manitoba and once for Saskatchewan. In a topological dataset such as an ArcInfo coverage, the boundary is drawn only once as a series of nodes and arcs. Each node and arc is attributed with information on surrounding features. A polygon is not a discrete object in a topological dataset. Instead, the nodes and arcs that mark its border define an area. A unique number identifies each node and arc. Figure 2.4 shows a simplified comparison of two layers - one with topology and one without.

Special tools are needed to edit topological data. Snapping to nodes or lines ensures that the nodes and lines of a new feature match the nodes and lines of existing features. That is, the locations of nodes occur at the same coordinates. Snapping tolerances can be set so that a new line will intersect an existing line at a node. The node may already exist or have to be created at the point of intersection. Another example of snapping to existing features is when a polygon is drawn next to an existing polygon and the boundaries of both polygons share nodes. Other editing tools used to clean datasets check for dangles, slivers and unclosed areas. Lines that overshoot the node of an intersecting line are dangles. Slivers occur when neighbouring polygons do not share the same nodes. An unclosed area occurs when arcs that make up a polygon do not share a common start and end point. Consequently, changing the location of one node affects the contents of many tables. Although topology is complex, it is needed to perform many types of spatial analysis.

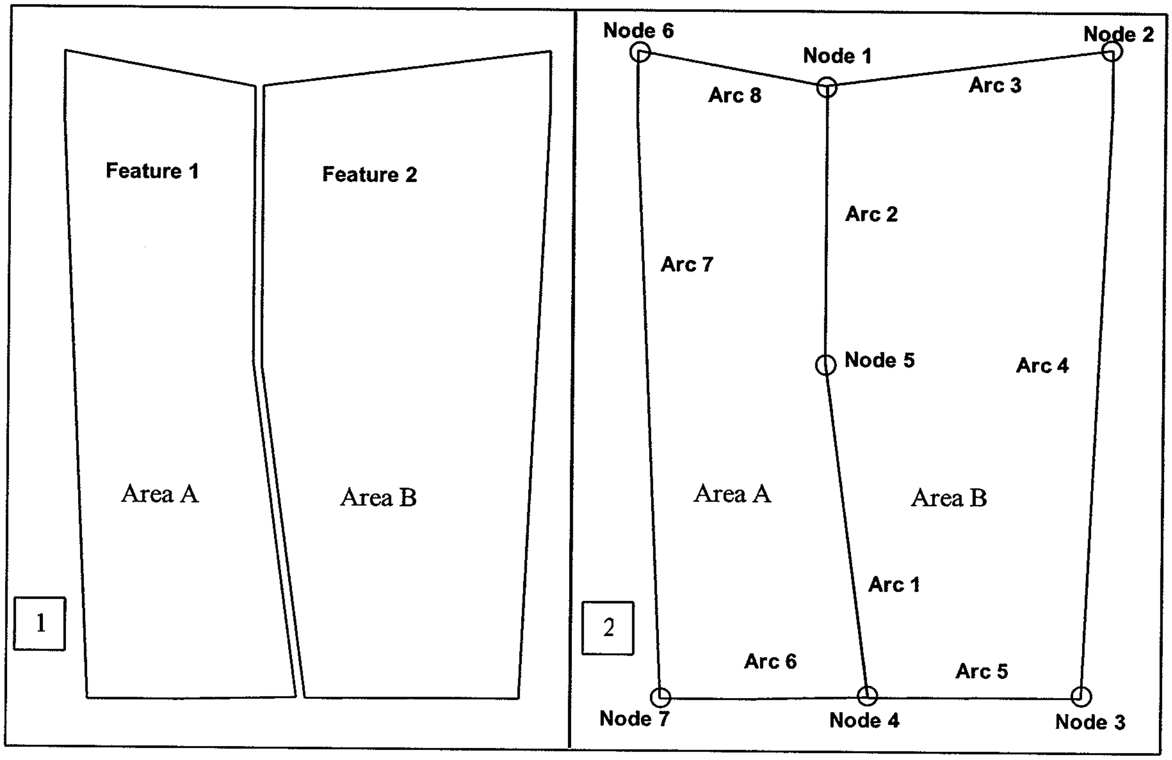


Figure 2.4: Panel 1 illustrates areas in a GIS without topology. Two polygons share a common boundary. A small space is shown between the shapes to illustrate the concept. The boundary has to be drawn for each polygon. Two features are discretely stored in the data table – Feature 1 and Feature 2.

Panel 2 represents areas in a GIS with topology. Each feature is made up of nodes and arcs. Area A is defined by Arcs 1, 2, 8, 7 and 6 and Nodes 1, 5, 4, 7 and 6. Area B is defined by Arcs 1, 2, 3, 4 and 5 and Nodes 1, 2, 3, 4 and 5.

2.4.4 ANALYSIS

Both topological and non-topological GISs software provide analytical tools. The database can be queried from both map and table interfaces. Queries are requests for specific information including place names, ranges in values and logical combinations. For example, a query can ask for all the samples of ultramafic rocks that were collected between 1997 and 2000 and that have been analyzed.

In a map interface, queries can be combined with buffering. A buffer is a region around a feature. The size of the region is determined by the size of the buffer. For example, a buffer of 10 m around a point would give a circular area with a diametre of 20 m. In combination with the query tool, it is possible to identify features meeting the query criteria within the buffer zone.

Overlay analysis uses the layer feature of the GIS. Features from one layer can be used to select or edit features from other layers. For instance, suppose a geologist may have a database of geochemical sample sites and that the only sites he is interested in are those that correspond to ultramafic rocks. The geologist could use the ultramafic polygons in the geology dataset to select corresponding geochemical samples from the geochemistry point database.

Data can be interpolated between features. A set of raw geophysical data can be plotted as points. The points are used to interpolate a surface. Different interpolation algorithms, such as Inverse Distance Weighting, kriging and splines, are incorporated in the software. Most vector-based GIS packages require add-on code to accommodate raster data. Digital elevation models (DEM) can also be created from a set of elevation points or contours.

Most vendors allow new code to be developed to meet specialized user requirements. The new tools can be designed from scratch or compiled from existing code. This aspect of GISs provides the user with the building blocks of the software. Very creative and powerful analytical tools can be developed to meet the needs of any project.

2.5 GIS AS A SYSTEM

Establishing a GIS project requires focused planning to produce successful results. Before an organization adopts a GIS, it must first assess current needs and resources and then project those needs into the future. A planning framework can be based on five components of a system: people, processes, data, software and hardware (PPDSH) (Donaldson Dewitz, 1996). In this framework, people perform processes on data using available software and hardware to meet project goals. The project is treated as a system designed to meet specific goals. The system determines how the units function together. Planning a GIS project using a systems framework allows planners to work with the organization at a range of scales.

The breakdown of system components into PPDSH can be applied at any scale, from international corporations to a Masters thesis. A GIS is not a stand-alone system - it must be incorporated into the existing working environment to be most effective. To effectively incorporate a GIS into an organization, components of the GIS should compliment the existing organizational system.

It is essential to understand the intended scope and extent of a GIS project to effectively manage its implementation. Consideration of resources at hand, including financial, human, time, hardware, software, training and attitude, provide a realistic guide by which to introduce a GIS. The management team must assess the required components of a GIS and determine how the components can fit into the existing organization. All aspects of introducing new technology and methods must be considered in order to make the transition smooth and effective. The GIS also must be allowed to evolve as needs, resources and technology change over time.

2.5.1 PEOPLE

The individuals involved in a system can be grouped into three categories: designer, manager and user. The designer develops the system to meet user needs. The manager ensures that system development is on schedule and when implemented, monitors the users productivity. The user implements the system into the working environment of the organization and produces results (Donaldson Dewitz 1996) Users can be further classified into basic, intermediate and expert sub-groups.

The expectations from each group of people must be considered when incorporating a GIS. A designer must have both technical knowledge of the programming language in addition to practical knowledge of what the system will be designed to do. In cases of mainstream software packages, the designer may be an individual who is responsible for purchasing the required product. The functionality of the program must be assessed relevant to the organizations needs.

The manager must understand how the GIS will contribute to productivity. Although management must support its use, the manager does not have to learn the software. Provisions may have to be made for interdepartmental access to data, software and hardware. This individual also oversees decisions such as authorization for training, resource acquisition and space allocation.

Determining the desired product and then managing people and resources to produce the deliverable can facilitate the introduction of a new system into an organization. The manager must be aware of the capabilities of the system at hand and incorporate the concept of PPDSH. If needed, outside sources can be contracted to perform functions unavailable inhouse. The manager may also assume the role of data manager, but in that capacity, must be intimately familiar with data sources, structure and models. A individual more appropriate for the data manager role may be an experienced user.

The user fills what is typically a hands-on position, working directly with data, software, hardware and filling requests from the project team. This individual performs a suite of other tasks and processes crucial to the organization and usually will not have experience with GISs or geographic concepts. However, with the ability to customize desktop GISs interfaces, any individual in the organization can be a user and therefore a hierarchy of users emerges. At the most basic level, a user may not have any spatial analysis skills but has a need for spatial information or may provide spatial information to others. At the basic level, all this user needs is an understanding of the capabilities of a GIS and a process to communicate these needs to a user who can fill them. Training is needed to introduce basic spatial techniques and geographic concepts. Additionally, all users need a common language to communicate their ideas. Defining words and concepts such that all users have a consistent language will allow communication of ideas. The language must also accommodate the users common language (jargon used in their field) so that the designer, manager and expert user learn concepts important to the organization.

An intermediate user has some ability to view and work with available datasets. An intermediate user can produce simple maps and perform basic GIS functions. They have acquired enough skills and understanding to be able to adopt, implement and utilize the new system and technology. For the GIS to evolve into an integral part of the organization, most users have to adapt to the new processes and accept them as part of their working environment (Obermeyer and Pinto, 1994; Sprague and McNurlin, 1993). This attitude of acceptance is most important at the basic and intermediate user level. Compared with other employees in the organization, they have the most to learn. Expert users, designers and managers implement the skills they were trained to use. Basic and intermediate users were hired to function in other capacities. Introducing a new tool as complex as a GIS is intimidating and must be supported by management through allocation of resources for ongoing training, technical support and equipment upgrades.

More advanced applications are handled by an expert user. This individual usually has a combination of experience and training specific to geomatics or infomatics. Complex modeling, problem solving and data conversions are typical tasks. Fluency in different operating systems and with a wide range of hardware facilitate the day to day operations in a GIS lab. This individual may be responsible for developing custom interfaces to meet the needs of basic to intermediate users in addition to designing the introduction of enterprise-scale GISs applications. Consulting with management on the purchase of technologies and the training of others to use software and hardware may be included in the expert user's job description.

With the right skill set and experience, the expert user can assume the role of designer. Although the first impression of an expert user may be "techie" in nature, the individual must possess communication skills and be approachable. It is the combination of expert user/designer in one individual or small team that holds great potential for an organization implementing a large GIS project. An understanding of the needs and abilities of the organization combined with those of the software allows for flexibility, adaptation and custom solutions to user defined problems.

2.5.2 PROCESSES

A process is a set of tasks that must be completed in a specific order to achieve a predetermined goal. A process can be established when:

- a set of common tasks is repeated often
- requests for uncommon tasks are received
- results must conform to internal standards
- results must conform to external standards
- many individuals are working on the same tasks at remote locations.

Consistent products are more likely to result from a defined process. Increased familiarity with the process and the expected results may also increase use of the results. In a small organization with close communication between all project members, processes conform to the working habits of those involved. Internal protocols do not seem important if the same people do the same job continuously. However, as soon as one individual leaves, the absence of defined processes becomes apparent. A process should not meet the needs of an individual performing the process. Rather, a process should meet the needs of the organization and all its members to ensure continuity and consistency in the event of staff turnover or transition to a new system.

For example, one process that can be affected by either scenario is data input from a GPS. The field staff must be aware of the projection parameters needed to collect data. The data input operator has to know the same information to integrate the new data with existing data. Perhaps the data needs to be appended to an existing project, or alternatively, it might be entered as a new project. Protocols for downloading the data must also be in place. The field staff may need to download data each day so it can be processed off-site. Alternatively, they may process their own data and submit a product to the data manager. New field staff and data compilers must be aware of these details to ensure the database is consistent. Should the hardware or method of acquisition change (for example, when selective availability was turned off May 1, 2000 (Office of Science and Technology Policy, 2000)) the process must be updated to reflect the changes. Processes must be updated when either internal or external conditions change.

Multi-site organizations that employ many people to perform the same task also require well-defined processes. The ability to merge data together and perform robust analysis is entirely dependent on the process of getting data into the system and the consistent use of recognized parameters. Robust analysis is crucial to ensuring that others can achieve the same results using the same procedures. Numerous processes exist in any organization. To improve the acceptance of a new system and new processes into an existing framework, the existing processes should be used as much as possible. Necessary changes should be fully detailed in terms of resources needed and predicted implementation times. Training and upgrades should occur in tandem. The time and money invested into process change is not as obvious as money spent on hardware or software. However, people perform processes with the resources available to them. Planning the implementation of a new process must consider the demands placed on existing people, hardware, software and data.

2.5.3 DATA

A common statement regarding data is "Garbage in, garbage out". The statement summarizes the importance of data quality and integrity. If bad data are used in a system, the results are meaningless. In order for a dataset to be bad, it must be inappropriate for the project. For example, data may have been collected at the wrong time of year or the field crew used different tools to measure the same observations. Therefore, data can be bad to begin with as a result of errors introduced through the collection method. Alternately, good data can become bad data if used inappropriately. An example of inappropriate data use can be illustrated from the Geochemistry dataset. One of the original goals of the CAMIRO TNB project was to create a 3-dimensional model of the belt. The drillhole data stored in the Geochemistry dataset were identified as the information source for this objective. After the dataset was examined, it was determined that the sampling density (i.e. the number of drillholes) was too coarse to create a 3-dimensional model of the belt. Drillhole data were unavailable for very large sections of the belt. Inconsistencies with lithologic descriptions (i.e. the language and terms used to describe lithologies) further limited use of the dataset for 3-dimensional modeling. Even if the people, processes, software and hardware are the best that money can buy, inappropriate data will compromise the entire system.

The difference between good and bad data is largely determined by how the data are used. That is, data should be used only if the original collection methods and scale are compatible with how the data is intended to be used. Therefore, a dataset may fit the needs of Project A but be inappropriate for Project B. Data are typically collected for a particular purpose. Sampling schemas, classification, scale and temporal factors influence the way information in a dataset is organized. The way data are organized determines how they can be used.

The largest part of a GIS project budget is the acquisition and integration of data into the GIS (Korte, 1997). Many projects fail because management can not imagine the amount of time needed to properly implement the data acquisition phase of the project. Many procedures at this stage take time and do not seem to produce results. For example, digitizing information from paper maps is not as simple as tracing the features from the map. Decisions regarding features, accuracy, direction, attributes, edge matching and how to deal with inconsistent legends across map sheets should be made before the puck hits the tablet.

These types of decisions need to be made in a planning session that involves the project team. The project team should host representatives from each interest group involved in the project. The representatives from a variety of areas of expertise will have valuable contributions. They each look at data from unique perspectives and place priority on different features of the area of interest. Discussions should result in a prioritized list of common data needs and special requests.

Data collection might be an option in some GIS projects. Questions that should be asked before a new data set is collected include:

- How will this data be used?
- How long will the data be valid?
- Will the data have to be updated regularly? If so, how often and at what expense?
- Can the data be used in another project? If so, are cost-sharing arrangements possible?
- What observations are needed and at what frequency?
- Can other observations not needed at this time but in the future be conducted?
- How will the quality of the data be ensured?
- Who will collect the data?
- What equipment is necessary for data collection and storage?
- In what format will the data be provided?
- Is that format compatible with existing hardware and software?

Planning the data collection phase of a project can not be done without considering the other components of the system. The objectives of the GIS must also be kept in mind to keep the project focused and on budget. Very detailed observations can be recorded, but if they surpass the needs of the project, then time and money are wasted. Alternatively, not collecting enough information when in the field can jeopardize the project if field crew have to return to the area of interest and incur additional expenses.

Data sharing is a useful way to reduce costs associated with data acquisition. Partnerships with organizations that need the same data are common. Establishing an agreement can foster benefits such as swapping technical expertise (why reinvent the wheel?), shared costs on a superior product and an additional perspective on the area of interest. Although the benefits can be great, the biggest obstacle remains to be proprietary attitudes towards datasets. Given the time, money and expertise invested in data collection, the dataset is a valuable resource that was collected with focus on the organizations project objective. Sensitive information is often found in these datasets. Although the obvious costs associated with data collection may be reduced with data sharing, it is difficult to put a value on potential losses if proprietary information is delivered to the wrong hands. Therefore, consideration must be given to both the dollar value and the knowledge value that is retained in a dataset. A well-developed dataset becomes a valuable resource to an organization. Just as efforts are made to retain valuable employees, organizations try to keep valuable data in-house.

Data dissemination is quick and efficient thanks to evolving technologies such as high speed internet connections, large capacity storage devices and mobile communication. Unfortunately, both good and bad data are disseminated quickly. Further improvements in transferability between software and operating platforms compounds the problem of transferring bad data. Before a data manager realizes a dataset is inappropriate for the project, the data has been ordered, received, paid for, translated and incorporated into the model used by decision-makers, sometimes in a matter of hours. When metadata is provided with a dataset, it is the responsibility of the data manager to read the document and note any discrepancies between what was expected from the dataset and what was delivered. Metadata is discussed further in Chapter 3, Data Management.

Reviewing the metadata document should be the first step when acquiring a dataset. At this point, information such as projection, datum, date of collection, geographical extents of the dataset and attribute descriptions are obtained from the metadata document. It is becoming more common to find metadata descriptions for commercial data products.

Standards for metadata documentation have been developed nationally and internationally and are invaluable when properly constructed (Burrough and McDonnell, 1998; CORINE, 1992; FGDC, 2001).

2.5.4 SOFTWARE

Many vendors provide off-the-shelf GIS packages. The software is designed to meet the needs of a very large user base. The typical GIS-in-a-box includes a generic set of tools found in most other GIS packages. In order to increase the flexibility of a product, users can create their own macros using the vendors programming language. When determining the needs of a new GIS project, it is essential that the software can handle the data and perform the required analysis. If the basic functions of the GIS are not sufficient, management must decide if it is better to purchase a customized GIS or develop the skills needed to customize the GIS in-house.

The needs of the project should define the criteria used to compare the benefits of competing GIS products. If users are creating data, the product should have editing tools specific to GIS tasks, such as snapping to nodes or lines. Quality control may be a useful built in feature if digitizing is a big part of the project. Compatibility with specific databases or spreadsheets may be important if a large amount of data are already maintained in that environment. The ability to create presentation quality maps and reports may be an important aspect of communicating results to stakeholders. Networking and operating system specifics must match existing hardware parameters.

A less tangible issue is the learning curve of the GIS. How much time and money needs to be invested in order to train people to use the software? What are the trade-offs between introducing an easy to use GIS with limited capabilities versus a technically challenging system that can handle all the spatial analysis needs of the project? If the user base is experienced and ready to implement the technology, then the higher end system would be appropriate. Otherwise, resistant users intimidated by complex, non-user friendly interfaces would be better matched with a simple system. Management must assess the attitude and skills of inhouse staff to make an effective choice. If a GIS is part of the long term operations of an organization, starting simple and letting the entire system evolve to more sophisticated levels may prove to be most economical.

Software upgrades are controversial events to many users. Just as users become functional with one version of software, the network administrator changes everything overnight. The interface may change. Transferring projects to the new version requires time and organization. Partners in remote locations may not have access to the upgrade and backwards compatibility becomes an issue, especially if enhanced features are included in the upgrade (Sprague and McNurlin, 1993). The vendor should be able to provide strategies to decrease the work required for the transition between versions of their product. A planned schedule of upgrades gives system administrators time to prepare their networks and people for the change. The vendor may offer additional training and support. Upgrading all licences may not be practical. Expert users might upgrade first, determine the challenges and recommend the appropriate time for intermediate and experienced users to make the change. This strategy allows in-house expertise to develop and filter through the organization. Changes that affect critical processes in the system can be flagged. Strategies to implement the change can be weighed against changing the process to accommodate the software.

GISs have the functionality to pull together all types of digital information. Images, data, graphics, statistics, tables and charts can be combined in a GIS. Each software product handles documents and processes data differently. The results from a selection of GIS products should be compared with each other to determine the relative ability of each product to perform specific functions. More vendors offer trial versions of their software for this purpose.

2.5.5 HARDWARE

The computer may seem to be the most obvious piece of hardware needed to run a GIS. It is possible to operate GIS software using only a computer. However, the functions will be limited and the results will not be available for others to see. A GIS lab should be equipped to support all functions of a GIS. These functions include data input, storage, manipulation, display and dissemination. Both onsite and offsite support may be necessary to facilitate project goals. In some cases where specialized tasks require unique equipment, it may be better to outsource the work to a company with appropriate equipment and trained staff. Repeated tasks, such as plotting large images or transferring data to compact disks, justify acquisition of plotters and high-speed compact disk read/write burners. Maintenance of equipment, supplies and upgrades should be factored into the cost of supporting specialized equipment.

Training staff to use the equipment is also necessary. Training for GIS hardware can include sessions on how to send documents to the plotter, download data from a GPS unit, insert pictures from a digital camera and customize the buttons on a digitizing puck. Peripheral devices usually come with software and uncommon graphical user interfaces

(GUIs). Software training and customization may be also be necessary. When everything is plugged in and online, the end result should support project goals and work with the rest of the system.

2.6 THE CAMIRO GIS PROJECT

The CAMIRO TNB project identified the need for a GIS at the outset. Project coordinators observed the emerging trend within the geologic community to move towards digital data compilations. GISs provided a means to collect, store, visualize, analyze and model a broad range of data as both raster and vector data had to be accommodated. A flexible tool was therefore needed as the data collection included observations from scales as small as the extent of the belt to as large as hand sample locales. Dimensions included the third and fourth, extending below surface and over billions of years.

The most important function of the system was to integrate all the data into one platform from where it could be displayed. The software had to be robust enough to manage the data demands yet still be accessible to the researchers. Internet mapping technologies were just emerging at the beginning of the CAMIRO project in 1998 but would require too much development for the scope of the TNB project. The people involved included representatives from industry, government and academia. Computing facilities and skills varied within this group. Data collection and dissemination processes were influenced as a result of the diverse user base. The Geographic Information System used by the CAMIRO TNB project was the result of the potential and limitations of the 5 components of a system – people, processes, data, software and hardware.

2.6.1 PEOPLE

Industry partnered with the provincial and federal governments to expand their knowledge of the TNB and develop new exploration methods for nickel based on research conducted by members of the academic community. Project goals were established by industry sponsors. Logistical support and scientists were provided by Manitoba Industry Trades and Mines through the Manitoba Geological Survey. Universities contributed researchers and analytical facilities. Each individual involved in the project contributed skills specific to his or her discipline. Table 2.1 lists the participating companies and institutions.

Table 2.1: Industry, government and academic partners in the CAMIRO TNB 97E-02 TNB project.

Industry	**Government**	**Academia**
INCO Ltd.	Manitoba Industry Trades and Mines	University of Alberta
Falconbridge Ltd.	Continental Geoscience Division - Geologic Survey of Canada	University of Saskatchewan
Billiton International Metals B.V.	Mineral Resources Division - Geologic Survey of Canada	University of Manitoba
Hudson Bay Exploration and Development Co. Ltd.	Geoscience Integration Section - Geologic Survey of Canada	Mineral Exploration Research Centre - Laurentian University
Western Mining International Ltd.		GEOTOP - Université du Québec à Montréal

Researchers specialized in geochronology, structural geology, petrography, mineralogy and geochemistry. For most of the group, GISs was a new technology. The geologists typically created maps with a desktop graphics package. Data were stored in spreadsheets and most often displayed as tables, charts or graphs. The GIS provided a new display medium for their data.

These geologists only saw the end products, maps, but were not involved in the technicalities of making the data GIS-ready. How the attributes are stored and displayed in a GIS affects their interpretation. Therefore, the researcher providing the data must understand how it will be used in the GIS.

In order for the system to be used by and useful to the researchers, they had to acquire a primary knowledge of spatial and geographical concepts. A series of presentations were given to the researchers to introduce GISs and its role in the project. The presentations were designed to foster an understanding of the technology and terms common to GIS applications. The dialogue allowed researchers to look at their data collections from a spatial perspective. Concepts such as projections and datums had to be covered to support data collected with Global Positioning Systems (GPS).

2.6.2 PROCESSES

The researchers completed most of the data entry process. Results from analyses were entered into spreadsheets, and in turn the spreadsheets were converted to point data in ArcView. The most difficult task involved providing spatial coordinates for sample sites. Although some researchers used GPS units, they were not provided training or parameters for data collection with the GPS. In particular, most did not set the datum or projection on the instruments. The data were also not corrected for the error introduced to each signal by the United States' Department of Defense. During the summer of 2000, the Department of Defence removed the positional error from each signal. Without correcting the data, the coordinates could be off by as much as 200 m. Also, the orientation and signal strength of each satellite affects the quality of the signal. The signal quality measure, or PDOP, must be within a certain threshold for the data to be useful, but the PDOP values were not noted for any GPS data.

Other researchers relied on paper maps to mark their sample locations. Years of sample sites could be found on one map sheet. Although these maps were large scale, coordinates had to be read off using two rulers. As the datum and projection of some of these maps were not available, this method of data entry left much opportunity for error that could not be quantified.

Data provided by company sponsors was acquired through project leaders. Questions important to a GIS specialist were not considered by the project leaders, and as contact with the companies was limited, these questions took a long time to answer. The answers also were filtered through too many individuals, whereas to facilitate the transfer of data from the companies, CAMIRO's GIS specialist should have been in direct contact with those from the companies. Communication was further restricted as individuals changed roles over the course of the project. Much confusion resulted from the lack of overlap in their roles and the lack of a defined process for data dissemination.

Each year, the CAMIRO project annual meeting was held in June, at which time, the yearly report was also provided to industry sponsors in addition to a compact disk compilation of data. Data would typically be provided to the GIS specialist at the last possible moment and consequently formatting of the data and compact disk was rushed. The compact disks were customized for two groups – researchers and industry sponsors. Researchers had access to all the data provided by the companies, but the industry sponsors did not have access to data from the other companies. A website with passwords to restricted data sets would have facilitated the distribution of data back to the sponsors and researchers more effectively than the use of compact disks.

2.6.3 DATA

The data collection assembled by the CAMIRO project is the result of combining existing datasets with new compilations. Industry sponsors provided geology and geochemistry datasets. The large scale geology maps from industry included more detail than any public dataset. Drill hole data were also provided by the companies and was used to populate the geochemistry dataset. The researchers created new datasets with the results of 3 seasons of field work. The following list in Table 2.2 describes each dataset available to researchers in the project and their origin.

Table 2.2: Datasets available to CAMIRO researchers.

Dataset	**Origin**	**Description**
INCO geology (includes structure	INCO	Restricted use by researchers. Geology of INCO area.
Hudson Bay Exploration and Development (HBED)geology	HBED	Restricted use by researchers. Geology of HBED area.
Falconbridge geology	Falconbridge	Restricted use by researchers. Geology of Falconbridge area.
GAC geology	Geologic Association of Canada (GSC) 1972	Includes area in exposed part of TNB.
Geology of South Pit, Thompson Mine	J. Kraus, D. Peck and W. Bleeker	Detailed mapping of South Pit.
Provincial Geology	Manitoba Industry Trades and Mines (MITM)	1:250,000 geology maps from the province.
Geochemistry	Researchers and Industry	Compilation of data from new field work, company sponsors, L. Hulbert (GSC), Laurentian University from previous CAMIRO project 94E04 and published data from literature. Includes geochemistry analysis, drillhole descriptions and identification fields.
Geochronology	N. Machado, A. Potrel, L. Heaman, K. Toope, D. Peck	Characterizes sediment provenance with U-PB age dates.
Geochmistry – PIXE and EM	J. Liwanag	Results from PIXE and Electron Microprobe analysis of sulfides. This collection is for specific underground locations. It is contains more detail than the larger geochemistry dataset.
Structure trends	N. Machado, A. Potrel	Dip and strike measurements from hard copy maps of the belt were averaged over 1km size grids. The results show regional trends in strike and dip for the northern portion of the belt.
Base Maps	MITM	Topographic coverage of the province including roads, towns, rivers, lakes, etc.
Aeromagnetics	Geologic Survey of Canada (GSC), MITM	Aeromagnetic coverage of the province with pixel size of 200m and 1000m. Very coarse resolution for provincial compilation of existing surveys.

2.6.4 SOFTWARE

The initial GIS used by the CAMIRO project was ArcView 3.1, upgraded later to ArcView 3.2. Some datasets were processed with ArcInfo by the GIS specialist. Microsoft Excel was a common spreadsheet package used by most of the researchers and data were often submitted as .xls files to the GIS expert. Requests for data from the GIS expert were filled by providing .xls files when the user did not need to see the data on a map. Although ArcView does not maintain topology, the software provided sufficient visualization and integration tools for the project goals. Although ArcView was too complicated for most researchers, it was compatible with industry data sets. Operating systems were not an issue as all partners ran Windows 95, NT or 98.

2.6.5 HARDWARE

All parties in the project used PCs and operated in a Windows environment. During both annual and research meetings, laptops and data projectors facilitated discussion and presentations. Reports would be distributed both as hard copy and digital documents and consequently, compact disk burners were an excellent resource. Digitizing and scanning were implemented sporadically and those who needed such services often had their own equipment. An HP 2500 design jet plotter proved indispensable for large format hard copy output. In the field, GPS was used to a limited extent and digital cameras were not used at all, although the incorporation of field photos could enhance the database.

CHAPTER 3

DATA MANAGEMENT

3.1 INTRODUCTION

Geoscientists are multidisciplinary by nature; geology, geophysics, geochronology, lithology, remote sensing, structural geology and geochemistry are just a few examples of the interrelated disciplines. Each area deals with a unique mix of temporal and spatial dimensions. The resultant collection of data falls on a continuum from microscopic to global scales that can be represented in tabular, image, map or text format. Integrating the volume and diversity of information is challenging. Developing a metadata format that encompasses relevant descriptors for all data types is intimidating.

Geoscientists study the current position, orientation, physical and chemical properties of lithologic units to determine how they were originally emplaced and subsequently deformed. Each analysis performed on an area of interest provides a piece to the complex puzzle that is being solved. The "puzzle pieces" can be in the form of drill hole data, geophysical interpretations, assay values, lithologic maps, structural interpretations, isotope age dates or remotely sensed images. Each data set has limitations in terms of interpretation and reliability. The limitations result from influences such as method of data collection and analysis, discretion of the data distributor and the basic nature of the data type.

A database is a repository into which observations and measurements are deposited. A data model provides the framework for a database, and data storage, use and manipulation are managed through this model. The data model is designed to optimize the use of different types of data and the types of observations and measurements must be considered when developing the model.

3.2 TYPES OF DATA

Data can be classified into four types: nominal, ordinal, interval and ratio. The type of data indicate how each dataset can be used. Nominal data are defined as discrete units that can not be compared quantitatively. They are descriptive values that can be grouped into classes. For example, the field geologist observes that an outcrop contains sedimentary rocks, iron formations and basalt. This observation represents nominal data. The geologist then comments on the texture of each lithology and notes that the sedimentary rocks are coarse grained, the iron formation is fine grained and the basalt is aphanitic. Since grain size can be ordered from smallest to largest, this set of observations is ordinal. Table 3.1 includes a data type column that categorizes data found in the Geochemistry data set.

One of the goals of the TNB field project was to create a geochronological map for the area. Samples of the rocks were taken back to the lab to determine age dates. These results provided interval-type data. When ages are considered in geological time, only the difference in age can be compared. A rock can not be twice as old as another because age is measured on a scale which does not include zero. However, a rock can be two billion years older than its neighbour (+/- a few Ma).

Finally, geochemical analyses are conducted to determine the presence of select elements. The results, measured as weight %, ppm or ppb, can be compared on a ratio scale. That is, one sample can have twice as much gold as another. A ratio can be assessed when a value of zero is a possibility. Comparing age values to assay values, nothing can be “0 units old” but an element may not be present. Multiplication and division operations can be used with ratio datasets, allowing descriptions such as "75% more", "half as much as" and "3 times the amount".

The data model design incorporates how observations about unique features are linked to a description of that feature. One feature can have many attributes. An attribute can be any of the four data types. Therefore, the data model has to accommodate both quantitative and qualitative descriptions. The data model states the parameters of each field (or column) in a dataset. Parameters can include the format of the data (numeric, string, floating point, real integer, boolean, date), the number of characters (usually limited to a maximum of 255 characters), the number of decimals (if numeric) and special formatting instructions (date fields may be dd/mm/yyyy or mon-dd-yy). Establishing these parameters restricts the characters that can be entered into the field. Data must be clean and conform to these parameters. If not, operations designed to work with the data will result in error. Often, the software will recognize inappropriate characters and not allow the user to enter the information. Table 3.1 is an example of a data model that could be used with the Geochemistry database, shown in Table 3.2. Examples of data not conforming to the data model are shown in Table 3.3.

Table 3.1: Data Model parameters and data types for the Geochemical dataset.

Field	**Data Model Parameters**	**Data Type**
HOLE_NUMBE	String	Nominal
LOCATION	String	Nominal
UM_BODY	String	Nominal
LOCAL_EAST	Numeric, 20, 0	Ordinal
LOCAL_NORT	Numeric, 20, 0	Ordinal
UTM_EASTIN	Floating point, 20, 16	Ordinal
UTM_NORTHI	Floating point, 20, 16	Ordinal
SAMPLE_NUM	String	Nominal
FROM	Numeric, 20, 8	Interval
TO	Numeric, 20, 8	Interval
LENGTH	Numeric, 20, 8	Ratio
ORIGINAL_C	String	Nominal
LITHOLOGIC	String	Nominal
CHEMICAL_C	String	Nominal
COMMENTS	String	Nominal
UNIT	String	Nominal
OTHER	String	Nominal
COMPANY	String	Nominal
SAMPLED_BY	String	Nominal
DATE	Date, yyyy	Ordinal
MAJOR_METH	String	Nominal
MINOR_METH	String	Nominal
TRACE_METH	String	Nominal
PGE_METHOD	String	Nominal
S_METHOD	String	Nominal
SEMI_METAL	String	Nominal
HALIDE_MET	String	Nominal
OS_METHOD	String	Nominal
SIO2	Floating point, 20, 16	Ratio
TIO2	Floating point, 20, 16	Ratio

Table 3.2: Excerpts from the Geochemical database. Note that the column headings are often truncated. The data were originally in an Excel spreadsheet and converted to dbf format for use in ArcView. The conversion is prone to truncating columns and data loss may result.

HOLE_NUMBE	LOCATION	UM_BODY	LOCAL_EAST	LOCAL_NORT	UTM_EASTIN	UTM_NORTHI	SAMPLE_NUM	FROM	TO	LENGTH	ORIGINAL_C	LITHOLOGIC
RP-91-01a	Winnipegosis				439150.0000	5871750.000	20037	454.10	454.5500000		flow bx.	
WL94-70		W-41			474677.0000	5959446.000	WA35949	425.00	425.2100000	0.210000000	Peridotite	6F
WL94-71					474634.0000	5976937.000	WA19701	213.00	213.1800000	0.180000000	Amphibolite	14
WL94-72					472482.0000	5967042.000	WA35954	519.65	519.9200000	0.270000000	Amphibolite	14
WL94-73					468531.0000	5973785.000	WA35955	283.45	283.5500000	0.100000000	Amphibolite	14
WL94-74					474781.0000	5979173.000	WA35960	479.38	479.8500000	0.470000000	Silicate IF	10K
WL95-102		W-77			465053.0000	5964694.000	WA22414	667.40	668.6500000	1.250000000	Dunite/Altered Ultramafic	6G
WL91-19	William Lake				475426.0000	5963300.000	WB09712	233.40	233.6500000	0.250000000		6E
WL91-19	William Lake				475426.0000	5963300.000	WB09710	241.80	242.0000000	0.200000000		6E.1
HW95-05	Halfway Lake				540592.0000	6103805.000	HW95-05-388.8	388.80	389.0000000	0.200000000	10c	6E.1
RP-96-23	Winnipegosis				449505.0000	5891250.000	51-97-025	547.29	547.6300000	0.340000000	sulphidic schist	10

HOLE_NUMBE	CHEMICAL_C	COMMENTS	UNIT	OTHER	COMPANY	SAMPLED_BY	DATE
RP-91-01a				B	Cominco	Cominco	1992
WL94-70				MBS	Falconbridge	Falconbridge	1994
WL94-71				MB	Falconbridge	Falconbridge	1994
WL94-72				MB	Falconbridge	Falconbridge	1994
WL94-73				M	Falconbridge	Falconbridge	1994
WL94-74				MBS	Falconbridge	Falconbridge	1994
WL95-102		Long sample		MBS	Falconbridge	Falconbridge	1995
WL91-19		porphblst w/ num serp vn. Alt ol w/ pss dendritic tex.		MTBS	Falconbridge	Burnham	1999
WL91-19		cg poik opx in fg trem metrix w/ abunt spongy serp ol porphblst, minor hm alt.		MTBS	Falconbridge	Burnham	1999
HW95-05		serp pxn pdt, ld$:1Q		MBRHPS	Falconbridge	Lesher/Keays	1996
RP-96-23		graphitic; 5% py	Grand Island Tholeiites	MTBRHS	Cominco	Peck	1997

HOLE_NUMBE	MAJOR_METH	MINOR_METH	TRACE_METH	PGE_METHOD	S_METHOD	SEMI_METAL	HALIDE_MET	OS_METHOD	SIO2	TIO2
RP-91-01a										
WL94-70									53.676766900	0.177530163
WL94-71									49.555346680	0.820727964
WL94-72									48.284812000	0.837072851
WL94-73									19.509296920	0.486692482
WL94-74									59.825377790	0.461773928
WL95-102									42.569713060	0.081977439
WL91-19	XRF CAF	XRF CAF			LECO CAF				34.040000000	1.150000000
WL91-19	XRF CAF	XRF CAF			LECO CAF				44.040000000	0.840000000
HW95-05	XWF97-0133	IAT97-0133	IMT97-0133	NISFA-LU	IRC97-133				43.840000000	0.260000000
RP-96-23	XRF CAF	XRF CAF	IMT97-0523		LECO CAF				63.953480660	0.433215788

Table 3.3: Examples of data entries that do not conform to the data model. Excerpts are from the Geochemistry database.

