Evaluation of Weights of Evidence to Predict Gold Occurrences in Northern Minnesota's Archean Greenstone Belts

By

Brian K. Hartley

A Thesis Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

August 2014

Table of Contents

List of Tables

List of Figures

Abstract

Much of northern Minnesota is underlain by rocks that make up the so-called Superior Province of the Canadian Shield -- the ancient core of the North American continent. These Superior Province rocks originated in the Archean Eon, between 4.0 and 2.5 billion years ago. Altogether, more than 60% of Earth's crust formed during this time period, making Archean terranes worldwide particularly rich in mineral resources.

Even among other Archean terranes, the Superior Province is exceptionally rich in gold. The Canadian province of Ontario, immediately north of Minnesota, is host to over 300 significant gold deposits, with 18 currently producing mines. However, no economically significant gold deposit has yet been discovered in Minnesota, despite several periods of intense exploration activity in the 1980s and early 1990s.

This project utilized public datasets representing geology, geophysics, and geochemistry to predict the likelihood of new gold occurrences in northern Minnesota's Archean bedrock, using a geospatial information systems (GIS) modeling technique called weights of evidence. The study area was ranked on a relative scale from low gold potential (Not permissive) to high gold potential (Favorable).

Comparison of model results to past and present gold exploration activity suggests that weights-of-evidence modeling is a useful tool for generating new exploration targets in northern Minnesota. Many of the tracts ranked as Favorable do not appear to have ever been drilled, so the bedrock in these areas has yet to be evaluated. Also, because most of the datasets used were either first published or significantly updated between 2004 and 2012, this project is likely the first to use them in a predictive model, and offers a new perspective on gold prospectivity in the region.

Chapter One: Introduction

Gold is an important commodity in the modern economy. Still widely valued as a form of currency, gold is also an important catalyst for certain industrial processes, and a conducting material in many advanced electronic components. Generally speaking, more than half of the gold produced each year is mined by just six countries (listed in order of contribution from greatest to least): China, Australia, the United States, Russia, Peru, and South Africa. Roughly 64 percent of the demand for gold is met by new mine production, and 36 percent is met by recycling. These activities annually contribute more than US$ 200 billion to the global economy (World Gold Council, 2013).

1.1 Gold Mining and Exploration

In terms of mining, gold is found in two broadly-defined settings: so-called *hard-rock* deposits and *placer* deposits. Hard-rock deposits occur in solid crustal rock, where gold has been concentrated by geologic processes. Generally, the gold in many of these deposits was transported in solution by superheated fluids (called hydrothermal fluids), then precipitated in faults, cracks, and other structures, forming the classic “veins of gold”. Globally, the average concentration of gold in the Earth's crust is approximately 4 parts per billion (ppb). In a typical hard-rock gold deposit, it has been concentrated by 1,000 to 10,000 times this value (Robb 2005).

Placer deposits occur where hard-rock gold has been weathered out, transported by some combination of water or wind, and concentrated by gravity in areas where it becomes “trapped” along with other heavy minerals. Many types of placer deposits are found worldwide in beach sands, river gravels, and wind-driven sediments. Because the common factor is transported material, placer deposits can occur at great distances from their bedrock source. For this project, only hard-rock gold deposits are of interest, and further use of the term *deposit* is in reference to this type, unless specifically noted otherwise.

Because gold deposits often occur deep in the Earth's crust, typically also covered by soils and other surficial cover, geologists generally use a deductive, multi-disciplinary approach to find

them. The process of searching for undiscovered mineral deposits is called mineral exploration, and is a type of scientific investigation. Empirical and theoretical knowledge of mineralizing systems is reviewed, observations are made, hypotheses are offered, and field investigations are used to confirm or reject them.

Carranza (2009a) identifies four phases of mineral exploration. First, *area selection* seeks to identify a region that is *permissive*, or capable of hosting a deposit based on empirical and theoretical criteria regarding rock types and the overall geological environment. Second, *target generation* refines permissive areas into *prospective* areas, that are also supported by evidence plausibly suggesting that a deposit may exist therein. Third, if a deposit is discovered, *resource evaluation* correlates evidence (gold intercepts and gold assay results) from drill holes, in order to gain an understanding of the deposit's size and mineral character. And finally, *reserve definition* employs rigorous metallurgical testing, in order to determine if or how much of the resource can be profitably extracted and processed into marketable commodities.

Note the difference between resource (gold in the ground) and reserve (gold in the ground that can be extracted and refined at a profit, using available technology and methods). The term *ore* refers to rock classified as reserve, and not just any rock that contains gold. In the same way, a gold *occurrence* only becomes a deposit if it is found to be of sufficient *tonnage* (bulk volume) and *grade* (concentration of gold per ton of ore) to be profitably mined.

Mineral exploration is data-intensive, and makes use of many kinds of geological information in addition to that from drill holes. Of particular importance is information acquired or enhanced by geophysics, where sensors measuring ambient gravity field strength, magnetic field strength, or electrical conductivity are deployed in order to gain an understanding of the rocks' physical properties (Moon, Whately, and Evans 2006). This project makes use of high-resolution aeromagnetic data (1:24,000), and small-scale geologic maps (1:500,000), published by the Minnesota Geological Survey (MGS).

Soil or stream-sediment samples are also often taken across an area, and their metal contents are analyzed in a laboratory process called assaying. This approach is called geochemical exploration, and can be a useful tool for locating hidden gold deposits (Harris, Wilkinson, and Bernier 2001; Muir, Schnieders, and Smyk 1995). This project utilizes regional soil geochemistry data, collected as part of a joint program between the MGS and the United States Geological Survey (USGS). It also utilizes results of till (glacial material) sampling conducted by the Minnesota Department of Natural Resources (MDNR).

If geophysical and geochemical surveys return interesting results, drilling might be conducted to collect either core or chip samples of the hidden bedrock for observation and assay (Marjoribanks 1997). Drilling is an integral part of mineral exploration, but it is expensive, and is generally only used if there is sufficient evidence of a possible gold deposit to justify its use. This project utilizes drill hole locations published by the MDNR, and selected drill hole records compiled by Frey (2012), Severson (2011), Frey and Hanson (2010), and others.

1.2 Geology and Gold Exploration History of Minnesota

The two most abundant elements in the earth's crust, by weight, are oxygen (O) and silicon (Si). Rocks are generally classified by silica (SiO_2) content; where rocks high in silica are termed *felsic*, rocks low in silica are termed *mafic*, and rocks in between are referred to as *intermediate*.

Rocks are also classified by the environment in which they form. *Plutonic* rocks form deep in the Earth's crust, from magma that cooled and crystallized slowly, and only reach the surface if they are exposed by uplift or erosion. *Volcanic* rocks are formed from liquid magma that cooled and crystallized on the Earth's surface, after being ejected during volcanic activity. Together, groups of volcanic and plutonic rocks are referred to as *volcanoplutonic*. Rocks that form from sediment (such as sand or mud), or eroded particles of any other rock type, are called *sedimentary* rocks.

Any rock that has been physically (but not chemically) modified by heat and pressure is called a *metamorphic* rock. A sedimentary rock that has been so modified is called a

metasedimentary rock. Volcanoplutonic rocks that have been metamorphosed, a favorite locale for gold, are sometimes are sometimes referred to as *greenstone*, due to the presence of metamorphic minerals that have a characteristically green color. Note that "greenstone" is a catch-all term, and not a reference to any specific rock type. Very intense metamorphism often separates and aligns light and dark minerals, producing a banded, *gneissic* texture, and such rocks are called *gneiss*.

The Earth began forming solid crustal rock (i.e. hard-rock) approximately four billion years ago. Some of this rock accreted to form the core of the North American Continent, referred to as the North American Craton (Figure 1). Today, the part of the craton that remains exposed, not covered by younger rocks, is called the Canadian Shield (Ojakangas 2009). The Canadian Shield is a complex amalgamation of many smaller cratons, which formed and accreted during an immense span of time called the Precambrian, which extends from 4.6 billion years ago to "just" 550 million years ago -- some 88% of Earth's history.

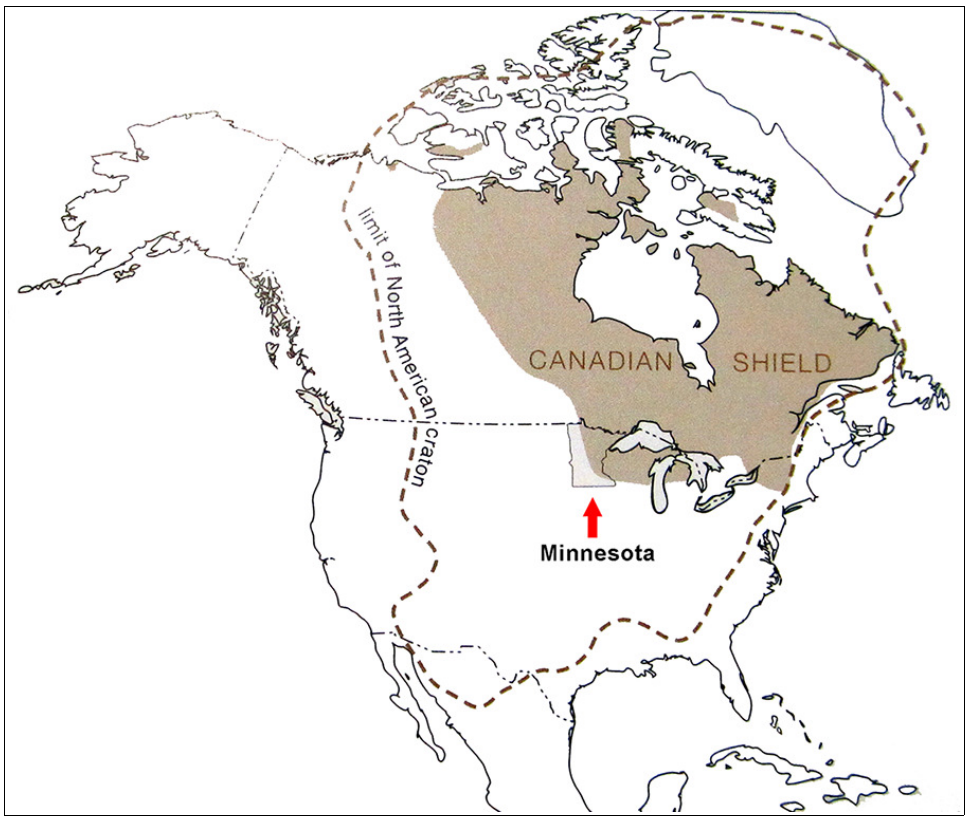


Figure 1: *Overview of the North American Craton, showing Minnesota in relation to the Canadian Shield. Modified from Ojakangas 2009.*

Underlying Minnesota is the subdivision of the Canadian Shield called the Superior Province (Figure 2). The Superior Province formed during the Archean Eon, which lasted from 4.0 to 2.5 billion years ago. Roughly 60 percent of Earth's crust formed during this time, making Archean rocks worldwide particularly rich in mineral resources (Robb 2005; Taylor and McLennan 1995).

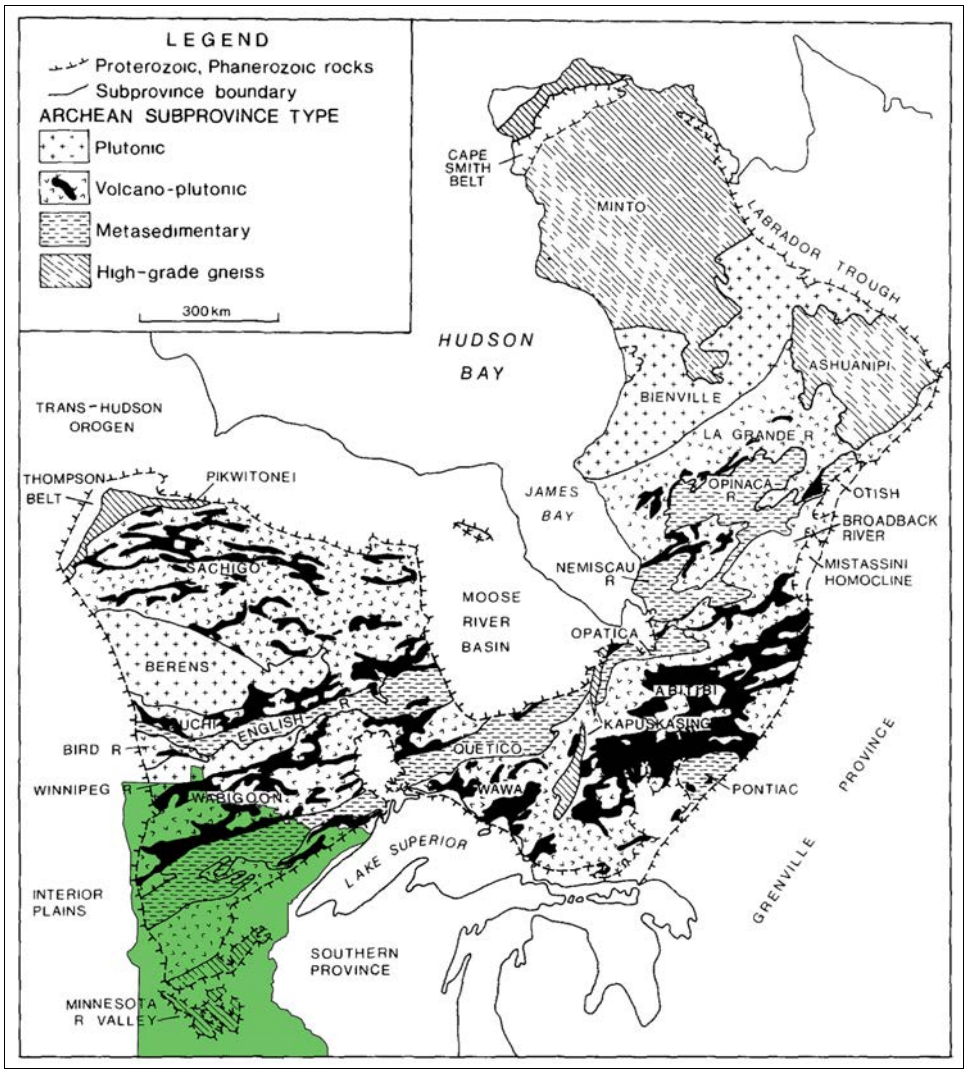


Figure 2: *Generalized geologic map of the Superior Province. Minnesota's position (in green) is approximate. Modified from Card 1990.*

Among other Archean terranes, the Superior Province is exceptionally rich in gold, with over 300 significant deposits presently known in neighboring Ontario, Canada (Poulsen, Card, and Franklin 1992). Gold occurrences have been identified in Minnesota, but to date no gold deposit of economic significance has yet been discovered.

Figure 3 presents a modern view of northern Minnesota, the region of the Superior province on which this thesis focuses. The first doumented "gold rush" in Minnesota was in 1865 in the Western Vermilion district, with another following in 1893 in the Rainy Lake area (Severson 2011). From the latter, one mine on Little America Island actually reached production, but was abandoned in 1895 due to high costs and low-grade ore.