[1] A negative number appears in a column designated for positive numbers.

[2] Text string appears in a column designated for numeric entries.

[3] The entered values revert to date notation in a column designated as positive numbers.

HOLE_NUMBE	**LOCATION**	**RB__XRF_**	**PB**	**BI**	**CD**	**HG**	**SN**	**W**	**MO**	**S**	ZP2O5
WL91-19	William Lake	-1.000000000[1]								0.005	0.01
HW95-05	Halfway Lake								15	0.5	N.D.[2]
RP-96-23	Winnipegosis	53.502149780	Jan-00[3]	Jan-00	Jan-00			Jan-00	Jan-00	Jan-00	0.04

The connection between a feature and its attributes in a relational database management system (RDBMS) can be one to one, one to many or many to many. A one to one relation means that each feature is attributed by unique descriptors. A one to many relation means that a feature shares attribute common to other descriptors. A many to many relation indicates that the feature can have many attributes and the attribute can be assigned to many features. Storage of common features in a database is often accommodated by a series of tables. For example, chemical analyses of drill hole and surface sample data are stored in a database. The collar and sample site location for each sample are stored as UTM coordinates and each sample is identified by a sample number, assigned by the field geologist. Each company uses a different coding system for identifying field samples and drill holes. A series of alpha-numeric characters can be used in the description and two fields can be used to identify each sample. However, depending on the scale of mapping, samples taken close to each other, for example, on either side of a contact, may display as identical points. Therefore, to eliminate possible confusion between the samples, each should be identified by a unique sample number. Table 3.4 contains examples from the Geochemistry database.

Table 3.4: Examples of sample identifications used in the Geochemistry database. The first column is the drill hole number. The ninth column lists the sample number from the drill hole. In order to find the sample both numbers are needed. The coding scheme is dependent on the researcher. Variations can include numerics only, text only and combinations thereof, including special characters such as “ – “ and “ . “

HOLE_NUMBE	LOCATION	UM_BODY	LOCAL_EAST	LOCAL NORTH	UTM_EASTIN	UTM_NORTHI	SAMPLE_NUM	FROM	TO
RP-91-01a	Winnipegosis				439150.0000	5871750.000	20037	454.10	454.5500000
RP-91-01a	Winnipegosis				439150.0000	5871750.000	20038	461.40	462.2500000
RP-91-01a	Winnipegosis				439150.0000	5871750.000	20039	466.00	467.0000000
RP-91-01a	Winnipegosis				439150.0000	5871750.000	20040	478.00	479.0000000
M77-30	Bucko Lake				522703.0000	6081297.000	WA20296	791.87	
M77-30	Bucko Lake				522703.0000	6081297.000	WA20297	792.48	
M77-30	Bucko Lake				522703.0000	6081297.000	WA20298	794.61	
MN95-86		MN-12			488624.0000	5999870.000	WA40909	102.40	104.4000000
MN95-86					488624.0000	5999870.000	WA40914	117.00	119.0000000
MN95-89		MN-12			486990.0000	5997631.000	WA40984	157.00	158.7500000
MN95-89					486990.0000	5997631.000	WA40987	164.00	165.5000000
MN95-90		MN-5			482418.0000	5993915.000	WA41004	150.70	152.0000000
MN95-90		MN-5			482418.0000	5993915.000	WA41012	175.00	177.0000000
MN95-90		MN-5			482418.0000	5993915.000	WA41038	223.65	225.6500000
MN96-128					487091.0000	5997663.000	WA42048	141.86	141.9700000
MN96-128					487091.0000	5997663.000	WA42065	181.50	181.6000000
13806	Ospwagan Lake		#######	#######	557635.0000	6156889.000	13806-640	640.00	
55752	Ospwagan Lake		-600.00	#######	558850.0000	6158218.000	55752-670	670.00	
81889	Lower Ospwagan Lake		-280.00	#######	558687.0000	6157780.000	81889-1372	1372.00	
81889	Lower Ospwagan Lake		-280.00	#######	558687.0000	6157780.000	81889-1410.5	1410.50	
86239	Ospwagan Lake		600.00	#######	559306.0000	6158253.000	86239-2425.5	2425.50	
86239	Ospwagan Lake		600.00	#######	559306.0000	6158253.000	86239-2503	2503.00	
86239	Ospwagan Lake		600.00	#######	559306.0000	6158253.000	86239-2614.5	2614.50	
86239	Ospwagan Lake		600.00	#######	559306.0000	6158253.000	86239-2886	2886.00	
86240	Ospwagan Lake		-100.00	#######	558504.0000	6157391.000	86240-1566	1566.00	
86240	Ospwagan Lake		-100.00	#######	558504.0000	6157391.000	86240-1584	1584.00	

In the case of drill hole, the collar identification is used to refer to all samples taken in the drill hole. Again, the location of each sample varies with depth, strike and dip and cannot be used to identify the unique observations. When the drill core is sampled, a unique id should also be applied and the depth noted. Depending on the sources of data for a specific project, information from drill core may only be identified by collar number and depth. Unique identifiers should be assigned to the dataset before being integrated with other data.

Observations may also be taken at the same location over time. The temporal dimension also must be incorporated into the data model. Time of day, season, date, duration and sequence of observations may be important observations.

3.3 DATA COLLECTION

Data collection is at an interesting crossroads. Affordable technologies, such as laptop computers, GPS and data loggers, have emerged in the field. Digital data collection offers an alternative to hand written notes in the field. As a result, the convenience of processing observations in the field on the fly must be weighed against the increased cost in equipment and training investment of the field staff. Batteries, weather conditions and ease of transport also affect the decision to take technology to an outcrop. Depending on the project goals, the accuracy of a GPS may degrade the effectiveness of a project due to increased costs of too much data collection. The sampling schema used to collect data with a GPS must consider the scale of other maps used in the project. It is also good practice to keep a written record of collected data.

Observations made in the field should be annotated legibly especially when the data are entered by a person different from the one who collected the information.

3.4 GEOREFERENCING

Each piece of data in a GIS must be georeferenced. Point data, such as sample sites, are relatively simple to map. Drill hole data present more challenges. The drill hole starts at the surface, or collar location, and follows a path into the crust, which involves changes in dip, azimuth and direction. Any sample taken from the drill core must be referenced to a 3-dimensional coordinate system. Ideally, each sample is recorded in a data table with all the information needed to plot the drill hole in 3D space. In reality, the crucial location information, dip and azimuth, may not be included. Exclusion of such information is largely a result of the costs and inability to collect accurate measurements.

Remotely sensed images can include georeferencing details. If not, the image can be "rubber-sheeted" to fit known control points. The accuracy of the fit depends on the amount of relief across the image area, choice and position of control points and the warping algorithm. One of the datasets provided by INCO was given with INCO's coordinate system. To use this dataset with others, it was warped to the UTM 14, NAD83 projection and datum. The accuracy of the warping procedure is not known since the process and control points used were not provided with the dataset. The importance of locational accuracy is compounded when information from remotely sensed images is combined with other layers of information. For example, combining structural lineaments interpreted from a synthetic aperture radar (SAR) scene with lithologic units determined through field work can help indicate major uncomformities and contact zones. Determining where the boundaries of continuous lithologies are intersected by apparent structural lineaments from the SAR scene can suggest discrepancies of boundary locations from field mapping. Questions arise as to which layer better represents reality. Further fieldwork may be needed to verify the relationship between structure and lithology. To place confidence in this type of analysis, the researcher must be confident in the positional accuracy of both data sets.

3.5 DATA INTERPRETATION

Data interpretation can introduce error if the interpreter is not familiar with the data set. Geochronology in particular is a discipline where interpretation of data requires specific knowledge about the method used to produce the dates. A single sample may be analyzed with a variety of dating methods, each producing a different age date. Certain dating methods can provide additional information regarding geologic events, for example, U-Pb dating of zircons can indicate crystallization ages. The researcher must understand how the dates are derived, the assumptions each method is based on and the relevance of interpreted geologic events in order to properly interpret the dataset. Without this specific knowledge, geochronology data is prone to be misrepresented.

Scale becomes an issue when maps of different scales are used together. Geochemical sample sites are point locations and the distribution of sample locations depends on the location of accessible outcrops and the number of samples taken. Interpreting geochemical data on its own provides detailed information about only the sample location. More value can be derived from the dataset if geochemistry is combined with other layers of information. For example, placing the sample on a map of regional geology can illustrate chemical relationships between lithologies, provided the lithologies are mapped with enough detail. If the units used in the geology map are too general, then the combination of geochemistry and geology data will provide limited results.

Raster data, such as satellite imagery or geophysical grids, introduce additional challenges related to scale. Pixel size determines the amount of detail represented in a raster layer. Combining coarse resolution raster layers with details such as point data or polylines reduces the ability to determine relationships between such layers, whereas fine resolution data can improve detection of local features.

Data collection in the geosciences often takes researchers into remote locations that are relatively unmapped. Working from air photos, outcrops are identified and sought out. Before GPS was available, locations and orientation were determined with map and compass. Many of the geology maps referred to today were created from observations made in the field and recorded in note books. As a result, errors can be introduced through the process of recording and transcribing field notes. At any given outcrop, the geologic interpretation can vary between geologists as a result of familiarity with the terrain, experience and note taking skills.

3.6 DATA ACQUISITION

A GIS needs data to work. The process of acquiring datasets and incorporating them into the GIS software is typically a resource-intensive process. Funds and time are needed to purchase basic data layers and integrate them with other datasets. In the past, the most difficult aspect of any GIS project was to find base datasets of sufficient scale and at affordable prices. Various vendors offered different products, most of which were available in only one software specific format. Basic data layers included information as found on paper topographic maps. The challenge of acquiring basic data has dissipated since the federal Government of Canada developed a digital version of the paper topographic map series. The National Topographic Data Base (NTDB) maps were organized based on the grid sections used by the paper map series at scales of 1:250,000 and 1:50,000. Although the product was available, many problems were associated with how it was developed.

The process of digitizing topographic maps included maintaining the paper map grid boundaries. Consequently, the area defined for a paper map sheet was retained in the digital version. The most immediate problem resulting from this approach was that the information found along the edge of a paper map did not match that of the neighbouring map sheets. Therefore, edge matching was a very obvious problem with the digital map collection.

The NTDB digital maps were created by scanning the paper maps and processing the scanned image to extract points, lines and polygons. When the digital maps are displayed, features crossing map sheet boundaries do not always line up. Features such as contour lines and polygons are cut off at the map edge. An additional problem arose as a result of map sheets being published at different times. The classification of certain features, such as wetlands, coastlines and dune areas, varied with publication year. Adjoining map sheets used different legends and land use classifiers, thereby further reducing consistency between map sheets. The contour interval also changed between map sheets, ranging from 25 to 30.1 metre intervals.

The NTDB map series is available in a variety of digital formats such as E00, shapefiles and mid/mif. Currently, the costs of acquiring digital map sheets are based on the number of features, that is points, lines and polygons, contained in each sheet. Over the last few years, new formats have been added to the list of available formats. A considerable amount of time must be invested to convert the files to a useable format if the software being used does not support the formats offered. The addition of new formats is a concern for those who invested in earlier products. The conversion process is lengthy, awkward and involves a considerable time investment. Any project that incurred the expense of acquiring NTDB map sheets also incurred the expense of converting the sheets to a useable format. The CAMIRO project could not afford to acquire the map sheets for all the project participants. Therefore, a consistent set of maps could not be used by researchers and industry.

Frequent updates and revisions to base map series incurs additional expenses to organizations that purchased the datasets. Recently, the Centre for Topographic Information (CTI) changed the data model for the 1:250,000 and 1:50,000 scale NTDB map series. The changes included renaming field headings. As CTI updates and corrects the existing product, licensed users receive the new files. However, the new files no longer conform to systems developed around the former naming conventions. Such a change requires making existing systems compatible to the new data model. Field headings are used to query the database. As this situation develops, it will be interesting to watch how CTI deals with the inconsistencies. Data monopolies foster negative attitudes in the industry and prevent the science from advancing due to repeated need to revamp systems and waste resources on inconsistent data products (Clarke, 1999).

In addition to monetary investment, legal issues have skewed the ability to freely transfer information. Purchase of data sets is not always an outright purchase. Often, only a license to use the data is available for purchase. The distributor owns the rights to the data and is entitled to collect royalties generated from its use. The great debate on the issue of data ownership revolves around subsequent changes to the dataset. If a product is changed in such a way so as to enhance its value or use, it is no longer the same product (Obermeyer and Pinto, 1994). As they currently exist, the laws do not address this issue sufficiently, therefore, the laws are easy to break and hard to enforce. The Manitoba Branch of the Canadian Institute of Geomatics invited local GIS professionals to attend a forum on the issue of free access to data (CIG, 2000). Participants in the discussion indicated their frustration with the copyright / ownership laws and the problems faced by researchers and the public to access publicly funded datasets in Canada while other countries are moving towards providing easily accessible and low to no cost data that meet international standards.

Within the scope of the CAMIRO TNB project, industry sponsors provided important geological datasets. Each dataset was created with different software. Legends and definitions for lithologies were not consistent across the INCO, Falconbridge and HBED datasets. The INCO dataset, is of particular interest to researchers as it covers most of the exposed portion of the belt and was originally created using INCO's coordinate system. Conversion to the project standard – UTM NAD83 – introduced positional error. Also noted were topological problems with polygons. Finally, polygons included in the hard copy version of the map were not included in the digital version.

Early in the project, a proposal for acquiring orthophotos of the belt area was submitted to a local company. The orthophoto collection would provide continuous coverage that could be combined with the existing base maps and data layers. The inclusion of such data would allow for verification of sample sites and outcrop locations. The dataset could also be used to quality control the entire geochemistry data set. Unfortunately, negotiations with the company proved fruitless as the cost of processing the data would have had to been entirely absorbed by the project. The processing cost included data collection, image processing and orthophoto conversion. The company selling the data was offsetting the entire cost of their orthophoto project onto their first client. This option was not within the budget and sponsors, although seeing the value in such a dataset, were not in a position to meet the asking price.

The issue evolved further when the province entered negotiations with other parties interested in orthophoto coverage of the Thompson area. A group of industry and government organizations, including some CAMIRO partners, formed a collective partnership to purchase the orthophoto collection. Unfortunately, CAMIRO itself was not included in the scope of the purchase and therefore could not use the data.

The scenario illustrates the expense of acquiring quality data. Individual parties can not afford to cover processing costs imposed by private companies involved in the business of providing data for profit. Further constraints are placed on how the data can be used. Purchases are designed as “license to use data” as opposed to “owning” the data outright. Efforts are underway in the province to make base data available at low to no cost, similar to the models used in the US, however, Canada is still a long way from mandating such provisions.

3.7 METADATA

Establishing a data collection involves considerable time and investment. Large collections can include data from many different sources and exist in different formats. Keeping track of data, especially in digital form, is an arduous task, given the dynamic nature of digital formats. Changes made to data should be documented when implemented. Documenting information about data provides reference material that can be accessed years after a data set has been acquired or used. Unfortunately, this task is often left to be completed at a later date and consequently decreases the accuracy of data history. The collection of information is referred to as metadata – commonly defined as data about data. An alternate definition examines the prefix "meta" with respect to its Greek origin, meaning change. Metadata details all the changes in terms of processing, maintenance and updates performed on a dataset.

Metadata format is designed to meet the needs of the user. It can be developed as a searchable database encompassing an entire data collection at one extreme or as a stand-alone document describing one data set at the other. The primary purpose of any metadata document is to include information that describes the dataset in order to determine how the data can be used. Other details describing physical parameters, extents, date and place of origin and contact information are useful. Large collections of data require extensive descriptions that must be contained within a searchable database.

International standards for metadata content and format are currently under review (FGDC 2001a). Standardizing issues include deciding on mandatory and optional fields, defining descriptive terms and ensuring file format compatibility. The standards are designed to be applicable to all data formats, across multiple hardware platforms and between software vendors. The standards also must consider future needs with respect to changes in technology.

The United States is a world leader in providing data freely while adhering to standards. The United States Geological Survey (USGS) is particularly instrumental in facilitating organizing efforts through the Federal Geographic Data Committee (FGDC). The FGDC coordinates the development of the National Spatial Data Infrastructure (NSDI) in the US. The NSDI encompasses policies, standards, and procedures for organizations to cooperatively produce and share geographic data. The 17 federal agencies that make up the FGDC are developing the NSDI in cooperation with organizations from state, local and tribal governments, the academic community, and the private sector (FGDC 2001b). An example of an FGDC product is the standard headings and layout for a metadata document. Figure 3.1 details the headings found in a document produced by the Metadata Collection Tool v2 ArcView extension.

Organizations may choose to develop their own format for content and file storage. The extensive documentation required by other metadata products may include details that are irrelevant or useless. These organizations must use caution when developing their own parameters and ensure that their product will be compatible with nationally and internationally accepted standards. In addition, the file format must be readable by available metadata tools. Ultimately, it is better to use a scaled down version of an accepted, standardized metadata product than to develop a new one. As the idea and importance of metadata becomes more integrated in the marketplace, the demand for user friendly products will encourage their development. At this early stage, the investment into a "beta-type" product is more than most companies are willing to support.

Metadata compilation is an intimidating task that should be well planned and funded. The resources needed to start a project from scratch are considerable. Time spent on assessing the current and future data needs of an organization and developing a strategy and metadata framework at the outset will be well invested. Datasets should be identified and prioritized for documentation. Metadata should be collected automatically when new datasets are generated.

Organizations must develop an alternative attitude with respect to the role data plays in every day operations. Important decisions are based on in-house databases that involved considerable investment of time and resources towards collection, maintenance and storage. To fully realize the potential of a data collection, the resource must be managed by a dedicated individual. A data manager must address issues of reliability and accountability with respect to products received and distributed. Anecdotal evidence from colleagues investigating methods of metadata collection indicate that there is a need to involve librarians in the process. Librarians are trained to collect metadata information about printed media and therefore they already have established keyword dictionaries and processes for documenting vast collections of information.

Often, an individual working with a dataset will possess a wealth of background knowledge needed to interpret the information. The expertise provided by an individual, or user knowledge, is not distributed with the dataset unless efforts have been made to extract and document important details. The data manager performs the function of retrieving user knowledge and documenting it as metadata.

Metadata should be included for reference with any distributed dataset. Responsibility for providing documentation lies with both the distributor and user. Data users must put pressure on data providers to supplement data with its history. Data distributors have to realize the importance of maintaining accurate and consistent records for their data collections. As more data becomes available in the global market, datasets with comprehensive histories will become more valuable. Data clearinghouses may require that metadata be provided in a standardized format that allows index and search functions (Hart and Phillips, 1998).

1. IDENTIFICATION_INFORMATION

```
Citation:
  Citation_Information:
    Originator:
    Publication_Date:
    Title:
    Edition:
    Geospatial_Data_Presentation_Form: Map
    Publication_Information:
      Publication_Place:
      Publisher:
    Other_Citation_Details:
    Online_Linkage:
    Larger_Work_Citation:
      Citation_Information:
        Originator:
        Publication_Date:
        Title:
        Publication_Information:
          Publication_Place:
          Publisher:
        Online_Linkage:
Description:
  Abstract:
  Purpose:
  Supplemental_Information:
Time_Period_of_Content:
  Time_Period_Information:
    Range_of_Dates/Times:
      Beginning_Date:
      Ending_Date:
  Currentness_Reference:
Status:
  Progress: Complete
  Maintenance_and_Update_Frequency: Continually
Spatial_Domain:
  Bounding_Coordinates:
    West_Bounding_Coordinate: -98.5583
    East_Bounding_Coordinate: -97.5159
    North_Bounding_Coordinate: 55.9982
    South_Bounding_Coordinate: 55.0406
Keywords:
  Theme:
    Theme_Keyword_Thesaurus: None
  Place:
    Place_Keyword_Thesaurus: None
Access_Constraints:
Use_Constraints:
Point_of_Contact:
  Contact_Information:
    Contact_Organization_Primary:
      Contact_Organization:
      Contact_Person:
    Contact_Position:
    Contact_Address:
      Address_Type: mailing and physical address
      Address:
      City:
      State_or_Province:
      Postal_Code:
      Country:
    Contact_Voice_Telephone:
    Contact_Facsimile_Telephone:
    Contact_Electronic_Mail_Address:
    Hours_of_Service:
Native_Data_Set_Environment:
```

2. DATA_QUALITY_INFORMATION

```
Attribute_Accuracy:
  Attribute_Accuracy_Report:
Logical_Consistency_Report:
Completeness_Report:
Positional_Accuracy:
  Horizontal_Positional_Accuracy:
    Horizontal_Positional_Accuracy_Report:
  Vertical_Positional_Accuracy:
    Vertical_Positional_Accuracy_Report:
Lineage:
  Source_Information:
    Source_Citation:
      Citation_Information:
        Originator:
        Publication_Date:
        Title:
        Edition:
        Geospatial_Data_Presentation_Form: map
        Publication_Information:
          Publication_Place:
          Publisher:
        Other_Citation_Details:
        Online_Linkage:
        Larger_Work_Citation:
          Citation_Information:
            Originator:
            Publication_Date:
            Title:
            Publication_Information:
              Publication_Place:
              Publisher:
            Online_Linkage:
    Source_Scale_Denominator:
    Type_of_Source_Media:
    Source_Time_Period_of_Content:
      Time_Period_Information:
        Range_of_Dates/Times:
          Beginning_Date:
          Ending_Date:
      Source_Currentness_Reference:
    Source_Citation_Abbreviation:
    Source_Contribution:
Process_Step:
  Process_Description:
  Source_Used_Citation_Abbreviation:
  Process_Date:
  Source_Produced_Citation_Abbreviation:
  Process_Contact:
    Contact_Information:
      Contact_Person_Primary:
        Contact_Organization:
        Contact_Person:
      Contact_Position:
      Contact_Address:
        Address_Type: mailing and physical address
        Address:
        City:
        State_or_Province:
        Postal_Code:
        Country:
      Contact_Voice_Telephone:
      Contact_Facsimile_Telephone:
      Contact_Electronic_Mail_Address:
      Hours_of_Service:
```

3. SPATIAL_DATA_ORGANIZATION_INFORMATION

Direct_Spatial_Reference_Method:
Point_and_Vector_Object_Information:
  SDTS_Terms_Description:
    SDTS_Point_and_Vector_Object_Type
    Point_and_Vector_Object_Count:

4. SPATIAL_REFERENCE_INFORMATION

Horizontal_Coordinate_System_Definition:
  Planar:
    Grid_Coordinate_System:
      Grid_Coordinate_System_Name:
      Universal_Transverse_Mercator:
        UTM_Zone_Number:
        Transverse_Mercator:
          Scale_Factor_at_Central_Meridian:
          Longitude_of_Central_Meridian:
          Latitude_of_Projection_Origin:
          False_Easting:
          False_Northing:
    Planar_Coordinate_Information:
      Planar_Coordinate_Encoding_Method: Coordinate pair
      Coordinate_Representation:
        Abscissa_Resolution:
        Ordinate_Resolution:
      Planar_Distance_Units:
  Geodetic_Model:
    Horizontal_Datum_Name:
    Ellipsoid_Name:
    Semi-major_Axis:
    Denominator_of_Flattening_Ratio:

5. ENTITY_AND_ATTRIBUTE_INFORMATION

Overview_Description:
  Entity_and_Attribute_Overview:
  Entity_and_Attribute_Detail_Citation:

6. DISTRIBUTION_INFORMATION

Distributor:
  Contact_Information:
    Contact_Organization_Primary:
      Contact_Organization:
      Contact_Person:
    Contact_Position:
    Contact_Address:
      Address_Type: mailing and physical address
      Address:
      City:
      State_or_Province:
      Postal_Code:
      Country:
    Contact_Voice_Telephone:
    Contact_Facsimile_Telephone:
    Contact_Electronic_Mail_Address:
    Hours_of_Service:
Resource_Description:
Distribution_Liability:

7. METADATA_REFERENCE_INFORMATION

Metadata_Date:
Metadata_Review_Date:
Metadata_Contact:
  Contact_Information:
    Contact_Organization_Primary:
      Contact_Organization:
      Contact_Person:
    Contact_Position:
    Contact_Address:
      Address_Type: Mailing and physical address
      Address:
      City:
      State_or_Province:
      Postal_Code:
      Country:
    Contact_Voice_Telephone:
    Contact_Facsimile_Telephone:
    Contact_Electronic_Mail_Address:
    Hours_of_Service:
Metadata_Standard_Name: FGDC CSDGM
Metadata_Standard_Version: FGDC-STD-001-1998

Figure 3.1: Metadata descriptors found in the FGDC metadata collector extension for ArcView. Seven categories of information can be collected. The list can be customized to include only the information deemed necessary for a project. When saved in this format, it is compatible with other FGDC metadata descriptions collected with different tools. Categories 3 and 4 are extracted from the datasets. All the other categories are entered manually or from previously saved documents.

CHAPTER 4

GEOLOGY

4.1 INTRODUCTION

The Thompson Nickel Belt (TNB) is located in central Manitoba; it occupies the eastern edge of the Trans-Hudson Orogenic terrane (THO). The THO formed during the collision of the Hearne, Rae and Superior provinces and the Sask craton in the Proterozoic (Figure 4.1; Hoffman, 1989).

The four areas that make up the THO are the North Quebec, Hudson Bay, Manitoba-Saskatchewan and Dakota segments. The TNB is located in the Manitoba-Saskatchewan segment along with the Fox River Belt and Split Lake Block (Figure 4.2). These three areas were also referred to as the Churchill-Superior Boundary Zone (CSBZ) before the Churchill province was subdivided into the Hearne, Rae and Burwell provinces (Lewry and Collerson, 1990).

The Hearne province is characterized by juvenile Late Archean rocks. Metamorphic grade ranges from greenschist-facies rocks in the south central core to higher grades in rocks radiating out towards the surrounding orogens. Granulite facies rocks are found in Saskatchewan, Northern Manitoba and the south Keewatin area next to the THO. The increase in metamorphic grade is attributed to Early Proterozoic crustal shortening, erosion deep in the Archean basement and increased heat due to plutonism in the Early Proterozoic.

The Reindeer Zone is located between the Superior and Hearne-Rae cratons as a result of a series of arc-continent collisions with the Hearne craton and final collision with the Superior craton. The rocks of the Reindeer Zone are composed of juvenile magmatic arcs and sedimentary basin rocks, ca. 1.9 – 1.8 Ga (Lewry and Stauffer, 1990). The most relevant belts in the Reindeer Zone to the TNB are the Flin Flon and Kisseynew belts since they bound the TNB to the west. Low-grade metavolcanoplutonic rocks are typical of the Flin Flon belt. These rocks are thought to be derived from arc or ocean-floor magmatism and are relatively undeformed and at low metamorphic grade although some deformation is evident in the central and western areas (White *et al.*, 1994 and references therein). The Kisseynew belt contains evidence suggesting a basin may have developed marginal to the Flin Flon and Lynn Lake arcs as interpreted by the presence of amphibolite-grade para- and orthogneisses. Mylonites are found along the south east portion of Kisseynew-TNB boundary whereas rocks along the Kisseynew-Flin Flon boundary indicate changes from metamorphic gradation to shear zones. Evidence supporting the terminal collision of the Reindeer Zone with the Superior craton is provided by coeval recumbent folding and shearing along the southern flank of the Kisseynew, north-trending, north-west verging folds and faults in the Flin Flon belt and anatectic granites and peak metamorphism in the TNB (White *et al.*, 1994 and references therein).

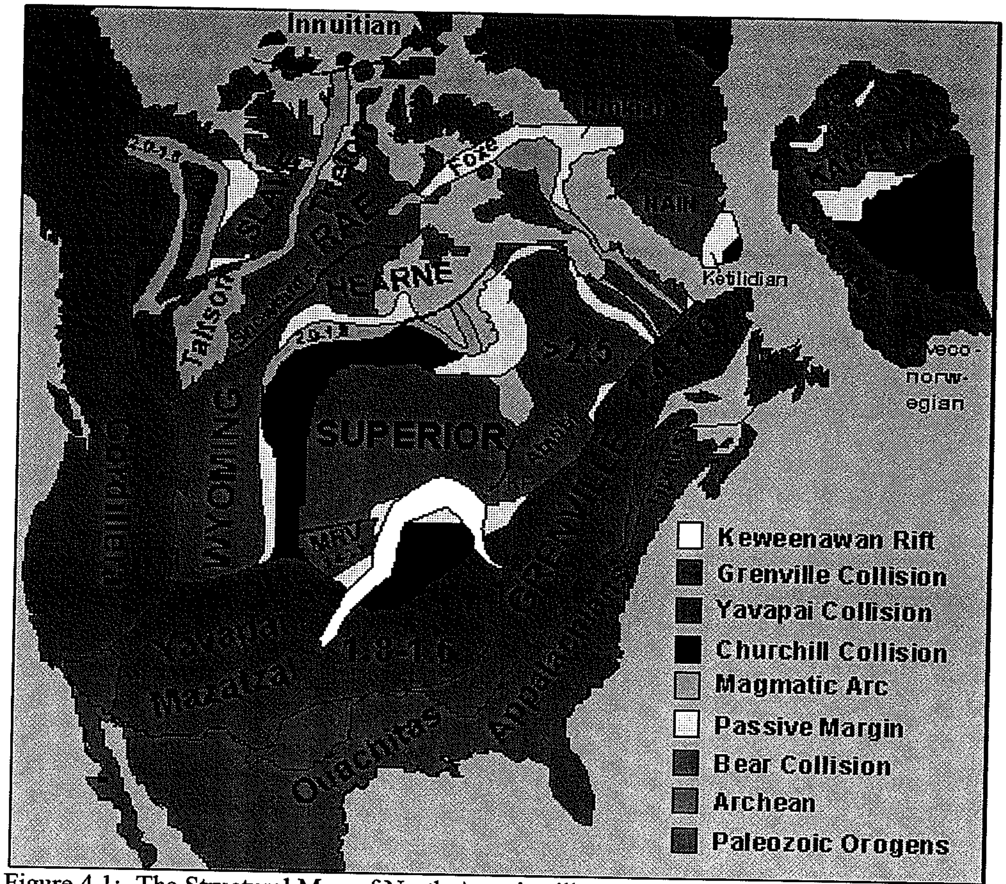


Figure 4.1: The Structural Map of North America illustrates the locations of the Hearne, Rae and Superior Provinces. Adapted from Church, 1998.

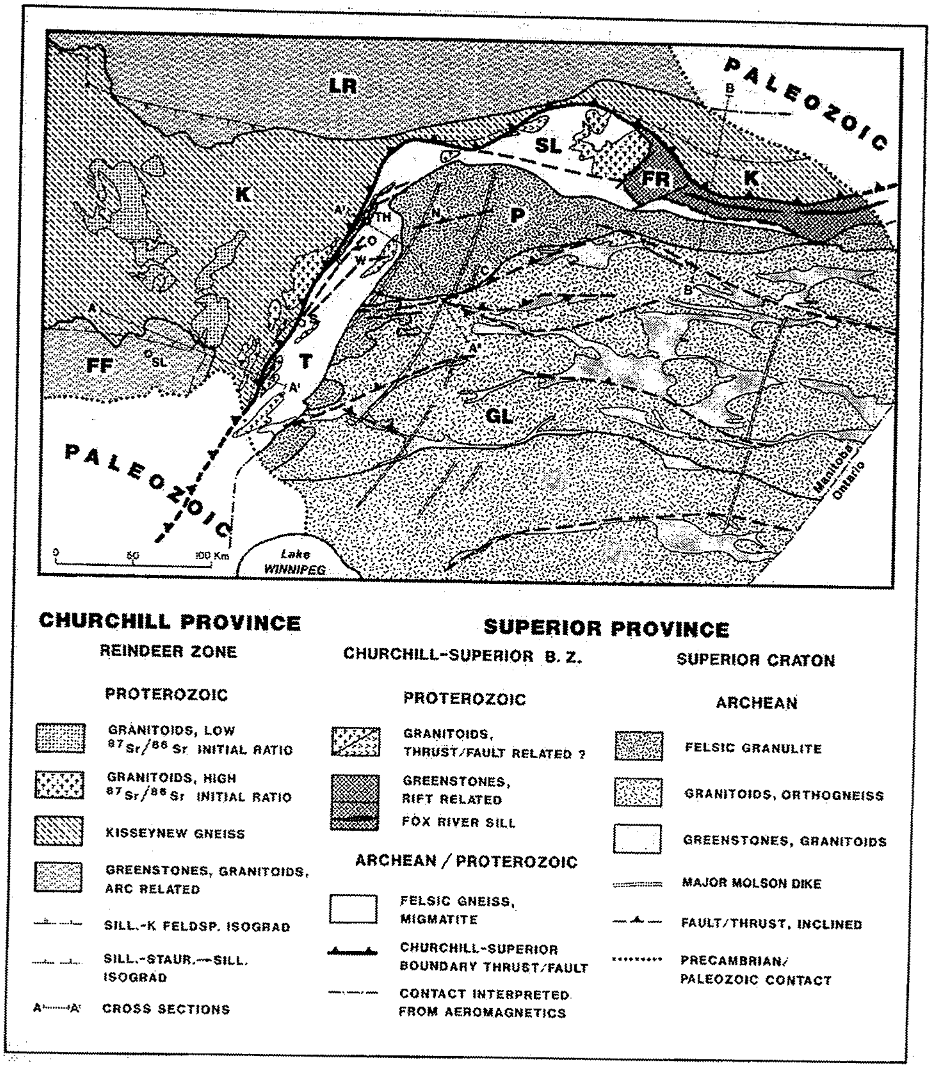


Figure 4.2: Former Churchill Superior Boundary Zone. Map illustrating locations of the Thompson Nickel Belt (T), Split Lake Block (SL) and Fox River Belt (FR). These areas were formerly considered the Churchill-Superior Boundary Zone. Other areas shown include the Flin Flon belt (FF), Kisseynew domain (K), Pikwitonei (P), Gods Lake (GL) and La Ronge (LR). Adapted from Weber (1990).

The Superior province is a large Archean craton composed of four major belt types - volcanic-plutonic terranes, metasedimentary belts, plutonic complexes and high-grade gneiss complexes. Ages of these terranes range from 3.6 Ga to 2.65 Ga. Of the thirteen terranes identified in the Superior province, the Pikwitonei is singled out as the most relevant with regards to the TNB (Figure 4.2). The boundary between the TNB and Pikwitonei terranes is distinguished by the location of where the Archean granulites of the TNB exhibit retrograde metamorphism and where the Molson dikes show ductile deformation and amphibolite facies metamorphism (Weber 1990).

According to Bleeker and Macek (1996), the TNB is characterized by retrograded Archean gneisses with infolded remnants of Ospwagan Group metasediments. Granite and pegmatites discordantly intruded into the gneisses and metasediments around 1.82 Ga and 1.77 Ga, respectively. Evidence indicates that ductile thrusting to the southeast occurred both prior and subsequent to intrusion of the Molson dikes around 1.88 Ga.

Nickel ore is associated with the Ospwagan Group metasedimentary rocks. The chronological order of the five formations within the Ospwagan Group has the Manasan at the base, followed by the Thompson, Pipe, Setting and Ospwagan (Figure 4.3). The Manasan and Thompson formations show an upward fining sequence of clastic sediments and carbonates. The upward fining of these sediments suggests a gradually deepening depositional basin, perhaps related to a rifting event. The first two members of the Pipe formation includes pelitic sedimentary rocks, iron formations and chert. The transition from carbonates to pelites suggests that the slope of the depositional surface leveled out; the transition may represent the bottom of a deep basin. Metaturbidites characterize the third member of the Pipe formation. Fine grained sediments are interspersed with coarse sediments. The appearance of turbidites suggests a nearby slope that would allow the development of turbidity currents. The Setting Formation hosts upper clastic sedimentary rocks that gradually coarsen, which may indicate basin closure. Finally, the Ospwagan formation hosts mafic to ultramafic subaqueous volcanics. The appearance of renewed mafic volcanism high in the sequence may be connected with impinging arc rocks during the collision. The sequence of basin closure followed by subaqueous volcanics hints at a possible subduction zone in an island arc setting (Bleeker and Macek, 1996).

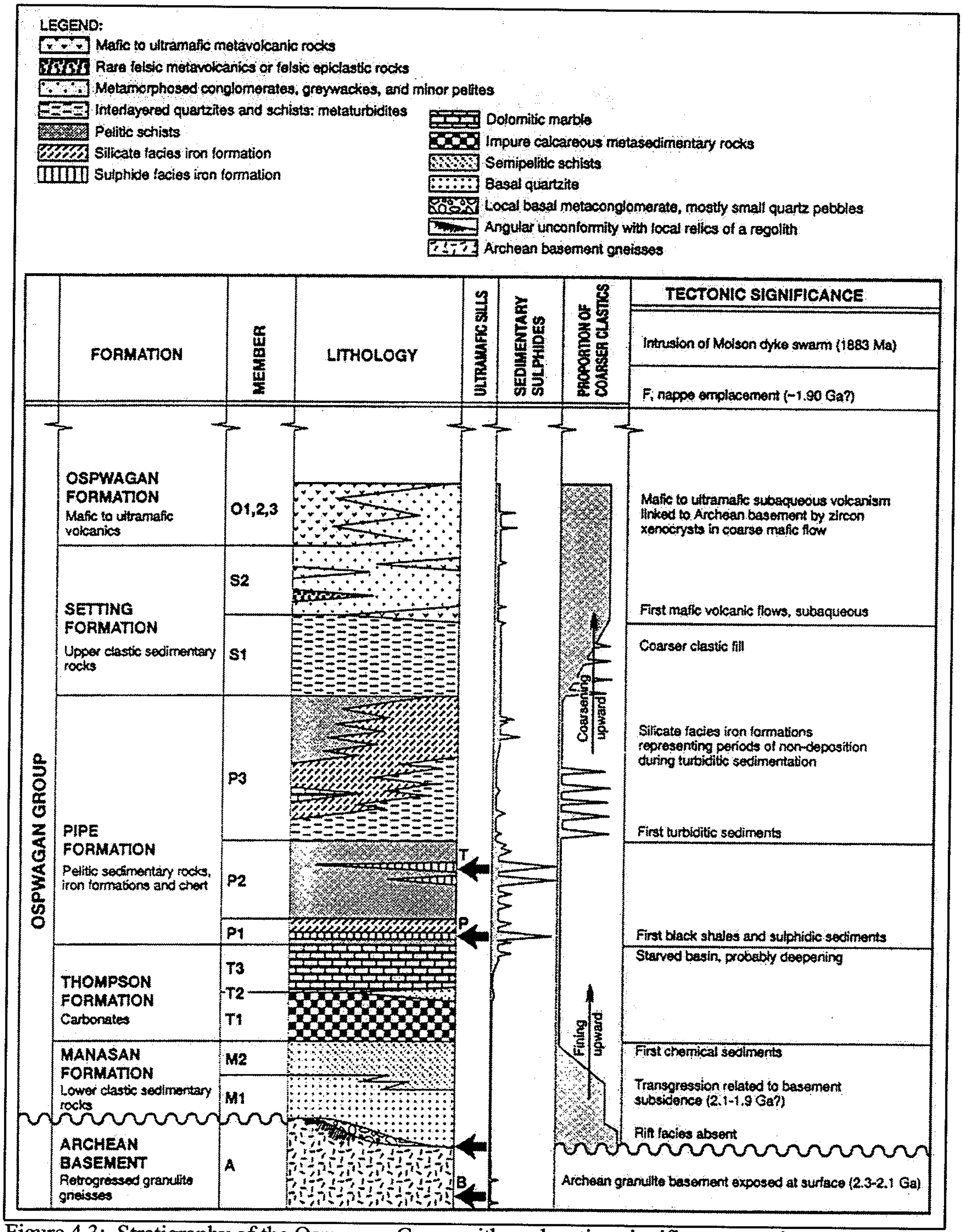


Figure 4.3: Stratigraphy of the Ospwagan Group with explanation significant tectonic events. Intrusions of ultramafic sills are indicated by arrows - Thompson (T), Pipe II (P), Bucko (B) and others. From Bleeker (1990a).

4.2 ISSUES

Interpretation of the TNB involves conflicting theories regarding tectonic environments and depositional settings. Unraveling the history of the belt is a necessary step towards understanding the emplacement and remobilization of nickel ore. By refining a model of ore genesis, exploration geologists have a framework with which they can work. The relationships between lithologies, structural features, metamorphic grade and geochronology are formed within this model. Theories regarding both the history of the belt and ore genesis are hotly debated, as noted at CAMIRO research meetings. The major issues involve the Archean basement, boundaries of the belt, a possible rifting episode and the interpretation of structural features as linked to deformation events.

4.2.1 ARCHEAN BASEMENT

The margin of the underlying Archean basement was previously believed to dip westward beneath Proterozoic metasediments (Ospwagan group) (Figures 4.4 and 4.5). A number of seismic reflection studies indicate otherwise. Interpretations from both the Lithoprobe and COCORP projects indicate the Proterozoic rocks extend eastward beneath the Archean rocks. Figure 4.6 shows the location of seismic transect lines from the Lithoprobe project. Figure 4.7 shows the location of seismic reflection profiles from the COCORP project, located over the southern extents of the SBZ in Montana and North Dakota. Figures 4.8 and 4.9 show the results and interpretations from the Lithoprobe project. Revised models of the character of the Archean basement based on results from the COCORP project are shown in Figure 4.10. (Nelson *et al.* 1993, Lucas *et al.* 1996 and White *et al.* 1994).

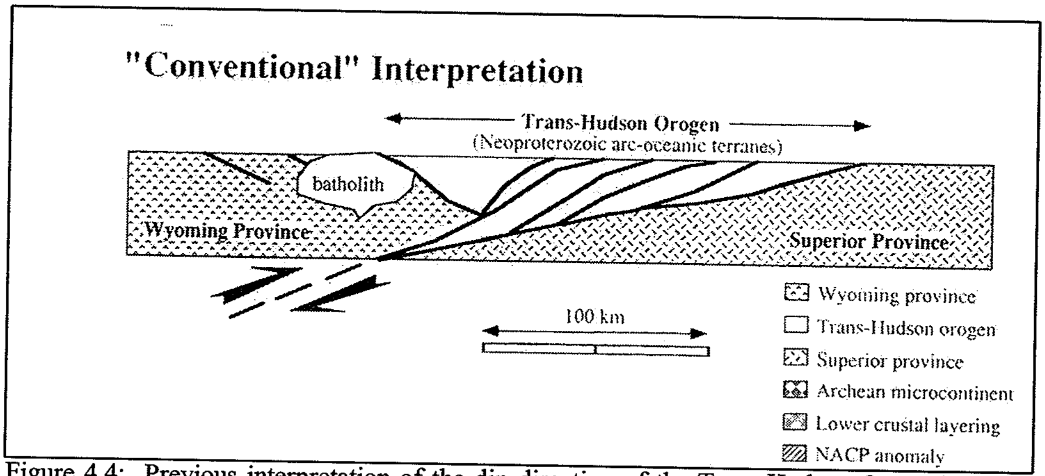


Figure 4.4: Previous interpretation of the dip direction of the Trans-Hudson Orogenic terrane. Adapted from Klasner and King (1990) and Nelson *et al.* (1993).

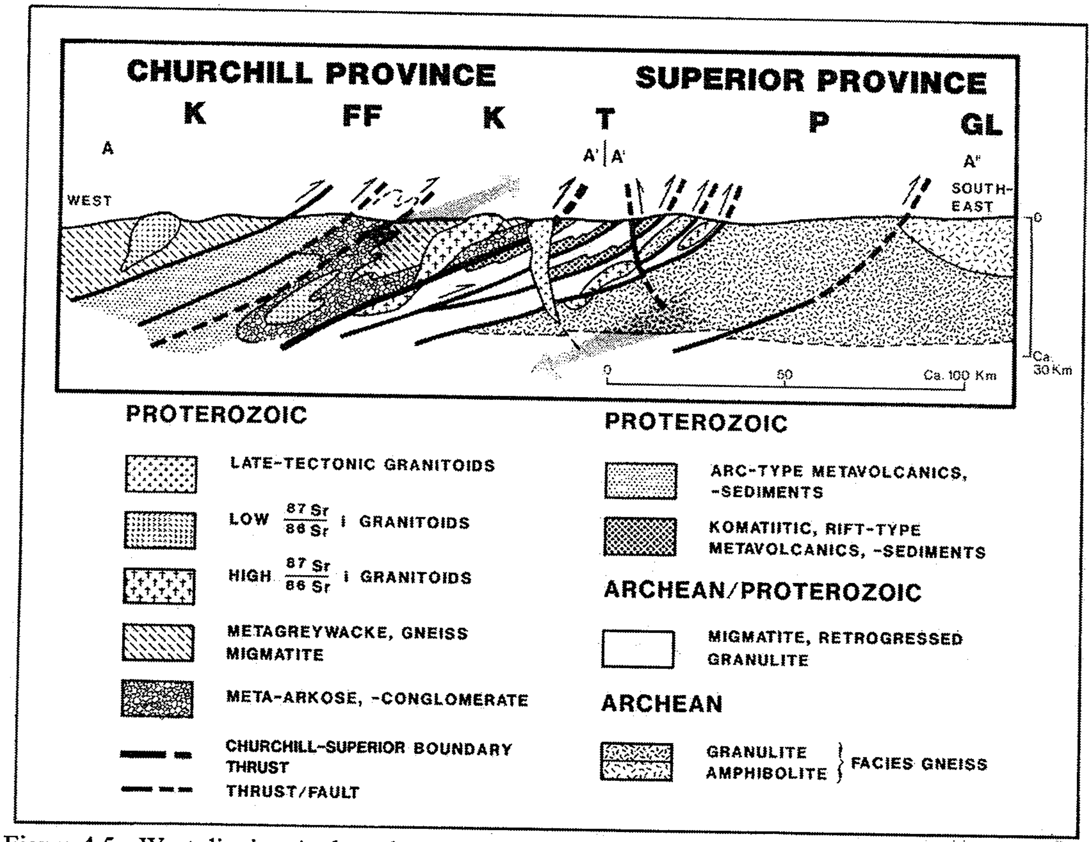


Figure 4.5: West dipping Archean basement schematic developed by Weber (1990). The dominant trend of west dipping structure fits previous interpretations of the belt. K – Kisseynew domain, FF – Flin Flon domain, T – Thompson Belt, P – Pikwitonei domain and GL – Gods Lake domain.

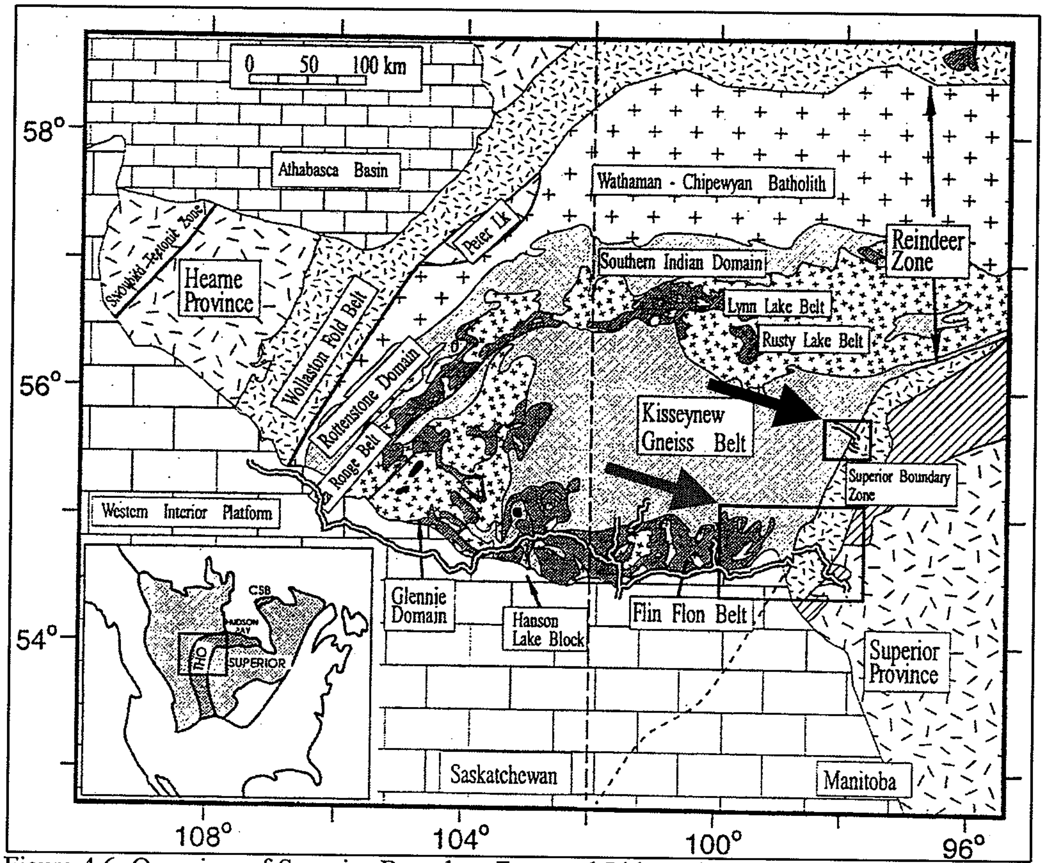


Figure 4.6: Overview of Superior Boundary Zone and Lithoprobe transect lines 1, 2 and 3 from the 1991 Lithoprobe seismic reflection project. The black arrow points to the area of line 1. The grey arrow points to the area of lines 2 and 3. From White *et al.* 1999.

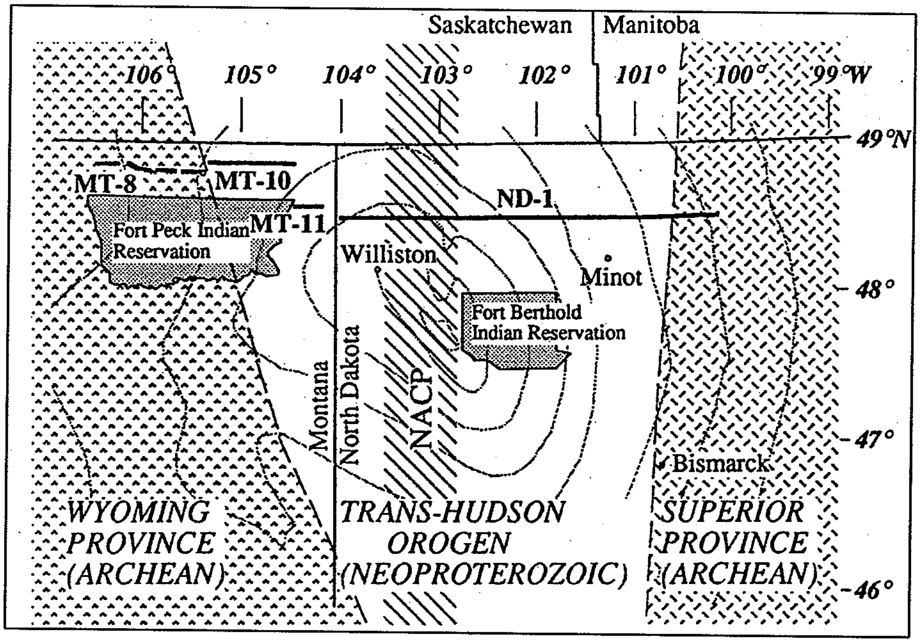


Figure 4.7: Location of COCORP seismic reflection profiles. From Nelson *et al.* 1993.

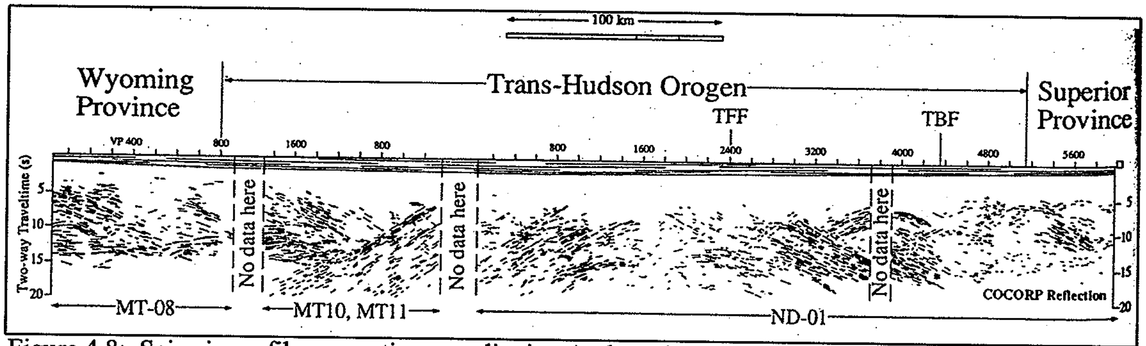


Figure 4.8: Seismic profile suggesting east dipping Archean basement under the TNB. Source data for line drawing from COCORP seismic reflection data. Adapted from Nelson *et al.* (1993).

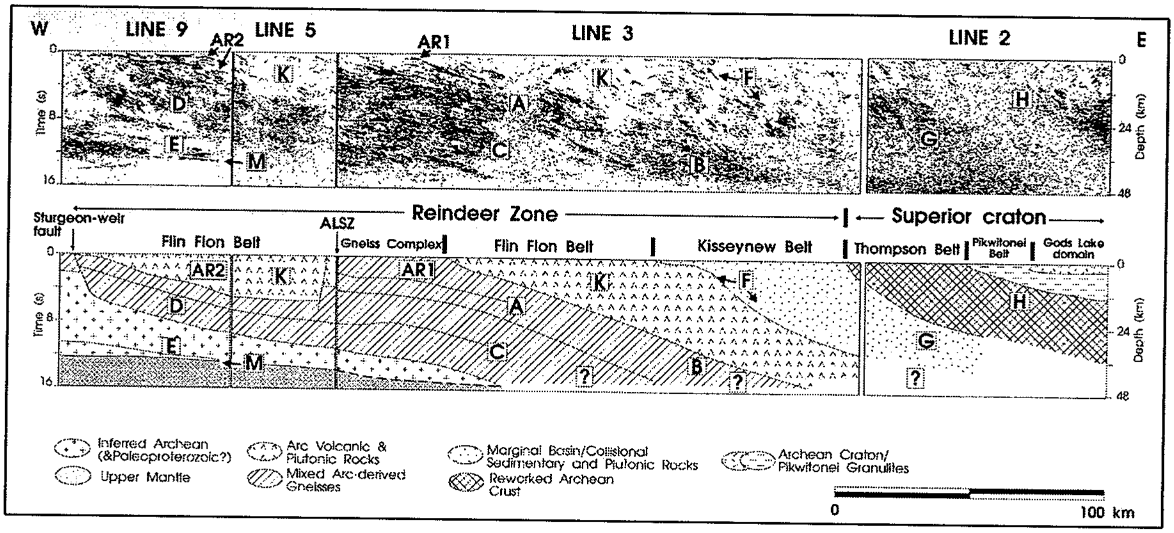


Figure 4.9: Seismic profile and interpreted cartoon model of east dipping TNB from White *et al.* (1994).

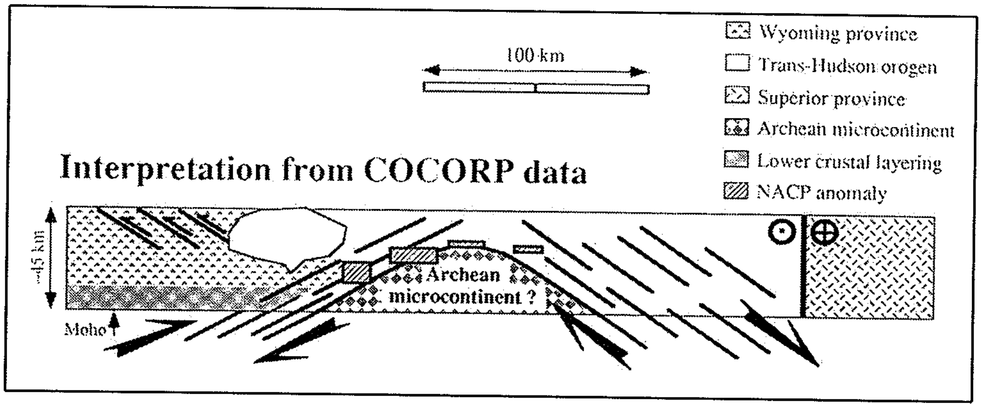


Figure 4.10: Revised model of THO terrane showing east dipping Archean basement. Adapted from Nelson *et al.* (1993).

The geometry of the Archean basement is an important factor when considering the tectonic history of the THO and TNB. Moreover, it will impact any 3D modelling derived from the GIS. The extent to which THO is underlain by Archean basement also contributes to understanding the development of the TNB. Archean terrane lies to the east of the belt. To the west, a complicated assemblage of island arcs and basins make up the Flin Flon Belt and Kisseynew Domain. The metamorphic grade of the western edge of the TNB is higher than the eastern boundary. The difference in grade indicates different temperatures and pressures acted on the rocks. Metamorphic grade is determined by classifying rocks by temperature and pressure conditions at time of metamorphism. The TNB exhibits different metamorphic conditions than its neighbours to the west in the Flin Flon belt and Kisseynew domain as well as to the east in the Pikwitonei and Gods Lake domain of the Superior Province. Refer to Figure 4.2 (Weber, 1990).

4.2.2 DEFINITION OF BOUNDARIES AND TERRANE

Though the TNB is recognized as an external belt of the Trans Hudson Orogen, the exact location of the belt is still very much questioned. The northeast trending feature emerges from under Phanerozoic cover near Wabowden. Just north of Moak Lake, the feature appears to curve to the east, continuing through the Split Lake block and the Fox River Belt (FRB) as shown by area 2 in Figure 4.11. Phanerozoic deposits cover the boundary in the area of the Hudson Bay Lowlands.

The Owl River Shear Zone (ORSZ), depicted as area 1 in Figure 4.11, which starts north of the eastward bend in the CSBZ, may prove to be a northern extension as a structural continuation of a sinistral fault. Examination of regional aeromagnetic images, as in Figure 4.11, reveal well defined linear structures along both ORSZ and FRB, similar to those found along the TNB.

However, more evidence is available to support the FRB as the eastern leg of the TNB due to the occurrence of similar lithologic assemblages. North-dipping Ospwagan Group supracrustals occur in the FRB. The formations that make up the Ospwagan Group may have been emplaced during different stages of a basin development event and consequently may explain why the Ospwagan Group appears to host nickel deposits within the TNB but not within the FRB .

4.2.3 STRUCTURAL INTERPRETATION

Structural interpretations of the Thompson Nickel Belt are complicated; what were once flat lying sedimentary and volcanic units have been turned on end and have undergone numerous periods of deformation and metamorphism (White *et al.,* 1994).

Deformation, intrusive igneous activity and metamorphism have obscured the relationship between rock units. In addition to this, there may have been docking and accretion of exotic terrane along the Superior Province. Consequently, the nature of the collision between the Superior and Churchill Provinces is difficult to model. There is minimal surface expression of the belt because of lake and vegetation cover. Examining the subsurface environment with geophysics and drilling is costly and prone to interpretation. These challenges aside, researchers have proposed a variety of scenarios that attempt to explain the belt as it exists today.

Early theories about the structure and tectonic history of the TNB were based on composites of aeromagnetic maps produced during the 1960s (Kornik 1971). Figure 4.11 is a modern example of aeromagnetic data. Regional aeromagnetic coverage provided researchers with a broad perspective from which to establish boundaries for the geologic provinces. Textures, distributions and patterns of magnetic anomalies were good indicators of magnetic features (Kornik and MacLaren 1966; Kornik 1969). Major structural features were identified. Apparent fault zones were marked for closer inspection. This approach, combined with field studies, produced maps defining the contact zone between the Superior (the Pikwitonei in particular) and Churchill provinces. Once boundaries were drawn, debate began regarding their correct location.

Aeromagnetic data for large portions of Manitoba provided a regional perspective of the TNB as the dataset included coverage of the belt and neighbouring areas. Early observations based on this aeromagnetic data suggested that the TNB might be a continuation of the ORSZ (Bell 1966). The data suggested the ORSZ was a sinistral fault with displacement of approximately 100 km. The strike of the ORSZ is sub-parallel to that of the TNB but there is no evidence to date that would support the view that the TNB is a sinistral fault. Furthermore, lithologies found in the two areas are quite different. Figure 4.11 displays the relative locations of the ORSZ, TNB and other features in the area. The spatial proximity and orientations of these features may indicate a cross-cutting relationship between the TNB and ORSZ. As both may have moved repeatedly during the orogeny, it is not clear which may predate the other.

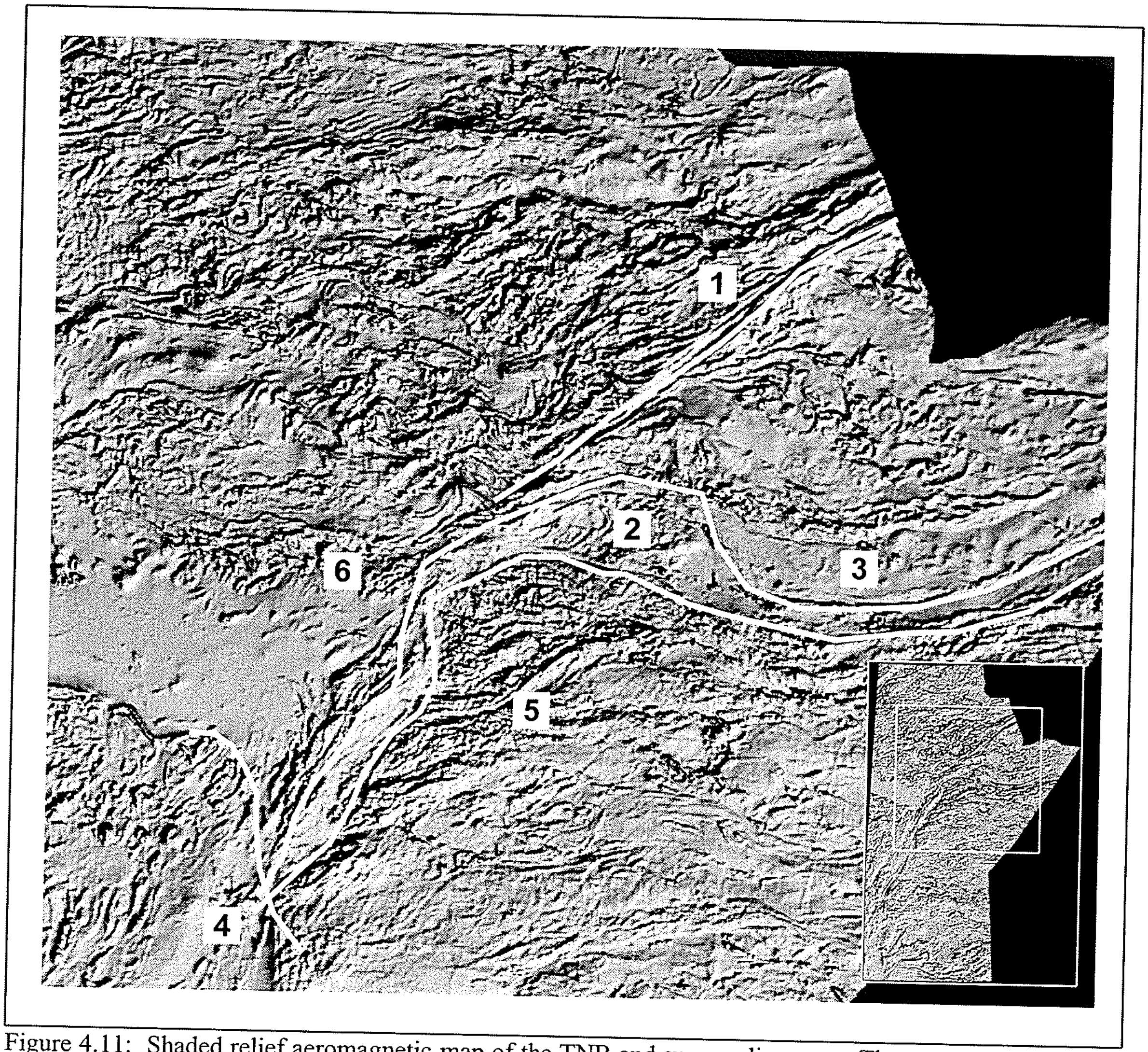


Figure 4.11: Shaded relief aeromagnetic map of the TNB and surrounding area. The aeromagnetic map is from Viljoen *et al.* (1990) and the boundaries were interpreted by Halden and Michalak. Thick white lines represent approximate boundaries between the TNB and the following areas:

1 Owl River Shear Zone
2 Split Lake Block
3 Fox River Belt
4 Paleozoic cover
5 Superior Province
6 Reindeer Zone

The TNB was subsequently interpreted to be related to the Split Lake Block (SLB) and the Fox River Belt (FRB). Polymetamorphic gneisses, similar to those in the TNB, are abundant in the SLB. Ospwagan group supracrustals and related intrusions appear in the FRB (Baragar and Scoates 1981, Scoates *et al.* 1977). Lithologic evidence suggests that the TNB is more likely to be related to the SLB and FRB than to the ORSZ.

Closer inspection of aeromagnetic data reveals interesting aspects of the TNB. Although the southern extension of the belt is overlain by Paleozoic rocks, expressions of it extend south into South Dakota. The boundary marking what is considered the eastern edge of the belt is also visible. Textures within the Superior Province are markedly different from the exposed area of the TNB. The strike of the TNB tends to the northeast whereas an east-west strike dominates the neighbouring Superior Province. The difference in the strikes becomes less evident as the Phanerozoic cover thickens. The western extent of the belt is delineated by a fault separating it from the neighbouring Reindeer Zone (White et al. 1994), where strike trends irregularly to the northeast (Zurbrigg 1963). The textures found west of the belt are also different. Throughout the area within and surrounding the belt, lenticular patterns are observed. These patterns are also noted in the geological maps provided by INCO, HBED and FB.

The TNB is composed of both reworked Archean gneisses from the Superior Province and paleo-Proterozoic continental margin rocks (Bleeker 1990a). This combination suggests that the belt hosts the boundary between the Churchill and Superior Province. The Superior Boundary Zone (SBZ; White and Lucas 1994) describes a region over which Archean rocks were interdigitated within Proterozoic rocks. It might be expected that the region of interaction is bounded by two control features, one to the east that marks the limit of the interaction in the Archean and one to the west of which only Proterozoic occurs. As the Orr Lake segment is now considered Archean, Proterozoic units may be found overlying Archean rocks, further obscuring Archean rocks at depth. Only when Archean and Proterozoic rocks are seen side by side is the actual contact observed. It may be unconformable-original and should be mapped as such. It also may be sheared, representing a late feature, and also mapped accordingly.

Early theories proposed the TNB was a transform fault (Lewry 1981), a suture associated with southeast directed subduction (Fountain and Salisbury 1981) and a suture associated with northwest directed subduction (Gibb 1983). Bleeker (1990b) and Weber (1990) proposed a scenario where the Superior province underthrust an island arc complex from the west and continued under the Reindeer Zone. These four theories were based on mapping, field observations, drill core and aeromagnetic data.

Bleeker's theory (1990b) was considered the most plausible as it accounted for all the existing data using a recognizable tectonic model. Figure 4.12 illustrates the main components of this evolutionary model based on surface observations, aeromagnetic data and drill core analysis. In block a, the Churchill Province advanced over the Superior foreland. Archean basement was incorporated into the Proterozoic lithologies as a result of thrusting and recumbent folding. Crustal shortening occurred in the lower levels of the overriding slab, producing F_1 - F_2 folds. The Archean foreland experienced strike-slip faulting after and perhaps during the Molson dyke intrusions. The thermal peak was reached after burial and the shear zones were overprinted with the signature of increased pressure and temperature.

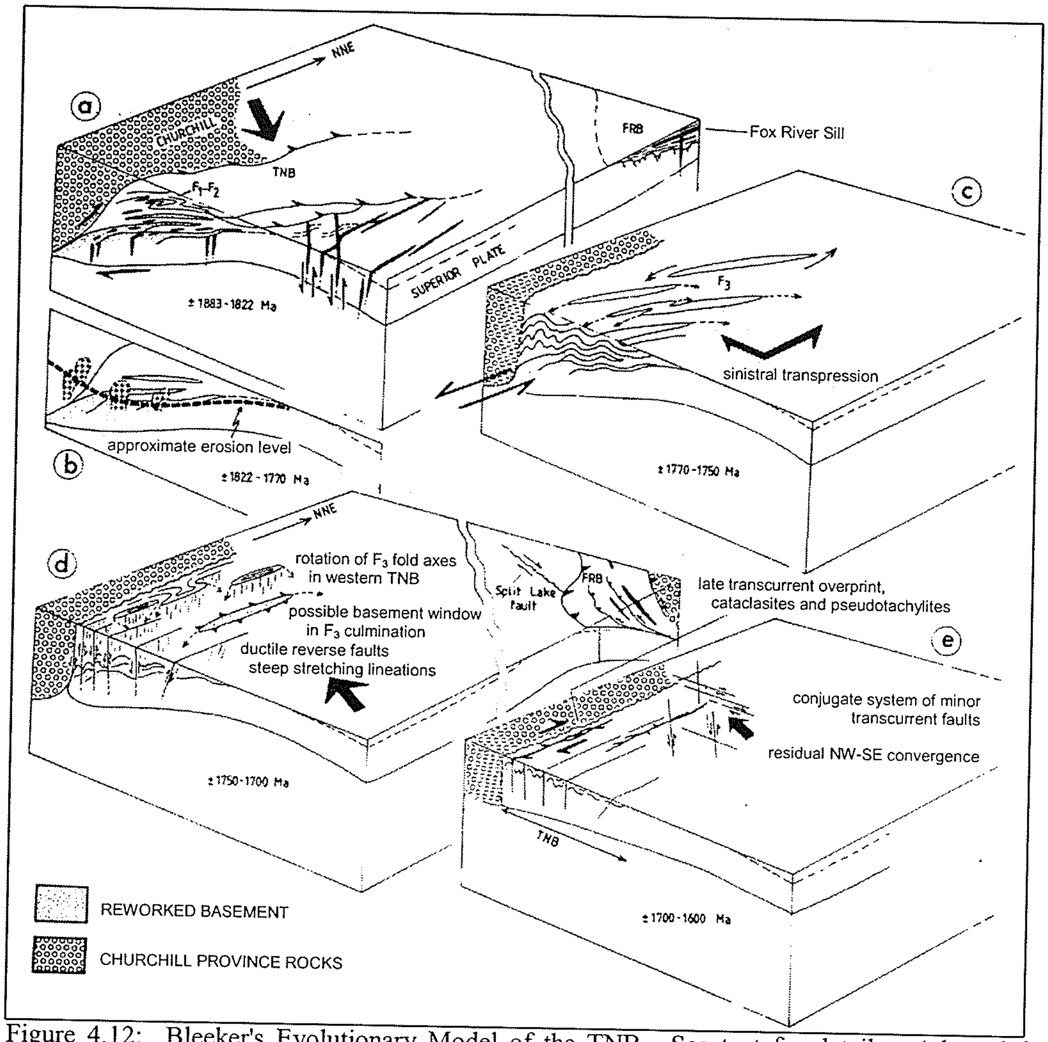


Figure 4.12: Bleeker's Evolutionary Model of the TNB. See text for details. Adapted from Bleeker (1990b).

Block b sketches the effect of metamorphism and anatexis on the lower levels of a thickened crust. Granitoid plutons are shown intruding into the lower levels. By about 1770-1750 Ma, oblique collision changed to sinistral transpression, resulting in en echelon F_3 folds. Over the next 50 million years, the Superior province moved slowly to the northeast, causing extensive ductile reverse faulting. F_3 folds were affected by the new orientation of movement. Plunges became steeper and as did the metamorphic gradient.

Throughout blocks c, d and e, isostatic rebound flexed the Archean rocks upward as erosion removed the overlying remnants of the previously overriding Churchill slab. The TNB is near its present configuration by block e. Dextral displacement becomes more prominent along the Churchill Superior suture, overprinting the effects of residual sinistral transpression. Strike-slip faults formed almost perpendicular to the belt in response to the continued northwest-southeast movement of the Superior and Churchill provinces. The Mackenzie dyke swarm, dated around 1270 Ma, crosscut the Superior Churchill boundary as northwest trending features with no displacement.

At the core of Bleeker's model is the original configuration of the Churchill and Superior Provinces. For his model to work, the Superior province must underthrust the Churchill province. The overriding slab is eroded and the underthrust slab rebounds. In the final configuration, an Archean basement still exists. White *et al.* (1999) proposed an alternate theory based on the results of Lithoprobe seismic reflection and magnetotelluric profiles.

As part of the Lithoprobe project, seismic data were acquired along two profiles in the Superior Boundary Zone to better constrain Trans-Hudson orogenic deformation and crustal geometry (White *et al.* 1999). An overview of the entire area is given in Figure 4.6. Figure 4.13 shows the location of Line 1; Lines 2 and 3 are shown in Figure 4.14. The two southern lines were processed and presented as a continuous line even though they were oriented differently and laterally shifted along strike. White *et al.* (1999) use the SBZ to describe the area Bleeker refers to as the TNB. Additionally, White *et al.* (1999) refers to the Reindeer Zone Collage, Snow Lake Arc, Kisseynew Basin and Flin Flon Belt specifically whereas Bleeker maintained the Churchill Province as a collective unit.