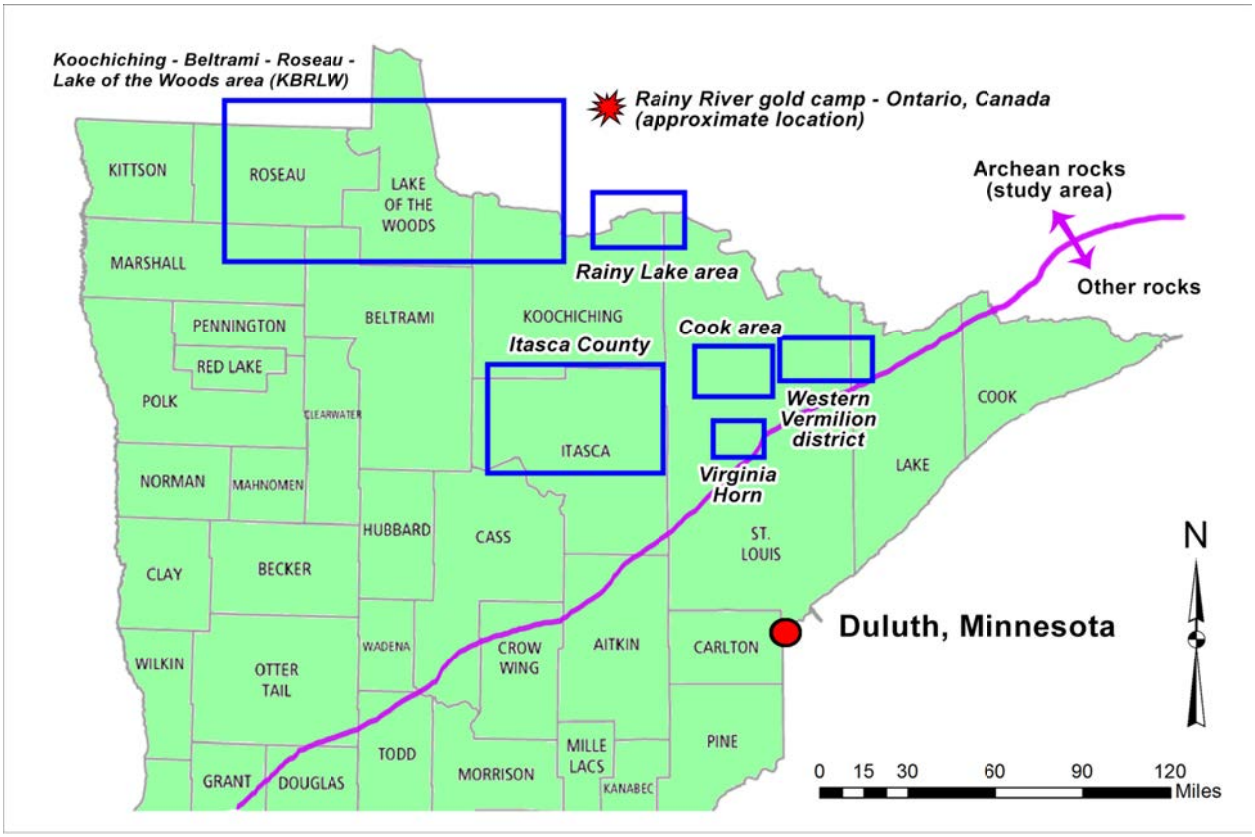


Figure 3: *County map of northern Minnesota, showing the approximate location of the Rainy River gold camp, and the six areas of gold exploration focus identified by Severson (2011).*

Not much exploration work was done in Minnesota after these short-lived periods of excitement, until 1981 when the Hemlo gold deposit was discovered in Ontario, approximately 350 miles to the northeast. The Hemlo discovery spurred a lot of interest because it did not "fit" any of the accepted thinking about Archean gold deposits at that time. As such, it had been overlooked despite close to one hundred years of exploration work in the surrounding area (Muir, Schnieders, and Smyk 1995). Figure 4 shows a timeline of exploratory drilling that was done in Minnesota, specifically for gold exploration, from 1980 to 2010.

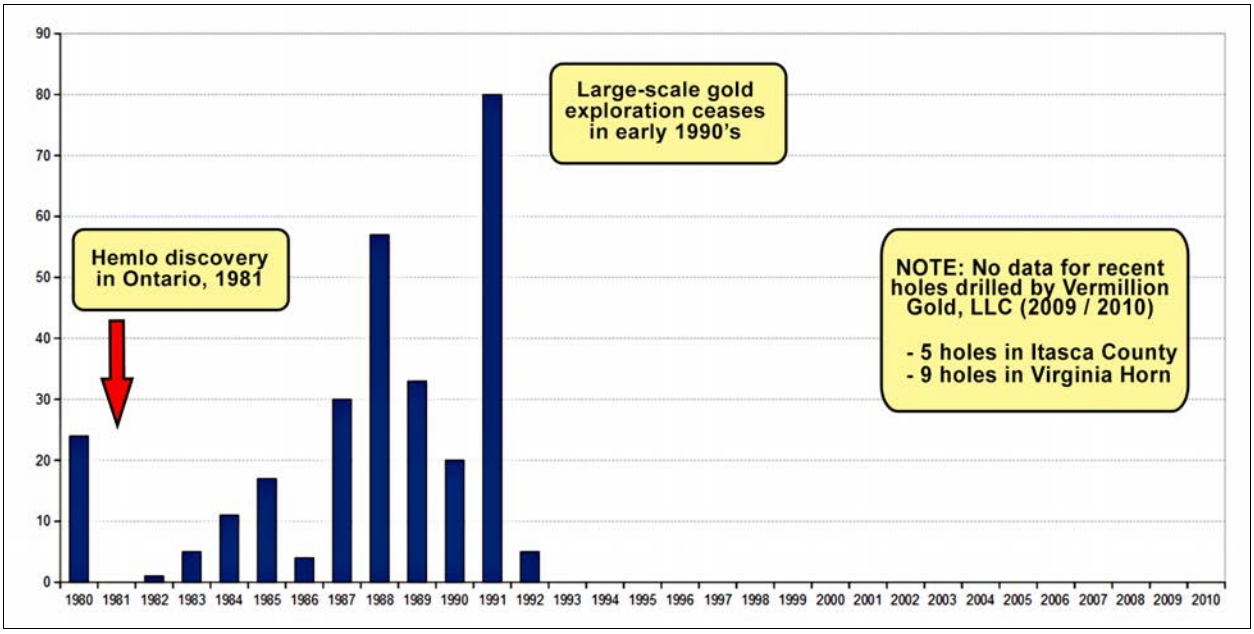


Figure 4: *Raw counts of gold-targeting exploratory drill holes in Minnesota, from 1980 to 2010.*

The most recent gold discovery in Ontario, the Rainy River gold camp (see Figure 3), hosts a deposit that contains as much as four million ounces of gold, along with ten million ounces of silver (Hardie et al. 2013). New Gold, Inc[1]. is currently working toward developing this deposit into a producing mine.

Notwithstanding the lack of new gold discoveries in Minnesota since the 1980s, the likelihood is that gold deposits remain to be found. Muir, Schnieders, and Smyk (1995) state "The presence of a major gold deposit in any [greenstone] belt carries with it the substantial possibility that additional deposits remain to be discovered." And Severson (2011) observes that "[previous explorers'] efforts have shown that zones with gold enrichment do indeed occur throughout the state and that the final prize of discovering a potential gold mine in Minnesota still awaits a company, or individual, with the fortitude to doggedly pursue such a venture."

1.3 Project Scope and Purpose

The purpose of this project is to explore the potential that undiscovered hard-rock gold deposits exist in northern Minnesota. This potential is qualitatively assessed by applying a geospatial information systems (GIS) modeling technique, called *weights of evidence*, to predict

1 New Gold, Inc. - Rainy River project: http://www.newgold.com/properties/projects/rainyriver/default.aspx

new gold occurrences.

The geographical scope of this project is the Archean bedrock underlying northern Minnesota: specifically, the southernmost Wabigoon, Quetico, and Wawa subprovinces of the Superior Province of the Canadian Shield (see Figure 2), that trend southwest from Canada into Minnesota. The study area covers some 91,540 km^2 (35,344 mi^2).

Although part of the Wawa subprovince, the Minnesota River Valley gneissic terrane (figure 2) was excluded from the study area. Most of the gold mineralization in the Superior Province occurred approximately 2.7 billion years ago in the late Archean (Poulsen, Card, and Franklin 1992; Kerrich and Cassidy 1994). By contrast, the Minnesota River Valley rocks formed deep in the Earth's crust around 3.8 billion years ago, and were later accreted to the Wawa subprovince during a period of uplift (Ojakangas 2009).

The operational scope of this project is the construction of a gold prospectivity model in a GIS, using available public datasets. This is an “office project,” similar to what might be done by an individual working on an early-stage or regional mineral exploration program. The main focus of this project is the GIS modeling operations, though some basic geology will be discussed during data preprocessing. No field work was conducted or scheduled, either to collect data or assess modeling results.

Chapter Two: Background

2.1 Mineral Prospectivity Modeling

Mineral prospectivity modeling is the subject of this project, and applies to the second phase of mineral exploration, target generation. This is a "desktop" activity, in which the goal is to analyze spatial relationships of known gold deposits or occurrences with available geological, geophysical, or geochemical datasets. From these spatial relationships, *patterns* can be identified in the datasets, that can be treated as *evidence* or *predictors* of undiscovered deposits. Where no patterns are present, new deposits are unlikely to be found; and conversely, where several patterns are present, new deposits more plausibly exist.

Areas where multiple patterns coincide strongly can be labeled *targets*, and field-based studies can be commissioned to further investigate them. In this way, mineral prospectivity modeling can be a useful tool in early-stage mineral exploration, for "due diligence" investigations of regions where a company may wish to establish new land positions.

Mineral prospectivity modeling is well suited for the geospatial information systems (GIS) environment, and is sometimes referred to as *predictive modeling* (Carranza 2009a). Because of the multi-disciplinary nature of gold exploration, the datasets involved are often a mix of vector data (points, lines, and polygons) and raster data (elevation, geophysics, geochemistry, or perhaps aerial imagery).

The methods for mineral prospectivity modeling may be classified as either *knowledge-driven* or *data-driven*. Knowledge-driven methods rely on a conceptual model of the spatial relationships between deposits and geologic evidence in other, well-explored areas. The conceptual model is applied to the area of interest based on expert judgment, and the relative importance of each piece of evidence is decided by the user. Examples of knowledge-driven approaches include Boolean logic methods, binary index overlay methods, and fuzzy logic methods. These are sometimes collectively referred to as "wildcat" modeling, in reference to "wildcat" drilling for oil,

where uncertainty due to lack of data is high (Carranza 2009a).

By contrast, the goal of data-driven methods is to remove dependence on expert judgment as much as possible, and draw insights of spatial relationships directly from the data. Some examples of data-driven methods include neural networks, logistic regression, and weights of evidence (Carranza 2009a). This project makes use of the weights-of-evidence method exclusively.

Data-driven methods are most easily applied to areas where many deposits or occurrences are known, and sufficient data exists from which to draw related evidence. Northern Minnesota only partially fits this requirement. There are no economically significant gold deposits presently known, but there are areas in which gold has been identified in drill holes, as well as in soil and till samples. This project is thus a data-driven predictive model of gold occurrences, and not gold deposits.

2.2 The Weights-of-Evidence Method

The weights-of-evidence method, first developed for medical diagnoses, was adapted for use in mineral prospectivity modeling in the 1980's. The mathematical base for weights of evidence is thoroughly described by Bonham-Carter (1994), briefly summarized by Carranza (2009a), and applied to gold exploration by Bonham-Carter (1988) and Raines (1999).

Weights of evidence works by assigning *weights* to each piece of evidence automatically, based on its correlation of that evidence and known occurrences of some objective, in this case gold occurrences. The basic idea is simple: if a particular combination of evidence can be associated with known gold deposits, then wherever that combination is present, there is increased likelihood of additional, presently-unknown gold. Importantly, weights of evidence also provides measures of confidence in its own model results.

2.3 Weights-Of-Evidence Terminology

For weights of evidence, as any other type of modeling, a *study area* must first be defined. This is the area of focus, in which all data and model calculations are confined to. In mathematical formulations, the study site is referred to as (S). The study area is comprised of a number of *unit*

cells, and there are N(S) unit cells within the study area.

Within the study area, a number of *training sites* are selected, that represent known locations of something that the model is attempting to predict. In mathematical formulations, training sites are represented as (T), and the number of training sites is designated as N(T).

The model considers the study area as a two-class raster, with one class being unit cells that are occupied by training sites, and the second class being unit cells are not occupied by training sites. An important assumption in weights-of-evidence modeling is that each training cell contains at most one training site.

The unit cells are often symmetrical (square), and their dimensions are chosen by the user. For this project, a unit cell size of 500 m^2 (0.25 km^2) was selected based on map scale, after consideration of Carranza (2009b). The goal is to select a unit cell size that reflects the scale of the final map product, the resolution of the evidence, and the areal extents of the objects represented by training sites.

Each piece of evidence, or *evidential theme*, is similarly presented in raster form, and reduced to a small number of *classes*. Each class represents a *pattern* of evidence. For example, anomalous magnetic highs are a pattern within the aeromagnetic evidential theme; and rocks of a certain type can also be considered to be a pattern in the geologic map evidential theme.

In a weights-of-evidence model, *weights* are assigned to each pattern of each evidential theme. These weights are calculated using two measures, *W+* and *W-*. The training sites located inside of the pattern are evaluated to obtain W+. If the pattern contains more training sites than would be expected by random chance, the W+ value is greater than 0. The process for obtaining W- is just the reverse. Here, if the number of points located outside the pattern is less than would be expected by random chance, the W- value is less than 0.

The two weights are combined into *contrast (C)*, a measure of overall correlation between a pattern and the training sites. The contrast is also assigned to each pattern, and is simply the

difference between the inside and outside weights ([W+] - [W-]). A large, positive C value indicates that the pattern is strongly correlated with the training sites, or that training sites are likely to occur within that pattern. A negative C indicates a negative correlation, and the training sites are unlikely to occur within the pattern. Both high and low C values can be useful in exploration.

The *Studentized contrast* is the ratio of the contrast to the standard deviation of its underlying weights ([C / σC]), a variant of the Student t-test (Kotz, Balakrishnan, and Johnson 2000). It is a measure of the confidence that the reported contrast is not zero. A pattern can be considered a useful predictor of training sites when it has a large positive contrast, and its Studentized contrast is also sufficiently large.

For this project, two types of weights are used, *categorical* and *cumulative*. The type of weights depend on the level of measurement represented by the patterns, specifically whether they are nominal, ordinal, interval, or ratio data (Stevens 1946).

Nominal data are unordered, such as the names of the general rock types in the four-class geologic map. The numerical values assigned to each class are simply codes for the rock types aggregated within it, and cannot be subjected to mathematical operations. These nominal patterns are evaluated with respect to the training sites using categorical weights.

Ordinal, interval, and ratio data, together referred to as ordered data, represent measured values (such as ambient magnetic field strength or gold concentration in soil) that are inherently scaled from low or small to high or large. Ordered data are evaluated using cumulative weights.

There are two ways to use cumulative weights. To test the hypothesis that the training sites should be more associated with a low value (such as feature proximity), *cumulative-ascending* weights are used. This helps the user to select a suitable zone of influence around a feature, that has suitably high W+ and C values, indicating a strong positive correlation with training sites. To test the hypothesis that training sites should be more associated with high values, such as a multi-class raster representing gold concentration in soils, *cumulative-descending* weights are used.

The *prior probability* of the study area is simply the number of unit cells that contain a training site, divided by the total number of unit cells within a study area, or N(T) / N(S). Thus, the prior probability is the likelihood of finding a training site before considering any of the evidence.

The *posterior probability* is the likelihood of finding whatever a training site represents in a given unit cell, after considering the evidence. Each unit cell in the study area is assigned a value of posterior probability that reflects the prior probability modified by the combined evidence present within the cell area, as considered important by the user.

2.4 Conditional Independence

One of the important assumptions of the weights-of-evidence method is that each evidential theme is *conditionally independent* of the others, i.e. that the presence of a pattern in one theme is not influenced by the presence or absence of patterns in other themes (Bonham-Carter 1994; Agterberg and Cheng 2002). Where conditional dependence is present, the posterior probability values can be inflated, or sometimes deflated.

In mineral exploration, conditional independence (CI) is not always possible, because patterns are chosen based on their empirical or theoretical spatial relationships with mineral deposits, and these patterns are often dependent on the same underlying geologic systems. Adding more patterns usually introduces related evidence, and thus decreases conditional independence. Raines (2006; 1999) suggests that where CI can not be eliminated or reduced to some acceptable level, the results of posterior probability should only be used to generate an ordinal *ranking* system. These ranks are to be considered a *relative*; ranging from "Not permissive" (least likely to host a gold occurrence) to "Favorable" (most likely to host a gold occurrence).

This project utilized the relative ranking system, so CI will not be considered further. Other data-driven modeling methods, such as logistic regression, are not affected by CI. With the same evidence and training sites, logistic regression usually produces ranks that are similar to those derived from the weights-of-evidence method (Wright and Bonham-Carter 1996).

Chapter Three: Software and Data Sources

This chapter provides an overview of the software used, and the datasets that were chosen for weights-of-evidence modeling of hard-rock gold occurrences in the Canadian shield region of northern Minnesota.

3.1 Software

This project made use of Esri® ArcGIS™ version 10.0 (not the most recent version, explained below), under a continuing educational license from Esri. The Esri ArcGIS Spatial Analyst and Geostatistical Analyst extensions were also used.

The core geological predictions were made using Spatial Data Modeler for ArcGIS (ArcSDM)[2] extension. This software was initially developed by Graeme Bonham-Carter at the Geological Survey of Canada (GSC) and Gary Raines at the U.S. Geological Survey (USGS), for use with Esri ArcView version 3. ArcSDM was subsequently ported to ArcGIS v9.x, and later v10.0, as a "toolbox" by Don Sawatzky (USGS) in consultation with Raines and Bonham-Carter, now all retired.

A plugin for ArcGIS was required to work with the aeromagnetic data in ArcGIS, as it was created by the MGS using Oasis Montaj™ software from Geosoft. This plugin is available from the Geosoft website[3] at no cost.

3.2 Data Sources

The data sets used in this project are all available at no cost to the general public, via the Web portals of three government agencies: the Minnesota Department of Natural Resources (MDNR), the Minnesota Geological Survey (MGS), and the United States Geological Survey (USGS). The types of data range from points, polylines, and polygons to raster files. Table 1 provides an overview of the datasets used from each source, and their purpose. All were reprojected to UTM Zone 15N, NAD83 datum.