The results from the Lithoprobe study provide a subsurface image of the TNB. Comparison of data from the northern and southern transects revealed important differences at depth. White *et al.* (1999) attributes Bleeker's hypothesis with new results from the Lithoprobe study.

Seismic profiles from Lines 2/3 reveal reflections that dip east/southeast beneath the SBZ. The presence of east dipping lithologies to depths of 25km is contrary to Bleeker's model which includes an underlying Archean basement. The reflections were projected to surface and contrasts tended to coincide with contacts between surficial lithologies (White *et al.* 1999).

Key differences exist between the models proposed by White and Bleeker as comparison of Figures 4.12 and 4.15 illustrates. Underthrusting of the Superior Craton beneath the Reindeer Zone occurs at 2.12 Ga in concert with continental rifting, shown in block a. Block b suggests an arc-continent collision takes place around 1.883-1.84 Ga. The collision might have involved the Snow Lake Arc, although that is not confirmed. This collision produced the thrust and fold belt, similar to block a of Bleeker's evolutionary model (Figure 4.12). Crustal thickening continues from 1.83 to 1.82 Ga as the Flin Flon Belt and Kisseynew Basin override the Reindeer Zone, causing east vergent folds and thrusting in the SBZ, shown in block c.

Block d is where the biggest differences occur. The Flin Flon Belt continues to converge and becomes wedged into the Superior Craton, manifesting as east side up folding and faulting on both sides of the Superior Boundary Fault (SBF). The SBF, which delineates the western edge of the SBZ and the eastern edge of the Churchill Province, is in place by 1.8 to 1.77 Ga (block d). Underlying the Flin Flon Belt is the Reindeer Zone, which also wedged eastward into the Superior Craton. Block e shows the current configuration of the area, resulting from intracontinental transpression from 1.77 to 1.72 Ga.

Figure 4.8 illustrates the seismic profiles generated for Lines 2/3. Not only did the Lithoprobe results offer new insights into the formation of the TNB, it also provided examples of differences that exist along the SBF. Seismic profiles generated for Line 1, about 100km north of Lines 2/3, are significantly different (Figure 4.16). The Reindeer Zone does not appear to underthrust the Superior Craton. The conclusion drawn from this observation is that the crustal structure of the SBZ varies along strike.

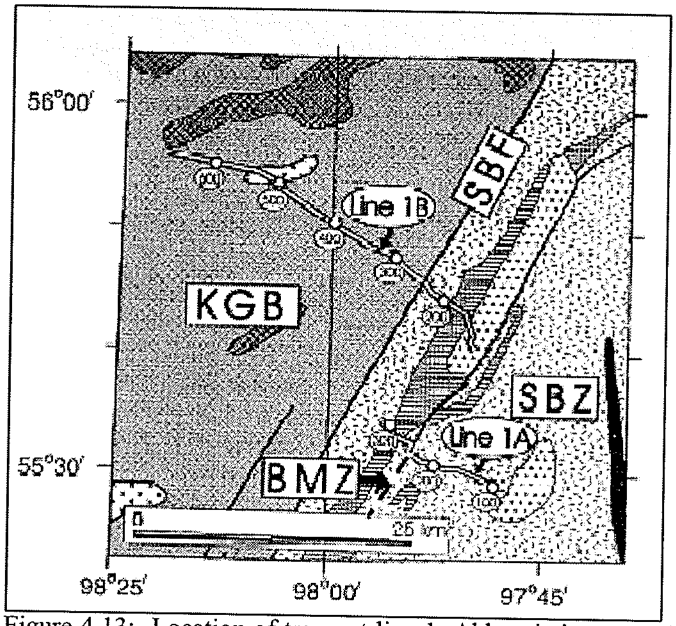


Figure 4.13: Location of transect line 1. Abbreviations are: KGB, Kisseynew Gneiss Belt; SBF, Superior Boundary Fault; SBZ, Superior Boundary Zone; BMZ , Burntwood-Missi Zone. From White *et al.* 1999.

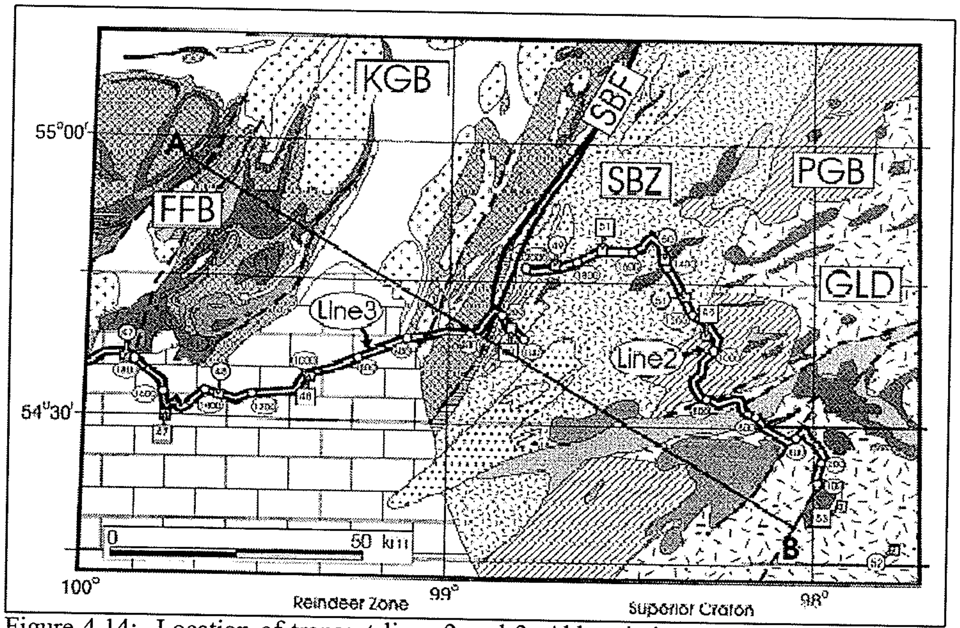


Figure 4.14: Location of transect lines 2 and 3. Abbreviations are: KGB, Kisseynew Gneiss Belt; SBF, Superior Boundary Fault; SBZ, Superior Boundary Zone; FFB, Flin Flon Belt; PGB, Pikwitonei Granulite Belt; GLD, Gods Lake Domain. From White *et al.* 1999.

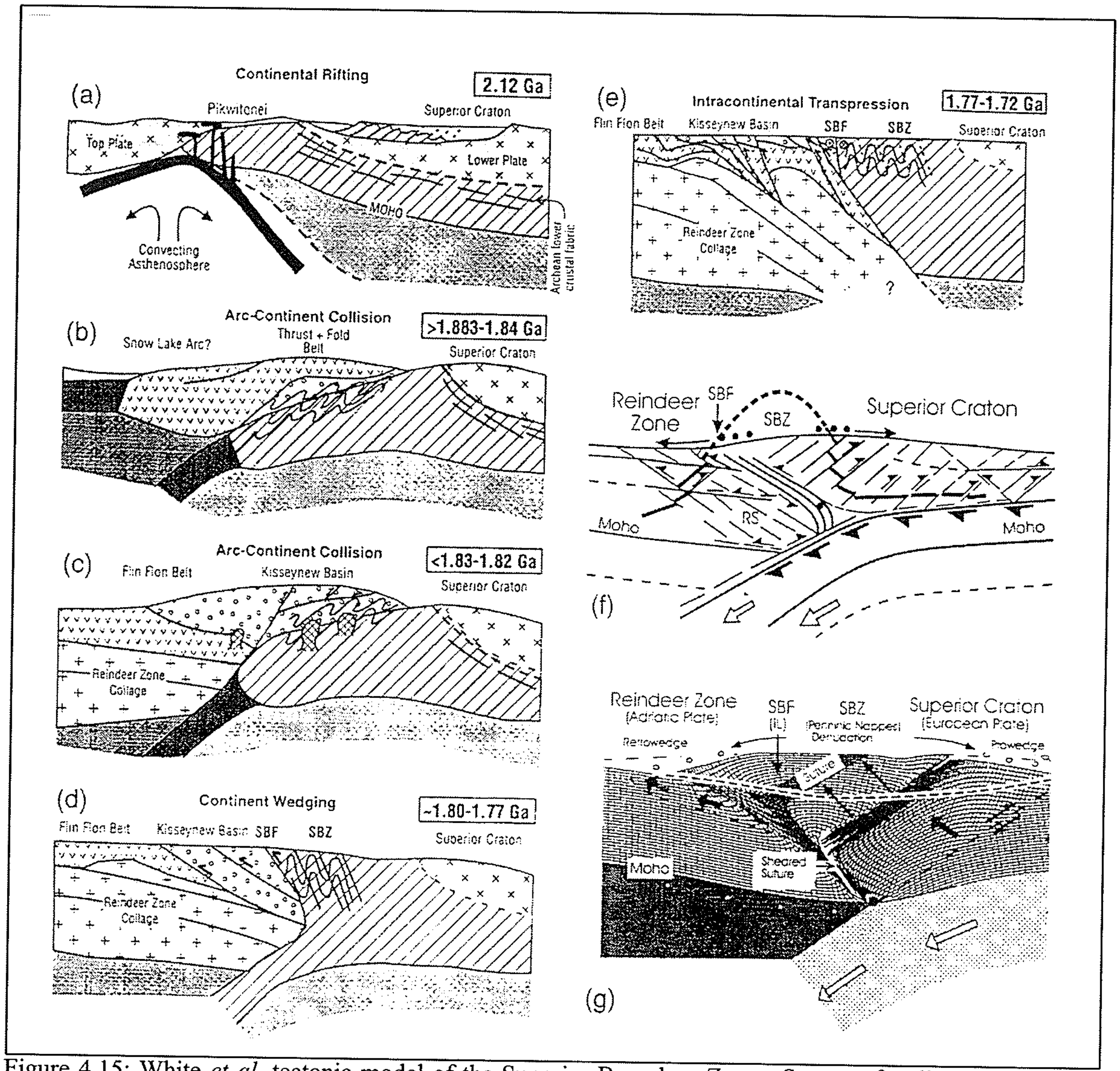


Figure 4.15: White *et al.* tectonic model of the Superior Boundary Zone. See text for discussion. From White *et al.* (1999).

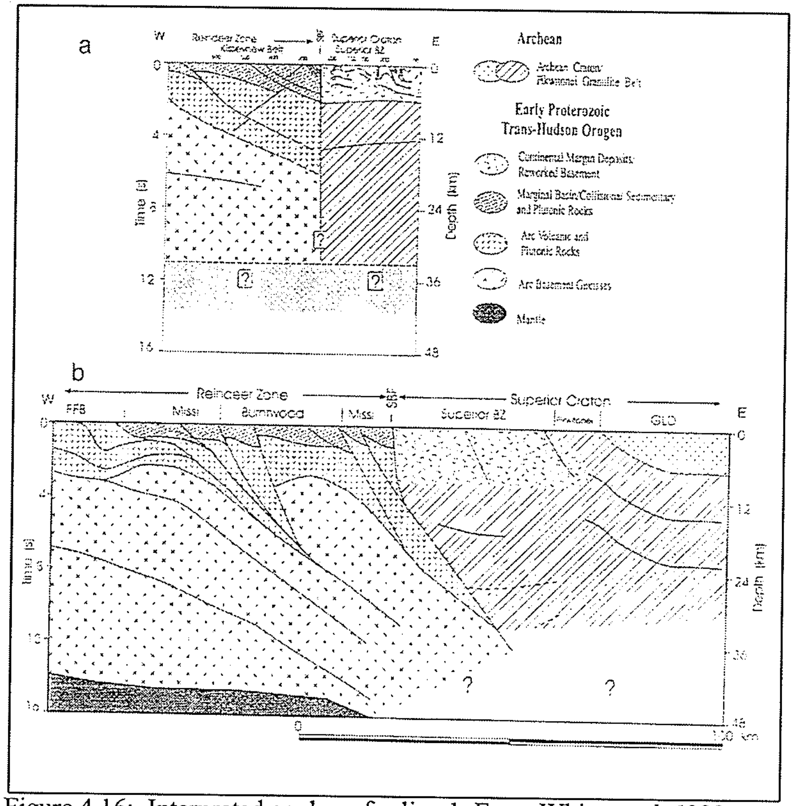


Figure 4.16: Interpreted geology for line 1. From White *et al.* 1999.

CHAPTER 5

SPATIAL ANALYSIS

5.1 INTRODUCTION

The INCO geology map package represents the results of years of field mapping in the TNB and was provided to the research group in a format compatible with the GIS. A series of layers was included in the package such as lakes, structure and geology. The geology map was the most important layer as it contained the most comprehensive collection of mapping for the belt. The map was provided in vector format and attributed with lithologic descriptions. Figure 5.1 shows the legend for the map; note that lithologies, formations and domains are all included in the legend.

The lithologies were mapped as polygons. They exist next to, within or near each other and the geological maps help to visualize these spatial relationships in two dimensions. Visual inspection of such maps indicate where the units exist next to or within each other. However, proximity to a particular unit is more difficult to discern visually. Proximity is a spatial measure of the distance between features. If a geological process created related features, than the distance between the features may suggest the scale of the process.

The spatial proximity of lithologic units to each other was investigated to determine if lithologies not currently considered influential to a nickel deposit may indeed have a role in emplacement or mobilization of nickel ore. Buffering and area tabulation functions available in ArcView were used to quantify the spatial relationships between lithologic units.

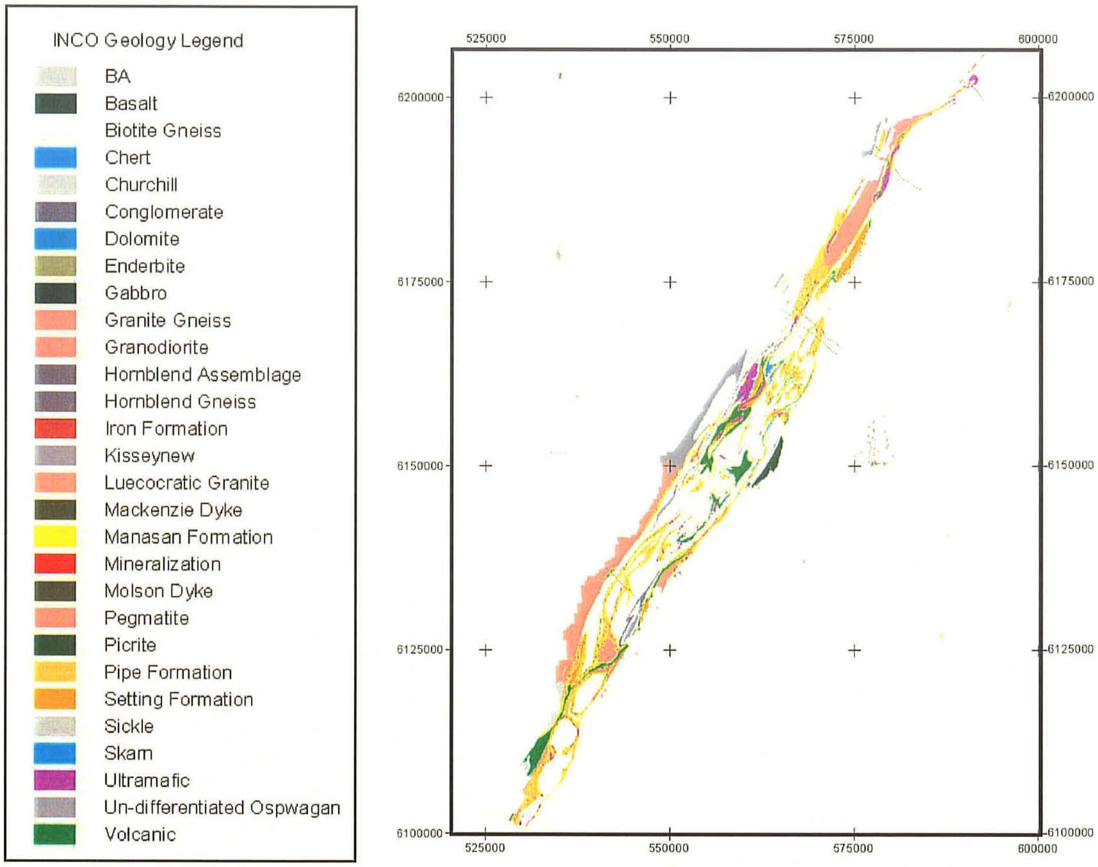


Figure 5.1: INCO Geology map and legend.

The existing model of nickel ore emplacement is based largely on Bleeker's model of stratigraphic relationships (Bleeker, 1990a; Figure 5.2). In this model, an ultramafic body intrudes into a sedimentary sequence. The ultramafic magma assimilates sediments, in particular the sulfide component. For a deposit to form, the ultramafic intrusion had to become saturated in sulfides. Sulfide facies banded iron formations are believed to be the sulfur source. Based on this model, three lithologies are important to the formation of a nickel deposit – ultramafic rocks, sulfide facies banded iron formations and mineralized bodies.

5.2 METHODS

The three lithologies were identified as targets based on observed spatial associations with nickel deposits. For example, the red polygon in Figure 5.3 represents ultramafics. That polygon was used to create a buffer zone extending out to 100m. The buffer zone was divided into 10m intervals, creating a series of concentric circles around the ultramafic polygon. Each circle was then identified as unique feature based on the distance from the ultramafic polygon. The first circle was labeled as the 10m zone, the second as the 20m zone and so forth. The area of each polygon from the geology map that fell within the circle was calculated. The results were stored in a table similar to that in Figure 5.3 which shows cartoons for each stage of the process.

The analysis excluded a series of polygons identified in the INCO legend. Lithologies classified as formations included the Manasan, Pipe and Setting Formations. The formations are defined by a group of members which were not mapped. The boundaries of members of each formation, especially the Pipe Formation, would be very useful for spatial analysis. However, the scale of mapping precludes the level of detailed needed to render the morphology of members.

Lithologies that covered very little area or were not metamorphic or igneous rocks were also excluded from the analysis and inlcude: BA, basalt, Churchill, conglomerate, enderbite, gabbro, hornblende assemblage, luecocratic granite and sickle. Finally, the Iron Formation lithology was renamed to Banded Iron.

The results were plotted as column charts displaying area in m^2 along the y-axis and buffer zone along the x-axis as shown in Figures 5.5 to 5.11. The legend for the charts is provided in Figure 5.4. The data are presented at two scales. The first scale shows the results for the entire TNB; the second shows results for the T1 mine site.

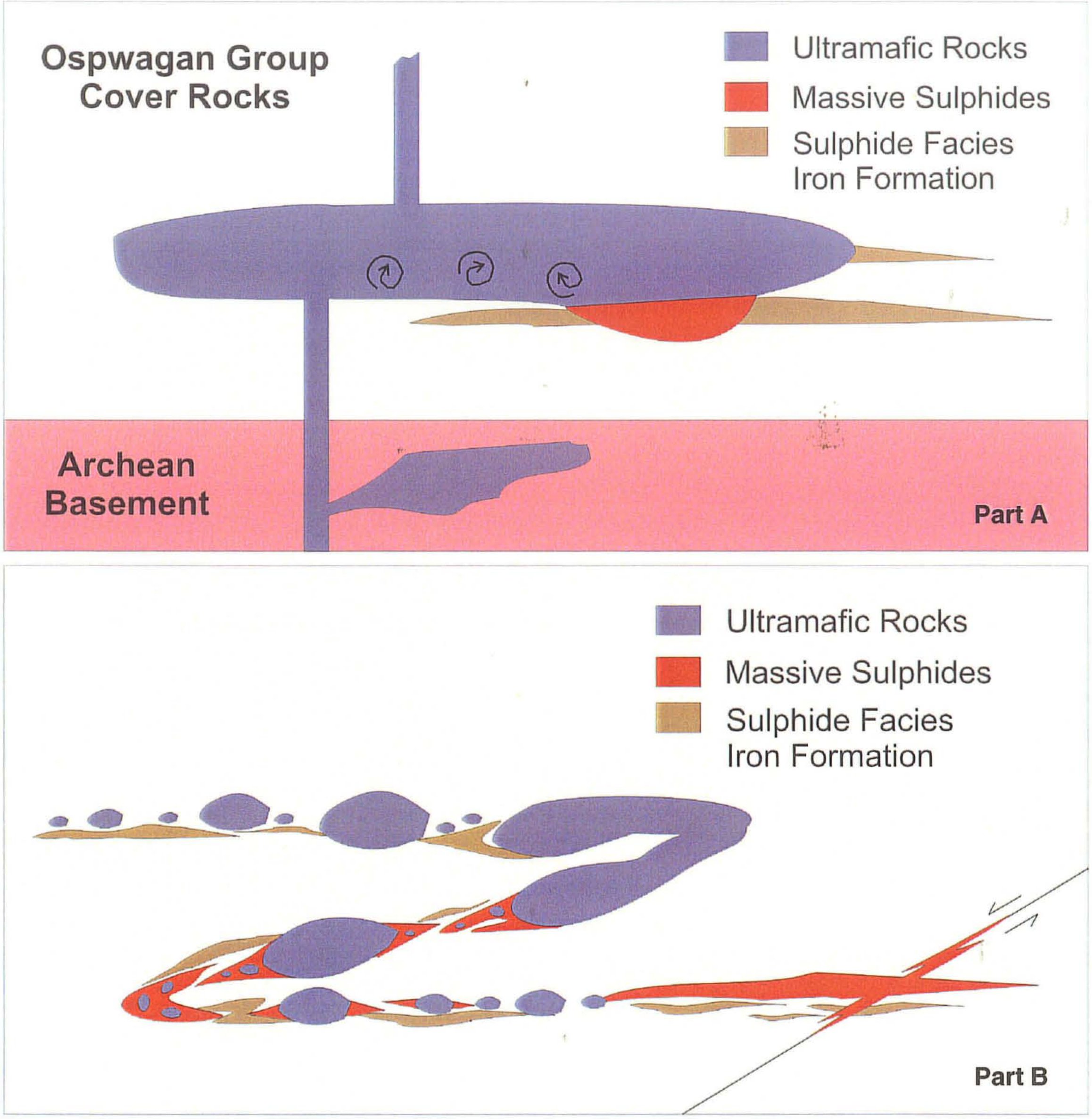
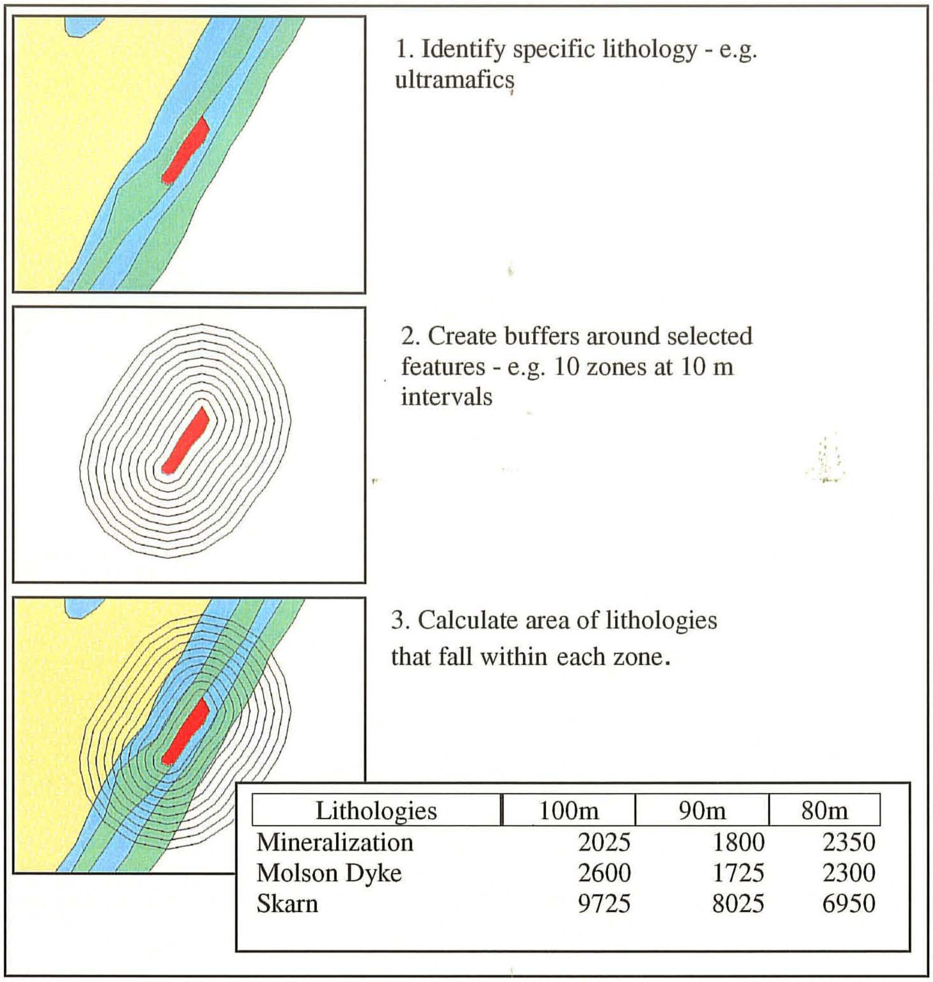


Figure 5.2: A model of ultramafic intrusion and ore deposition is illustrated in Part A. Part B illustrates a model of ore remobilization resulting from folding, faulting and shearing. Adapted from Bleeker (1990a) and Liwanag (2001).



1. Identify specific lithology - e.g. ultramafics

2. Create buffers around selected features - e.g. 10 zones at 10 m intervals

3. Calculate area of lithologies that fall within each zone.

Lithologies	100m	90m	80m
Mineralization	2025	1800	2350
Molson Dyke	2600	1725	2300
Skarn	9725	8025	6950

Figure 5.3: Buffer analysis process. Step 1 selects the lithology of interest. 10 buffer zones, each at 10m wide, are created in Step 2. The area of each lithology is calculated within successive 10m wide zones around the lithology in Step 3. The table shows examples of the area calculation (m^2) for three 10m zones - 100, 90 and 80m.

5.3 RESULTS

The legend for the spatial analysis figures is provided in Figure 5.4 and includes only the lithologies used in the analysis. Figure 5.5 shows the proximity of all lithologies to mineralization across the belt. Skarns, shown in bright blue, commonly occupy the area enclosed by the 30m buffer. This result suggests that skarns and mineralized zones may be related. Skarns form as products of hydrothermal cells. The circulating fluids of a hydrothermal cell transport elements via convection currents. The mobilization of elements in superheated fluids, or metasomatism, may play a role in mobilizing sulphides to interact with ultramafics.

These results are not conclusive, but may suggest that a hydrothermal cell producing skarns occurring near banded iron formations may be related to mineral deposits. This observation may indicate the approximate scale of nickel deposits and the environment in which they originally formed. The results for a mine site, shown in Figure 5.6, also indicate that skarns appear in the 30 and 40m buffer zone more than the other buffer zones. These observations are incorporated into the data fusion models discussed in Chapter 6.

Figures 5.7 and 5.8 show lithologic proximity to banded iron. There are no obvious changes in proximity to different units across the belt. It is interesting to note the mine site results in Figure 5.8. Skarns occupy more area in the 70m buffer zone. Future work may investigate at the relationship between banded iron, skarns and mineralized bodies to better determine the scale of nickel ore formation.

Proximity to ultramafics was divided into north and south regions because calculating buffer zones and areas within the buffer zones was restricted due to the large number of polygons in the dataset. Figures 5.9 and 5.10 show the results for the entire belt. Note the difference in area of ultramafics covered by skarns between north and south regions. Volcanics, shown in red, occupy more area in the south as compared to the north. Also note the difference in the y-axis scale. Larger areas are covered by lithologies in the south. However, the relative amount of each lithology is similar if the contribution from volcanics is excluded. Mine site results for proximity to ultramafics, shown in Figure 5.11, is unremarkable.

🟥 Volcanic
🟧 Banded Iron
🟦 Chert
🟦 Skarn
🟦 Dolomite
🟦 Picrite
🟨 Molson Dyke
🟪 Mackenzie Dyke
🟥 Biotite Gneiss
🟩 Hornblend Gneiss
🟧 Pegmatite
🟪 Granite Gneiss
🟩 Granodiorite
🟫 Kisseynew
🟩 Undifferentiated Ospwagan
🟨 Ultramafic
🟦 Mineralization

Figure 5.4: Legend for Figures 5.5 to 5.11.

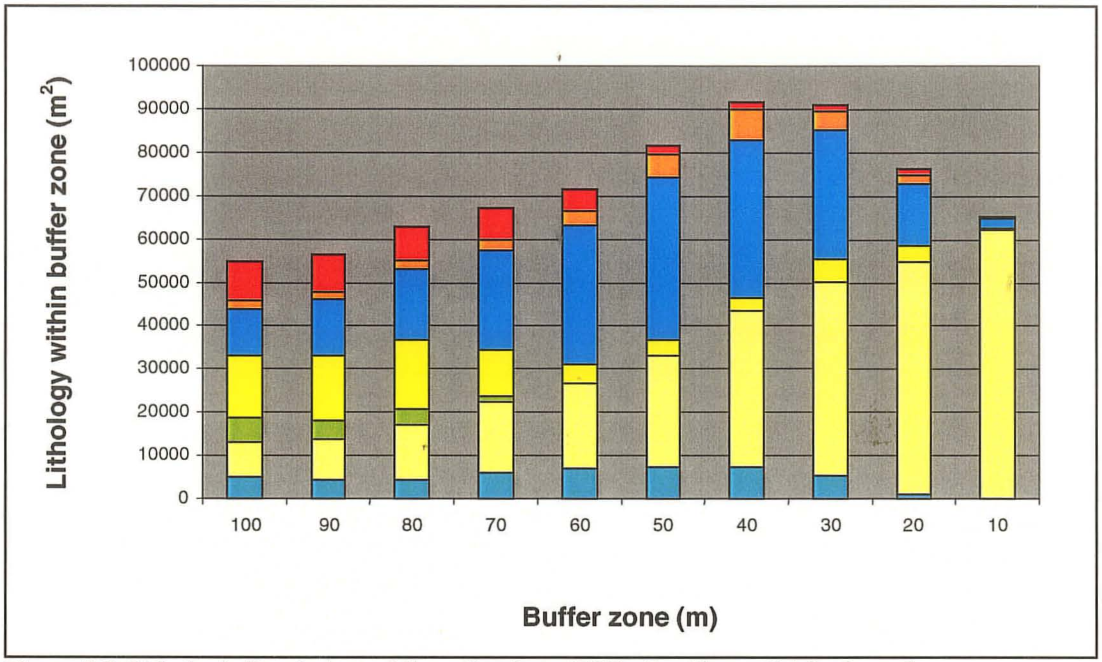


Figure 5.5: Lithologic Proximity to Mineralization – TNB. See Figure 5.4 for legend.

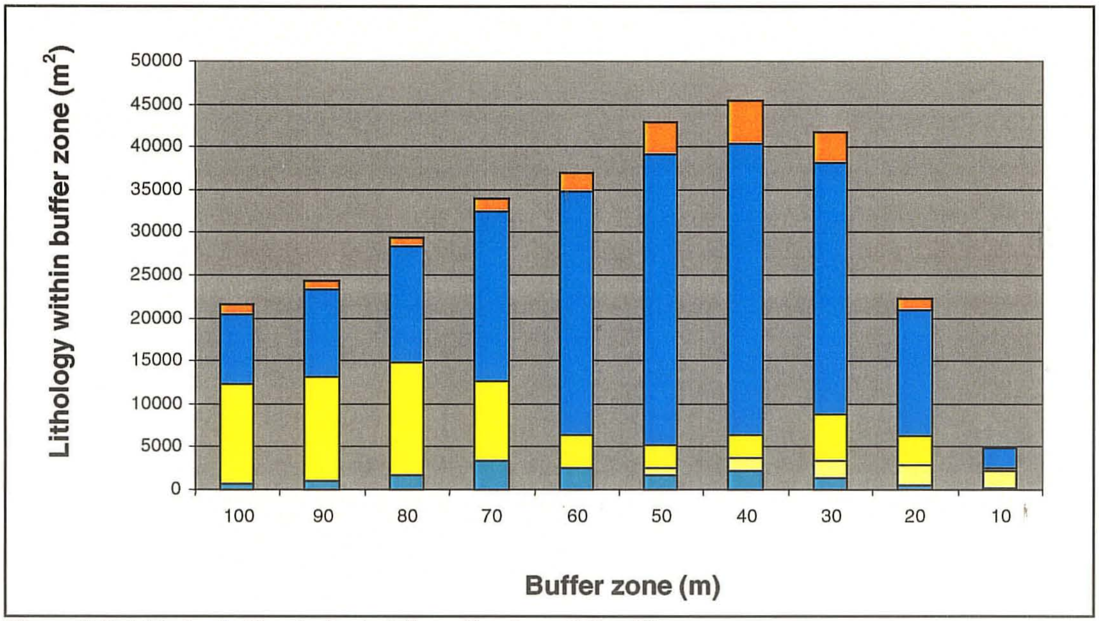


Figure 5.6: Lithologic Proximity to Mineralization – Mine Site.

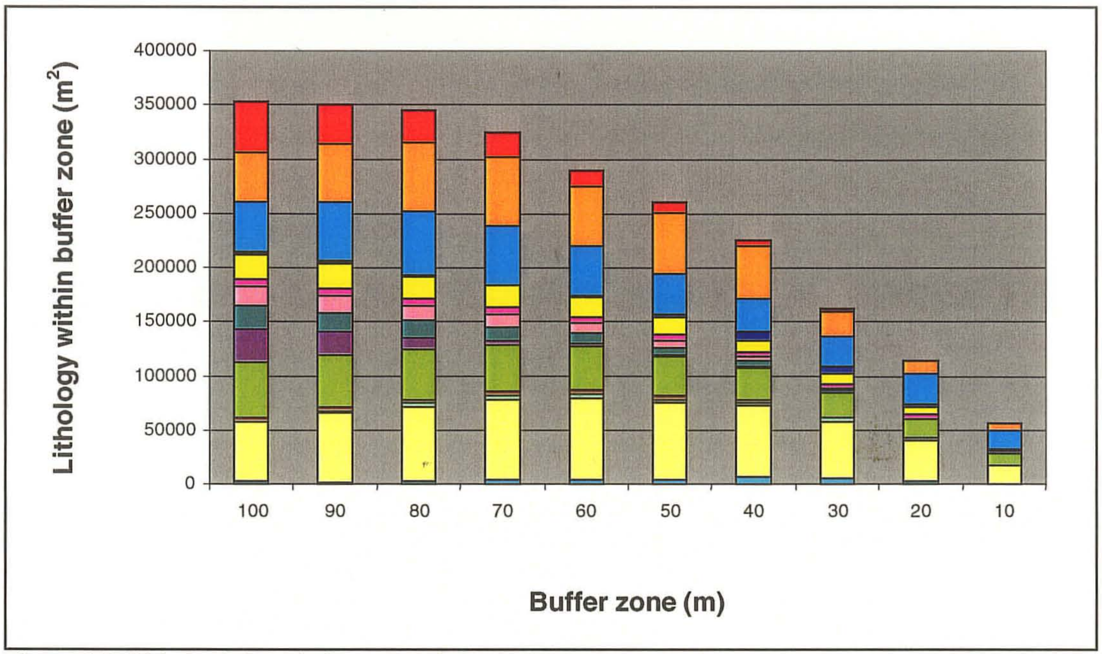


Figure 5.7: Lithologic Proximity to Banded Iron – TNB.

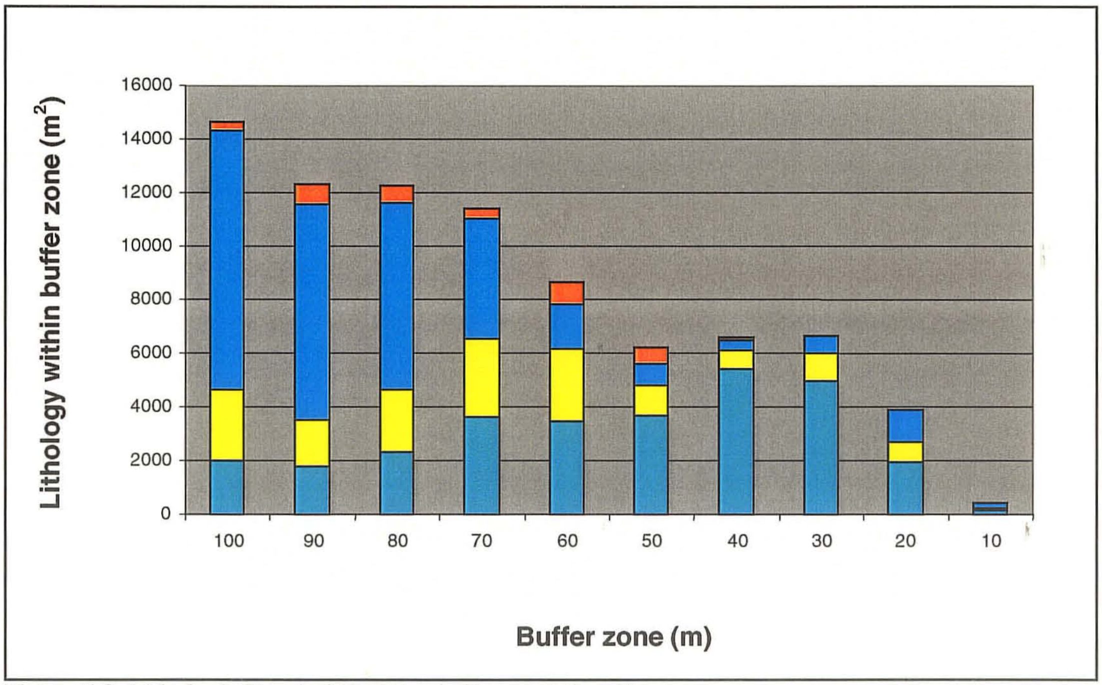


Figure 5.8: Lithologic Proximity to Banded Iron – Mine Site.

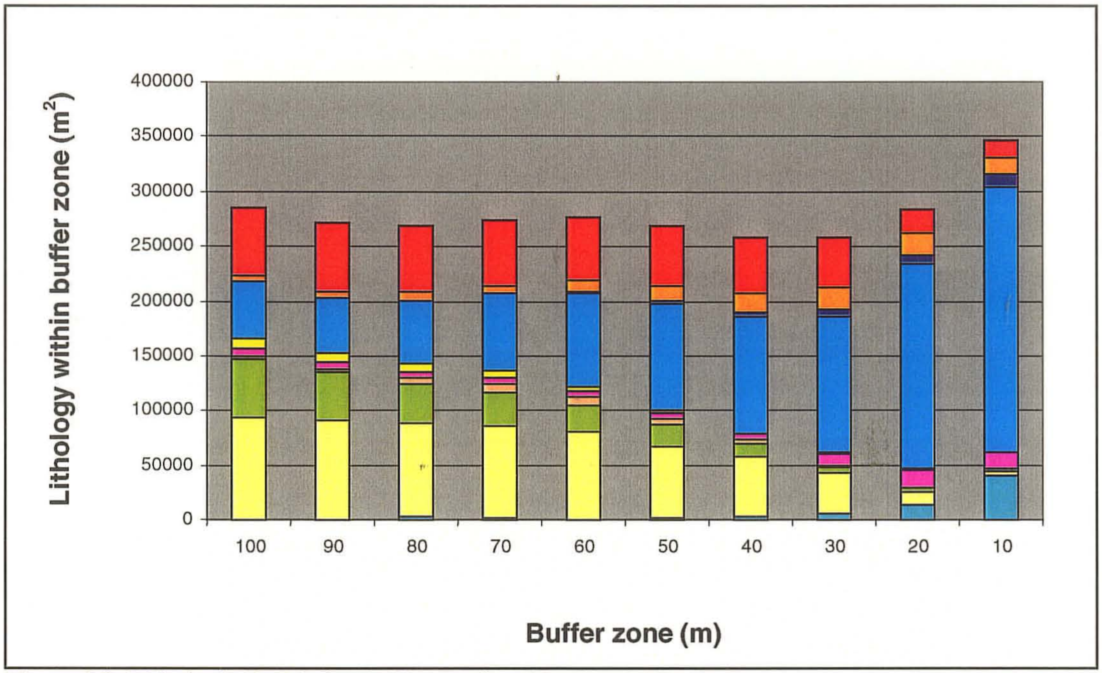


Figure 5.9: Lithologic Proximity to Ultramafics (North) – TNB.

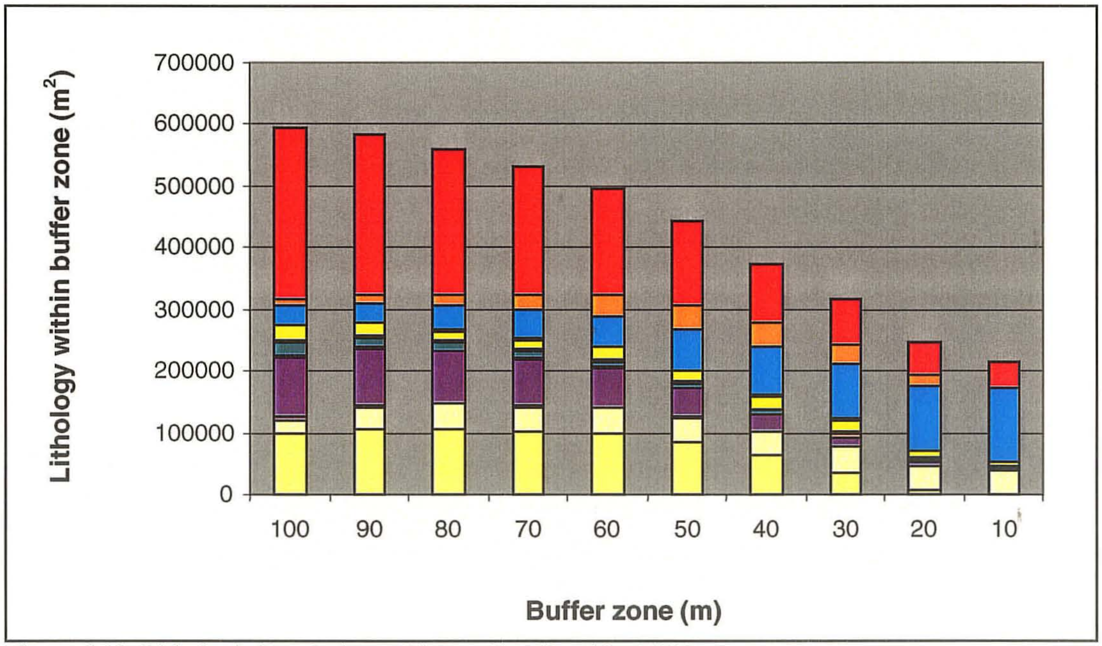


Figure 5.10: Lithologic Proximity to Ultramafics (South) – TNB.

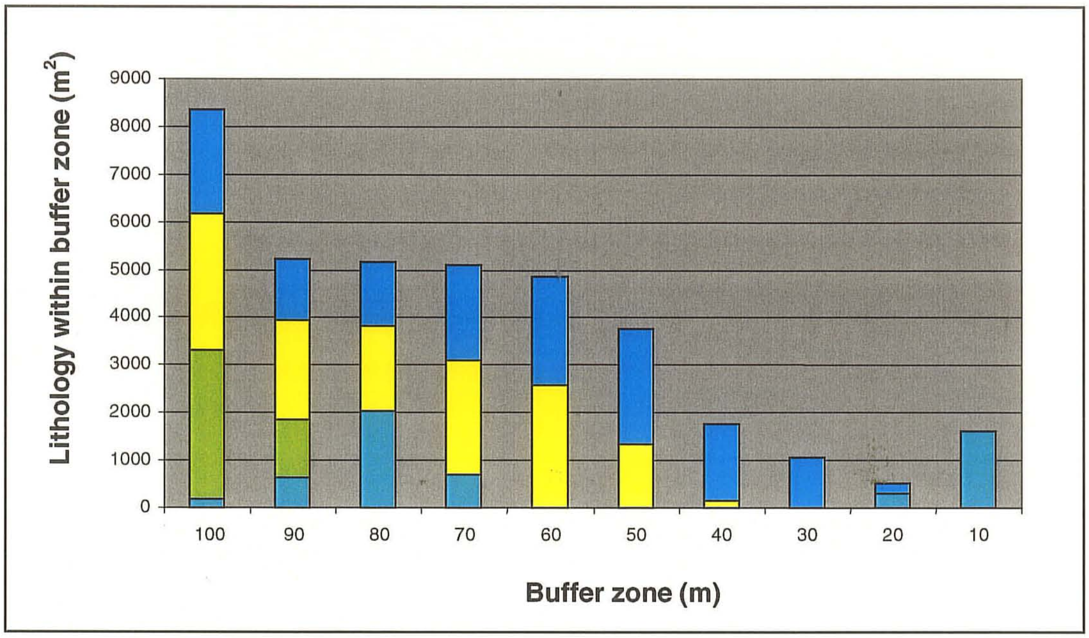


Figure 5.11: Lithologic Proximity to Ultramafics – Mine Site.

CHAPTER 6

DATA FUSION

6.1 INTRODUCTION

INCO's nickel and copper operation in Thompson resulted from the detection of the ore body by aeromagnetic sensors in 1956, wrapping up ten years of exploration in the area (Zubrigg 1963). Other mine sites were established along the TNB in following years. The Moak Lake deposit was active before the Thompson deposit was discovered. The Thompson discovery occurred as a result of investigating aeromagnetic and conductivity anomalies in the area. INCO based their exploration parameters for delineating a potential ore body on conditions found at Moak Lake.

Many geologic features are examined to understand ore emplacement and redistribution of ore in the context of a tectonic setting. Emplacement and redistribution have both spatial and temporal implications. An obvious goal of the CAMIRO project is to locate other nickel deposits. Data fusion techniques were employed to create mineral potential maps and contribute to the underlying goal of developing new exploration methods for the TNB.

Theories and models of ore emplacement are generated from observations made in the field, from geologic maps and from geochemical, geophysical, mineralogical and isotopic data. Intrusions, deformation and metamorphism alter the spatial relationship between adjacent lithologies. Techniques used to interpret previous relationships between units include analysis of geochemistry, structure, petrology, geochronology and geophysics. In order to determine where a nickel deposit may occur, it is useful to examine the environment in which known nickel deposits occur. Knowing where nickel does not occur is also important. Exploration models and constraints on where ore might be distributed are primary deliverables in this project. However, the first important step is to describe how a nickel deposit occurs.

6.2 MINERAL DEPOSIT MODELS

A mineral deposit model combines descriptions and criteria that characterize a deposit. Criteria can be determined by either empirical observations or conceptual models. Empirical models are derived from data that describe physical parameters of known deposits. Conceptual models attempt to explain how a deposit formed using geological processes. A model for any ore deposit typically incorporates elements of both conceptual and empirical models using data derived from known deposits (Hodgson 1993).

After years of study, researchers proposed that structural and petrographic evidence suggest tectonic events remobilized the original deposit consistent with the protracted deformational history discussed earlier. These tectonic events obscured details of the original relations between rocks. Unwrapping the formations to determine their original relations proved to be difficult. However, by studying exposed contacts and drill core, certain lithologies appeared to be associated with ore deposits. The Ospwagan group was identified as the most likely to host a deposit. Mapping done by Bleeker and Macek has shown that ultramafic rocks are found within the Archean basement and in the P1 and P2 formations. So a combination of anomalies occurring in consortium with P1 and P2 will likely form the foundation of any model.

Considering the TNB, one mineral deposit model may not be sufficient. Current thought promotes the idea that known deposits may have been emplaced and remobilized under different conditions. A number of models may be needed to establish the environments in which nickel was deposited and for this research, two models were used. These models incorporated events and processes that are believed to lead to nickel deposition and remobilization. The two models distinguished between structural and stratigraphic influences. Figure 5.1 illustrates cartoons of stratigraphic and structural models.

6.2.1 Stratigraphic Model

A fairly well accepted concept is that nickel ore is associated with ultramafic rocks (Bleeker, 1990a). The basis for this concept is that known deposits have been found in or adjacent to ultramafic sills. Another important factor is that in order to reach sulphide saturation the ultramafic magma must assimilate sedimentary sulphides (Bleeker, 1990a). Therefore, metamorphosed sedimentary rocks containing sulphur are also identified as important criteria for a nickel deposit. Some additional questions include 1.) What was the original composition of the sedimentary rocks? 2.) Where did the sediments originate? 3.) How were they deposited? 4.) What was the environment of deposition? In turn, similar questions can be posed about the ultramafic intrusions, for example, 1.) What was the temperature of the magma? 2.) What was the composition of the magma? By extracting details specific to a known ore deposit, a set of characteristics can be developed and used to suggest locations of new deposits.

An example of this approach is seen in the work of Bleeker (1990a). Volcanic massive sulfide (VMS) deposits are believed to form as a result of the reaction between host rocks and intruding magma. A sulphide source is necessary to produce such a deposit. Studies are underway to determine if the sulfur needed to form the various deposits in the TNB were provided by the magma, by a lithology the magma passed through prior to intrusion in the host rock, the host rock itself or through a process such as metasomatism which occured post intrusion.

6.2.2 Structural Model

Liwanag (2001) expanded on ore emplacement research by examining processes that mobilized ore. Five areas were sampled in the 1998 field season, including the 1C and 1D ore bodies from T3, T1, Birchtree and William Lake. Each of the five areas hosted economical nickel ore deposits. Liwanag grouped the sulfides into five types: disseminated sulfides in metasedimentary rocks, remobilized sulfides in metasedimentary rocks, magmatic sulfides in ultramafic rocks, remobilized sulfides in ultramafic rocks and sulfides in chemical sedimentary rocks.

Results from her research indicate that the emplacement of nickel ore and the resulting tenor may be sensitive to different processes. Consequently, the second model would incorporate the severity of deformation and metamorphism. Structural and geophysical evidence seems to support this scenario. The environments and processes that formed the existing expression of the TNB may have varied along strike and over time. This knowledge can be incorporated into exploration methods using data fusion techniques.

6.3 DATA FUSION

Data fusion describes a group of methods that combine quantitative data with qualitative user knowledge by assigning weights to data (Wald 1999). The weights applied to the data can be determined either by statistical methods (Bayesian approach) or by an expert in the area of study (Fuzzy Logic and Dempster-Shafer Belief Theory (DSBT)). The weights are chosen with respect to a specific question, or hypothesis. The question is posed with the intent of extracting information from the data collection. The information produced is influenced by both the weighting scheme and the actual data.

The hypothesis is a statement that describes the objective of the decision maker. It guides the assignment of belief values to each piece of evidence. For example, mineral potential maps can be created for any number of minerals. Lithologic units are an important piece of evidence when a particular mineral is associated with a lithology. Gold and nickel will occur in different lithologies, so if the hypothesis is "A gold deposit occurs", the weights assigned to each lithology will be different from those assigned if the hypothesis was "A nickel deposit occurs".

The weights assigned using the Fuzzy Logic method are measures of support for the hypothesis. Another term used to indicate support is belief. Complimenting belief is disbelief. For example, on a scale from 0 to 1, 0.8 would be assigned to a piece of evidence that is highly supportive of the hypothesis. The disbelief is 0.2, or 1 - belief. The relationship is depicted graphically in Figure 6.1. The thick grey line is the belief function used to represent values assigned to individual pieces of evidence from one data layer, indicated on the x-axis. The shape of the function is also determined by the relative ordering of evidence value. The belief value is plotted along the y-axis. The solid vertical black line represents a piece of evidence within a layer of data.

Experts assign belief values based on their interpretation of how the piece of information supports the hypothesis. Often, there is some hesitation regarding both the data quality and how well the data support the hypothesis. While the Fuzzy Logic method only handles belief values, the DSBT incorporates measures of uncertainty. If the expert is uncertain about some aspect of the data, a quantitative value can be assigned to indicate the degree of uncertainty. Compared with the Boolean ideology of Fuzzy Logic belief/disbelief, DSBT allows an expert some flexibility with the assignment of support values. Figure 6.2 illustrates the relationship between belief, disbelief and uncertainty. When uncertainty equals zero, the DSBT is the same as Fuzzy Logic.

A new line appears on the graph representing plausibility. The belief function now indicates a conservative measure of support whereas plausibility represents an optimistic measure. Uncertainty occupies the area between belief and disbelief. Belief no longer equals one minus disbelief. Plausibility now equals one minus disbelief. Other relationships that can be extracted include:

$$1 = \text{Disbelief} + \text{Uncertainty} + \text{Belief} \qquad \text{Equation 1}$$

$$\text{Plausibility} = 1 - \text{Disbelief} \qquad \text{Equation 2}$$

$$\text{Disbelief} = 1 - \text{Plausibility} \qquad \text{Equation 3}$$

$$\text{Uncertainty} = \text{Plausibility} - \text{Belief} \qquad \text{Equation 4}$$

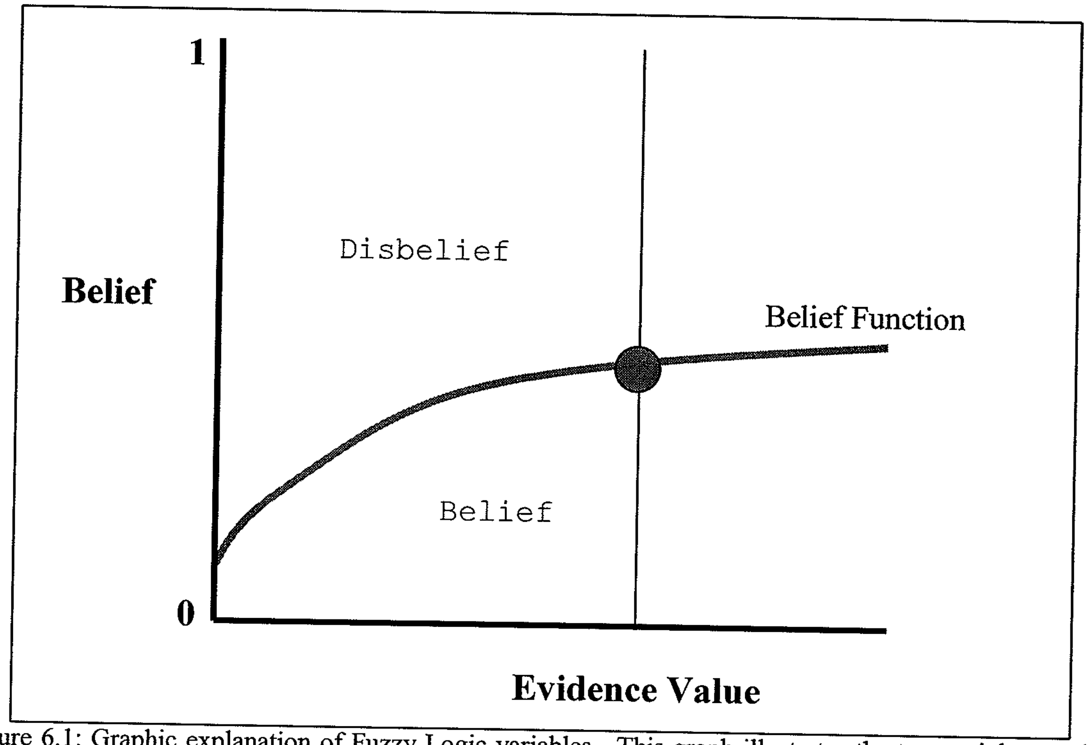


Figure 6.1: Graphic explanation of Fuzzy Logic variables. This graph illustrates the two weights used in the Fuzzy Logic method. The Evidence Value - or the attribute which is assigned a weight - lies along the x-axis. The range of values is dependant upon the dataset used in the model. The y-axis represents the weight assigned to each attribute and falls on a continuous scale from 0 to 1. A belief value of 0 indicates that the attribute does not contribute to the model. A belief value of 1 indicates the attribute strongly contributes to the model. The grey dot indicates the belief value assigned to an attribute. The thick grey line represents a function created by connecting all the belief values assigned to each piece of evidence in the dataset. Disbelief is calculated as 1 minus Belief. Only one value is assigned to the attributes. Adapted from Bonham-Carter (1996).

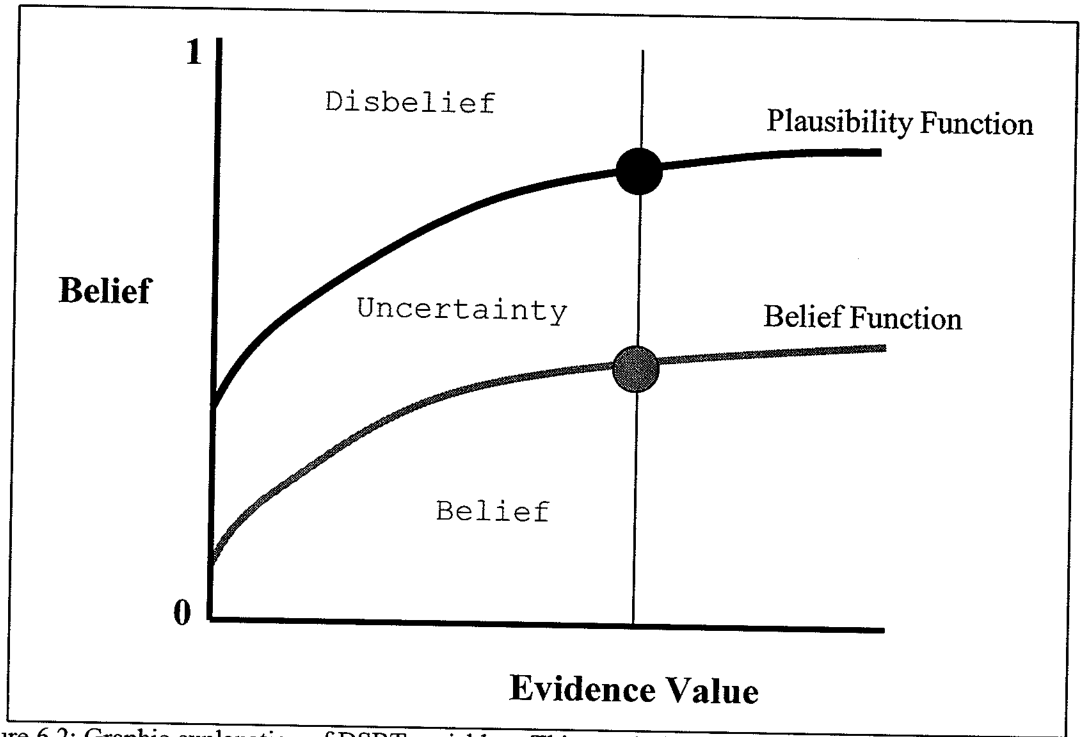


Figure 6.2: Graphic explanation of DSBT variables. This graph illustrates how the four weights used in the DSBT are related to each other. The Evidence Value - or the attribute which is assigned a weight - lies along the x-axis. The range of values is dependant upon the dataset used in the model. The y-axis represents the weight assigned to each attribute and falls on a continuous scale from 0 to 1. A belief value of 0 indicates that the attribute does not contribute to the model. A belief value of 1 indicates the attribute strongly contributes to the model. The grey dot indicates the belief value assigned to the attribute. The thick grey line represents a function created by connecting all the belief values assigned to each piece of evidence in the dataset. Likewise, the black dot represents the plausibility value assigned to each attribute. The black line is the function that represents plausibility of the dataset. Uncertainty is calculated as Plausibility minus Belief. Disbelief is calculated as 1 minus Plausibility. Only two values are assigned to the attributes. The remaining two are calculated. Adapted from Bonham-Carter (1996).

6.4 WEIGHTS

The hypothesis used in the application of the DSBT to determine areas of potential nickel deposits was: *A nickel deposit is likely to occur*. The DSBT was applied using datasets provided by the CAMIRO project including layers of aeromagnetics, geochemistry, structural features, and geology. ArcView and Spatial Analyst provided the tools needed to handle the data in a GIS.

The weighting of each layer depends on how likely the type of data is able to indicate a nickel deposit. The concept of "how likely" has to be a quantitative value. The expert assigning the weight to a particular attribute must believe that the occurrence of the attribute indicates some level of potential for a nickel deposit. The mineral deposit model also guides the assignment of weights such that attributes believed to contribute to the existence of a nickel deposit will be assigned high levels of support. Table 6.1 describes both the stratigraphic and structural mineral deposit models and how weights were assigned to each layer of data.

Table 6.1: Description of mineral deposit models, data and assignment of weights.

	Fuzzy Logic		Dempster-Shafer Belief Theory	
	Stratigraphic Model	**Structural Model**	**Stratigraphic Model**	**Structural Model**
Model Summary	High belief is placed on stratigraphic relationships that express a contact between magma and sulfide (S) iron formations (IF). The SIF provide S to ultramafic (UM) intrusions.	Significant influence from the D3 event is represented by assigning high belief to structure. The D3 event is the primary influence, but source rocks are still needed and are represented in this model by assigning medium belief to lithology.	High belief is placed on stratigraphic relationships that express a contact between magma and SIF. The SIF provide S to UM intrusions.	Significant influence from the D3 event is represented by assigning high belief to structure. The D3 event is the primary influence, but source rocks are still needed and are represented in this model by assigning medium belief to lithology.
Data				
Lithology	High belief values assigned equally to the following lithologies: UMs, skarns, marbles, IF, mineralization an volcanic rocks. Higher belief values assigned to IF proximal to UMs in the Pipe Formation.	Medium belief values assigned equally to the following lithologies: UMs, skarns, marbles, IF, mineralization an volcanic rocks. Higher belief values assigned to IF proximal to UMs in the Pipe Formation.	High belief values assigned equally to the following lithologies: UMs, skarns, marbles, IF, mineralization an volcanic rocks. Higher belief values assigned to IF proximal to UMs in the Pipe Formation. Small uncertainty assigned to all.	Medium belief values assigned equally to the following lithologies: UMs, skarns, marbles, IF, mineralization an volcanic rocks. Higher belief values assigned to IF proximal to UMs in the Pipe Formation. High uncertainty assigned to all.
Aeromagnetics	Medium belief values assigned to large digital numbers, decreasing linearly.	Low belief values assigned to large digital numbers, decreasing linearly.	Medium belief and uncertainty values assigned to large digital numbers, decreasing linearly.	Low belief and medium uncertainty values assigned to large digital numbers, decreasing linearly.
Lineaments	Low belief values assigned to areas proximal to lineaments.	High belief values assigned to areas proximal to lineaments.	Low belief values assigned to areas proximal to lineaments.	High belief and low uncertainty values assigned to areas proximal to lineaments.
Faults	Low belief values assigned to areas proximal to faults.	High belief values assigned to areas proximal to faults.	Low belief values assigned to areas proximal to faults.	High belief and low uncertainty values assigned to areas proximal to faults.
Strike	Low belief values assigned to areas with strike near 30°.	High belief values assigned to areas with strike near 30°.	Low belief values assigned to areas with strike near 30°.	High belief and low uncertainty values assigned to areas with strike near 30°.
Skarns	High belief values assigned to areas within 20 to 40 metres of skarns. IF within these areas are weighted higher.	Medium belief values assigned to areas within 20 to 40 metres of skarns. IF within these areas are weighted higher.	High belief values assigned to areas within 20 to 40 metres of skarns. IF within these areas are weighted higher. Small uncertainty assigned to all.	Medium belief and uncertainty values assigned to areas within 20 to 40 metres of skarns. IF within these areas are weighted higher.

6.5 DATA

The data fusion methods were applied using datasets provided by the CAMIRO project including layers of aeromagnetic data, structural features and lithology. Each dataset will be described in terms of its scale, origin, processing and rationale for inclusion in the models. Figure 6.3 shows the extents of the area around the TNB used in the data fusion calculations. Major towns and active or previously active mine sites are also indicated. The inset map locates the area of interest with respect to the province of Manitoba.

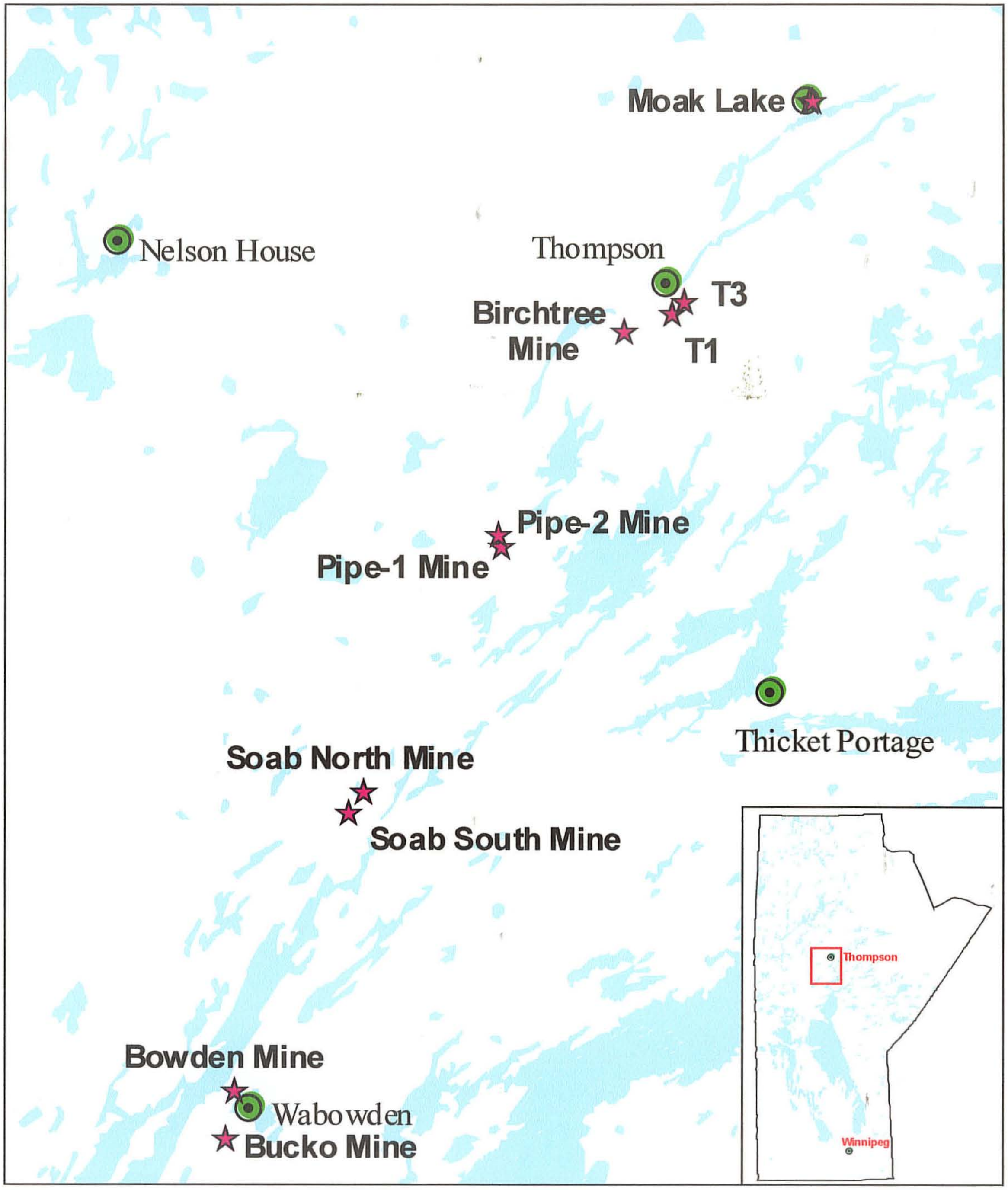


Figure 6.3: The area of interest showing lakes, mine locations and towns. The inset map of Manitoba shows the outline of the area used in the data fusion application. The pink stars indicate the surface location of known deposits. Towns are represented by green circles. North is up.

6.5.1 LITHOLOGY

The geology map provided by INCO in June 2000 includes the most detailed mapping available for the belt. The map is a work in progress and represents a portion of INCO's regional geological mapping. The version released to the CAMIRO project was restricted to areas under consideration in the project. The map was created for use at a scale of 1:20,000 and was also used in Chapter 5 for the spatial analysis application.

Figure 6.4 shows the extents of the map used in the data fusion process. A large portion of the map does not contain data. In order for the Fuzzy Logic method to be applied, areas with no data were classified as "no data" to avoid null values as they complicate the calculation of the Fuzzy Logic equations.

The weights assigned to each lithology are determined by the role the lithology plays in the formation of a nickel deposit as well as the quality of the map itself. A number of issues influence the quality of the lithology map including feature morphology, classification of lithologic features, errors inherent to the dataset and absence of georeferencing information.

The morphology of the polygon features tends to be oriented 30° NE and are typically much longer along strike than across, resulting in cigar shaped units. The distance across strike ranges from 20m to 1km. The morphology is an important consideration as this layer is combined with others to calculate the mineral potential maps.

The relatively short distance across strike means that when the polygons are converted to pixels to facilitate analysis, the size of the pixels must be small enough to render the narrow features. For example, creating a grid with a pixel size of 50m would not correctly render a polygon feature that is only 20m wide. However, a 10m pixel would be better able to render the 20m wide polygon into grid cells. As a result of the conversion from vector to raster-based maps, error is introduced to the dataset. Where polygon boundaries were defined by lines as a vector file, in raster format they are defined by pixels and the line becomes jagged. The areal extent of units consequently changes, sometimes to the point where small polygons do not appear in the grid. Therefore, the choice of map scale is very important as it affects subsequent analysis.

The map included both lithologic units and formations. However, members of formations were not mapped. The location and morphology of the members can improve interpretation of lithological relationships in the belt. In addition, the topography of the formations and intrusions is determined from outcrops. The position of ultramafic intrusions along limbs of folds or within nose regions can not be determined from this geology map. Since the position and volume of ultramafic rock proximal to sulfide facies banded iron formations may indicate a potential deposit, the 2D representation is not sufficient to make conclusions about deposits (personal communication with Lescher and Burnham, Sudbury 2001). However, given the scope of the project and lack of other data, this dataset was incorporated into the data fusion models.

Errors found in the INCO geology dataset include labeling and topological errors. These errors introduce a high level of uncertainty to the dataset. The topology errors include instances of duplicate polygons positioned in the same location. There are also instances where large polygons cover smaller polygons. Again, one location can be identified by two lithologies. The small polygons should have “cut” into the larger polygons. The large polygon would be like a donut and the hole in the donut would be where the small polygon would show through. With regards to the labeling errors, there are instances where a single polygon is identified as two different lithologies, compounding the topology problems. These errors indicate the dataset was not subjected to quality control measures and suggests further errors may exist within the dataset.

Finally, the dataset was originally provided in a coordinate system created by INCO. The conversion to NAD27 UTM zone 14 was done using a warping algorithm using obvious landmarks as ground control points (GCPs). The process, algorithm and GCPs were not provided in a metadata document. Therefore, positional accuracy with respect to other datasets may also introduce error.

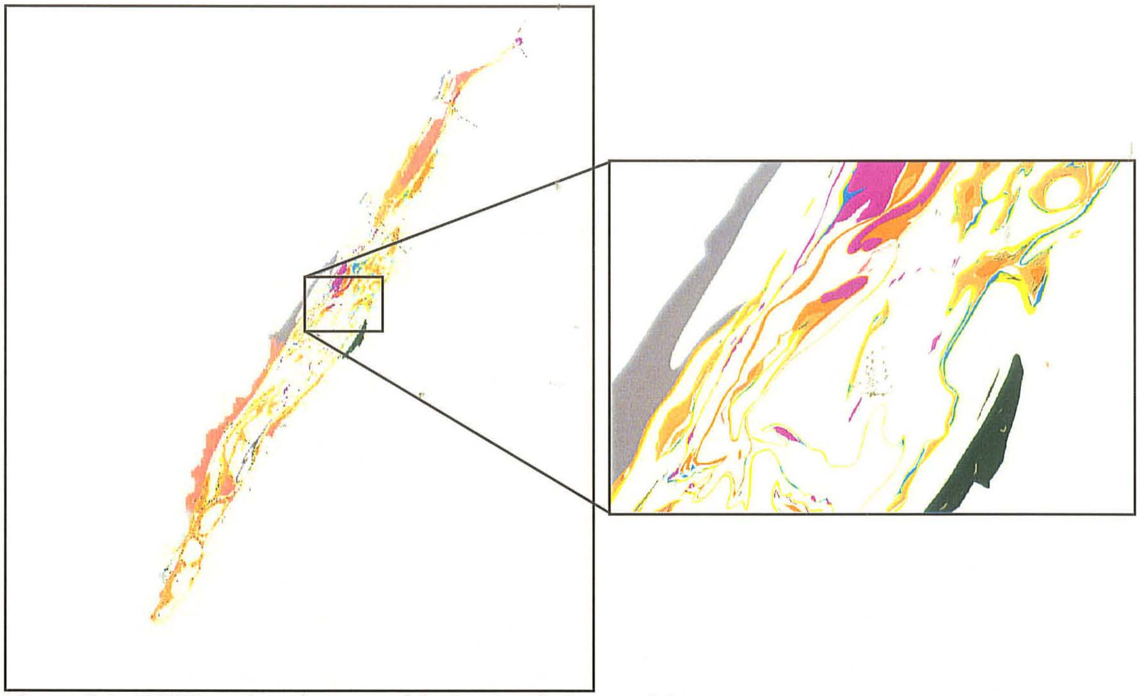


Figure 6.4: INCO geology layer used in the data fusion model.