2 ArcSDM: Spatial Data Modeler for ArcGIS 10 (and earlier versions): http://www.ige.unicamp.br/sdm/default_e.htm

3 Geosoft plugin for ArcGIS: http://www.geosoft.com/support/downloads/plug-ins/plug-arcgis

Table 1: *Overview of data sets used, and their sources.*

Source	Description	Type	Purpose
MDNR	County boundaries	Polygon shapefile	Basemap
	Drill hole records	Point shapefile	Training sites
	Major watersheds	Polygon shapefile	Thematic evidence
	Open-file project #378 (Drill core re-evaluation)	Point shapefile	Training sites
	Open-file project #373 (Drill core re-evaluation)	Point shapefile	Training sites
	Open-file project #392 (Gold-in-till survey)	Point shapefile	Training sites
	Open-file project #379 (Gold-in-till survey)	Point shapefile	Training sites
MGS	Map #S-22, Precambrian bedrock geology	Point, line, and polygon shapefiles	Thematic evidence
	Aeromagnetics (1:24,000 scale)	Raster (GRID)	Thematic evidence
USGS	Regional soil geochemistry - MN	Point shapefile	Thematic evidence

3.2.1 Minnesota Department of Natural Resources (MDNR)

The first dataset obtained from MDNR was a polygon shapefile representing Minnesota county boundaries, to be used for geographic reference. The second was a point shapefile containing surface locations of all drill holes in Minnesota, dating back to 1907. Included are drill holes for engineering, scientific, and mineral exploration purposes; these serve both as a source of training sites and also as a source of validation sites, comparing model results to past gold exploration activity. The third dataset was a polygon shapefile representing major watersheds, to provide basic hydrologic information about the study area and a means to aggregate the soil geochemical samples into a representative surface for modeling.

The fourth MDNR dataset was a re-evaluation of 13 historical drill holes in the Rainy Lake area (see Figure 3), performed as part of open-file project #378. From the 13 selected drill cores, 217 new samples were taken for gold assay, partially guided by 2,023 semi-quantitative analyses conducted with a hand-held x-ray fluorescence (XRF) unit. This project is complete, and the final report (Frey 2012) is available on the project Web page[7]. The spreadsheet 'p378_chem_shape_ab6' was downloaded in October of 2013 and converted to a point shapefile in ArcGIS.

7 MDNR open-file project #378: http://www.dnr.state.mn.us/lands_minerals/mpes_projects/project378.html

The fifth MDNR dataset was another drill core re-evaluation, for 12 historical drill holes in the Western Vermilion district (see Figure 3), performed as part of open-file project #373. A total of 65 samples were collected for assay, guided by the use of XRF. This project is also complete, and the final report (Frey and Hanson 2010) is available from the project Web page[8]. The spreadsheet 'p373_ddh_list' was downloaded in October of 2013, and converted to a point shapefile in ArcGIS.

The sixth dataset obtained from the MDNR was a gold-in-till survey of the Cook area (see Figure 3), conducted as part of its open-file project #392. Till is an unconsolidated mixture of soil, sand, gravel, and boulders deposited on top of bedrock by glaciers. This study is ongoing at the time of this writing, so no final report is yet available, but a spreadsheet containing sample locations and laboratory results was released to the public[5] in December of 2013. The spreadsheet 'cats_attribute_table' was downloaded in January of 2014, and converted to a point shapefile in ArcGIS.

The seventh and final MDNR dataset was another gold-in-till survey, conducted as part of its open-file project #379. For this study, 133 samples were collected from a project area named “Big Fork East,” in northeastern Itasca County (see Figure 3), and analyzed using scanning electron microscopy (SEM). The total counts of gold grains recovered from these samples are the highest of all MDNR gold-in-till surveys.

In both of the gold-in-till surveys, the tiny gold grains were extracted from each of the samples, and analyzed using scanning electron microscopy. Because gold is a soft, malleable metal, where the grains show signs of significant deformation it can be assumed that they were transported some distance from where they were originally weathered out of bedrock (Figure 5). Where they are described as “pristine,” it is likely that the bedrock source is nearby, generally within a few hundred meters. As such, examining gold grains in till samples can be a useful tool in searching for hidden gold deposits.

8 MDNR open-file project #373: http://www.dnr.state.mn.us/lands_minerals/mpes_projects/project373.html

5 MDNR open-file project #392: http://www.dnr.state.mn.us/lands_minerals/mpes_projects/project392.html

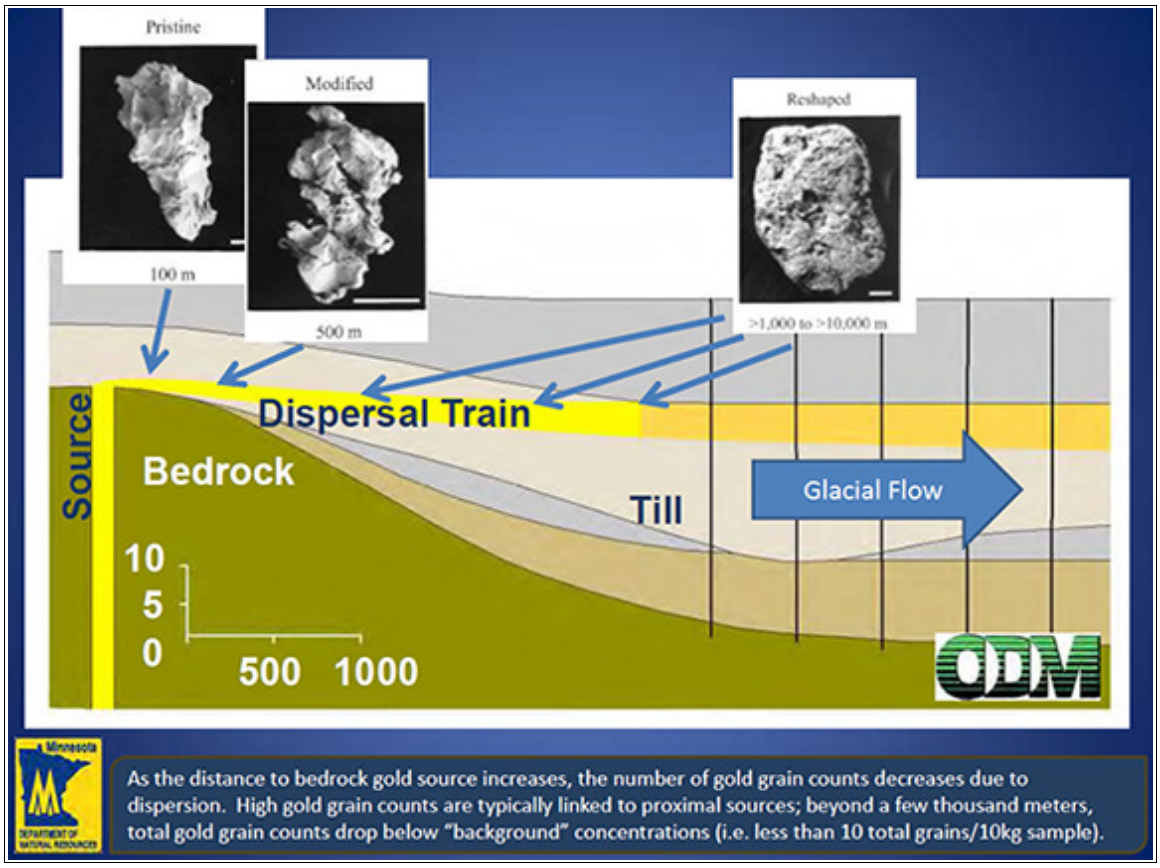


Figure 5: *Explanation of gold grain characteristics and transport distance, from a slide accompanying the data for open-file project #379.*

Very few of the gold grains in open-file project #379 were classified as pristine, but their shapes suggest they may have been part of a placer deposit before being redistributed by glacial ice. By contrast, the gold grains in open-file project #392 yielded several samples having more than five pristine gold grains. The latter project is also ongoing at the time of this writing, but the shapefile 'bfe_till_march22_2011.shp' (containing sample locations and laboratory results) was released to the public[6] in March of 2011, and downloaded in October of 2013.

3.2.2 Minnesota Geological Survey (MGS)

The first dataset obtained from the MGS was Map #S-22, Geologic Map of Minnesota, Precambrian Bedrock Geology. The metadata states that the published map scale is 1:500,000, although much of the northeastern portion was digitized from maps at larger scales. Overall, it is suggested that the map accuracy is close to 1:100,000, but variable, with horizontal accuracies

6 MDNR open-file project #379: http://www.dnr.state.mn.us/lands_minerals/mpes_projects/project379.html

varying from less than one to several hundred meters. Much of the variation comes from the western and southwestern parts of the state (i.e. far outside of the study area), where due to thickness of surficial cover, geological interpretations were made largely from geophysical surveys. The map contains several shapefile layers; the two used in this project are the bedrock unit polygons and fault polylines. This data was downloaded in October of 2013 from the MGS data portal[9].

The second MGS dataset used was the high-resolution aeromagnetics. This data was collected in the 1970s and 1980s, and originally published in 1992 (Chandler 2007). The data was significantly improved by line leveling and lost data recovery using Geosoft's Oasis Montaj software, and re-published in 2007. Grid files for the entire state were downloaded from the MGS aeromagnetic data portal[10] in October of 2013, imported into ArcGIS using the Geosoft plugin, and converted to Esri GRID format.

3.2.3 United States Geological Survey (USGS)

The USGS is in the process of completing a nation-wide geochemical survey, Open-File Report 2004-1001 (USGS 2004). Version 5 of this survey, which contains updated data for Minnesota among other locales, was released in September 2008. The survey consists mostly of stream-sediment, soil, and till samples that were subject to several laboratory analytical methods to describe their elemental components. A point shapefile containing the sample locations and results for the state of Minnesota was downloaded[11] in October of 2013 and imported into ArcGIS for processing.

9 MGS Map #S-22 download page: http://conservancy.umn.edu/handle/11299/154540

10 MGS aeromagnetic data, ftp download site: ftp://mgssun6.mngs.umn.edu/aero2007/GRID_DATA/

11 USGS soil geochemistry for MN: http://tin.er.usgs.gov/geochem/select.php?place=fUS27&div=fips

Chapter Four: Methods

The weights-of-evidence modeling method in ArcSDM is based in raster processing for all evidential layers. Accordingly, most vector datasets obtained from MDNR, MGS, and USGS, needed to be pre-processed into rasters (with the exception of the training sites, which are handled as points). During preprocessing, a detailed review of each dataset was performed, and specific data sub-selected as described below.

4.1 Data Preprocessing

4.1.1 Bedrock Geology

To define the study area, as described in the first chapter, polygons not assigned to the Wabigoon, Quetico, or Wawa subprovinces were excluded by selective query. The Minnesota River Valley rocks were similarly excluded from the map (see Figure 6). Most of the "holes" present in the resulting mask are mapped rock units (such as small intrusive bodies) that are either younger than the Archean bedrock or were formed as a result of unrelated processes; thus, were assigned as *not* belonging to the aforementioned subprovinces by the map creators.

The bedrock geology map was reduced to four classes, by assigning mapped rock units to the general rock-type categories that were introduced in Chapter 1. This reduction was deemed necessary because there are a limited number of training sites, and a smaller number of classes can help provide more stable weights (Bonham-Carter 1994). In a true data-driven approach, such classification would be done only after analyses of the weights tables. In this case, later analyses of the weights tables supported this knowledge-driven classification.

The bedrock unit polygons were then converted to an integer raster, using the Polygon to Raster tool in ArcGIS, with a raster cell size of 500 meters. The choice of cell size was based on the metadata for the bedrock units, which states that horizontal accuracy, especially in the western portion of the study area, is variable from less than 10 to several hundred meters. Much of the mapping in this portion was done based on geophysical interpretations, due to the depth of cover

and lack of outcrop, while the mapping in the eastern portion of the study area is much more accurate. A cell size of 500 meters provides a reasonable compromise.

Table 2: *Classification scheme for segregating bedrock units into four broad terranes: Greenstone (green), granitic plutons (pink), metasedimentary rocks (gray), and iron-formation (red).*

VALUE	MAPCODE	DESCRIPTION	TERRANE
1	Amv	Mafic metavolcanics; volcaniclastic and hypabyssal intrusions	Greenstone
2	Asd	Syenitic, monzodioritic, or dioritic pluton	Gran. Plut.
3	Agr	Granitic intrusion	Gran. Plut.
4	Avs	Volcanic and volcaniclastic rocks; felsic to intermediate	Greenstone
5	Agt	Tonalite, diorite and granodiorite	Gran. Plut.
6	Agn	Granitic to granodioritic orthogneiss	Greenstone
7	Agp	Gabbro, pyroxenite, peridotite, lamprophyre intrusion	Greenstone
8	Ags	Schist and tonalite- to granodiorite-bearing paragneiss	Metased.
9	Agm	Granite to granodiorite, variably magnetic	Gran. Plut.
10	Agd	Granodioritic intrusion	Gran. Plut.
11	Agu	Granitoid intrusion, undiff, poorly constrained by core / outcrop	Gran. Plut.
12	Ams	Schist of sedimentary protolith	Metased.
13	Aag	Mafic to ultramafic hypabyssal intrusives; gabbro, anorthosite	Greenstone
14	Asc	Conglomerate, lithic sandstone, graywacke, mudstone	Metased.
15	Aqs	Biotite schist, paragneiss, and schist-rich migmatite	Metased.
16	Aif	Iron-formation	Iron Fm.
17	Aql	Lac La Croix Granite; locally pegmatitic and magnetic	Gran. Plut.
18	Aqa	Amphibolitic schist and gneiss	Metased.
19	Aqg	Granite-rich migmatite; locally magnetic	Metased.
20	Adt	Tonalite, diorite and granodiorite	Gran. Plut.
21	Ast	Saganaga Tonalite	Gran. Plut.
22	Aqt	Tonalite- to granodiorite-rich migmatite	Gran. Plut.
23	Aks	KLG volcanogenic sandstone, siltstone, conglomerate, slate	Metased.
24	Akc	KLG volcanic conglomerate, breccia; alk., hornblende-bearing	Metased.
25	Auv	Ultramafic to mafic volcanic and hypabyssal intrusive rocks	Greenstone
26	Aqp	Porphyritic quartzofeldspathic dike	Gran. Plut.
27	Akv	KLG volcanic flows, breccia, and tuff; hornblende-bearing	Greenstone
28	Acv	Calc-alkalic volcanic and volcaniclastic rocks	Greenstone
29	Aqm	Quartz monzonite, monzonite, granodiorite; non-magnetic	Gran. Plut.
30	Asg	Graywacke and mudstone; typically greenschist facies	Metased.
31	Amm	Interlayered volcanic and volcaniclastics; amphibolite grade	Greenstone

4.1.2 Selection of Training Sites

The training sites for this project represent known gold occurrences, interpreted from two types of field investigations: drill holes and till samples. Table 3 lists the selected training sites, and

their associated gold showings.

Gold intercepts in drill holes were identified using the reports by Severson (2011), Frey (2012), and Frey and Hanson (2010). Gold in till samples were selected from MDNR open-file projects 392 and 379, where till samples contained high counts of either total or pristine gold grains (normalized to a 10-kg sample). Only one training site per till sampling program was selected, because realistically, using these to represent gold in bedrock involves significant and questionable assumptions.

Table 3: *Training sites selected from drill holes and gold-in-till samples.*

DDH / Sample #	Area / prospect	Intercepts / Au assays
SH-1	Western Vermilion District / Raspberry	Several, up to 6,240 ppb
6314-36-2	Western Vermilion District / Foss Lake	Several, up to 3,110 ppb
LL-1	Cook area / Lost Lake (Newmont)	1 unspecified, 10,000 ppb
LL-87-13	Itasca County / Lost Lake	6 >1,000 ppb, max 5,400 ppb
GBD-1	Itasca County / Gale Brook	1 interval, 360 ppb over 10'
RR-1	Rainy Lake / R.L. – Seine River fault zone	2 intervals, max 3,560 ppb
SS-2A	Rainy Lake / Seattle Slew	1 at 2,780 ppb
FT-14	KBRLW	4 intervals, up to 2,745 ppb
FT-21	KBRLW / Hero Deposit	3 intervals, up to 1,440 ppb
IND-1	KBRLW	Several, >500 ppb
WF-2	KBRLW / Winterfire	Several, >500 ppb
ND-2	Rainy Lake (MDNR open-file project #378)	799 ppb (XRF)
EN-5	Eagle's Nest Shear (MDNR open-file project #373)	4,000 ppb (XRF)
SXL-4	Murray Shear (MDNR open-file project #373)	40,000 ppb (XRF)
CATS-202	Cook area (gold-in-till, MDNR project #392)	13.8 pristine gold grains (norm.)
BF-08	"Big Fork East" (gold-in-till, MDNR project #379)	143.5 total gold grains (norm.)

In total, there are sixteen training sites, each reasonably indicating the location of anomalous bedrock gold mineralization. Five of the six gold exploration areas depicted in Figure 3 are represented, Virginia horn being the only area excluded. There were three drill holes with suitable gold intercepts in this area, but their descriptions did not match the mapped geology, and were excluded to preserve relational accuracy between the training sites and the geologic map. Figure 6 shows the final training site locations against the four-class geologic map of the study area.