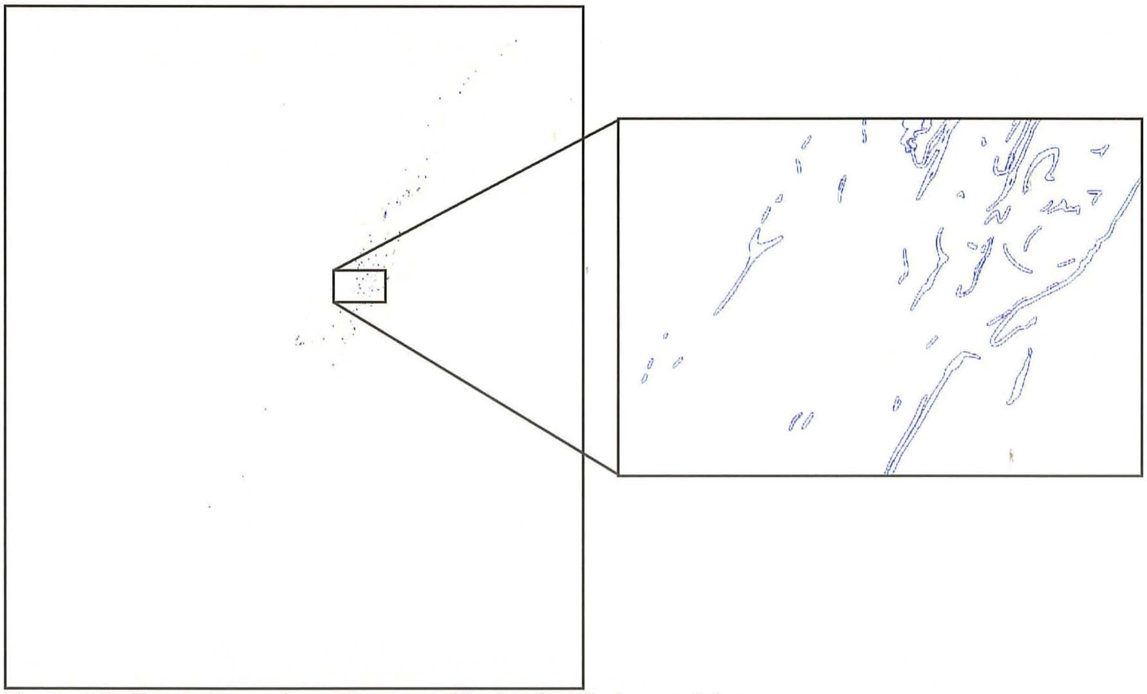


Figure 6.5: Proximity to skarns layer used in the data fusion model.

6.5.2 PROXIMITY TO SKARNS

The proximity to skarns layer is the result of a preliminary investigation of possible spatial relationships between lithologies as mapped in INCO's geology map. Geological processes occur within a continuum of scales, though the scale of events that lead to development of nickel deposits is largely unknown. By examining the current positions of specific lithologies with respect to those occuring within a buffer of 100 m, it may be possible to indirectly determine important processes. The preliminary spatial investigation revealed that skarns often occurred within 20 to 40m of mineralized bodies.

The results were used to attempt to locate other potential mineralized bodies by creating a 50m buffer zone divided into 10m intervals around all units mapped as skarns. The 20 to 40m intervals from skarns were used to select iron formations, the rational being that the proximity of skarns to iron formations suggests that a sulphide source was proximal to a hydrothermal cell. This scenario formed the basis for the Stratigraphic Model.

The iron formations within 20 to 40 m of skarns were attributed in the lithology map database and assigned higher belief values than other iron formations. The 20 to 40m intervals were saved as a new file to be used in the models to identify stratigraphic relationships. Areas with no data were reclassified to "no data". Figure 6.5 shows the proximity to skarns dataset.

6.5.3 AEROMAGNETICS

Geophysical data used in the present analysis were restricted to the data provided by the Province of Manitoba and Geological Survey of Canada on the compact disk, "Geology, Magnetic and Gravity Maps of Manitoba: A Digital Perspective" (1999). The shaded-relief high-resolution tif image of vertical gradient was used as the aeromagnetic dataset. The dataset represents a collection of surveys flown across the province over a number of years. The "data" is provided as a tif image and therefore the raw data values are not known. The digital numbers ranging from 0 to 254 represent the normalized data. The data were provided with UTM zone 14 coordinates using the NAD27 datum; the image covered the entire province. To reduce processing time, only the area of interest was used in the calculations. The tif image was saved as a grid using Spatial Analyst and the area of interest was clipped from this grid. Figure 6.6 shows the extents of the aeromagnetic dataset.

Resolution of the image was quite coarse at a pixel size of 200m. For this reason, the dataset was not appropriate for the data fusion application. The size of the pixels was much larger than features in other datasets. The problem became very apparent when the aeromagnetic data were overlain with lithology. One pixel covered many different lithologies. If the data fusion process were processed with grid cells at 200m, the results would be meaningless. Therefore, the image was resampled to 10m. All other grids were created using the parameters of the aeromagnetic data such as extents and grid cell size. The range of digital numbers was too great for weights to be assigned so the dataset was also reclassified into 10 classes of about 25 digital numbers each. After reclassification, the largest values 10, 9 and 8 were given the highest weights. Weights assigned to the remaining classes decreased linearly.

6.5.4 PROXIMITY TO FAULTS

The fault dataset was also from the “Geology, Magnetic and Gravity Maps of Manitoba: A Digital Perspective” (1999) compact disk. It was prepared as a 1:1 million scale map using UTM zone 14, NAD 27. Buffers were created around the faults at 100m intervals to 1km. Areas closest to the fault were assigned the highest belief values. Areas with no data were reclassified to “no data”. Figure 6.7 illustrates the dataset.

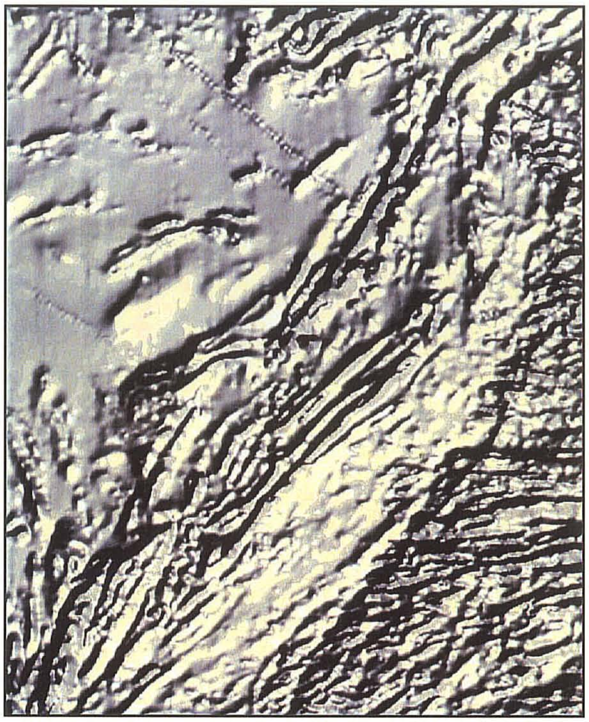


Figure 6.6: Aeromagnetics data used in the data fusion model.

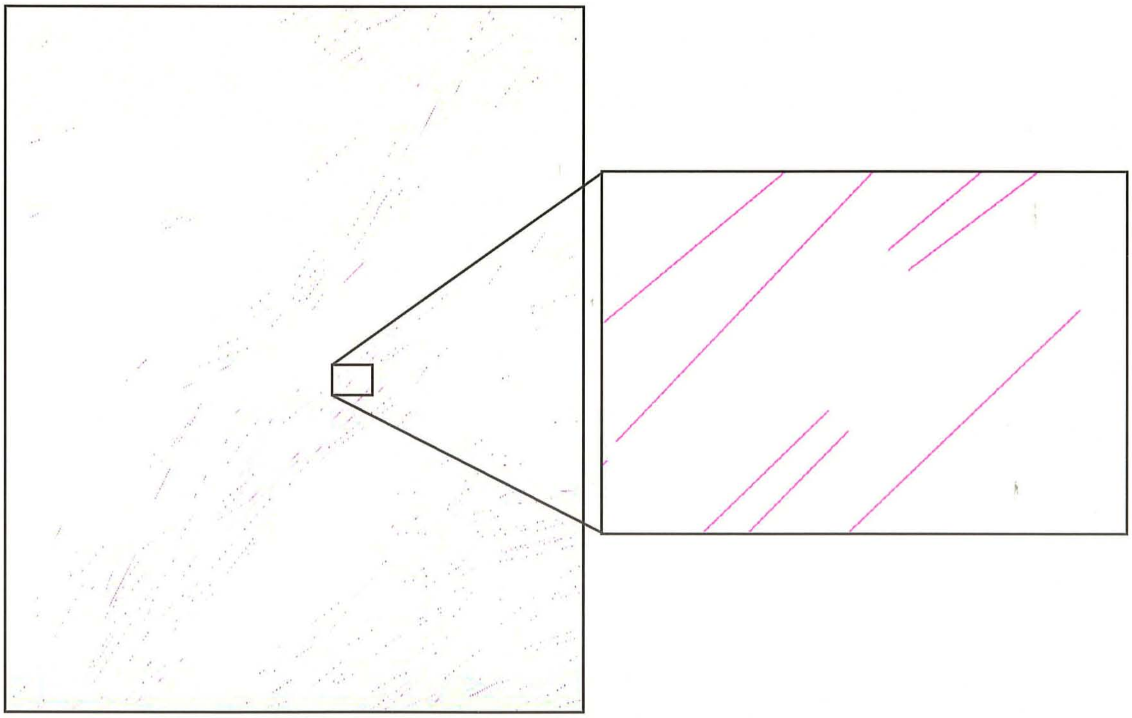


Figure 6.7: Proximity to faults layer used in the data fusion model.

6.5.5 STRIKE

The researchers from UQAM, Machado and Potrel, undertook a project to determine trends in strike and dip in the area of interest (Machado and Potrel, 2000). Working off geological maps, all measurements of strike and dip were averaged over one square kilometre. The resulting strike and dip were drawn on a new map by hand and then digitized. The dip values were manually entered and strike was calculated using ArcView. Areas with strikes around 30° were given high weights, reflecting the overall strike of the belt and consequently, the structural influence. Figure 6.8 illustrates the dataset.

6.5.6 LINEATIONS

The lineation dataset was derived from the aeromagnetic dataset. The aeromagnetic dataset displayed patterns that could be interpreted as structural features. PCI is a powerful image analysis package and the Line module was used to extract linear features from the dataset. The results were exported from PCI as a shapefile and opened in ArcView. A buffer was created around the line segments at intervals of 10m extending to 100m. Areas closest to the lineation were weighted highest, with belief values decreasing linearly. Areas with no data were reclassified to “no data”. Figure 6.9 illustrates the dataset.

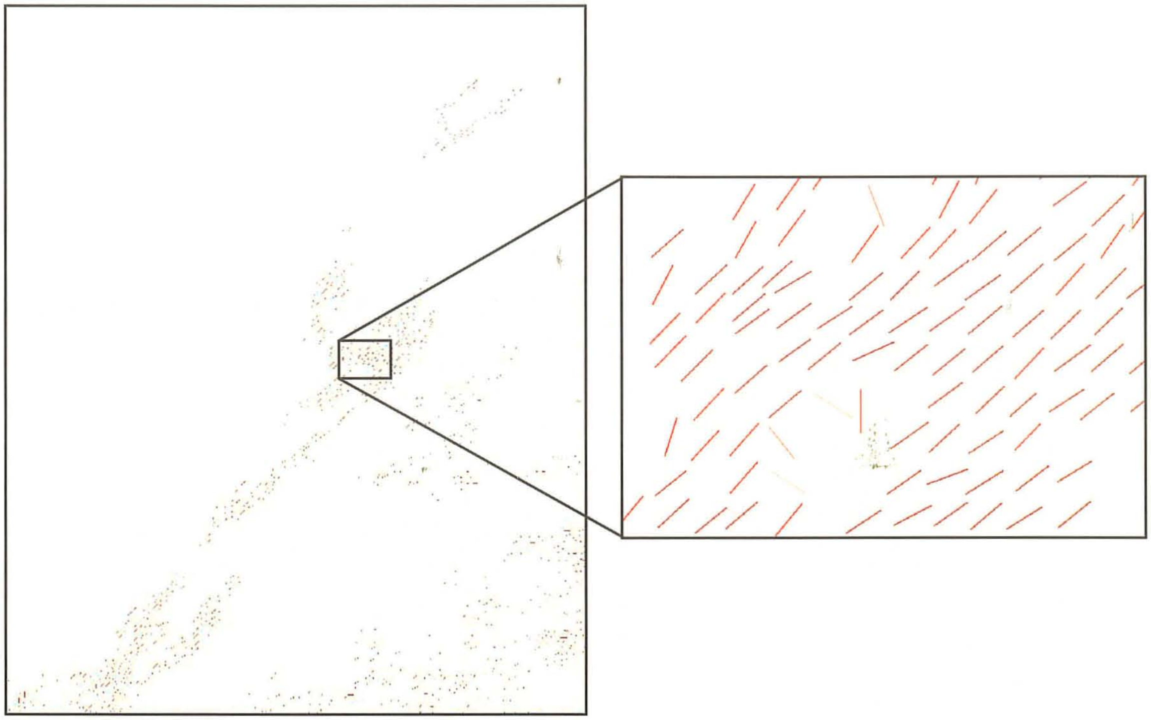


Figure 6.8: Strike layer used in the data fusion model.

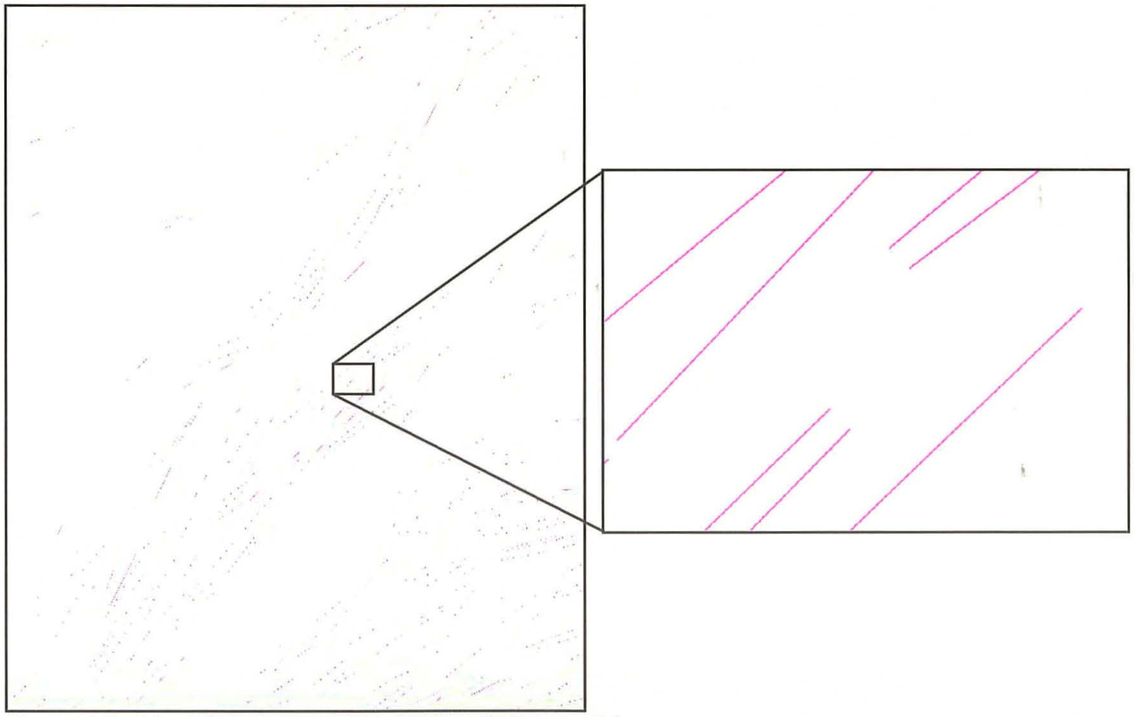


Figure 6.9: Lineation layer used in the data fusion model.

The weights assigned to each layer for each model are graphed in Figures 6.10 to 6.13. The weights assigned to data input into the DSBT calculations are shown in Figures 6.10 and 6.11. The pink line graphs the support values assigned to each attribute and the blue line graphs the plausibility value. Figures 6.12 and 6.13 show the weights assigned to attributes used in the Fuzzy Logic method. The blue line indicates support.

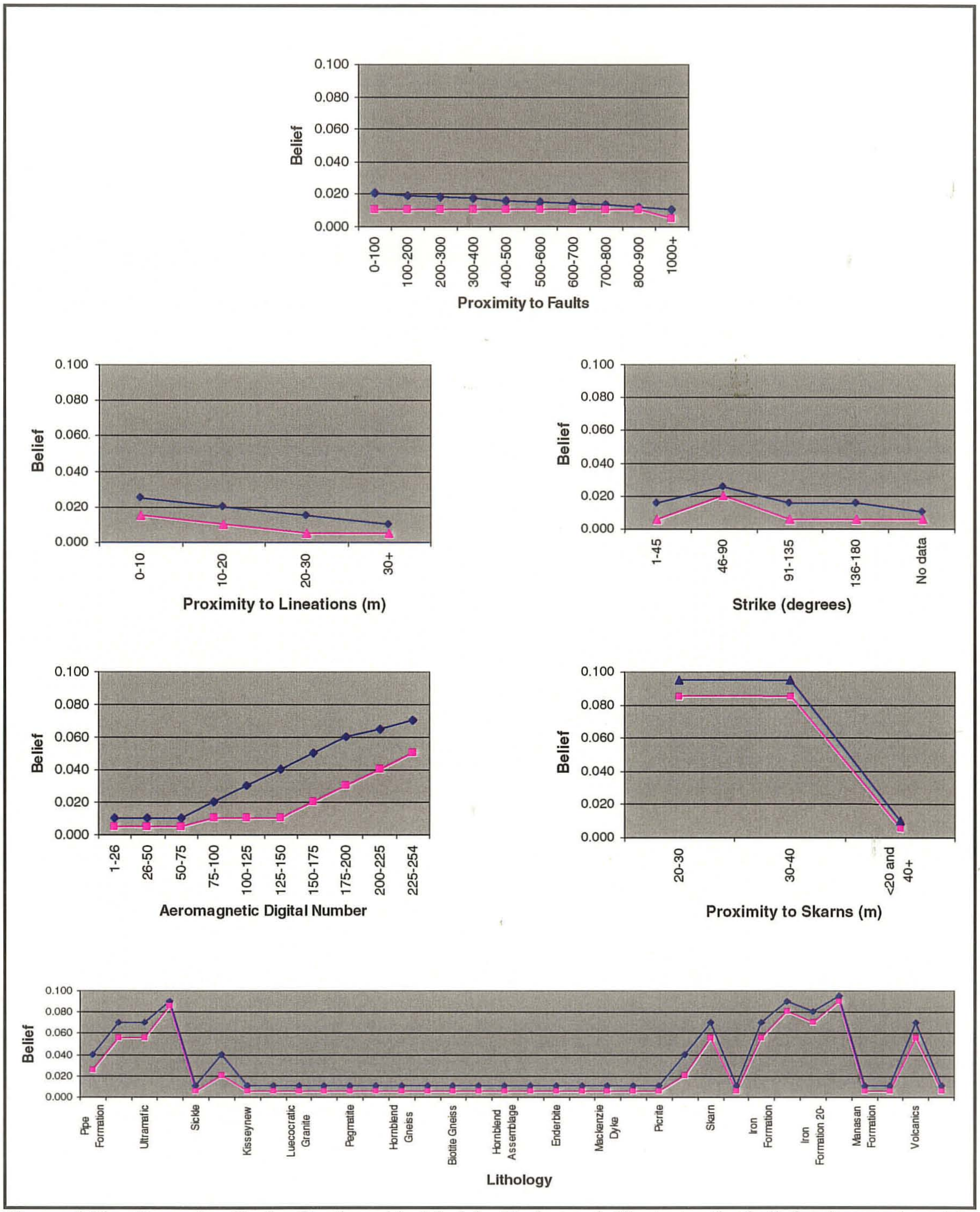


Figure 6.10: Dempster Shafer Stratigraphic Model. Each graph illustrates the belief values assigned to each attribute. According to the DSBT Stratigraphic Model, high belief is assigned to areas 20 to 40 m from skarns and Iron Formation proximal to the Pipe Formation; medium belief to aeromagnetics; low belief to faults, lineaments and strike. Values along the y-axis range from 0 to 0.1. Values along the x-axis represent the attributes. The plausibility function is shown is blue and the belief function is shown in pink.

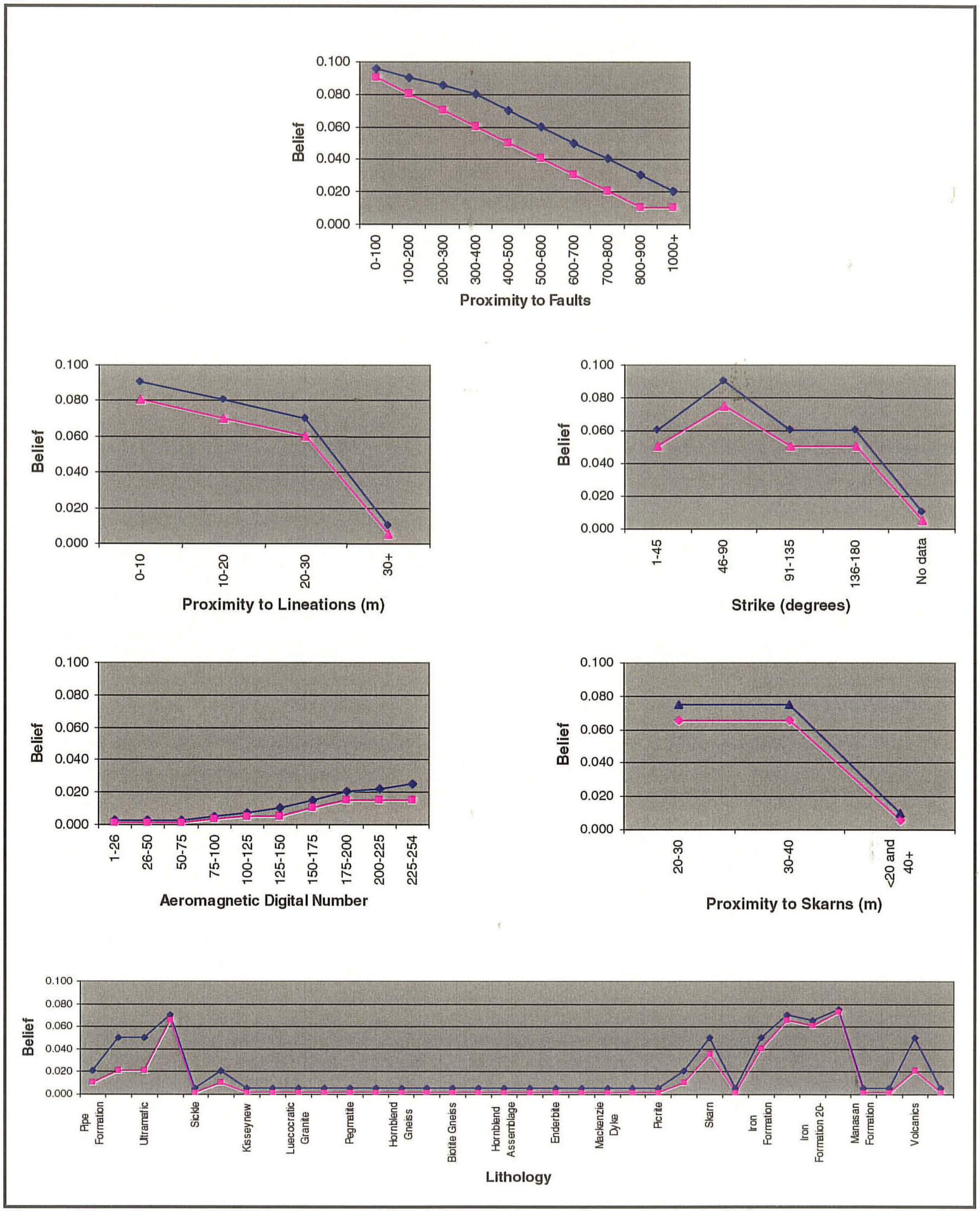


Figure 6.11: Dempster Shafer Structural Model. Each graph illustrates the belief values assigned to each attribute. According to the DSBT Structural Model, high belief is assigned to faults, lineaments and strike; medium belief to areas 20 to 40 m from skarns and Iron Formation proximal to the Pipe Formation; and low belief to aeromagnetics. Values along the y-axis range from 0 to 0.1. Values along the x-axis represent the attributes. The plausibility function is shown is blue and the belief function is shown in pink.

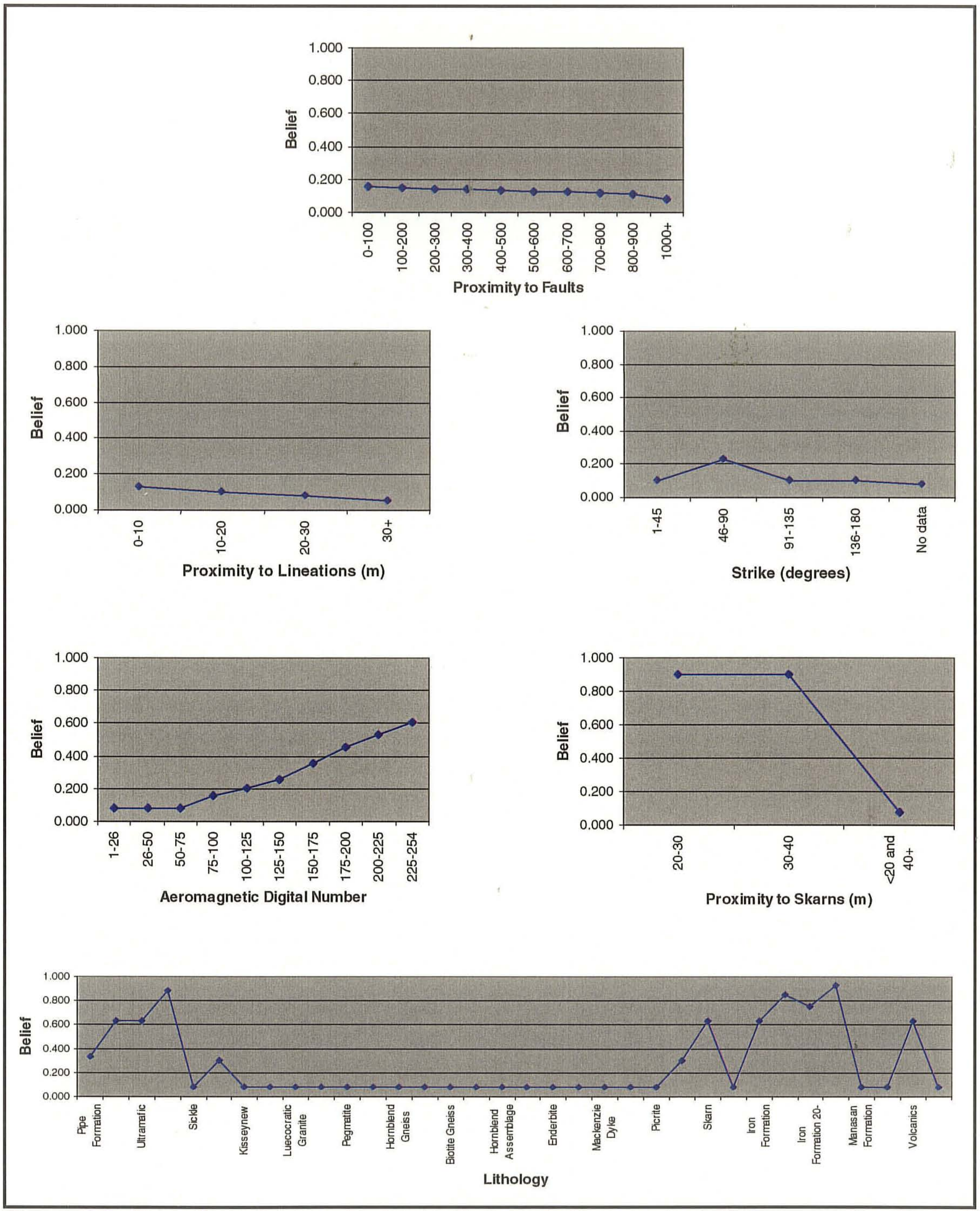


Figure 6.12: Fuzzy Logic Stratigraphic Model. Each graph illustrates the belief values assigned to each attribute. According to the FL Stratigraphic Model, high belief is assigned to areas 20 to 40 m from skarns and Iron Formation proximal to the Pipe Formation; medium belief to aeromagnetics; and low belief to faults, lineaments and strike. Values along the y-axis range from 0 to 1. Values along the x-axis represent the attributes. The belief function is shown in blue.

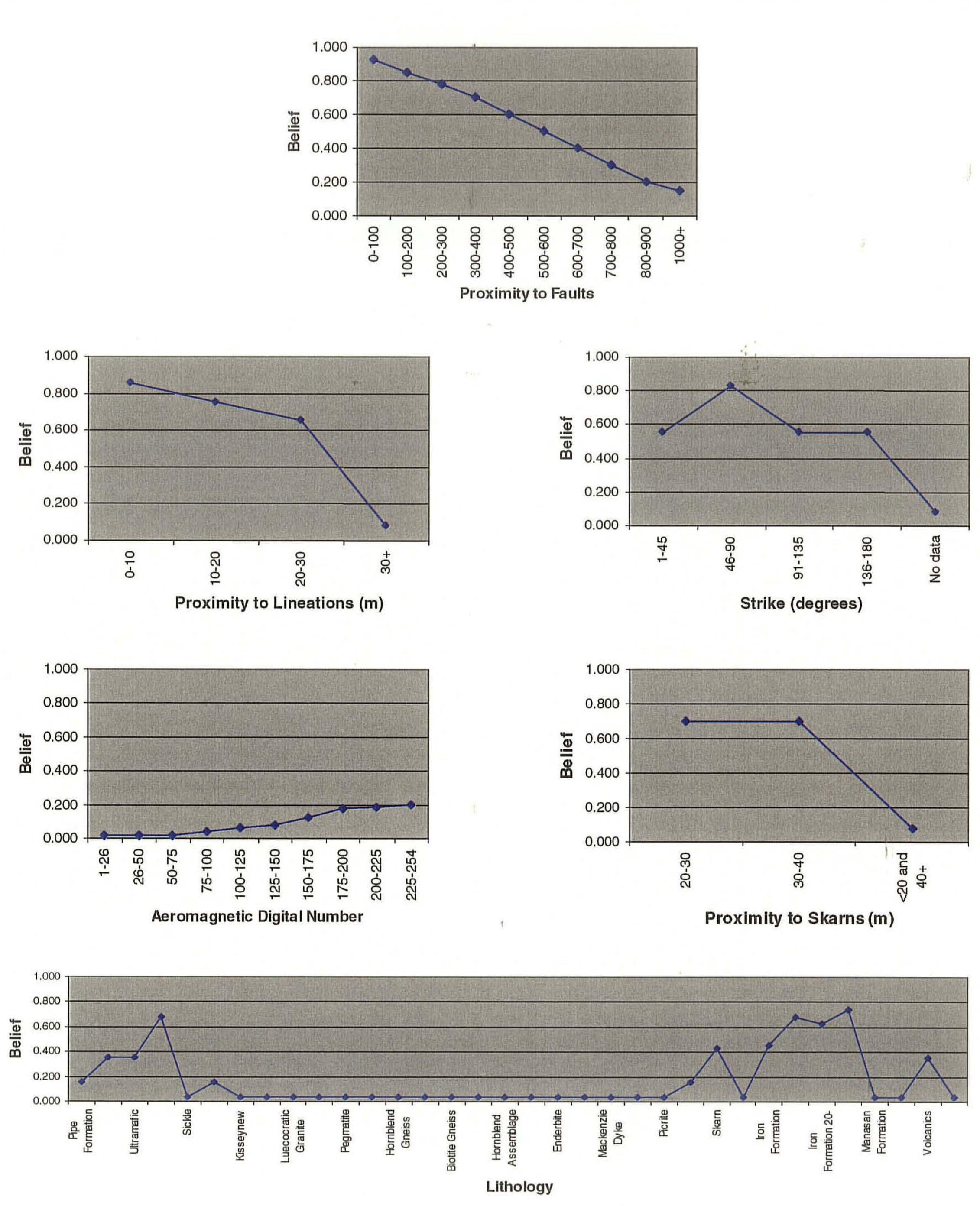


Figure 6.13: Fuzzy Logic Structural Model. Each graph illustrates the belief values assigned to each attribute. According to the FL Structural Model, high belief is assigned to faults, lineaments and strike; medium belief to areas 20 to 40 m from skarns and Iron Formation proximal to the Pipe Formation; and low belief to aeromagnetics. Values along the y-axis range from 0 to 1. Values along the x-axis represent the attributes. The belief function is shown in blue.

6.6 RESULTS

All the processing was done using ArcView 3.2 and Spatial Analyst. Four sets of mineral potential maps resulted from this process: DSBT Structural Model, DSBT Stratigraphic Model, Fuzzy Logic Structural Model and Fuzzy Logic Stratigraphic Model. Results from all four models will be presented with areas within 200m of known deposits considered first, followed by other areas of high potential.

The data fusion process produced maps displaying areas that have the potential to host nickel deposits. The potential ranges from low to high. Areas of high potential indicate locations that have a high potential of hosting a nickel deposit. High potential is determined by classifying the range of results from the support map into 10 classes of equal intervals. Classes 1 to 3 are considered low potential whereas classes 4 to 10 are considered high potential. The total area of high potential is very small relative to the area of low potential therefore seven classes are needed to visualize areas of high potential. Interpreting the results is easier when only areas of high potential are displayed.

This approach will also help determine if the data fusion methods and mineral deposit models can predict areas of known nickel deposits. Figure 6.14 shows the area of interest and known nickel deposits. The known nickel deposits are either active or previously active mines. Each area is identified by a number that will identify the known deposit in each result map. The figures in Section 6.6.1 will compare the results of each model with the location of known nickel deposits.

6.6.1 KNOWN DEPOSITS

All known deposit locations are based on the surface location of the corresponding mine shaft. The nickel deposit typically exists near the mine shaft and usually at some depth. Additionally, the geophysics data were provided with 200m pixels. The coarseness of this data limits the accuracy of the results. Therefore, areas of high potential that occur within 200m of a known deposit might be considered meaningful at this scale. As the location of known deposits is indicated by a point, it is necessary to assign a reasonable area proximal to the point to verify the ability of the models to predict known deposits.

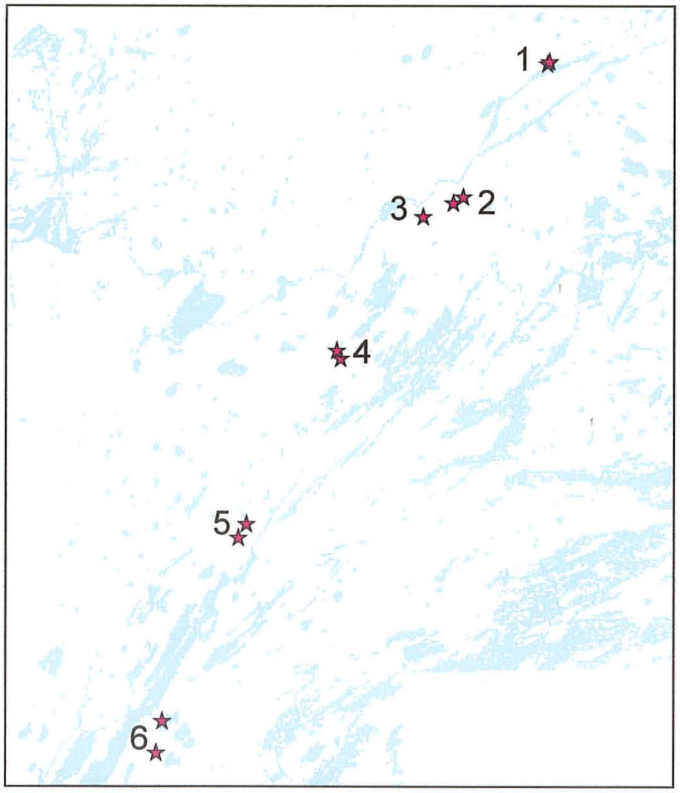


Figure 6.14: Locations of known deposits used to validate results from the Dempster-Shafer Belief Theory and Fuzzy Logic methods of data fusion. 1 – Moak Lake Mines; 2 - Thompson Mines; 3 – Birchtree Mine; 4 – Pipe Pit Mines; 5 – Soab Mines; 6 – Bowden and Bucko Mines

6.6.2 MOAK LAKE MINES – LOCATION 1

All four models produce areas of high potential near the Moak Lake Mines but beyond the 200m buffer. Figure 6.15 shows the results for the structural models produced using Fuzzy Logic and the DSBT. The one area of high potential corresponds with areas proximal to lineaments and ultramafics. Areas of high potential are shown in black.

The stratigraphic models produce more areas indicating high potential for a nickel deposit than the structural models as shown in Figure 6.16. These areas appear because of the high weights assigned to ultramafics within 20 to 40 metres from skarns and areas proximal to lineaments. Areas of high potential do not fall within 200m of the mines. However, the Moak Lake region would be flagged as having high potential by all four models.

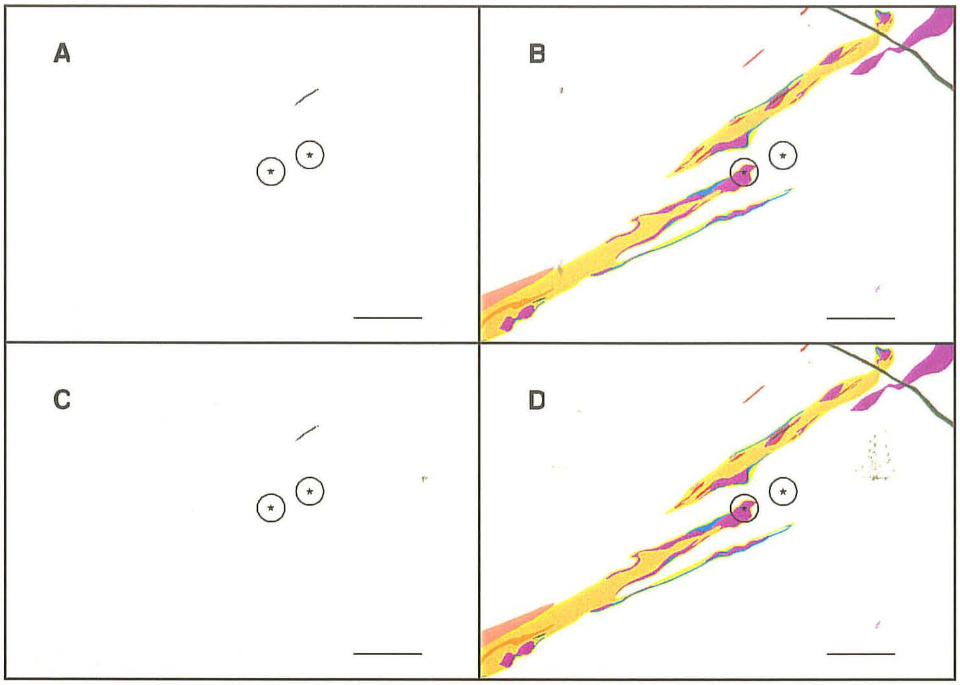


Figure 6.15: Moak Lake Structural Model – Panel A shows areas of high potential produced by the Fuzzy Logic Method. Panel C shows areas of high potential produced by the DSBT method. Panels B and D show the lithology map for reference. The black line in the bottom right corner is 1 km long. The black stars represent the surface location of the mines. The circle around the stars marks 200m. Areas of high potential are shown in black. Areas of high potential exist beyond 200m in both panels A and C.

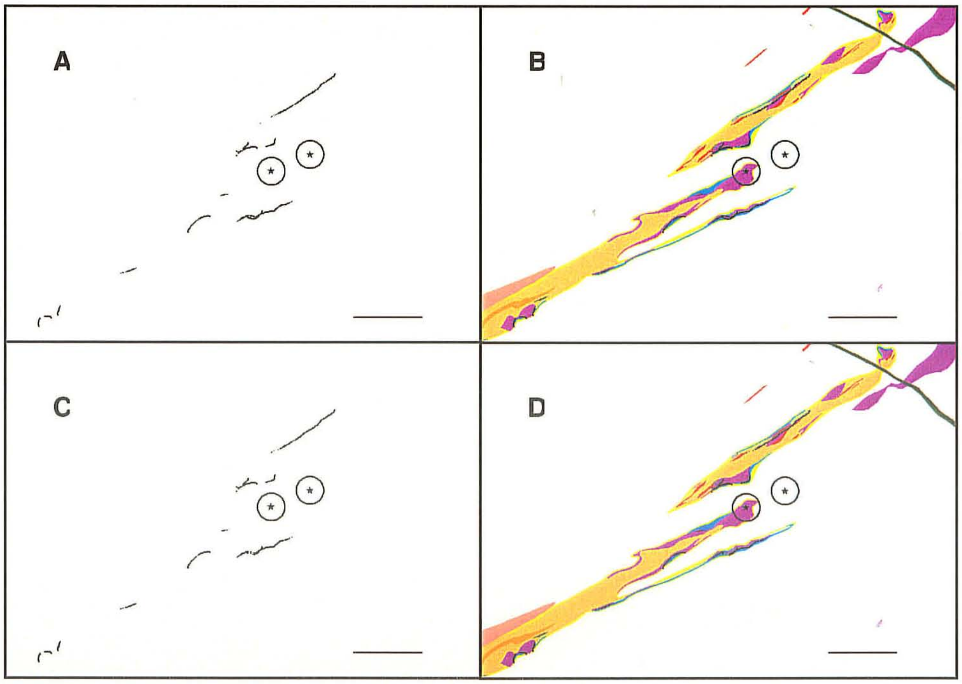


Figure 6.16: Moak Lake Stratigraphic Model – Areas of high potential exist beyond 200m in both panels A and C.

6.6.3 THOMPSON MINES – LOCATION 2

The T1 and T3 mines are believed to be controlled by structural features. T1 is located southwest of T3. Results from all models indicate high potential beyond 200m of the T1 mine. Figure 6.17 shows the results from the structural models. Figure 6.18 shows areas of high potential from the stratigraphic models.

The areas of high potential for all models are a result of assigning high weights to areas within 20 to 40 m of skarns and areas proximal to faults. None of the models are able to detect areas of high potential near T3.

6.6.4 BIRCHTREE MINE – LOCATION 3

The structural models do not return any areas of high potential as shown in Figure 6.19. The stratigraphic models do detect areas of high potential near Birchtree Mine, but the areas fall well beyond the 200m radius. Figure 6.20 show the results.

Examining the lithology map and the location of the Birchtree Mine in Figure 6.20 reveals that there is no lithologic information to the west of Birchtree Mine. Perhaps more information about the lithology would produce more areas of high potential.

6.6.5 PIPE PIT MINES – LOCATION 4

The stratigraphic models for both the Fuzzy Logic and DSBT methods produce areas of high potential well beyond 200m of Pipe Pit 2, shown in panels A and C of Figure 6.21. However, panel A also shows that the Fuzzy Logic model produced a small area of high potential between the two mines. The structural models did not predict any areas of high potential and are not included.

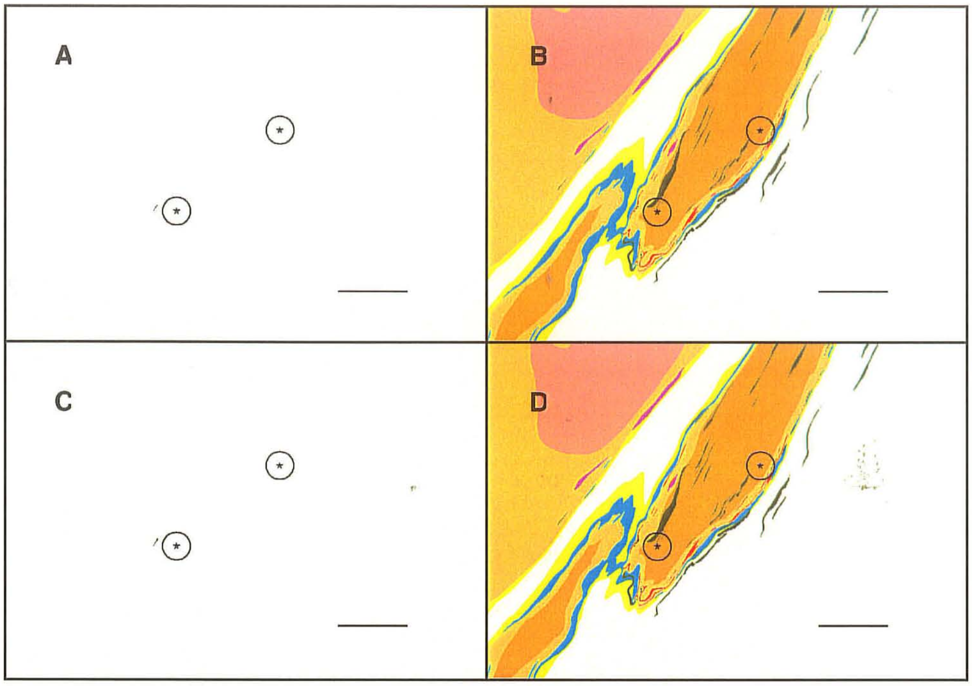


Figure 6.17: Thompson Mines Structural Model - Areas of high potential exist beyond 200m of T1 in both panels A and C.

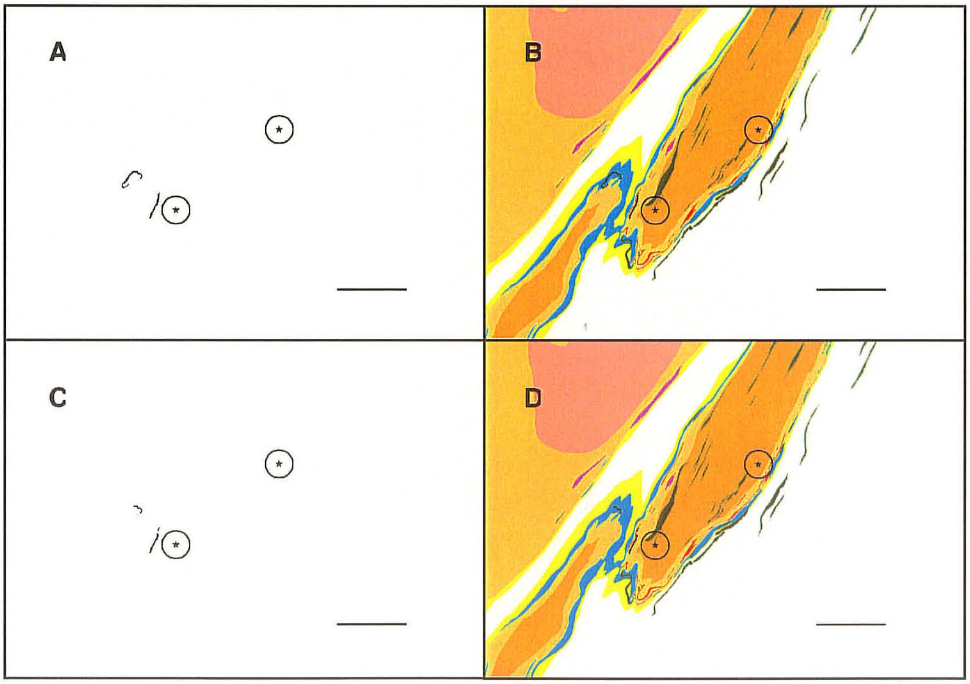


Figure 6.18: Thompson Mines Stratigraphic Model - Areas of high potential exist beyond 200m of T1 in both panels A and C.

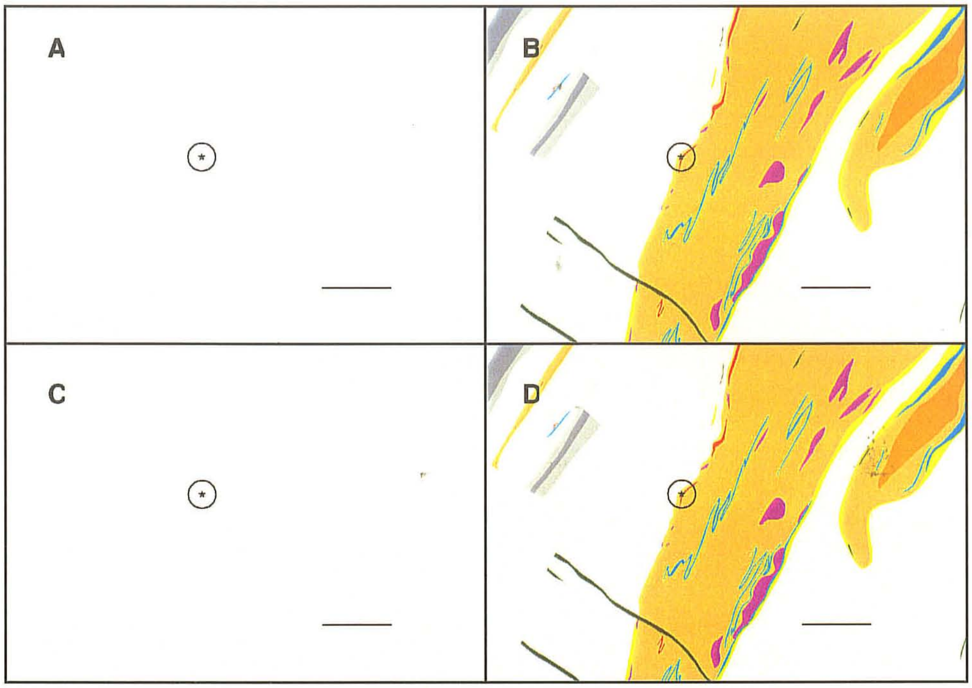


Figure 6.19: Birchtree Mine Structural Model – The structural models did not produce any areas of high potential near Birchtree Mine.

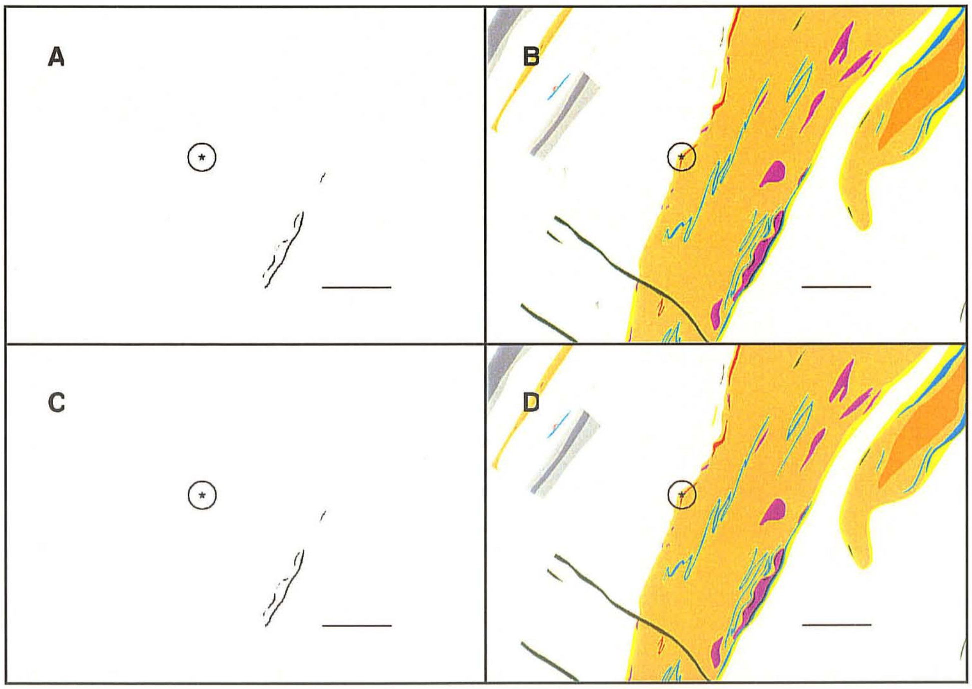


Figure 6.20: Birchtree Mine Stratigraphy Model - Areas of high potential exist well beyond 200m in both panels A and C.

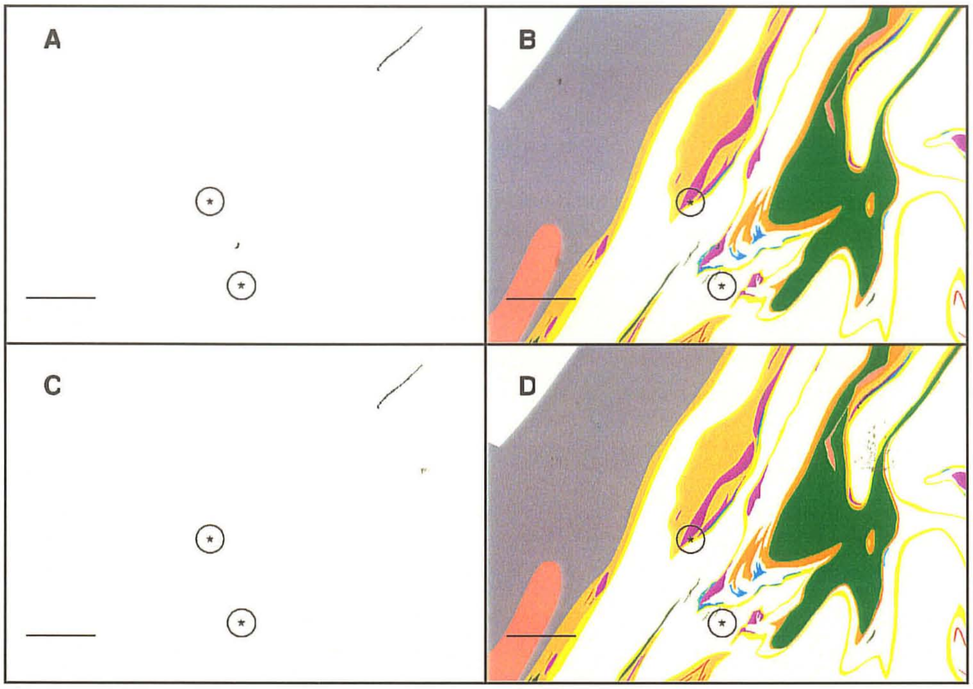


Figure 6.21: Pipe Pit Mines Stratigraphic Model - An area of high potential exists well beyond 200m of both mines. A small area of high potential is detected by the Fuzzy Logic method, shown in panel A.

6.6.6 SOAB LAKE MINES – AREA 5

The structural models present areas of high potential near the Soab Lake Mines, but beyond 200m. Figure 6.22 shows the results. The stratigraphic models do not produce any results near Soab Lake Mines and are therefore not included.

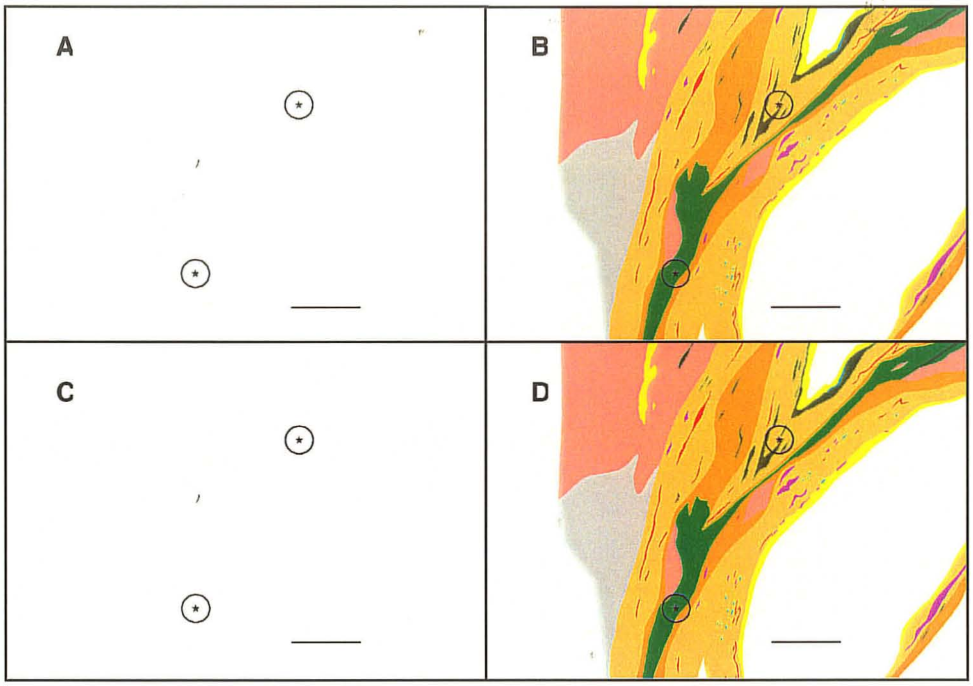


Figure 6.22: Soab Lake Mines Structural Model – An area of high potential exists beyond 200m in both panels A and C.

6.7 AREAS OF HIGH POTENTIAL

Other areas of high potential were predicted by the models. The regions are identified by lakes close to the area of high potential as there are few other distinguishing landforms. Figure 6.23 shows the approximate locations of these other areas. The results will be compared to locations of drill hole collars and hand sample locations taken from the Geochemistry database to determine if the companies have been active in these areas.

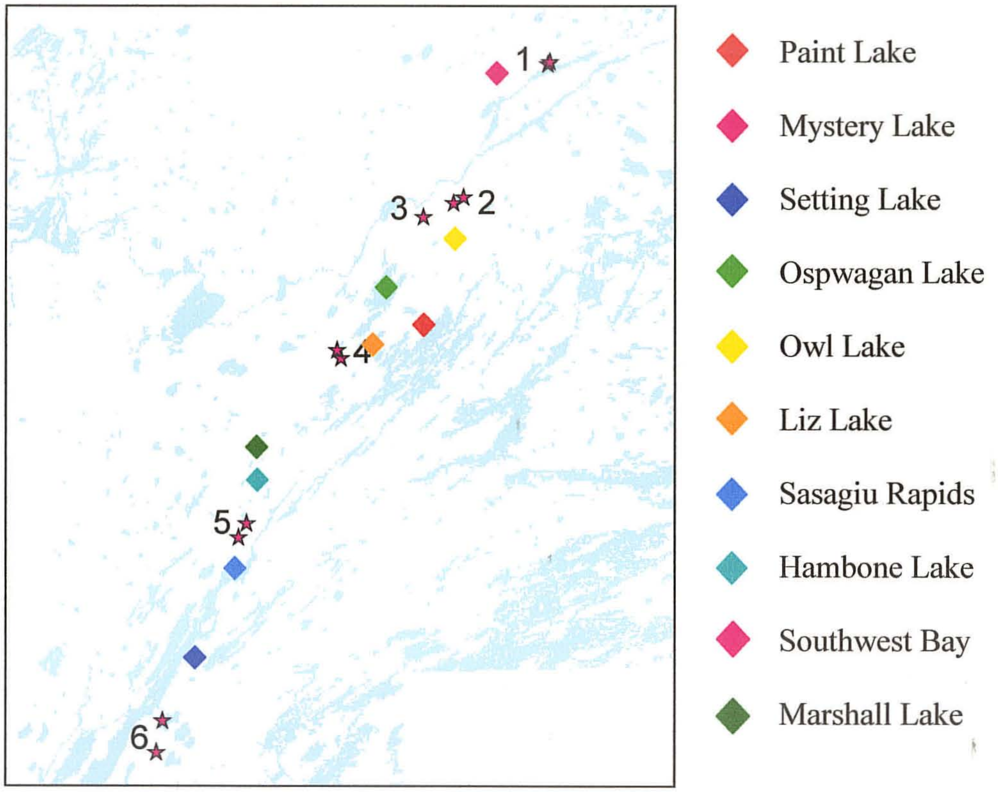


Figure 6.23: Other areas of high potential predicted by the Fuzzy Logic and DSBT methods.

6.7.1 Paint Lake

The structural models indicate areas of high potential south of Nichols Lake, shown in Figure 6.24. A smaller area of high potential also occurs just west of Paint Lake. The stratigraphic models produce more areas of high potential as a result of the high weighting given to skarns proximal to ultramafic rocks, shown in Figure 6.25. There were no drill hole collars or hand sample locations in the areas shown in Figures 6.24 and 6.25.

6.7.2 Mystery Lake

A small areas of high potential is predicted by the Structural Model on the east shore of Mystery Lake, shown in Figure 6.26. The area corresponds to an iron formation within an ultramafic body on the geology map. The Stratigraphic Model predicts two areas of high potential, one to the west of Mystery Lake, the other on the east shore of the lake. Both areas correspond to iron formations that are proximal to ultramafic rocks.

Panel D shows the surface location of drillhole collars as red triangles. The area of high potential that occurs on the east shore of Mystery Lake is just south and along strike of the sampling area. The larger area to the west of the lake is approximately 400m from another drillhole. The sponsors did identify the area as an exploration target as indicated by the locations of the drillholes.

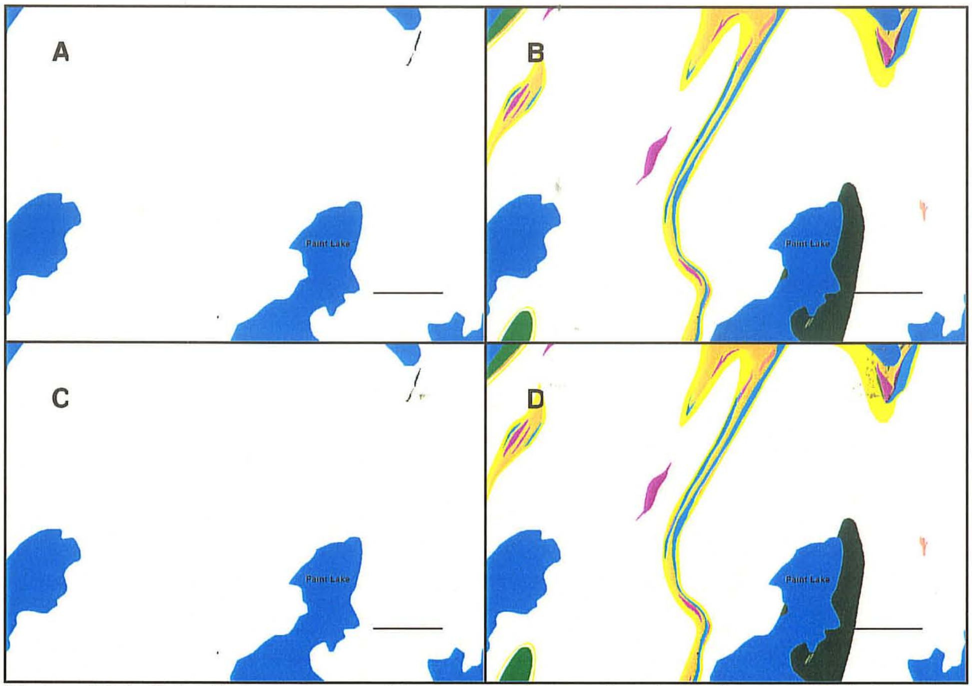


Figure 6.24: Paint Lake - Areas of high potential predicted the Structural Model. A small area west of Paint Lake is shown in black. The area is proximal to a skarn on the geology map. A larger area south of Nichols Lake occurs between an ultramafic body and a skarn.

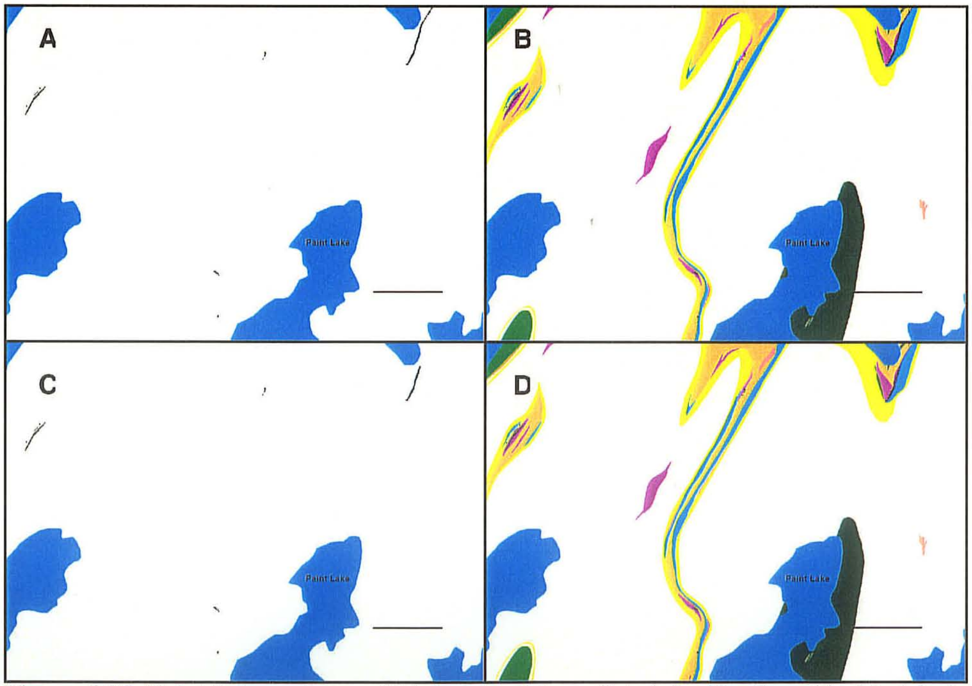


Figure 6.25: Paint Lake - Areas of high potential predicted the Stratigraphic Model. Areas west of Paint Lake are proximal to skarns and ultramafics on the geology map. Areas within and south of Nichols Lake occur between an ultramafic body and a skarn.

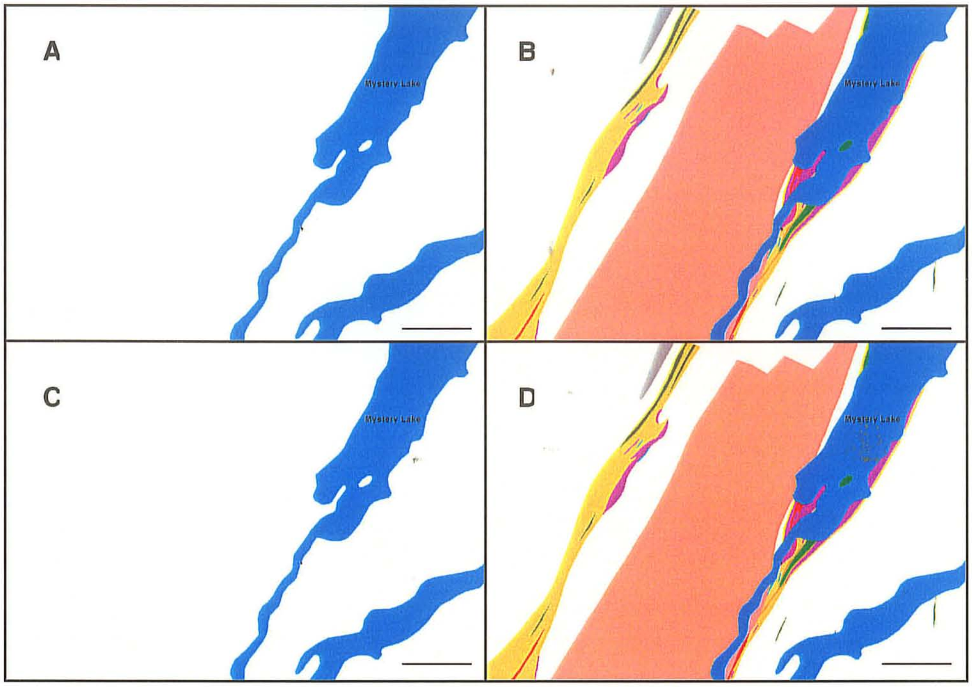


Figure 6.26: Mystery Lake - Areas of high potential predicted by the Structural Model. A very small area on the east shore of Mystery corresponds to an iron formation within an ultramafic body on the geology map.

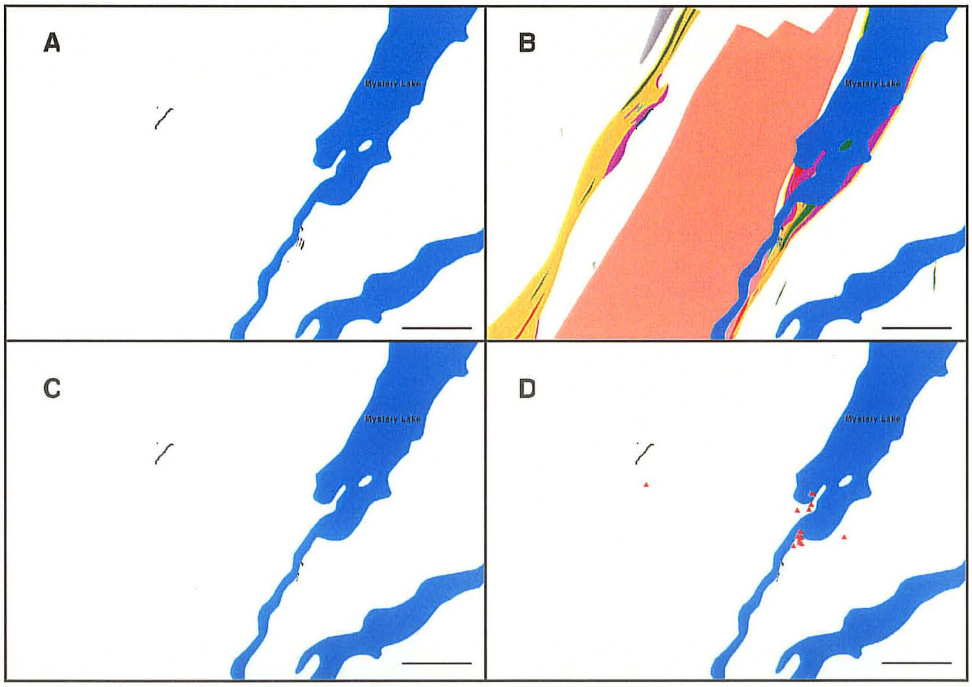


Figure 6.27: Mystery Lake - Areas of high potential predicted by the Stratigraphic Model. Both areas of high potential correspond to an iron formation within an ultramafic body on the geology map. The red triangles in Panel D show the surface locations of drillhole collars.

6.7.3 Setting Lake

Only the Structural models produce areas of high potential in the Setting Lake area. The areas result from weights assigned to structural features as the lithology map does not extend into this area. Also, no drillhole or hand sample locations occur within the view. Figure 6.28 shows the area of high potential with results from the Fuzzy Logic method in panel A and the DSBT method in panel B.

6.7.4 Ospwagan Lake

The Stratigraphic Model produces areas of high potential for both data fusion methods. Figure 6.29 indicates the many areas of high potential occur proximal to ultramafics and skarns. A drillhole collar does occur near one area of high potential in the eastern part of the figure of panel D.

6.7.5 Owl Lake

Figure 6.30 shows areas of high potential predicted by the Stratigraphic Model in the Owl Lake region. The areas in the west are proximal to skarns and ultramafics. Areas in the east are proximal to iron formations. Drillhole collars and hand sample locations do not occur in the area around Owl Lake.

6.7.6 Liz Lake

The Stratigraphic Model predicts areas of high potential to the west of Liz Lake. These occur proximal to skarns and ultramafics, shown in Figure 6.31. A drillhole collar is located in the northeast area of the panel D, but does not fall within a reasonable distance of any areas of high potential.

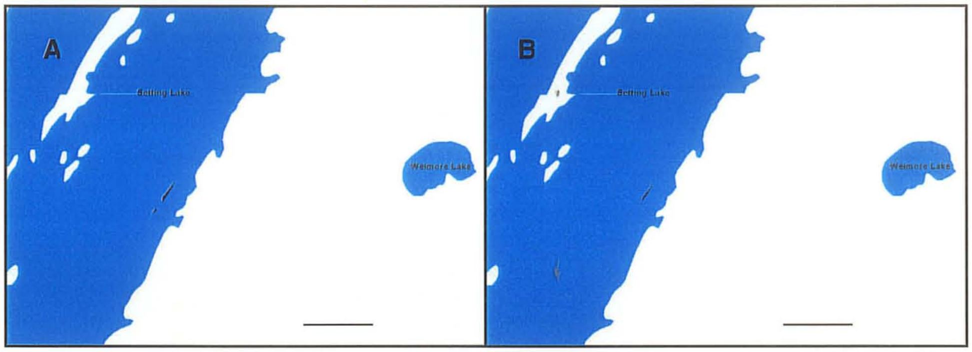


Figure 6.28: Setting Lake - Areas of high potential predicted by the Structural Model. A small area lying near the east shore of Setting Lake corresponds to structural features. Panel A displays results from the Fuzzy Logic method and panel B shows results from the DSBT method.

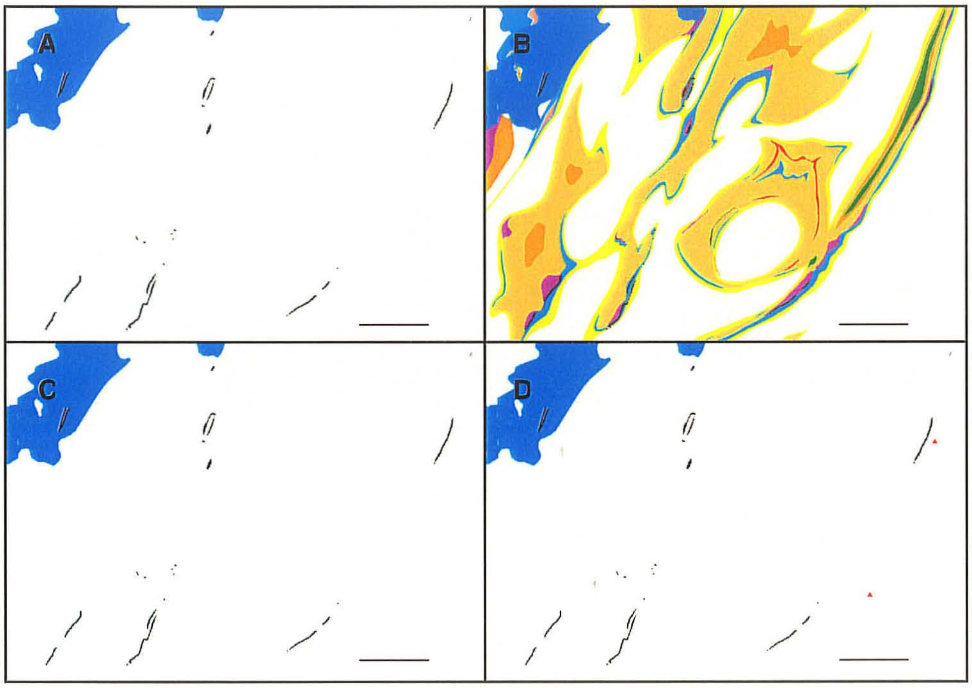


Figure 6.29: Ospwagan Lake - Areas of high potential predicted by the Stratigraphic Model. All the areas of high potential occur proximal to ultramafics and skarns. Panel D indicates a drillhole collar within 200m of the eastern most area of high potential.