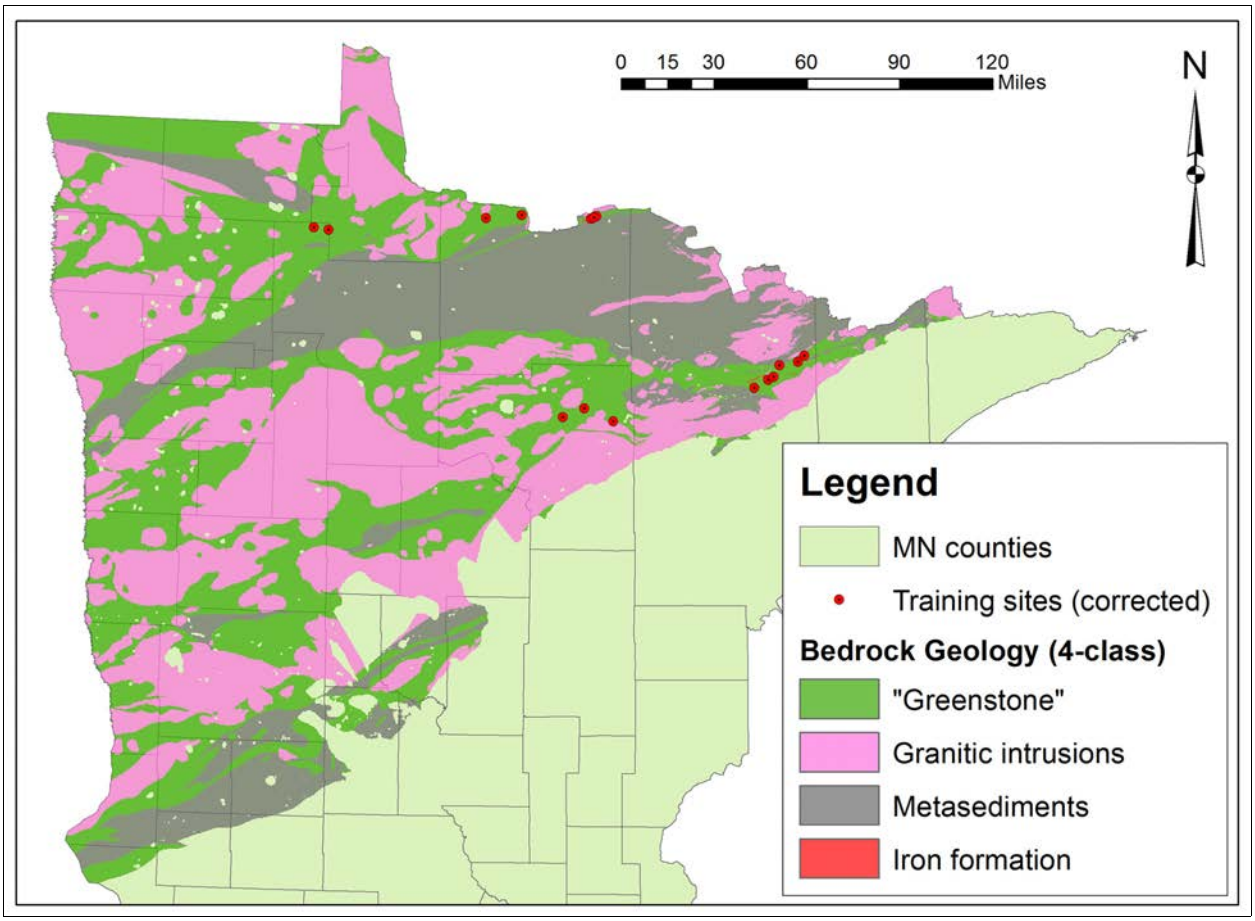


Figure 6: *Selected training sites, shown with the four-class geologic map raster.*

4.1.3 Major Fault Proximity

Among the best prospectivity criteria for Archean gold deposits is proximity to large, regional-scale, steeply-dipping fault and shear zones (Klein and Day 1994). These significant structures were extracted from the geologic map, and buffered for use as evidence.

Major faults were extracted simply by selecting those that had been assigned names in the attribute table. While this is was a simplistic method, it provided fairly accurate results, in that the named faults are large, and roughly outline the subprovince and major terrane boundaries (see Figure 2). For a more thorough approach, it would probably be best to consult an expert on northern Minnesota's Precambrian geology.

The major faults were buffered using the Feature Proximity tool in ArcSDM, which creates an integer raster representing distance from each fault line. The result was a raster with 100-meter cell size, segregated into 15 classes, each representing a proximity of 5 kilometers (Figure 7).

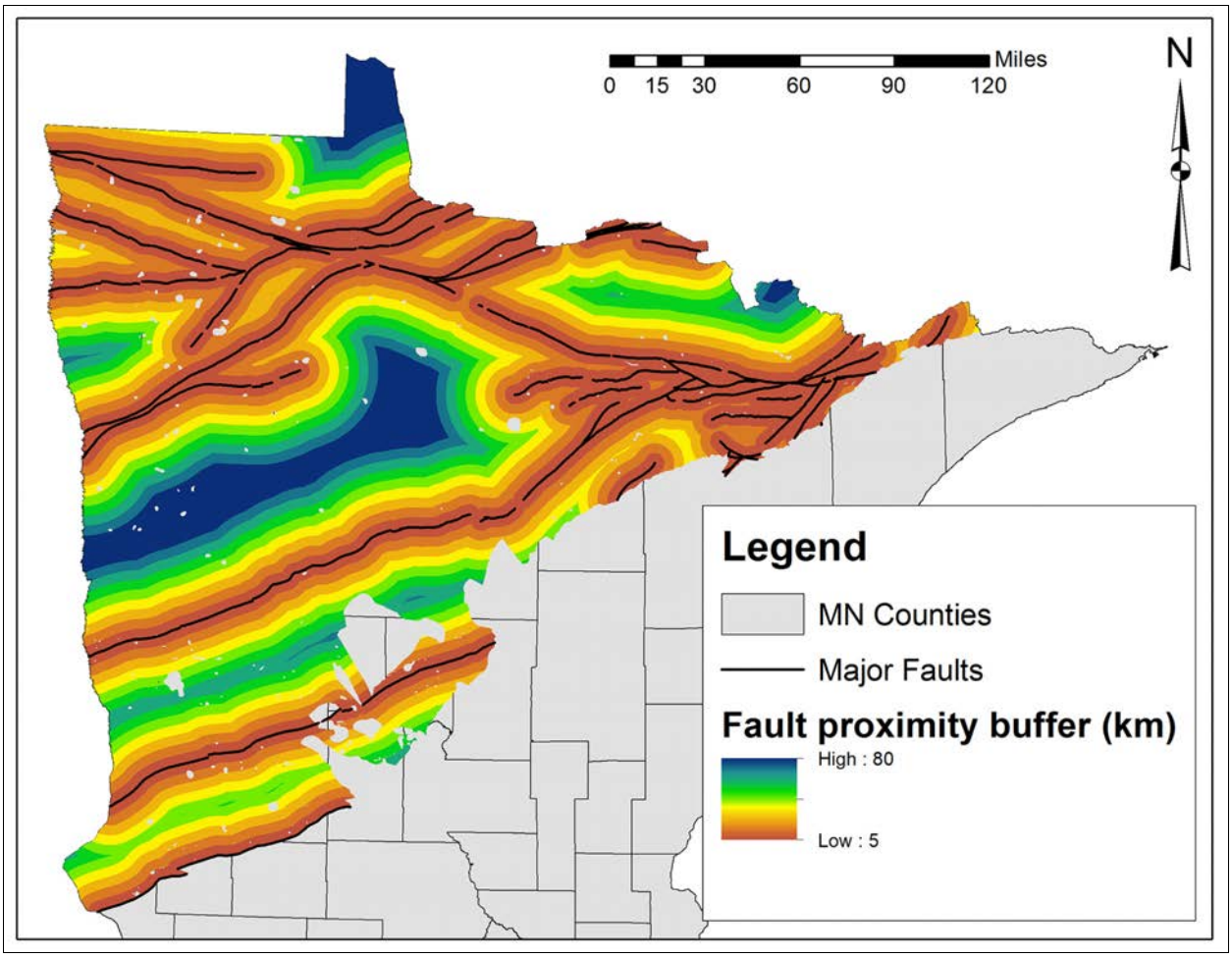


Figure 7: *Major faults and shears within the study area, and proximity buffers in 5-km increments.*

4.1.4 Aeromagnetics

Of particular interest to this project is the high-resolution (1:24,000) aeromagnetic data collected by the MGS. Because it was first published in 1991, and significantly updated in 2007, this data was not available during the intensive gold exploration campaigns in the 1980s.

Anomalous magnetic lows may indicate areas where magnetic minerals have been destroyed by hydrothermal fluids, such as those that are sometimes driving mechanisms of gold mineralization (Robb 2005). On the other hand, magnetic highs may indicate iron formations, or other types of rocks which are sometimes favorable hosts of gold mineralization.

The aeromagnetic dataset (once converted to Esri GRID format) was masked to the study area and subjected to the Zonal Statistics tool in ArcGIS (Spatial Analyst), using the MEAN

parameter. The vector geologic map was used, with the ObjectID (OID) of each polygon, to define zones coincident with bedrock units. The zonal mean raster was then subtracted from the original aeromagnetic raster (Figure 8), using the Raster Calculator tool in ArcGIS (Spatial Analyst).

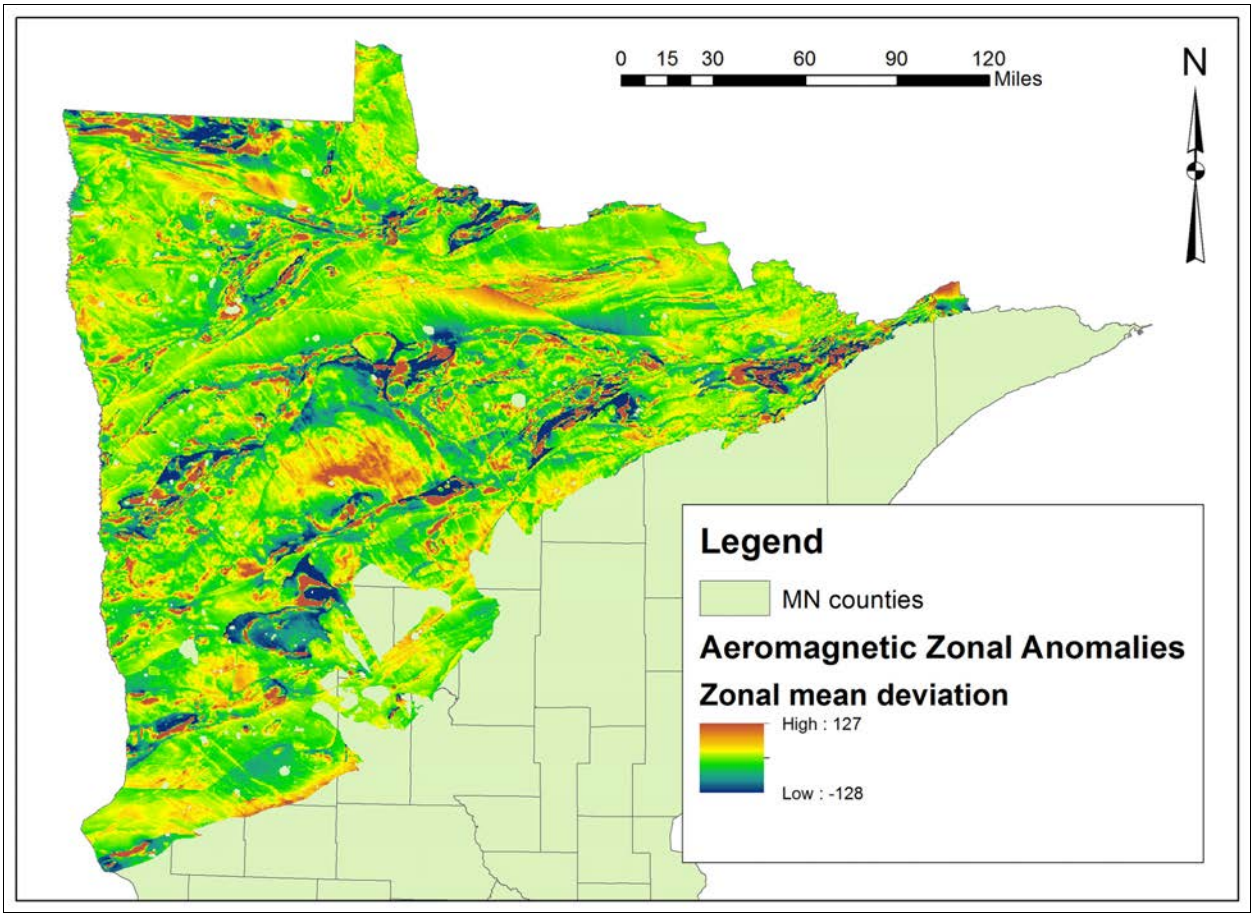


Figure 8: *Aeromagnetic zonal anomaly map, converted to integer form.*

Because ArcSDM requires an integer raster, the zonal anomaly map was reclassified with the Reclassify tool in ArcGIS (Spatial Analyst). Negative values represent cells lower than the zonal mean, and positive values represent cells higher than the zonal mean. Overall, the range of values goes from -128 to +127, in increments of 1 (256 total values).

4.1.5 Geochemistry

As part of a large, national effort, geochemical samples for Minnesota were collected in four programs, each as a joint effort between the USGS and the MGS[12]. The "Gold Analysis" column in Table 4 is in reference to the analytical method used to analyze each sample for gold. The acronym

12 USGS description of sampling for Minnesota: http://mrdata.usgs.gov/geochem/doc/groups-cats.htm#states2004

FA-AAS refers to "fire assay – atomic absorption spectrometry," which has a lower detection limit (LDL) of 5 ppb for gold[13]. All but 28 of the 243 soil samples report values below the LDL, so for this project the LDL was ignored.

Table 4: *Overview of Minnesota sample points in the USGS geochemistry dataset.*

Sampling Program	Num. Samples	Sample Type	Gold Analysis
States_2004	379	Soil, stream sediments	FA-AAS
States_2006a	563	Soil, stream sediments	FA-AAS
States_2006b	230	Soil, stream sediments	FA-AAS
States_2007	221	Till	FA-AAS

The 2007 program was entirely focused on glacial till sampling. The 2004, 2006a, and 2006b programs were a mix of soil and stream sediment samples. Because most of the samples were soil samples, these points were extracted, joined, and clipped to the study area (Figure 9).

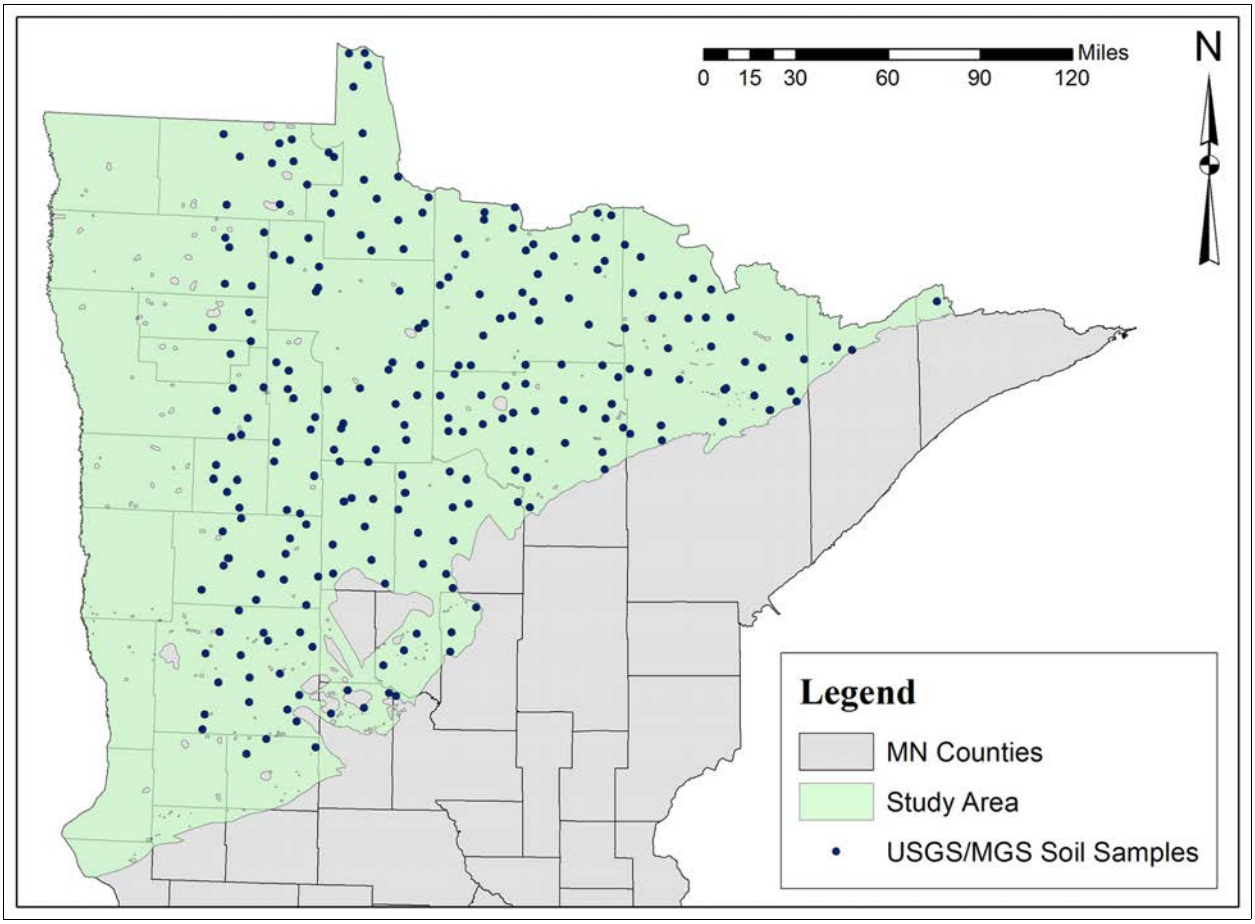


Figure 9: *Locations of soil samples in the USGS geochemistry dataset.*

13 SGS description of FA-AAS analytical method: http://tin.er.usgs.gov/geochem/doc/pge.htm

For most sample points, two or more samples were collected, at different depths. This can cause problems with analyses, as shallow samples can indicate higher gold values than deeper samples, due to the scavenging and concentrating of metals due to organic processes (Carranza, 2009a; Moon, Whately, and Evans 2006). For consistency, samples taken from a depth of less than 12 inches (30.5 cm) were removed.

Carranza (2009a) describes two types of methods for creating surfaces from point data: *interpolated* and *non-interpolated*. Interpolated methods rely on basic techniques such as contouring, or geostatistical techniques using weighted-moving-average methods, such as inverse-distance weighted (IDW) and kriging. By contrast, non-interpolated methods are used to create surfaces without interpolation, either by aggregating point samples into existing vector polygons, or creating polygons around the points using common GIS tools.

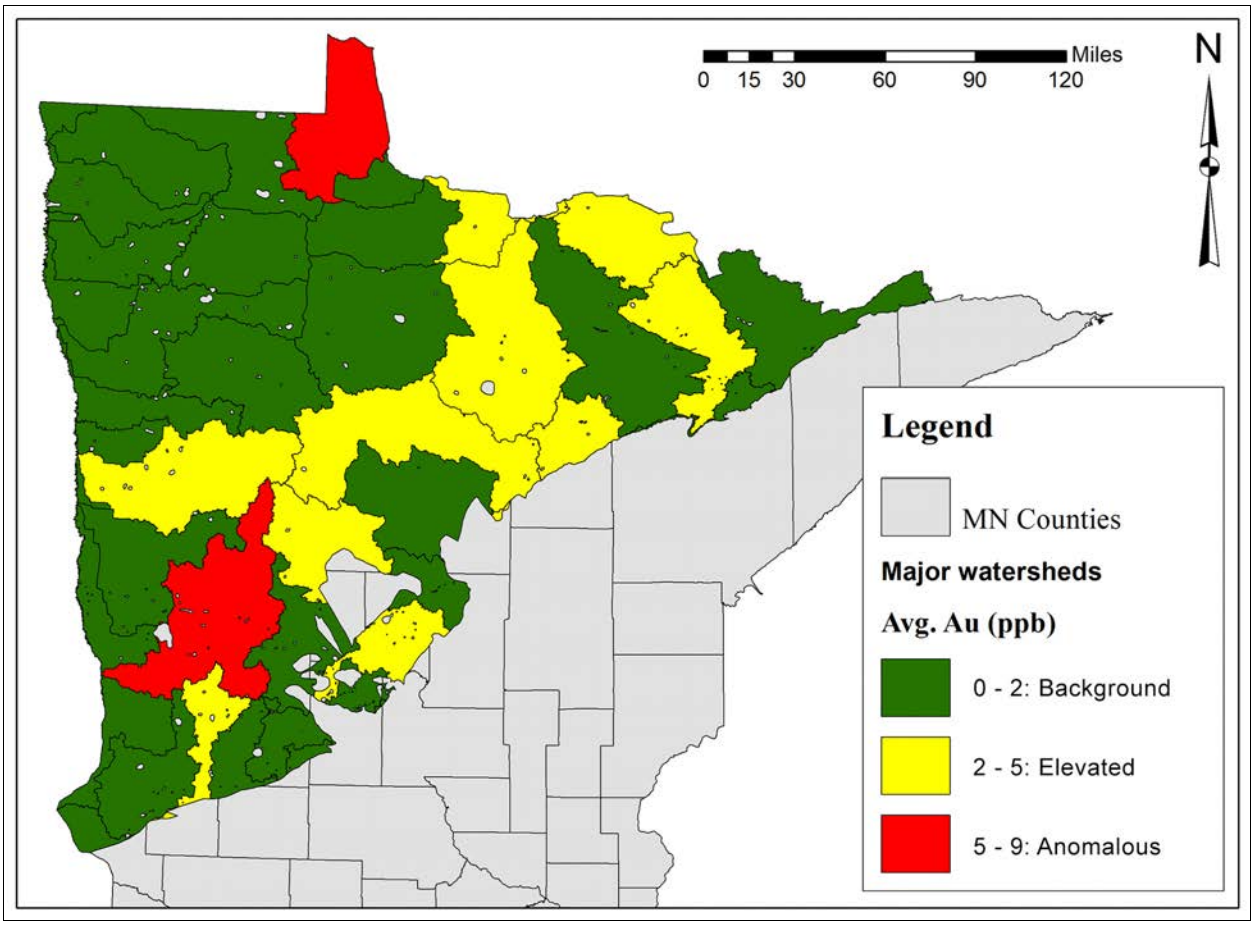


Figure 10: *Gold-averaged major watersheds, organized into three classes.*

Both IDW and kriging were tested, but due to low sample density, the coarse resolution of the interpolated raster was deemed insufficient for model input. So for this project, two non-interpolated methods were used to create evidential themes from geochemistry data. The vector polygons used here are major watersheds obtained from MDNR, and *Theissen polygons,* which are created automatically based on the spatial distribution of the points themselves.

The map in Figure 10 shows the major watersheds within the study area, arbitrarily classified into three categories based on the average gold values in the soil samples. Several watersheds had no samples, and were lumped into the first category (in green).

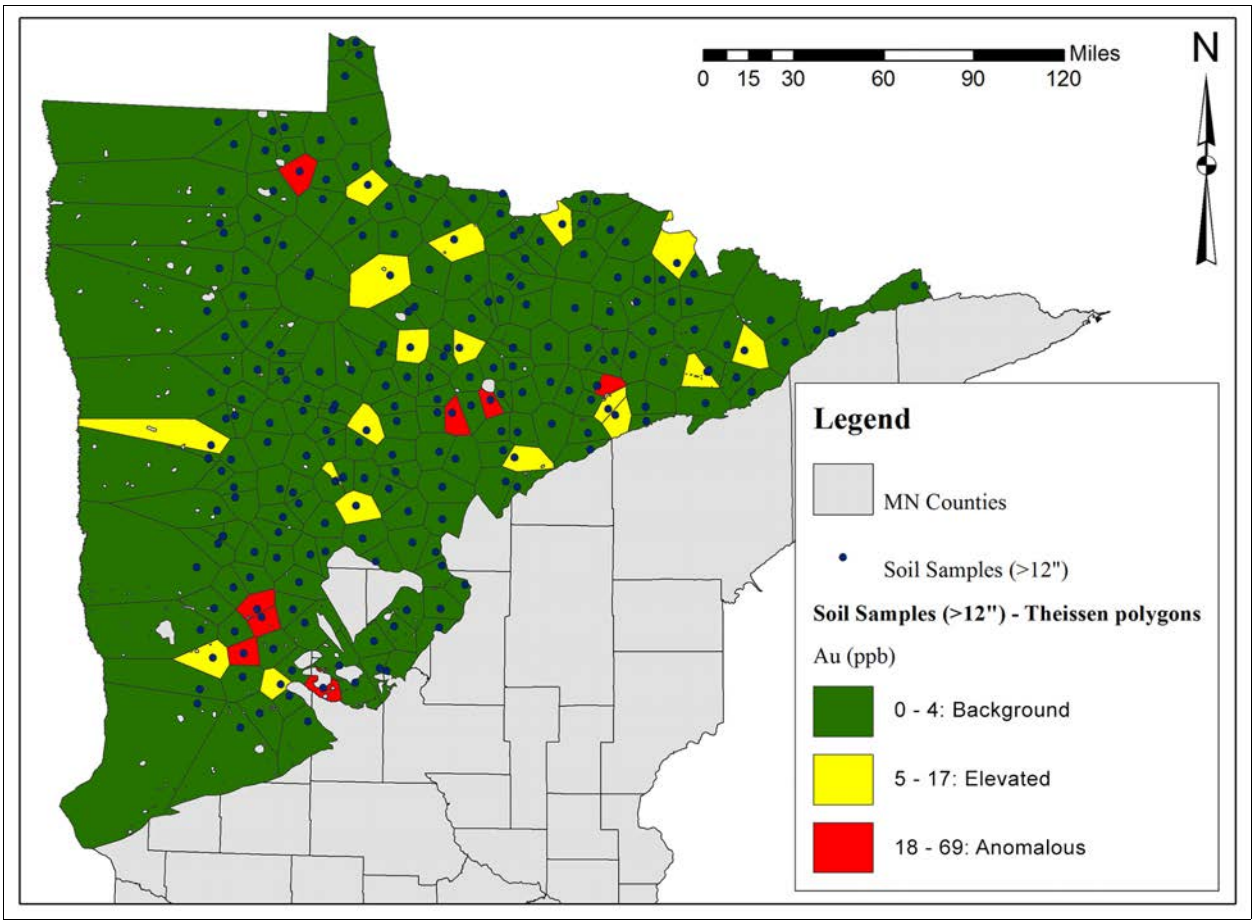


Figure 11: *Geochemistry theme derived by creating Theissen polygons around each sample point, organized into three classes based on box-plot statistics.*

In contrast, figure 11 shows the result of the Theissen polygon method, again classified into three categories. The first category (in green) represents sample points with gold values below the 5-ppb LDL. The second category (in yellow) represents gold values above the LDL, but below the

18-ppb anomaly threshold. This threshold was chosen by averaging the *upper inner fence* (UIF) values, derived through the application of box-plot statistics from exploratory data analysis (Carranza 2009a) to each set of soil samples (Guiterrez et al. 2012). The third category (in red) represents samples having gold values at or above the anomaly threshold. Both geochemistry themes were converted to integer rasters having a 100-meter cell size.

4.2 Methods

Each evidential theme was analyzed in order to quantify the spatial relationships between it and the training sites, using the Calculate Weights tool in ArcSDM. Then, the evidential themes were combined using the Calculate Response tool in ArcSDM, producing maps of posterior probability; i.e. the probability of unknown gold occurrences considering all the evidence.

A weights-of-evidence modeling term that is appropriate to discuss here is the *confidence level of Studentized contrast,* an input chosen by the user when running the Calculate Weights tool. Recall that the Studentized contrast is the ratio of the contrast (C) between W+ and W-, to the standard deviation of the underlying weights, or [C / σC]. For this project, a regional-scale ranking of prospectivity based on a limited number of training sites, a low confidence level is acceptable. For a larger-scale project with more specific requirements and suitable data, a higher confidence level might be selected. Here, a 70% level of confidence was arbitrarily chosen as a threshold, which roughly equates to a Studentized contrast of 0.542 (Sawatzky et. al. 2010b).

For cumulative weights measures, such as those calculated for the fault proximity raster, the Calculate Weights tool assigns a GEN_CLASS value of 2 to patterns that are associated with training sites, or have positive W+ values. Conversely, it assigns a GEN_CLASS value of 1 to patterns that are not associated with training sites, or have negative W- values. Either can be acceptable for the model as long as the contrast satisfies the Studentized contrast confidence level. GEN_CLASS is abbreviated as “G_C” in the tables shown here in this chapter.

For categorical data, such as the four-class geologic map, the GEN_CLASS value is the same

as the class value in the evidential theme, where this confidence criteria is met. For those classes where the confidence criteria is not met, the GEN_CLASS value is either 99 or some other value that differs from any class value in the raster. All patterns or areas that do not meet the confidence level are treated as a single class. Where a class does not contain any training sites, it is arbitrarily given a small fraction of a training site for the purpose of calculating a weight, so the program does not encounter a division by zero error, and the patterns can still be used in the model.

The key WEIGHT and WEIGHT_STD (standard deviation of WEIGHT) columns are used later by the Calculate Response tool. In the following tables, these weights are either positive where the pattern is associated with training sites, or negative where it is not. The W+, W-, Contrast (C), and Studentized contrast (displayed as "St. C." here) are useful for making interpretations.

4.2.1 Analysis of Weights Tables

The geochemical rasters were reduced to three classes based on expert decisions; they were analyzed using cumulative-descending weights because they still represent ordinal data, a type of ordered data. The hypothesis was that training sites should be associated with areas of elevated or anomalous gold (class 2 and/or 3), and not areas classified as background (class 1).

In both the gold-averaged watersheds and the Theissen polygon themes, no training sites occupied areas classified as anomalous because there were only a few training sites in a large study area, of which the Anomaly class is a small portion. For modeling purposes, the Anomaly was lumped together with the Elevated class, which does contain training sites. These two classes lumped together are positively correlated with the training sites, as indicated by a positive contrast that exceeds the confidence limit of 0.542 (70%) for both analysis themes (Tables 5 and 6). Based on its higher contrast, the gold-averaged watersheds theme is the stronger predictor of training sites.

Table 5: *Cumulative-descending weights for the gold-averaged major watersheds theme.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
3	Anomaly	7833.05	0	0	0	0	0	2	0.5043	0.3162
2	Elevated	34553.03	10	0.5043	-0.5069	1.0112	1.9582	2	0.5043	0.3162
1	Background	91541.5	16	-0.0006	10.0382	-10.0388	-0.7097	1	-0.5069	0.4083

Table 6: *Cumulative-descending weights for the Theissen polygon geochemistry theme.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
3	Anomaly	1480.87	0	0	0	0	0	2	0.8723	0.5774
2	Elevated	7174.86	3	0.8723	-0.1260	0.9983	1.5586	2	0.8723	0.5774
1	Background	91541.51	16	-0.0006	10.0382	-10.0388	-0.7097	1	-0.1260	0.2774

The hypothesis for the aeromagnetic zonal anomaly map was that gold occurrences are associated with both high and low anomalies, as explained in Chapter 3. Therefore, the aeromagnetic zonal anomaly raster was analyzed twice: once with cumulative-descending weights to test the association between the training sites and high anomalies (Table 7), and again with cumulative-ascending weights to test the association between the training sites and low anomalies (Table 8). In both cases, only the extreme highest and extreme lowest zonal anomaly classes appear to be associated with training sites. So, the results are in support of the hypothesis about this data.

Table 7: *Cumulative-descending weights for the aeromagnetic zonal anomaly theme. Only the top two rows are shown, as all but the highest zonal anomaly class have negative weights.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
127	Highest anomaly	1000.25	4	3.1254	-0.2766	3.4020	5.8903	2	3.1254	0.5003
126	Next highest	1012.75	4	3.1130	-0.2765	3.3895	5.8685	1	-0.2766	0.2887

Table 8: *Cumulative-ascending weights for the aeromagnetic zonal anomaly theme. Only the top 10 rows are shown, as all but the nine lowest zonal anomaly classes have negative weights.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
-128	Lowest anomaly	685.5	3	3.2157	-0.2001	3.4157	5.3304	2	3.3447	0.5003
-127	Next lowest	696	3	3.2004	-0.2000	3.4004	5.3065	2	3.3447	0.5003
-126	Next lowest	711.25	3	3.1787	-0.1998	3.3785	5.2725	2	3.3447	0.5003
-125	Next lowest	724.25	3	3.1606	-0.1997	3.3603	5.2440	2	3.3447	0.5003
-124	Next lowest	737.25	3	3.1428	-0.1995	3.3423	5.2160	2	3.3447	0.5003
-123	Next lowest	751.5	3	3.1236	-0.1994	3.3230	5.1859	2	3.3447	0.5003
-122	Next lowest	768.5	3	3.1012	-0.1992	3.3004	5.1507	2	3.3447	0.5003
-121	Next lowest	785.5	3	3.0793	-0.1990	3.2783	5.1163	2	3.3447	0.5003
-120	Next lowest	803.5	4	3.3447	-0.2788	3.6235	6.2731	2	3.3447	0.5003
-119	Next lowest	819.5	4	3.3249	-0.2786	3.6036	6.2387	1	-0.2788	0.2887

The aeromagnetic anomaly map also was reduced to three classes based on the results of the cumulative weights tables. Class 1 represented low anomalies, class 2 represented “background”

data, and class 3 represented high anomalies. Once completed, the resulting three-class raster (Figure 12) was analyzed again using categorical weights (Table 9). This allowed both significant high and low anomalies to be included in the model, as both appear to have some association with the training sites. This fact could be reflective of the fact that no specific type of gold occurrence was sought while selecting training sites. As such, the training sites may represent diverse styles of gold mineralization that occur in a variety of rock types, having different magnetic properties.