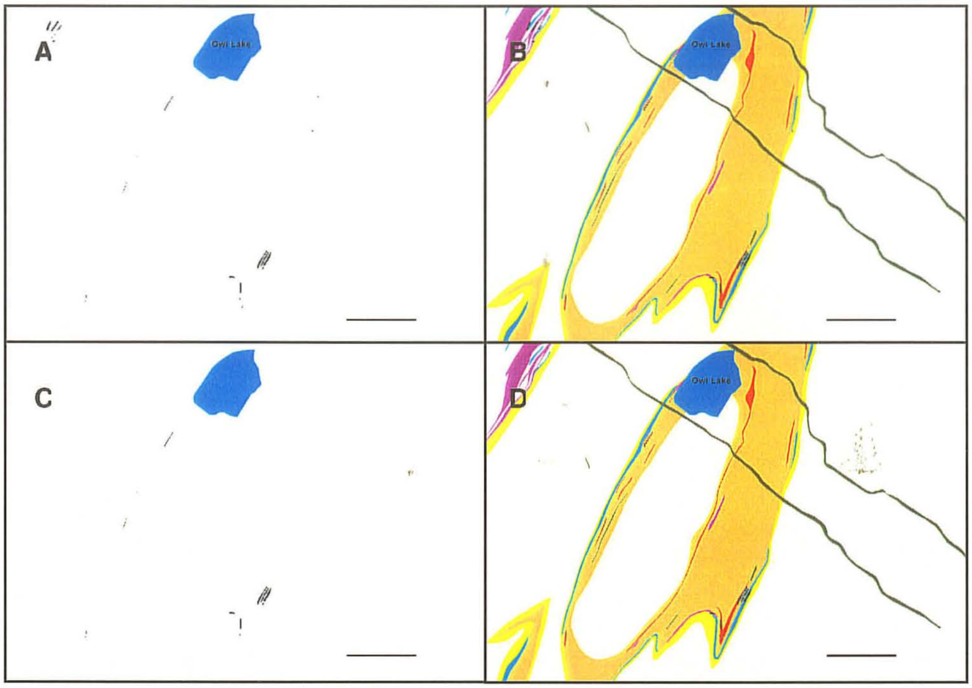


Figure 6.30: Owl Lake - Areas of high potential predicted by the Stratigraphic Model. All the areas of high potential in the west occur proximal to ultramafics and skarns. Areas of high potential in the east occur proximal to iron formations.

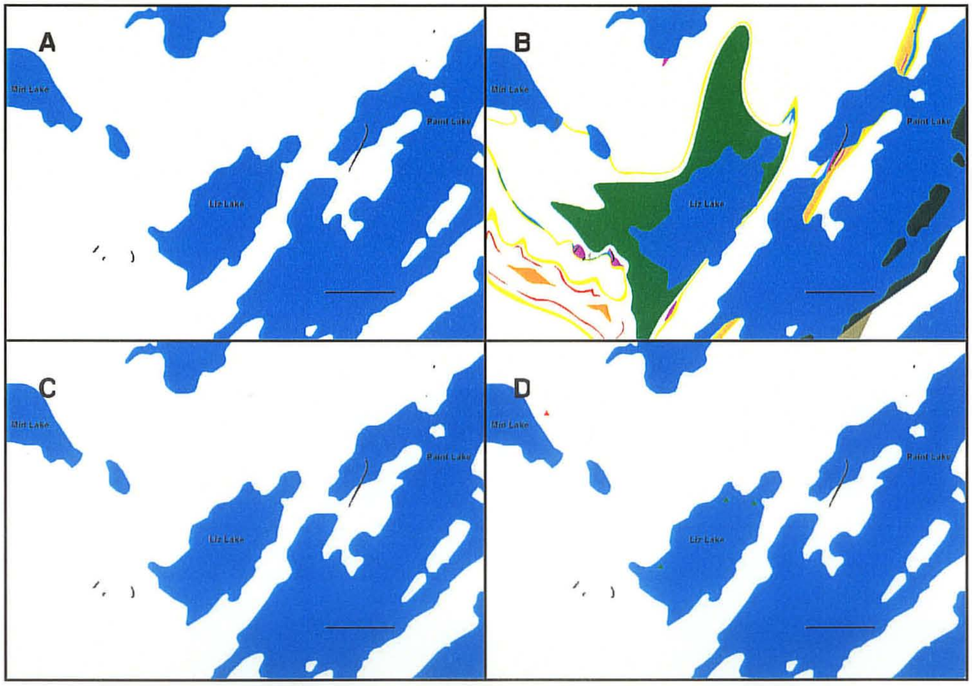


Figure 6.31: Liz Lake - Areas of high potential predicted by the Stratigraphic Model. All the areas of high potential occur proximal to ultramafics and skarns. The drillhole collar in shown in panel D falls well beyond the areas of high potential.

6.7.7 Sasagiu Rapids

Figure 6.32 displays areas of high potential predicted by the Stratigraphic Model for Sasagiu Rapids. These areas occur proximal to skarns. One drillhole collar location occurs in the northwest corner of panel D, but not within areas of high potential.

6.7.8 Hambone Lake

Figure 6.33 displays areas of high potential predicted by the Stratigraphic Model for the Hambone Lake area. These areas occur proximal to ultramafics. One drillhole collar falls within 500m of an area of high potential.

6.7.9 Southwest Bay

Figure 6.34 displays areas of high potential predicted by the Stratigraphic Model for Southwest Bay. All the areas of high potential occur proximal to ultramafics. One drillhole collar shown south central in panel D falls within 500m of an area of high potential.

6.7.10 Marshall Lake

Figure 6.35 displays areas of high potential predicted by the Stratigraphic Model for the Marshall Lake area. The area of high potential corresponds to an ultramafic body on the geology map.

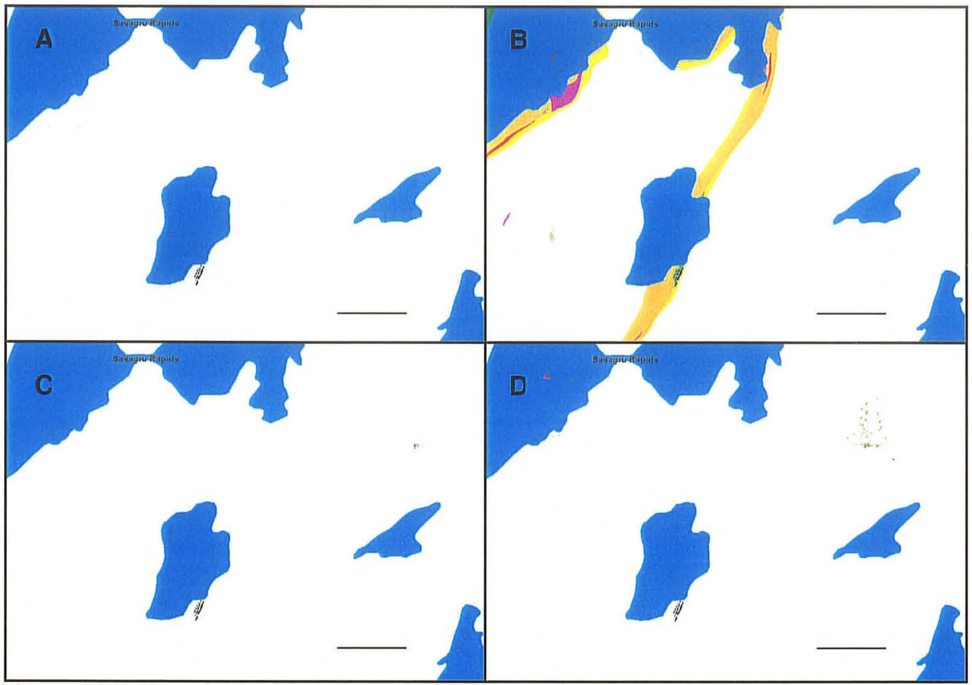


Figure 6.32: Sasagiu Rapids – Areas of high potential predicted by the Stratigraphic Model. All the areas of high potential occur proximal to ultramafics and skarns. The drillhole collar in shown in panel D falls well beyond the areas of high potential.

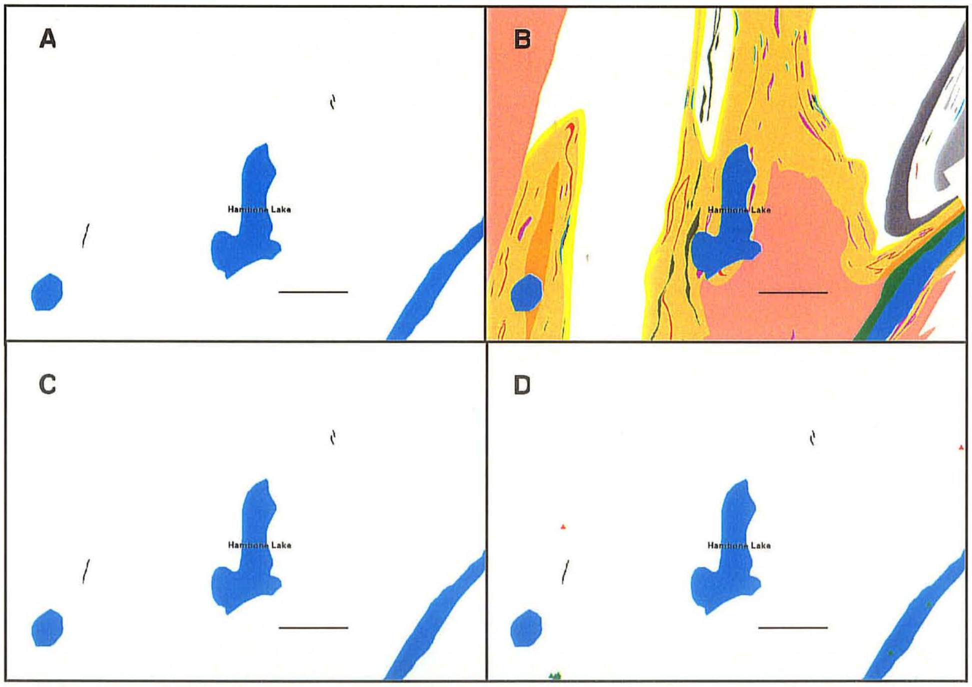


Figure 6.33: Hambone Lake – Areas of high potential predicted by the Stratigraphic Model. All the areas of high potential occur proximal to ultramafics. One drillhole collar in shown in panel D falls within 500m of an area of high potential. The green triangles in the southwest corner mark hand sample locations.

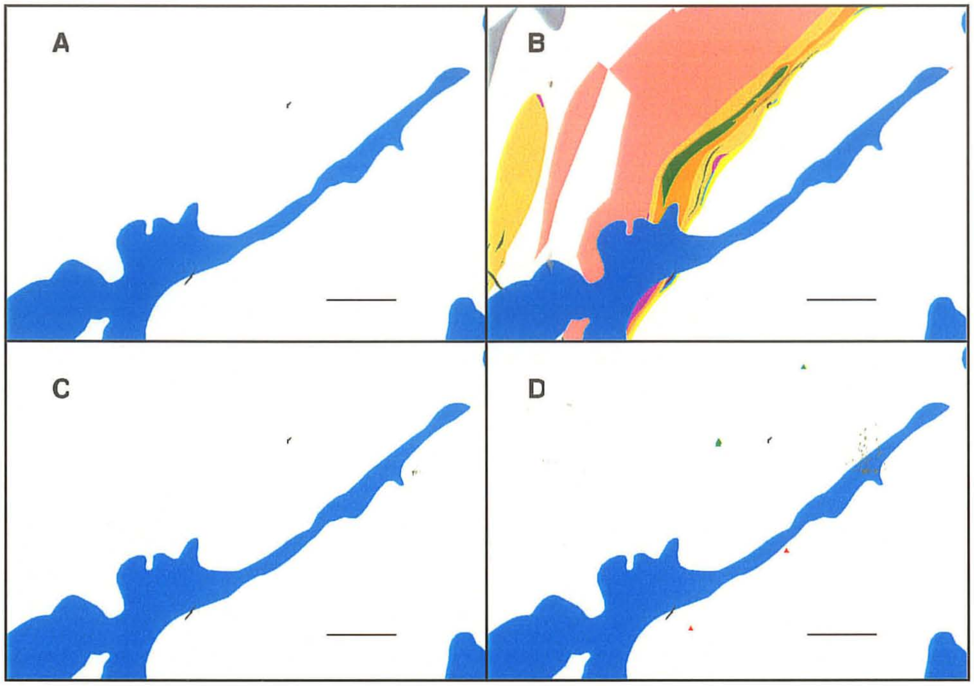


Figure 6.34: Southwest Bay – Areas of high potential predicted by the Stratigraphic Model. All the areas of high potential occur proximal to ultramafics. One drillhole collar in shown south central in panel D falls within 500m of an area of high potential. The green triangles in the north half of the panel mark hand sample locations.

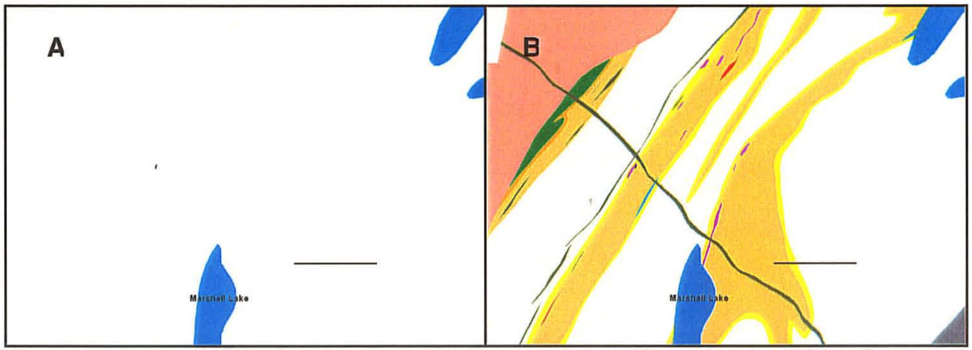


Figure 6.35: Marshall Lake – Areas of high potential predicted by the DSBT Structural Model. A small area north west of Marshall Lake corresponds to an ultramafic body on the geology map.

6.8 MODEL SUMMARIES

6.8.1 STRUCTURAL MODEL SUMMARY

The DSBT Structural Model identified four regions with high potential for hosting a nickel deposit. The largest areas are found near Paint Lake and Setting Lake, shown in Figures 6.24 and 6.28. Mystery Lake and Marshall Lake also host areas of high potential as illustrated by Figures 6.26 and 6.35.

The Fuzzy Logic Structural Model identified areas of high potential for nickel deposits near Paint Lake, Mystery Lake and Setting Lake. Figures 6.24, 6.26 and 6.28 show the results. The overall area predicted by the structural models was relatively smaller than the areas predicted by the stratigraphic model. This is presumed to be the effect of sparse structural data.

6.8.2 STRATIGRAPHY MODEL SUMMARY

The DSBT Stratigraphy Model predicts areas of high potential in many locations. The area west of Ospwagan Lake and north of Mystery Lake hosts a relatively large collection of areas with high potential shown in Figures 6.29 and 6.27. Areas near Paint Lake and Owl Lake also indicate clusters of high potential illustrated in Figures 6.25 and 6.30. The remaining areas show small areas of high potential and include Mystery Lake (Figure 6.27), Liz Lake (Figure 6.31), Sasagiu Rapids (Figure 6.32), Hambone Lake (Figure 6.33) and Southwest Bay (Figure 6.34). The same figures can be used to refer to results from the Fuzzy Logic Stratigraphy Models .

6.9 DISCUSSION

Prediction of known deposits was deemed successful if an area of high potential fell within 200m of the known deposit. The models were unable to predict areas of high potential within 200m of known deposits. However, the models did locate areas of high potential for nickel deposits within 1km of the Moak Lake Mines and Thompson T1 Mine. In addition to areas proximal to known deposits, other areas of high potential were detected across the belt. These areas may present options for future research or exploration.

Some locations, such as the area near Birchtree Mine, may produce better results if the datasets included more information. Further mapping projects by the companies could include defining members of formations in the lithology map. Additionally, results could greatly improve by incorporating higher resolution geophysical data. The stratigraphic models tended to predict larger areas of high potential than the structural models. This is a direct result of the sparseness of structural data for the belt. However, given the scale of mapping, the results support use of data fusion for regional scale mineral potential mapping to predict prospective areas for exploration. The results must be interpreted based on the quality and appropriateness of the datasets that went into the models. Additionally, the mineral potential maps are intended to focus on locations that present features that may host a nickel deposit. Finally, the sponsoring companies can test the merit of these methods to predict areas with high potential for nickel deposits with their own datasets.

CHAPTER 7

DISCUSSION

7.1 GIS EVOLUTION CHALLENGES

Integrating a GIS into an organization requires balancing needs with available resources. The CAMIRO TNB project began the process of integrating their data into a GIS in 1998. The original goals at the time were to enter the data, visualize data with one software, analyze spatial relationships and create a 3D model of the TNB. The project successfully completed 2 of the 4 goals. Although efforts were made to analyze spatial relationships in the belt, only preliminary analysis was possible due to the limitations of key datasets. Modeling the belt in three dimensions was never a realistic goal given the spatial distribution of sample sites and restrictions on drill hole dip and strike measurements.

During the CAMIRO TNB project's lifespan, data had to be collected, processed, entered into the GIS, displayed and cross-referenced with analysis common to each researchers specialty. People were spread across the country and involved in many other projects and did not have the time needed to develop the technical ability to use a GIS. Realistically, the only time most of the researchers saw their data displayed in the GIS was at yearly research meetings. Access to the files was limited in between the busy meetings because only people with previous GIS experience made use of the compact disk compilations distributed at these meetings.

Due to the time and technical limitations, the GIS dataset was not used or updated as frequently as possible. A suggestion for future projects is to set up a website that is dedicated to hosting the data collection. The data can be displayed with a very simple viewer that allows visualization, simple queries and printing. The learning curve would be much shallower and shorter than that required for introducing a specialized GIS package. The other benefit to online access is that only one version of the data exists. Individuals would no longer be tasked with keeping track of compact disks. Also, the large number of participants limited the number of compact disks distributed. Everyone would have access at all times. Viewers are free so no cost would be involved in acquiring the software. Installation is simple and the interface contains only a few tools common to many windows applications and browsers.

Certain datasets were provided by the companies with the understanding that the contents would not be revealed to any individual not involved in the research. Research participants signed memorandums of understanding (MOUs). The MOUs protected the companies from losing proprietary rights to their data. Access to these datasets – specifically, large scale geological maps – was not provided until the last year of the project. Before this time, data were integrated with maps created by the province at scales of 1 : 250,000. These maps were not able to render the detail needed for spatial analysis of lithologies in the belt since the company maps were created at a much larger scale.

The process of disseminating data from the companies to the researchers was not effective. The datasets were also prone to errors. Up to 10% of the polygons in the INCO dataset were identified by more than one lithology. Unfortunately, the data in GIS format were not quality controlled by INCO before distribution nor at the receiving end by the CAMIRO group. This oversight underlines the real need for maintaining metadata records. The data were also delivered with INCO's own grid system. The data were converted from the secret grid to UTM zone 14 NAD 83 by MITM. The process and ground control points were not provided when individuals who did the conversion moved on to other projects.

With respect to the data fusion application, geophysical data were a very important dataset in all the models. Verbal agreements with the companies indicated that aeromagnetic data with better resolution than the provincial dataset would be available to researchers. This agreement was not fulfilled. Consequently, the data fusion results provide only an indication of what can be done with appropriate data.

Introducing terminology such as belief, plausibility, uncertainty and disbelief into discussions of geological problems typically met some resistance. The idea of assigning a value describing the uncertainty a researcher hosts towards a piece of data seemed contradictory. However, in any piece of geological literature, authors incorporate caution through their choice of words. Phrases such as "the occurrence of x within the 123 zone suggests that the temperature and pressure ranged between +/-a and b" lends itself to measures of uncertainty. "Suggests" implies that there are other possibilities, that nothing is certain, and therefore, there is uncertainty. Qualifying statements such as "within" incorporate restrictions on when an occurrence is possible. Plausibility and belief are inherent on the upper and lower limits through the "+/-" qualification. The statement is true only within these limits. Geologists already use these concepts; they just do not realize it.

7.2 PERCEPTIONS AND DATA

A geologist is familiar with tools of the trade, both mechanical and analytical. Plotting R-values and mapping with a compass are techniques with which they are relatively comfortable. However, a common phenomenon occurs with comfort and familiarity - that is oversight. Unless something is obviously abnormal, scientists are comfortable noting expected trends and are prepared to attribute the observed trends to a physical process. In contrast, when new techniques are introduced, the scientist must pay closer attention to both the data and the analytical method because everything is new and nothing is expected. The data are therefore seen from a new perspective.

There are two important phases in the process of mapping an area – field observations and lab analysis. Both phases require interpretation. The quality of the interpretation can be affected by the scientists experience, resources available to the project (including time and equipment) and *a priori* knowledge of the area. Lack of experience, resources and knowledge can hinder the scientific merit of a project. Therefore, project managers must develop well defined objectives that can be met with the resources at hand.

When an exploration project reaches the local mapping scale, detailed field mapping begins. Field geologists enter an area with some expectations about the rocks. The rocks mapped should meet the criteria of the exploration model, otherwise, why map the area? These expectations can cloud the observations of the field crew. They expect to find specific lithologies that meet the model. When faced with a potential lithology, the call made can be preferential to a preconceived conclusion. Exploring other options becomes less of a choice. In fact, the observations meet the model specifications rather than the optimal situation of re-evaluating the model with new observations.

7.3 PROPRIETARY DATA

The CAMIRO project sponsors insisted that certain collections of information remained proprietary. Individual companies invested time and money into creating the geology maps and were not in a position to share this data with competitors. Their situation can be appreciated as they developed their maps to meet their specific needs. Other organizations develop commercial products to meet the needs of clients. They approach the market from a profit oriented perspective and provide custom solutions. Before clients implement custom mapping solutions, they need base maps.

Base maps include general information like provincial boundaries, roads, rivers and other topographical features. The features are common to everyone and have been mapped by the federal government. These maps can be considered infrastructure, similar to roads, utilities and communication infrastructure that society depends on. As infrastructure, map development was supported by the government. Due to low population densities and expanses of undeveloped land in many areas of Canada, the federal government has the most comprehensive map collection of Canada. However, use of the maps in digital format are restricted through licensing agreements with considerable cost. Consequently, there are no other alternatives to digital 1:50,000 and 1:250,000 scale base maps of Canada.

If the federal government considers base maps as necessary infrastructure, they should reduce the high cost of obtaining the data. By eliminating the cost of obtaining base maps, the mapping industry can move towards developing value added mapping products and web-based mapping applications and thereby increase the potential spin-offs from the industry.

Rather than simple presentation of features, the relation between features will become more important. Questions of "where is" will be replaced with "why is?" and "Where else does it occur?" The answer will not lie in noting the red dot lies next to the blue line but in the determining why many red dots fall within a distance of the blue line.

7.4 CONCLUSION

Although geologists are already asking why something exists where it does, they are just starting to harness the potential of digital spatial analysis. Features do not have to be next to each other to be important. Scale of relationships can be incorporated in their analysis. Multi-feature evaluation, not easily possible with visual inspection, can become a routine practice with a GIS. Data can be collected specifically for spatial analysis. What seemed to small to map previously will be invaluable to future analysis. Additionally, more demand for 3D GISs will push the technology to meet more needs from the diversified user community.

CHAPTER 8

CONCLUSION

8.1 CONCLUSION

The CAMIRO TNB project is nearing completion. The GIS database still has to be compiled. As a working product, the database met limited success. Restrictions on data access and use as set out in MOUs with industry reduced the effectiveness of data display and integration. At research meetings, maps presented as focal points of discussion had to be carefully screened for content so as not to conflict with the MOU. Consequently, the impact of having spatial data for analysis and display was compromised.

Many sources of error were identified in key datasets late in the project. The lack of detailed aeromagnetic data and incomplete geology maps reduced the ability of the data fusion models to locate potential deposits. As a result, only regional assessments could be completed. Areas of high potential were noted for further discussion.

The process of accessing and using many datasets with a complicated GIS proved to big a challenge for the project as a whole. However, the software and data were compatible. GIS experts were able to integrate all available data with provisions for deadlines and busy schedules. Compiling a collection of data to the TNB will prove invaluable for future researchers.

At this time, the geologic community is moving towards the digital environment. Many still have to be convinced that data management is as necessary as chemical analysis. In fact, assessing the data and software needs of a project from a systems perspective is ideally suited to a GIS expert. The expert is able to assemble very detailed datasets into a comprehensive overview of a region. The interaction between people and data has to be monitored to determine the success of a system. Identifying limitations in terms of processes and software and adapting solutions to meet the challenge is also part of the system approach.

This project has identified issues that should be addressed in future multidisciplinary projects. A realistic assessment of participants technical resources and timelines for data acquisition can increase use of data in GISs, hopefully evolving from simple visualization to more integrated analysis.

References

An, P., Moon, W. and Bonham-Carter, G.F., 1994. Uncertainty Management in Integration of Exploration Data using the Belief Function. Nonrenewable Resources. 3, p. 60-71.

Ansdell, K., Lucas, S.B., Connors, K. and Stern, R., 1995. Kisseynew Metasedimentary Gneiss Belt, Trans-Hudson Orogen (Canada): Back-arc Origin and Collisional Inversion. Geology, 23, p. 1039-1043.

Aronoff, S., 1989. Geographic Information Systems: A Management Perspective. WDL Publications, Ottawa, Canada.

Baragar, D.A. and Scoates, R.F.J., 1981. The Circum-Superior Belt: A Proterozoic Plate Margin? In: Precambrian Plate Tectonics. Edited by A. Kroner. Elsevier, Amsterdam, p. 297-330.

Bell, C.K., 1966. Churchill-Superior Province Boundary in Northeastern Manitoba. Geological Survey of Canada, Paper 66-1, p. 133-136.

Bleeker, W., 1990a. Evolution of the Thompson Nickel Belt and its Nickel Deposits, Manitoba, Canada. Ph.D. thesis, University of New Brunswick, Fredericton, N.B., Canada, 400 p. and maps.

Bleeker, W., 1990b. New Structural-Metamorphic Constraints on Early Proterozoic Oblique Collision along the Thompson Nickel Belt, Manitoba, Canada. In: The Early Proterozoic Trans-Hudson Orogen of North America. Edited by J. F. Lewry and M. R. Stauffer, Geological Association of Canada Special Paper, 37, p. 57-73.

Bleeker, W. and Macek, J., 1996. Evolution of the Thompson Nickel Belt, Manitoba: Setting of Ni-Cu Deposits in the Western Part of the Circum Superior Boundary Zone – Field Trip Guidebook A1. Geological Association of Canada/Mineralogical Association of Canada Annual Meeting, Winnipeg, Manitoba, May 27-29.

Bonham-Carter, G.,1994. Geographic Information Systems for Geoscientists: Modeling with GIS; Elsevier Science Inc., Tarrytown, New York, 398 p.

Brandstätter, G. and Sharov, A., 1998. Environmental Monitoring in the High Arctic using Different Types of High-resolution Satellite Imagery. International Archives of Photogrammetry and Remote Sensing, 32, p. 201-216.

Buchroithner, M., 1998. Geodata Interrelations: Inventory and Structuring Attempt of Taxonomic Diversity. In: Proceedings of the 2nd Conference of Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images, p. 11-15.

Burrough, P.A., 1986. Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press, Oxford.

Burrough, P.A. and McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press Inc., New York, 333 p.

CIG, 2000. Manitoba Branch Fall Forum, November 2000. Winnipeg Chamber of Commerce Club Meeting Room.

Clarke, K.C., 1997. Getting Started with Geographic Information Systems. Prentice Hall, Upper Saddle River, New Jersey, p. 338.

CORINE, 1992. Corine Land Cover Project, Environmental Agency, Copenhagen.

Church, W.R., 1998. Structural Provinces of North America from http://publish.uwo.ca/~wrchurch/300/14strprv.gif. University of Western Ontario.

Csathó, B. and Schenk, T., 1998. Multisensor Data Fusion for Automatic Scene Interpretation. International Archives of Photogrammetry and Remote Sensing, 32, p. 336-341.

Donaldson Dewitz, S., 1996. Systems Analysis and Design and the Transition to Objects. The McGraw-Hill Companies, Inc., Toronto, Ontario, p. 557.

Fabbri, A.G., 1984. Image Processing of Geological Data. Van Nostrand Reinhold Company, Agincourt, Ontario p. 244.

FGDC, 2001a. fgdc.er.usgs.gov/metadata/metadata.html

FGDC, 2001b. fgdc.er.usgs.gov/

Fountain, D.M. and Salisbury, M.H., 1981. Exposed Cross-sections Through the Continental Crust: Implications for Crustal Structure, Petrology And Evolution. Earth Planetary Science Letters, 56, p. 263-277.

Gibb, R.A., 1983. A Model for Suturing of Superior and Churchill Plates: An Example of Double Indentation Tectonics. Geology, 11, p. 413-417.

Goodchild, M.F., 1992. Geographical Information Science. International Journal of Geographical Information Systems, 6.

Green, W.R., 1991. Exploration with a Computer - Geoscience Data Analysis Applications. Pergamon Press, Toronto, Ontario, p. 225.

Hodgson, C.J., 1993. Uses (and Abuses) of Ore Deposit Models in Mineral Exploration. In Ore Deposit Models Volume II, Geoscience Canada Reprint Series 6, p. 1-11.

Hoffman, P.F., 1989. Precambrian Geology and Tectonic History of North America. In: The Geology of North America - An Overview. Edited by A.W. Bally and A.R. Palmer. Geological Society of America, Boulder, Colorado, p. 447-512.

Klasner, J.S. and King, E.R., 1990. A Model of Tectonic Evolution of the Trans-Hudson Orogen in North and South Dakota. In: The Early Proterozoic Trans-Hudson Orogen of North America. Edited by J. F. Lewry and M. R. Stauffer, Geological Association of Canada Special Paper. 37, p. 271-286.

Korte, G.B., 1997. The GIS Book. OnWord Press, Sante Fe, New Mexico, p. 414.

Kornik, L.J., 1969. An Aeromagnetic Study of the Moak Lake-Setting Lake Structure in Northern Manitoba. Canadian Journal of Earth Science, 6, p. 373-381.

Kornik, L.J., 1971. Magnetic Subdivision of Precambrian Rocks in Manitoba. Geological Association of Canada, Special Paper Number 9, p. 51-60.

Kornik, L.J. and MacLaren, A.S., 1966. Aeromagnetic Study of the Churchill-Superior Boundary in Northern Manitoba. Canadian Journal of Earth Science, 3, p. 547-557.

Lewry, J.F., 1981. Lower Proterozoic Arc-microcontinent Collisional Tectonics in the Western Churchill Province, Nature, 294, p. 69-72.

Lewry, J.F. and Collerson, K.D., 1990. The Trans-Hudson Orogen: Extent, Subdivisions and Problems. In: The Early Proterozoic Trans-Hudson Orogen of North America. Edited by J. F. Lewry and M. R. Stauffer, Geological Association of Canada Special Paper. 37, p. 1-14.

Lillesand, T.M. and Kiefer, R.W., 1994. Remote Sensing and Image Interpretation, Third Edition. John Wiley & Sons, Inc., Toronto, p. 750.

Liwanag, J., 2001. Post-magmatic Modification of the Thompson Nickel Belt Deposits, Manitoba, Canada. MSc Thesis, University of Manitoba, Canada, p. 103.

Lucas, S.B., White, D., Bleeker, W., Hajnal, Z., Lewry, J. and Weber, W., 1996. Crustal Structure of the Superior Boundary Zone (Thompson Belt) from New Lithoprobe Seismic Reflection Data. In: Trans-Hudson Orogen Transect. Lithoprobe Rep.. 55, p. 82-94.

Lucas, S.B., Green, A., Hajnal, Z., White, D., Lewry, J., Ashton, K., Weber, W. and Clowes, R., 1993. Deep Seismic Profile Across a Proterozoic Collision Zone. Surprises at Depth. Nature, 363, p. 339-342.

Machado. N., 1990. Timing of Collisional Events in the Trans-Hudson Orogen: Evidence from U-Pb Geochronology for the New Quebec Orogen, the Thompson Belt, and the Reindeer Zone (Manitoba and Saskatchewan). In: The Early Proterozoic Trans-Hudson Orogen of North America. Edited by J. F. Lewry and M. R. Stauffer, Geological Association of Canada Special Paper, 37, p. 433-447.

Machado, N. and Potrel, A., 2000. Personnel communication. Research trip, UQAM, Montreal.

Mangolini, M., 1997. Apport de la fusion d'images satellitaires multicapteurs au inveau pixel en télédetection et photo interpretation. Thèse de doctorat, Universite Nice-Sophia, France, p. 174.

Meredith, J.R. and Mantel Jr., S.J., 1995. Project Management – A Managerial Approach Third Edition. John Wiley and Sons, Inc., Toronto, Ontario, p. 767.

Merriam-Webster Online, 2001. www.m-w.com .

Moran, S. and Lopez Baros, S., 1996. Raster Imagery in Geographic Information Systems. OnWord Press, Sante Fe, New Mexico, p. 495.

Nelson, K.D., Baird, D.J., Walters, J.J., Hauck, M., Brown, L.D., Oliver, J.E., Ahern, J.L., Hajnal, Z., Jones, A.G. and Sloss, L.L., 1993. Trans-Hudson Orogen and Williston Basin in Montana and North Dakota: New COCORP Deep Profiling Results. Geology, 21, p. 447-450.

Obermeyer, N.J. and Pinto, J.K., 1994. Managing Geographic Information Systems. The Guilford Press, New York, New York, p. 226.

Office of Science and Technology Policy, 2000. www.ostp.gov/html/0053_4.html .

Pohl, C. and van Genderen, J.L., 1998. Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications. International Journal of Remote Sensing, 19, p. 823-854.

Richards, J.A., 1993. Remote Sensing Digital Image Analysis – An Introduction. Springer-Verlag, New York, p. 340.

Schuurman, N., 1999. Critical GIS: Theorizing an Emerging Science. Cartographica, 36, Monograph 53.

Scoates, R.F.J., Macek, J.J. and Russell, J.K., 1977. Thompson Nickel Belt Project: Manitoba Resources Division, Report of Field Activities 1977, p. 47-54.

Sprague Jr., R.H. and McNurlin, B.C., 1993. Information Systems Management in Practice, Third Edition. Prentice Hall, Englewood Cliffs, New Jersey, p. 576.

Star, J. and Estes, J.E., 1990. Geographic Information Systems: An Introduction. Prentice Hall, Upper Saddle River, New Jersey.

Tsai, A.Y.H., 1988. Database Systems – Management and Use. Prentice-Hall Canada, Inc., Scarborough, Ontario, p. 572.

Viljoen, D., Chackowsky, L., Lenton, P. and Broome, H.J., 1999. Geology , Magnetic and Gravity Maps of Manitoba: A Digital Perspective: Manitoba Energy and Mines Geological Services MEM OF99-12 CD-ROM.

Wald, L., 1998. Data Fusion: A Conceptual Approach for an Efficient Exploitation of Remote Sensing Images. In: Proceedings of the 2nd Conference of Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images, p. 17-23.

Wald, L., 1999. Some Terms of Reference in Data Fusion. IEEE Transactions on Geoscience and Remote Sensing. 37.

Weber, W., 1990. The Churchill-Superior Boundary Zone, Southeast Margin of the Trans-Hudson Orogen: A review. In: The Early Proterozoic Trans-Hudson Orogen of North America. Edited by J. F. Lewry and M. R. Stauffer, Geological Association of Canada Special Paper, 37, p. 42-55.

White, D.J. and Lucas, S.B., 1994. A Closer Look at the Superior Boundary Zone. 1994 Lithoprobe Trans-Hudson Orogen Transect Workshop Report, No. 38, p. 35-41.

White, D.J., Jones, A.G., Lucas, S.B. and Hajnal, Z., 1999. Tectonic Evolution of the Superior Boundary Zone from Coincident Seismic Reflection and Magnetotelluric Profiles. Tectonics, 18, p. 430-451.

White, D.J., Lucas, S.B., Hajnal, Z., Green, A.G., Lewry, J.F., Weber, W., Bailes, A.H., Syme, E.C. and Ashton, K. 1994. Paleo-Proterozoic Thick-skinned Tectonics: Lithoprobe Seismic Reflection Results from the Eastern Trans-Hudson Orogen, Canadian Journal of Earth Science, 31, p. 458-469.

Zurbrigg, H.F., 1963. Thompson Mine Geology. Canadian Institute of Mining and Metallurgy Bulletin, 56, p. 451-460.