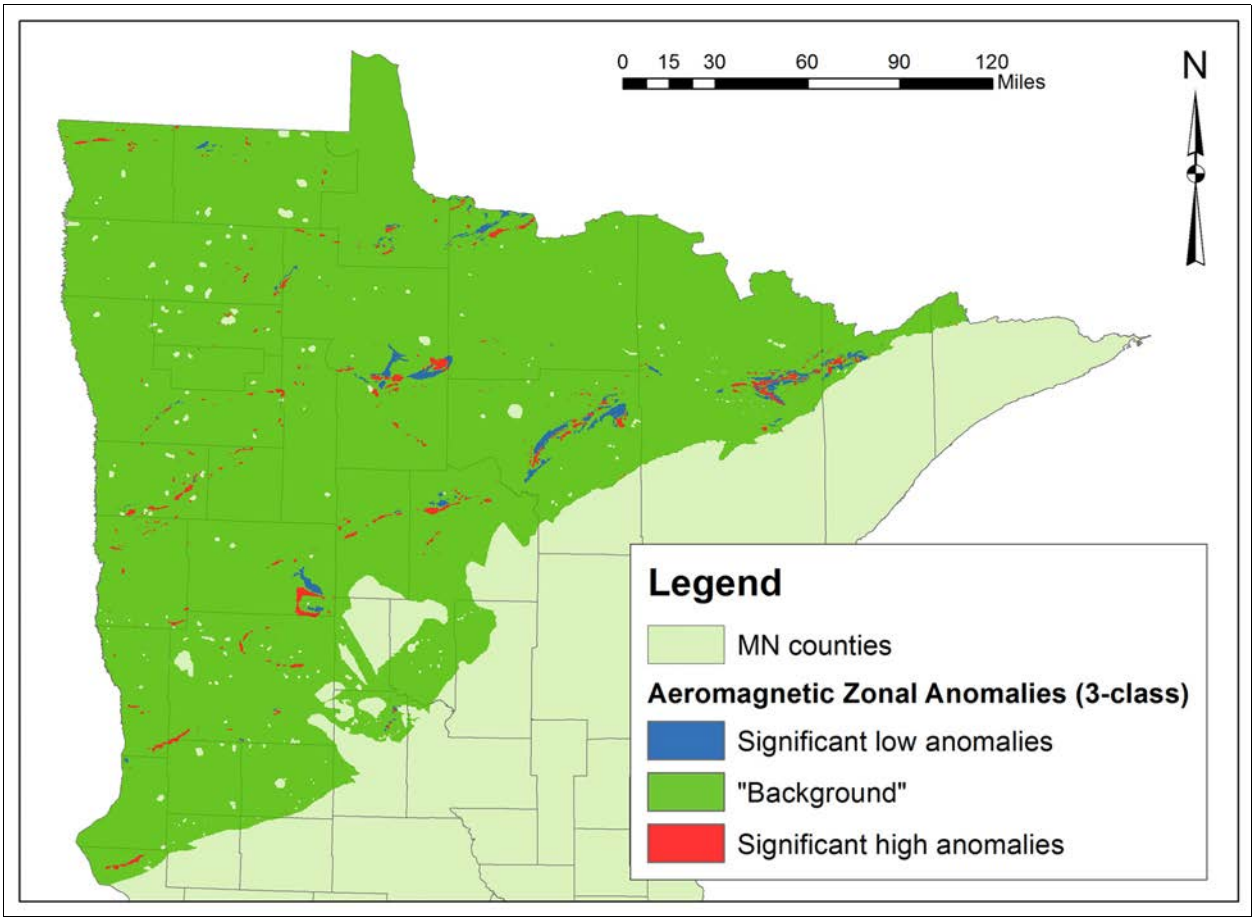


Figure 12: *Improved three-class aeromagnetic anomaly map of the study area.*

The categorical weights derived for the three-class aeromagnetic zonal anomaly raster (Table 9) confirm that, indeed, both high and low anomalies are positively associated with the training sites, the low anomalies somewhat more so, indicated by a slightly higher contrast. The analytical process of considering both cumulative ascending and descending weights allowed the data to define how to deal with ordered evidence where a correlation was expected with both high and low values.

Table 9: *Categorical weights for the three-class aeromagnetic anomaly raster.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
1	Low anomaly	803.5	4	3.3449	-0.2788	3.6237	6.2735	1	3.3449	0.5003
2	“Background”	89225.25	8	-0.6732	3.2292	-3.9024	-7.8026	2	-0.6732	0.3536
3	High anomaly	1000.25	4	3.1256	-0.2766	3.4022	5.8906	3	3.1256	0.5003

The four-class bedrock geologic map was evaluated using categorical weights, as the map units are nominal data. The class values assigned are based on a category of rock types, which are not orderable as any ranked or measured value. Greenstone and iron formation both have a strong positive correlation with the training sites, while granitic plutons and metasediments do not (Table 10). The very small area of mapped iron formation provide a high density of training points, and thus a high W+, because this measure is dependent on area and number of training sites.

Table 10: *Categorical weights for the four-class geologic map raster.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
1	“Greenstone”	30781.5	13	0.8793	-1.2626	2.1419	3.3440	1	0.8793	0.2774
2	Granitic plutons	38448.25	0	0	0	0	0	99	-6.9660	10
3	Metasediments	22011.75	0	0	0	0	0	99	-6.9660	10
4	Iron formation	23.75	3	6.6120	-0.2074	0.2774	10.5084	4	6.6120	0.5867

The zero value of Contrast (C) does not mean, in practice, that granitic intrusions are not prospective. For example, the so-called Viking Porphyry in the Virginia Horn area is host to a small, low-grade gold deposit (Severson, 2011). This would be labeled as a granitic intrusion according to the four-class scheme, but there were no training sites in this area.

The fault proximity theme was a 15-class raster, each representing a buffered distance of 5 km from the major faults extracted in chapter 4. The distance classes are ordinal data, in that they are ordered from nearest to furthest from the major faults and shears. As such, cumulative-ascending weights were calculated. Table 11 shows the weights for the proximity patterns calculated using this method.

Table 11: *Cumulative-ascending weights for the fault proximity theme.*

CLASS	DESCRIPTION	AREA km2	# PTS	W+	W-	C	St. C	G_C	WEIGHT	WEIGHT STD
5	Prox 0-5 km	23665.76	13	1.1452	-1.3749	2.5201	3.9345	2	0.5444	0.2501
10	Prox 5-10 km	40042.28	15	0.7624	-2.1974	2.9598	2.8658	2	0.5444	0.2501
15	Prox 10-15 km	53079.82	16	0.5444	-6.5107	7.0551	0.7053	2	0.5444	0.2501
20	Prox 15-20 km	62968.84	16	0.3735	-6.2135	6.5870	0.6585	1	-6.5107	10
25	Prox 20-25 km	71327.27	16	0.2489	-5.8674	6.1163	0.6114	1	-6.5107	10
30	Prox 25-30 km	77794.18	16	0.1621	-5.4819	5.6440	0.5642	1	-6.5107	10
35	Prox 30-35 km	82685.26	16	0.1011	-5.0421	5.1433	0.5142	1	-6.5107	10
40	Prox 35-40 km	85608.05	16	0.0664	-4.6416	4.7080	0.4707	1	-6.5107	10
45	Prox 40-45 km	88215.06	16	0.0364	-4.0629	4.0993	0.4098	1	-6.5107	10
50	Prox 45-50 km	90120.19	16	0.0150	-3.2126	3.2276	0.3227	1	-6.5107	10
55	Prox 50-55 km	90955.93	16	0.0058	-2.3259	2.3316	0.2331	1	-6.5107	10
60	Prox 55-60 km	91143.6	16	0.0037	-1.9395	1.9432	0.1943	1	-6.5107	10
65	Prox 60-65 km	91282.13	16	0.0022	-1.5115	1.5138	0.1513	1	-6.5107	10
70	Prox 65-70 km	91413.02	16	0.0008	-0.8091	0.8099	0.0810	1	-6.5107	10
75	Prox 70-75 km	91515.41	16	-0.0003	0.7849	-0.7852	-0.0785	1	-6.5107	10
80	Prox 75-80 km	91541.51	16	-0.0006	10.0382	-10.039	-0.7097	1	-6.5107	10

Since these faults are controls on gold concentration (Klein & Day, 1994), a “cutoff” distance class can be selected by analysis of Table 11. Exploration efforts can generally be limited to lie within this cutoff distance. One way to visualize this cutoff range is with a contrast curve, such as described by Bonham-Carter (1994). Figure 13 shows the contrast curve from Table 11. The contrast value ([W+ - W-]) is plotted on the y-axis, against ascending distance class on the x-axis. The contrast peaks at 15 km, and also exceeds the confidence criterion of 0.542 (70%) specified for the Studentized contrast. It is important to note, however, that the highest contrast does not always satisfy the confidence limit, as it does here.

Where the magnitudes, i.e. absolute values, of W+ and W-, are approximately equal, as in class 5, proximity is doing a good job of differentiating between areas associated and not associated with the training sites. At classes 10 and 15, W- is much larger than W+, indicating these areas are less associated, but still contain some training sites. At even higher classes, beyond 15 km, W+ is small, W- is large, and the WEIGHT values are negative, so areas beyond this range will be down-weighted by this piece of evidence.

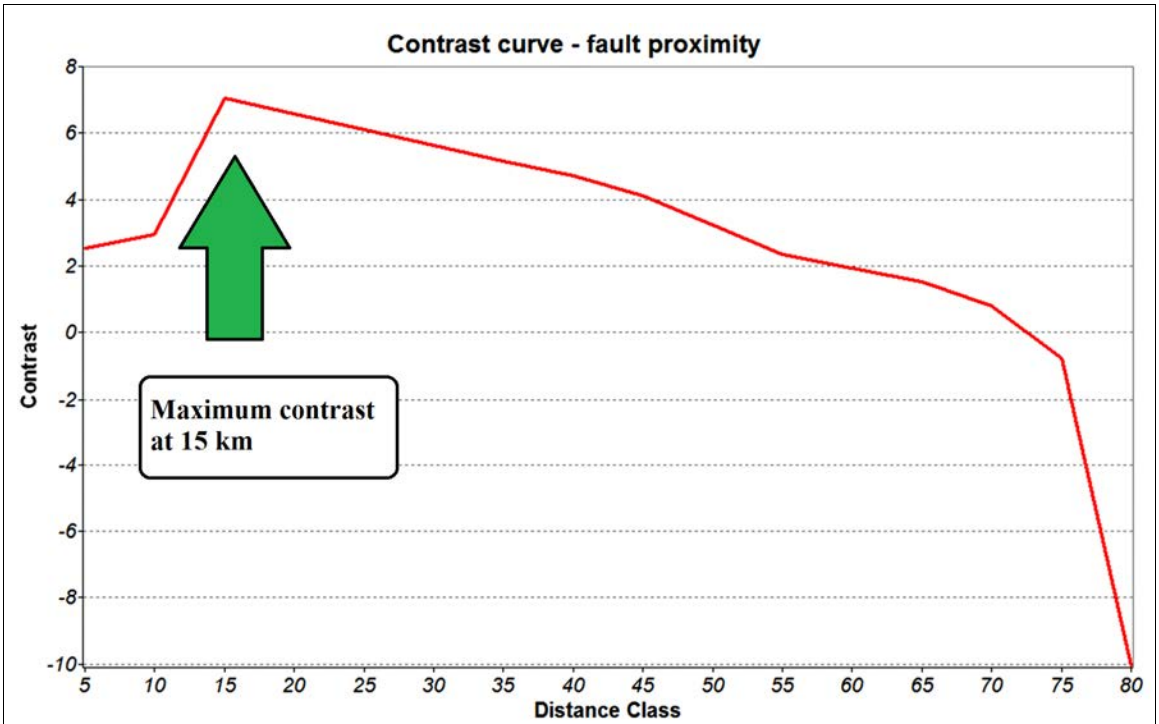


Figure 13: *Contrast curve for fault proximity, cumulative-ascending weights.*

Table 12 provides a summary of prospectivity criteria for gold in the study area, derived from analysis of the weights tables for each evidential theme. The table is sorted by contrast, and thus from top to bottom, lists the evidence from strongest to weakest correlation with the training sites.

Table 12: *Summary of weights tables, providing prospectivity criteria for gold within the study area.*

PROSPECTIVITY CRITERIA	AREA km2	C	St. C	WEIGHT	WEIGHT STD
Within 15 km of faults	53079.82	7.0551	0.7053	0.5444	0.2501
Magnetic low anomalies	803.5	3.6237	6.2735	3.3449	0.5003
Magnetic high anomalies	1000.25	3.4022	5.8906	3.1256	0.5003
Bedrock: "Greenstone"	30781.5	2.1419	3.3440	0.8793	0.2774
Watershed Au: Elevated +	34553.03	1.0112	1.9582	0.5043	0.3162
Theissen Au: elevated +	7174.86	0.9983	1.5586	0.8723	0.5774
Bedrock: Iron formation	23.75	0.2774	10.5084	6.6120	0.5867

Fault proximity is the most strongly correlated, but since the area within 15 km is over 53,000 km^2, or approximately 58% of the study area, this weight is the second weakest of the set. In this way, the fault proximity evidence suggests where exploration efforts should not be focused

(beyond 15 km), but does not strongly suggest where efforts should be directed. Thus, fault proximity can be considered as a type of *negative evidence*.

By comparison, aeromagnetic anomalies have a lower contrast than fault proximity, but because the area covered by high and low anomalies is approximately 2% of the study area, the weights for these are very strong. Aeromagnetic anomalies are thus areas where exploration efforts should be focused, and can be considered *positive evidence*.

Similarly, the highest weight of the set comes from iron formations in the four-class geologic map theme, but only because the area covered by this rock type is very small, at just under 24 km^2. However, the W+ is significantly larger than W-, so this evidence suggests that while iron formations may be good places to look for gold occurrences, other areas are not necessarily bad places to look. The other rock type that is associated with training sites, the greenstone class, has a W- that is much larger than W+ in magnitude. So the bedrock geology map suggests that iron formations and greenstones are good places to look for gold occurrences, and the other two rock types are not.

The gold-averaged major watersheds theme has W+ and W- values that are approximately equal in magnitude, so this evidence does a good job at differentiating between areas that are associated with training sites and areas that are not. The Theissen polygon theme has a W+ that is much higher in magnitude than W-, suggesting that areas classified as elevated or anomalous are good places to look, but areas outside of these are not necessarily bad. As such, the gold-averaged major watersheds provides the most useful evidence of the two geochemistry themes.

4.2.2 Model Creation

Two weights-of-evidence models were created, each using the fault proximity theme, three-class aeromagnetic anomaly theme, and four-class geologic map theme. The models differed only in geochemistry, with Model 1 using the gold-averaged major watersheds theme, and Model 2 using the Theissen polygon theme.

In weights-of-evidence modeling, the *efficiency of classification* is the rate at which the model accumulates area, starting with cells having the highest posterior probability, versus the rate at which it accumulates training sites. If the model accumulates area and training sites at an equal rate, the efficiency of classification is 50%, or no better than a random process (Chung and Fabbri 2003). If the model accumulates training sites at a faster rate than area, the efficiency measure is above 50%, indicating that the model performed well, or at least better than random.

The *efficiency of prediction* is an analogous test, where instead of training sites, a set of validation sites, which the model had not "seen," is considered. In this case, a set of 145 gold-targeting exploratory drill holes was used in a *blind test*. These holes were drilled specifically by private companies seeking gold, and the set does not include holes used as training sites. While it is not known whether these holes actually intercepted gold, the idea is private companies will only expend capital on drilling if there is sufficient evidence to justify its use. However, some of the companies might have been exploring for something quite different than those represented by the training sites. As such, while not a “true” blind test, these drill holes provide a simple means of testing the validity of model results.

The additional Spread column is simply the difference between the efficiencies of classification and prediction, intended as a means of testing the consistency of each model's results. Large differences between the classification and prediction rates may indicate a problem with the model, or simply that the blind-test drill holes were chasing some type of target other than represented by the training sites and evidence used here.

Table 13: *Overview of model results and variable geochemistry themes.*

Model #	Geochem. Theme	Eff. Of Classification (%)	Eff. Of Prediction (%)	Spread
1	Major watersheds	96.6	65.8	30.8
2	Theissen polygons	96.5	65.5	31.0

The efficiency measures are generated by the Area-Frequency Table tool in ArcSDM. Because the limit of precision for posterior probability values in ArcSDM is 0.000001 (10^{-6}) and

there are large areas with posterior probabilities less than this. The Con tool in ArcGIS was used to convert all values below the precision limit to zero, for post-modeling analysis purposes. This ensures that the area-frequency tables are not populated with inaccurate values caused by values below the precision limit. If ignored, values for efficiency of classification and prediction can become skewed. The areas with posterior probability less than 10^{-6} also had extremely low Confidence levels, indicating that these extremely low posterior probabilities could not be differentiated from a value of zero.

From the area-frequency table, a *cumulative area-posterior probability curve (*CAPP curve*)* can be created, by plotting the RASTERVALU of posterior probability on the y-axis against the CAPP_CumAr (cumulative area) on the x-axis (Figure 14). The CAPP curve provides guidance on how to classify the posterior probability map for display.

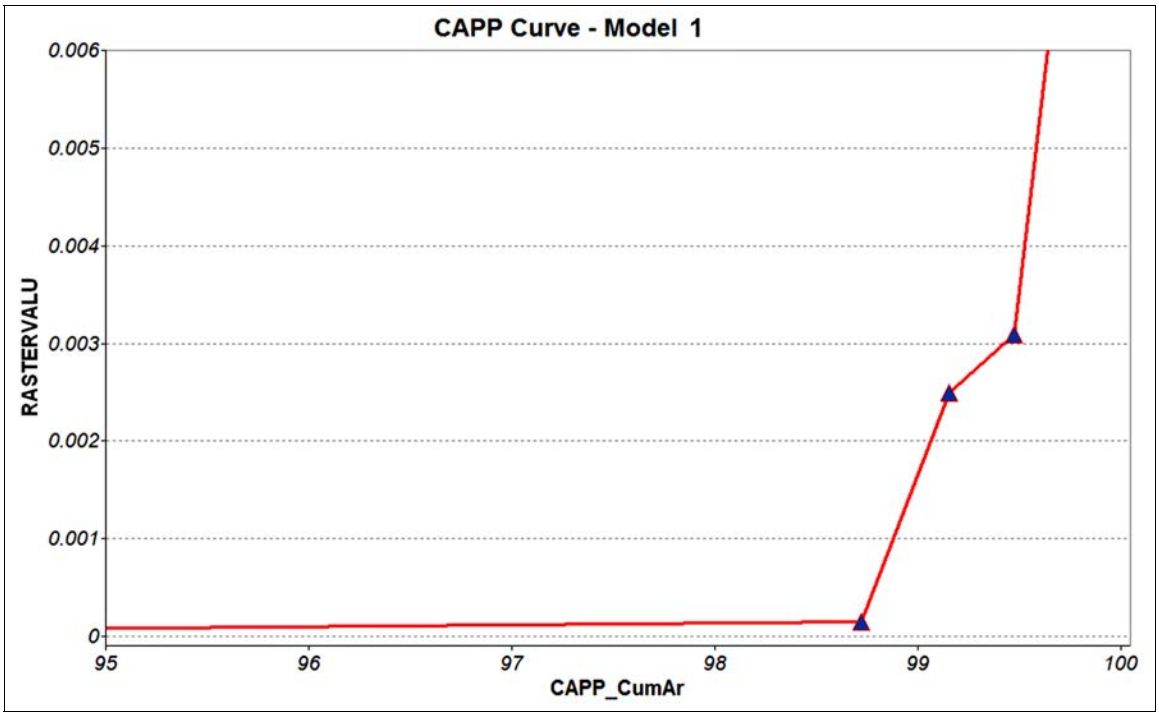


Figure 14: *Cumulative area-posterior probability (CAPP) curve for Model 1.*

Natural breakpoints, corresponding to prediction thresholds, can be selected where the CAPP curve significantly changes slope (Sawatzky et al. 2010b). A common starting point is the prior probability, in this case 0.000044. However, for Model 1, breakpoints were selected at 0.000082 to separate *Not permissive* from *Permissive,* and 0.00453 to separate Permissive from *Favorable.*

Figure 15 shows the resulting prospectivity map, along with the training sites and the drill holes used for the blind test.

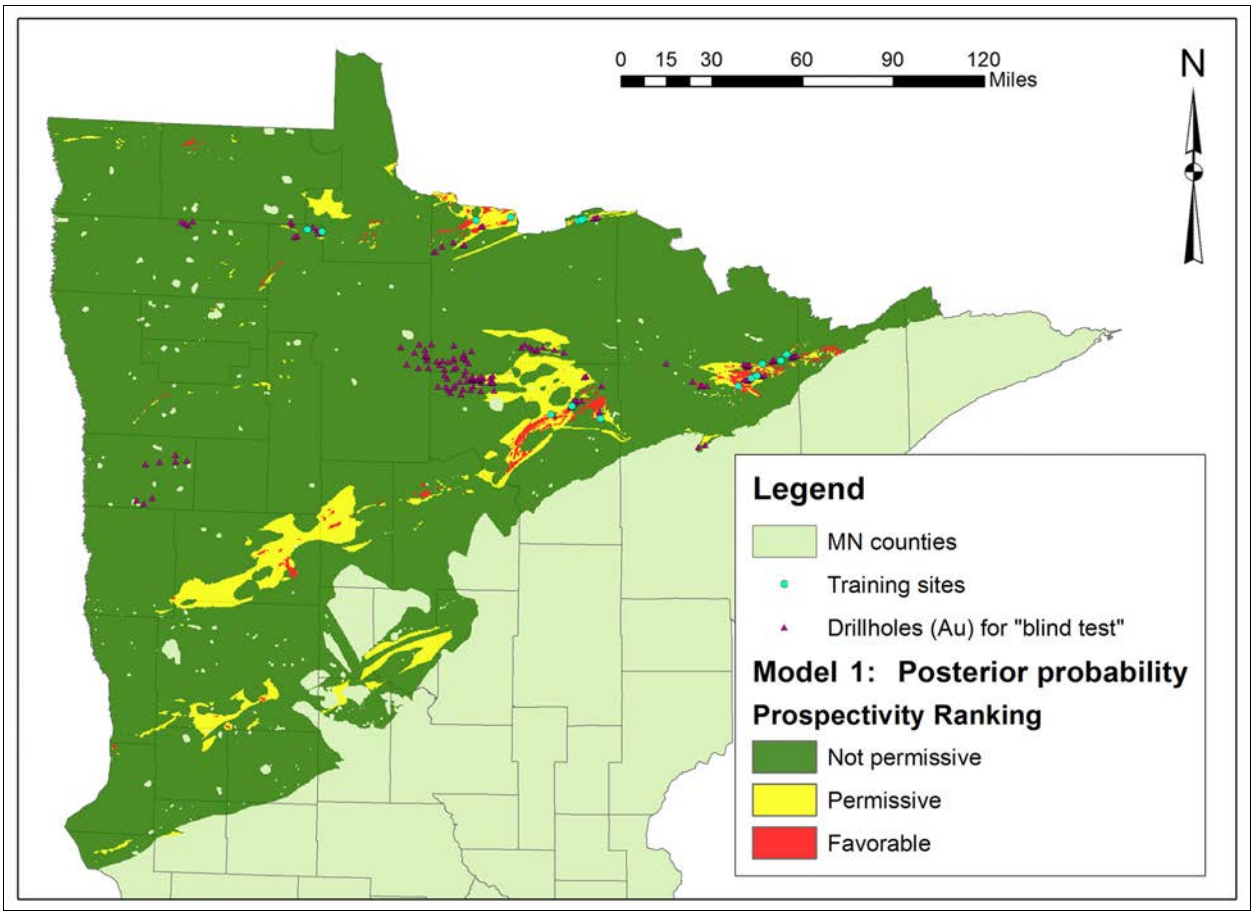


Figure 15: *Posterior probability map for Model 1, using the gold-averaged watersheds geochemistry theme. Model 1 had an efficiency of classification of 96.6%, and an efficiency of prediction of 65.8%.*

Ultimately, the breakpoints and consequently prediction thresholds are subjective decisions by the user, taking into account the quality and quantity of both evidence layers and training points. Changing the breakpoints results in different maps of posterior probability, so following the standard methodology is useful. In this study, the same breakpoints are chosen for both models, and justification for each choice is provided. But every model is different, and the results of each should be considered individually.

Model 2 used the same sets of evidence, differing only in the geochemistry theme; whereas Model 1 used gold-averaged major watersheds theme, Model 2 used the Theissen polygons.

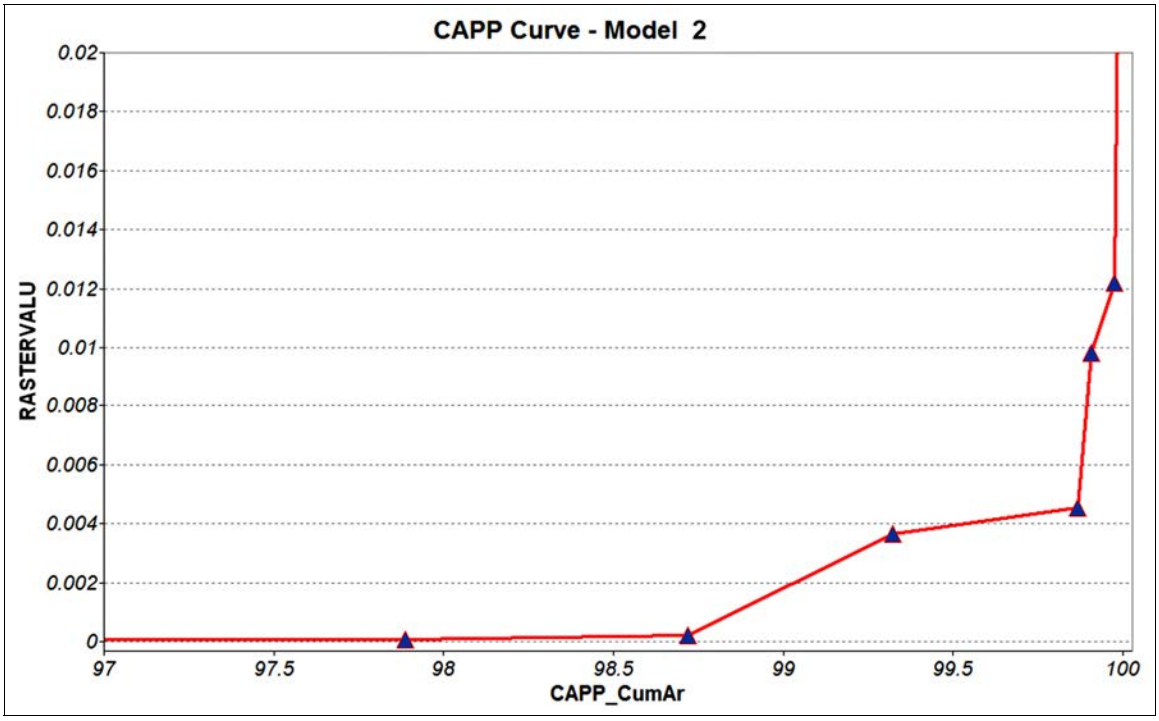


Figure 16: *Cumulative area-posterior probability (CAPP) curve for Model 2.*

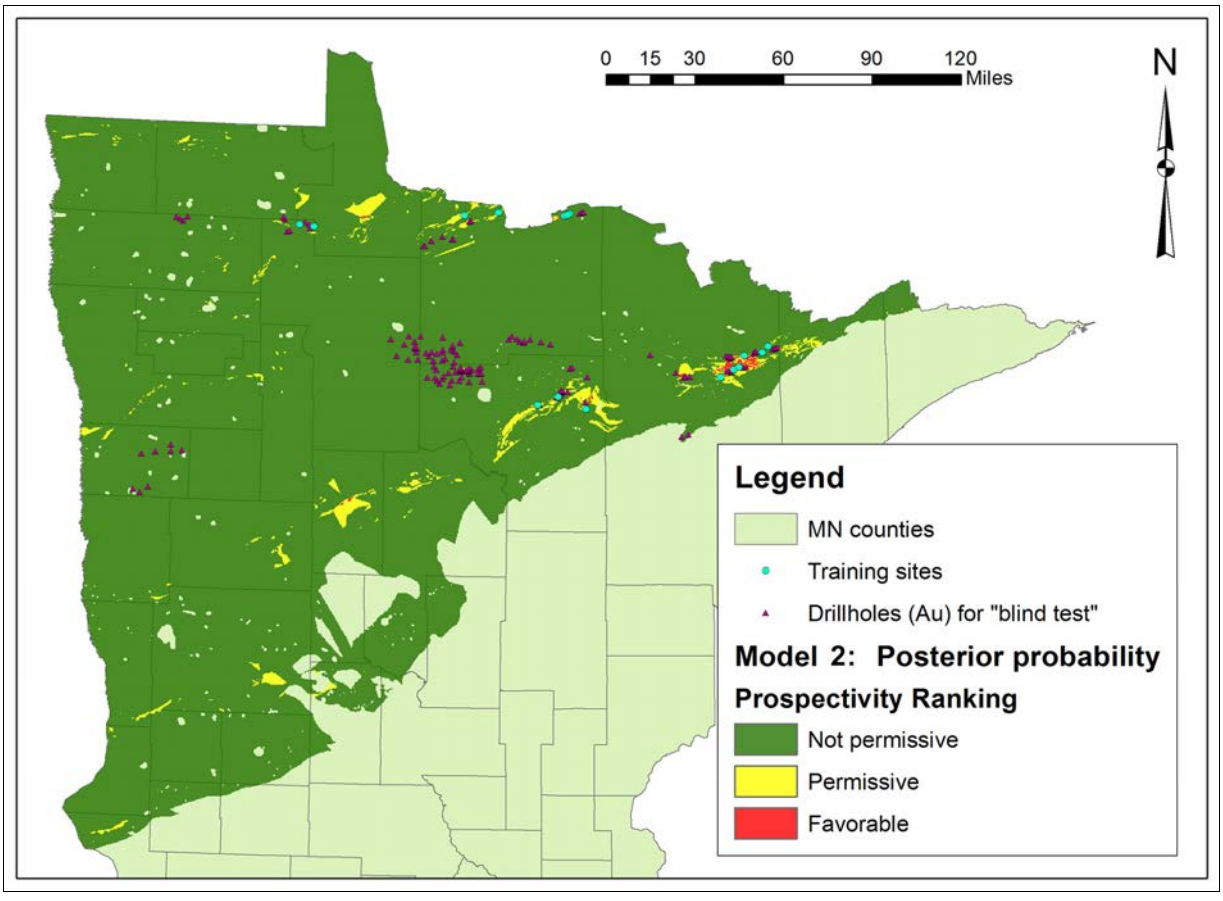


Figure 17: *Posterior probability map for Model 2, using the Theissen polygon geochemistry theme. Model 2 achieved an efficiency of classification of 96.5%, and an efficiency of prediction of 65.5%.*

The CAPP curve for Model 2 (Figure 16) suggested breakpoints at 0.000056 to separate Not permissive from Permissive, and at 0.003099 to separate Permissive from Favorable. Figure 17 shows the resulting prospectivity map.

Noticeable is a large cluster of blind test drill holes in the center of the study area, occupying an area classified as not permissive by both of the models created here. All of these were drilled by the same company during the 1980s; however, records of what this company was targeting were not included in the historical documents voluntarily submitted to the MDNR (Severson 2011). Clearly, the efficiency of prediction would be increased if these sites were excluded.

Before being used to generate exploration targets, the model results and the associated weights tables should be analyzed using expert judgment, to ensure that areas classified as favorable make sense with regard to the underlying geology. Because this project is part of a geography program, the focus is on the GIS operations and modeling methods, and this expert judgment is not applied here. However, a brief evaluation of the models in their geologic context is presented in the next chapter, in the "Comparison with Past Exploration" section.

Chapter Five: Results

Weights of evidence presents a plethora of results. Arguably, the model prospectivity maps, Figures 15 and 17, are already results. Those maps are re-presented side-by-side below (Figure 18) for ease of comparison. Discussed afterward are estimates of confidence in those models and those results, which are equally important - perhaps even more important, for economic reasons.

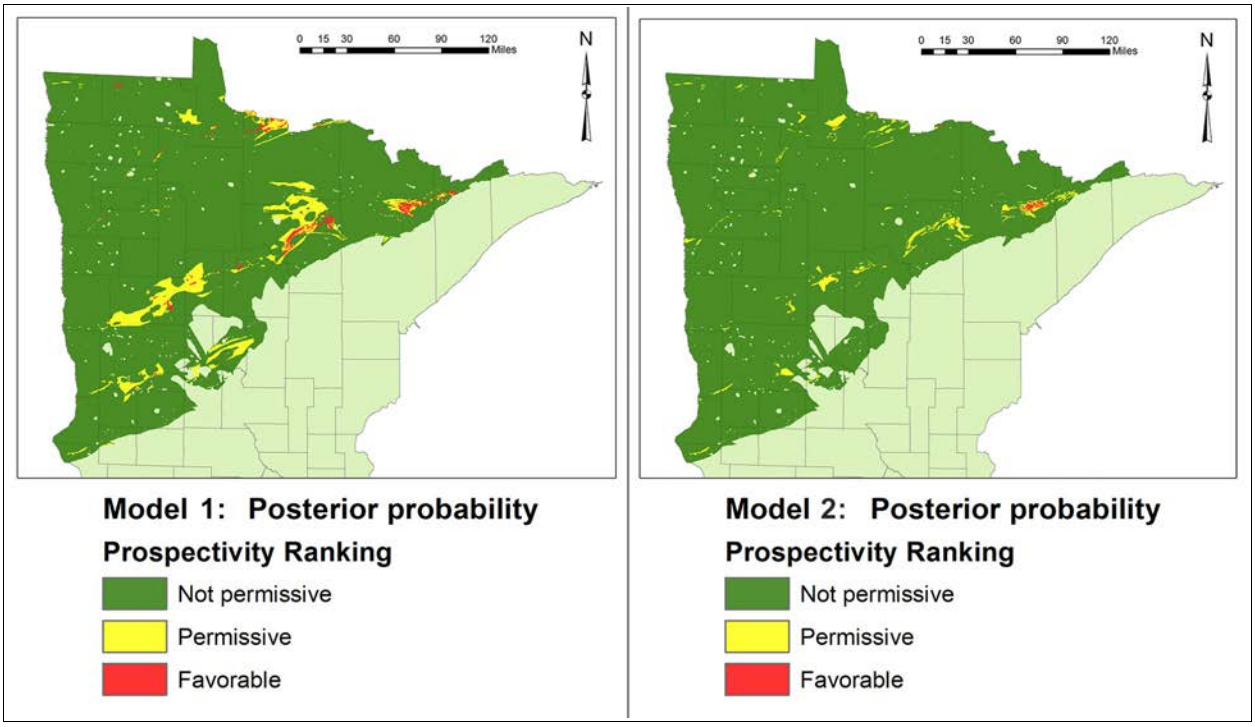


Figure 18: *Prospectivity maps from Model 1 (left) and Model 2 (right).*

5.1 Confidence Maps

The Calculate Response tool in ArcSDM also generates a *confidence map*, which returns the posterior probability in each unit cell divided by the standard deviation of all posterior probability values. This is a test of the confidence that the reported value of posterior probability is other than zero, and is a useful tool for post-modeling analyses.

For this project, where the prospectivity maps indicate a Favorable ranking and the reported confidence meets the 70% confidence criterion, the ranking is valid and these areas may be a new gold exploration targets. Where the 70% confidence criterion is not met, any Permissive or Favorable ranking is considered invalid, and should be changed to Not Permissive to disqualify

them as potential new gold exploration targets.

To test the validity of model results, each of the confidence maps were reclassified into binary form, where class 1 represents cells where the confidence is less than 0.542, or the approximately 70% limit of acceptable confidence that was chosen for this study. Class 2 represents all values above this limit (Figure 19).

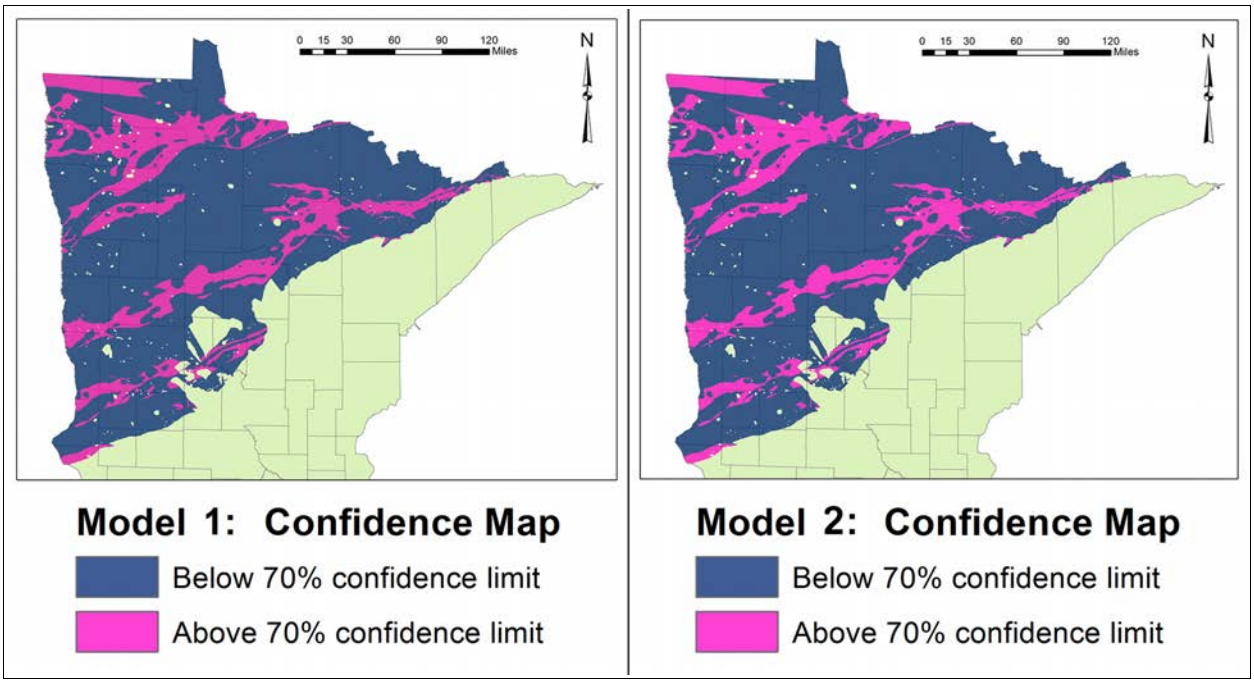


Figure 19: *Confidence maps for both models, reclassified to show areas above and below the 70% level of confidence criterion.*

Each of the prospectivity models was then reclassified, so that class 1 represents Not permissive ranks, class 3 represents Permissive ranks, and class 5 represents Favorable ranks. The binary confidence raster was subtracted from the reclassified posterior probability raster, using the Minus tool in ArcGIS (Spatial Analyst). This provided a set of unique conditions that identify potential issues. The idea was to check that areas ranked as permissive or favorable also occurred in areas where the confidence was acceptable. Table 14 illustrates this simple operation.

None of the Minus operations returned values of 2 or 4 (for either model), so all areas ranked as permissive or favorable are in areas of acceptable confidence. If problems had been identified, another reclassification could be performed, converting -1, 2, and 4 to zeroes. This would result in a "corrected" three-class map of posterior probability, with 0 representing Not permissive

areas, 1 representing Permissive areas, and 3 representing Favorable areas.

Table 14: *Illustration of unique conditions resulting from the subtraction of the binary confidence map from the reclassified map of posterior probability. Results having values of 2 or 4, highlighted in red, must be changed to "Not permissive" as they do not meet the 70% confidence requirement.*

Model Class	Conf. Class	Subtraction Result	Explanation
1	1	0	Not permissive in model, low confidence
1	2	-1	Not permissive in model, acceptable confidence
3	1	2	Permissive in model, low confidence
3	2	1	Permissive in model, acceptable confidence
5	1	4	Favorable in model, low confidence
5	2	3	Favorable in model, acceptable confidence

5.2 Comparison of Ranked Areas

Measuring how much area was assigned to each prospectivity rank is another useful way to compare model results. Table 15 lists the area in square kilometers that each class occupies.

Table 15: *Area (in square kilometers) occupied by each ranked prospectivity class.*

Model #	Geochem.	Class 1 (km^2)	Class 1 (%)	Class 2 (km^2)	Class 2 (%)	Class 3 (km^2)	Class 3 (%)
1	Watersheds	85,094.75	93.49	5,163.75	5.67	770.50	0.84
2	Theissen	89,105.25	97.89	1,802.00	1.98	121.75	0.13

In both models, most of the study area is ranked as Not permissive, with less than 1 percent ranked as Favorable. Model 1 ranked more total area as Permissive and Favorable. Since each model used the same evidence, the variation is caused by the geochemistry themes. The Theissen polygon theme had 45.96 percent of the study area classified as Elevated and Anomalous (42,072.5 km^2), and the gold-averaged major watersheds theme had 64.83 percent of the study area classified as such (59,346.7 km^2). The Elevated and Anomalous classes in the gold-averaged major watersheds theme had a higher contrast than their counterparts in the Theissen polygon theme, and was thus more strongly associated with the training sites (see Table 12).

5.3 Success and Prediction-Rate Curves

Efficiency of classification can also be represented graphically, as a so-called *success-rate*

curve (Figure 20). The success-rate curve plots the proportional number of training sites, accumulated from highest to lowest posterior probability, on the y-axis. On the x-axis, the proportional area is plotted, accumulated from highest to lowest posterior probability.

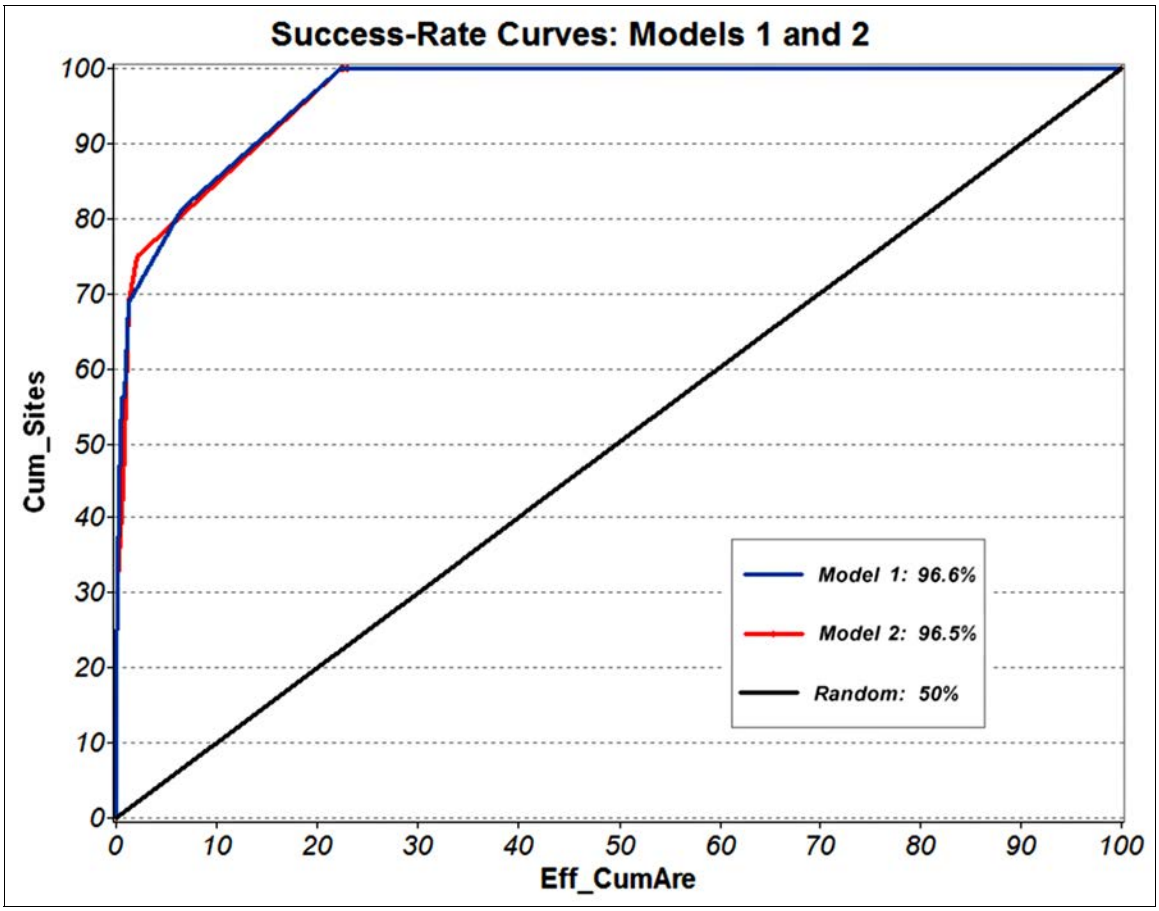


Figure 20: *Success-rate curves for prospectivity models 1 and 2.*

The efficiency of classification is really a measure of the total area underneath the success-rate curve. With 50% of area underneath the curve, the classification results can not be considered any better than random (Chung and Fabbri 2008). The black line in Figures 20 and 21 represent this 50% value. The area under the curve can be found by summing the Eff_AUC column in the area-frequency table.

The same applies to the efficiency of prediction; when represented graphically, it is referred to as a *prediction-rate curve* (Figure 21). In both cases, the models performed better than random. Because Model 1 performed slightly better in terms of success and prediction rates, it was selected as the overall "best" model.

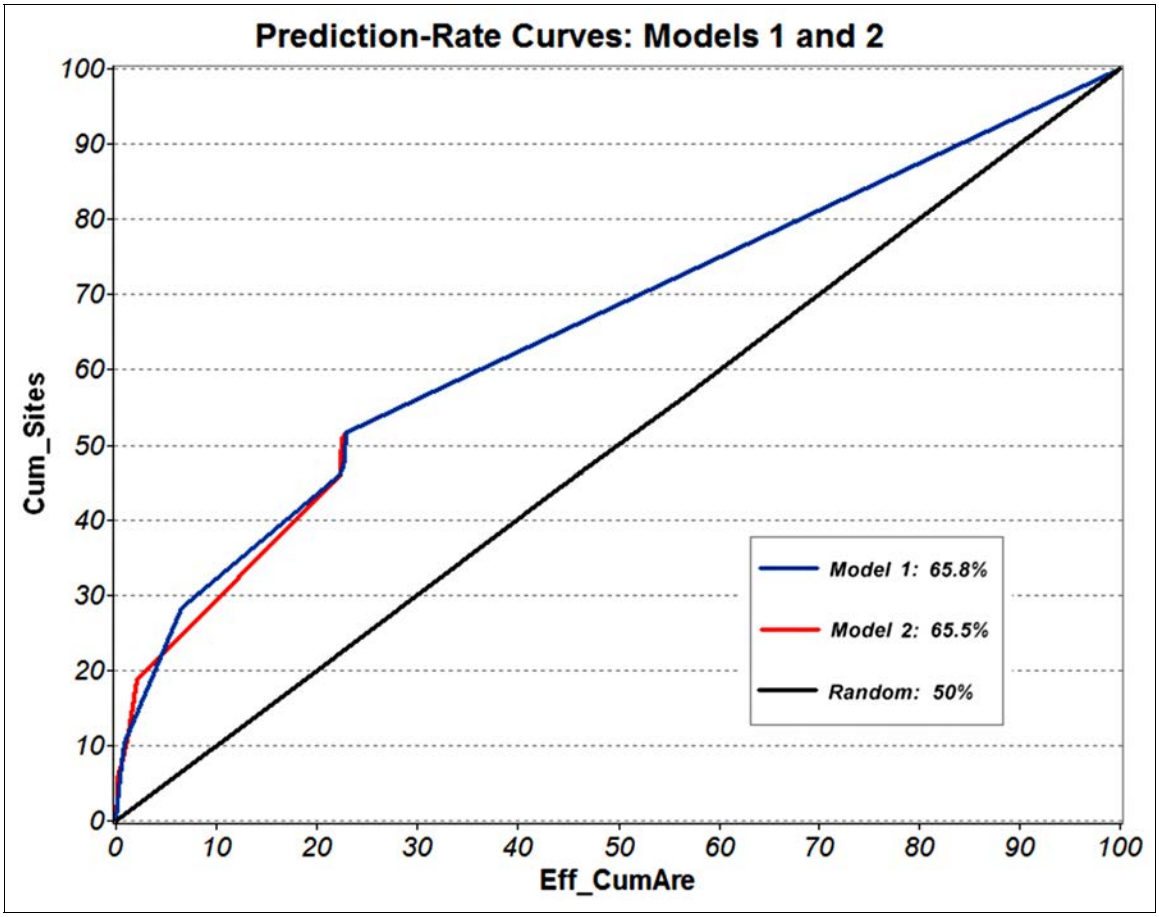


Figure 21: *Prediction-rate curves for prospectivity models 1 and 2.*

5.4 Comparison with Past Exploration Activity

Five of the six areas of gold exploration in Minnesota, identified by Severson (2011), contain unit cells ranked as Favorable for gold occurrences by Model 1, and all six of these areas contain unit cells ranked as Permissive (Figure 22). Only the Virginia Horn does not contain area ranked as Favorable. However, recall from Chapter 4 that there is a small, low-grade gold deposit in the Virginia Horn, hosted in the so-called Viking Porphyry (Severson 2011). This porphyry has only been partially explored at the time of this writing.

Thus, the results of Model 1 are interesting, even without detailed analysis of the geological context of the permissive and favorable cells. The model did a good job of identifying areas of past gold exploration activity, and thus, weights-of-evidence modeling appears to be useful for generating new gold exploration targets within the study area. But of course, the ultimate test of this project's findings will come from future field investigations.

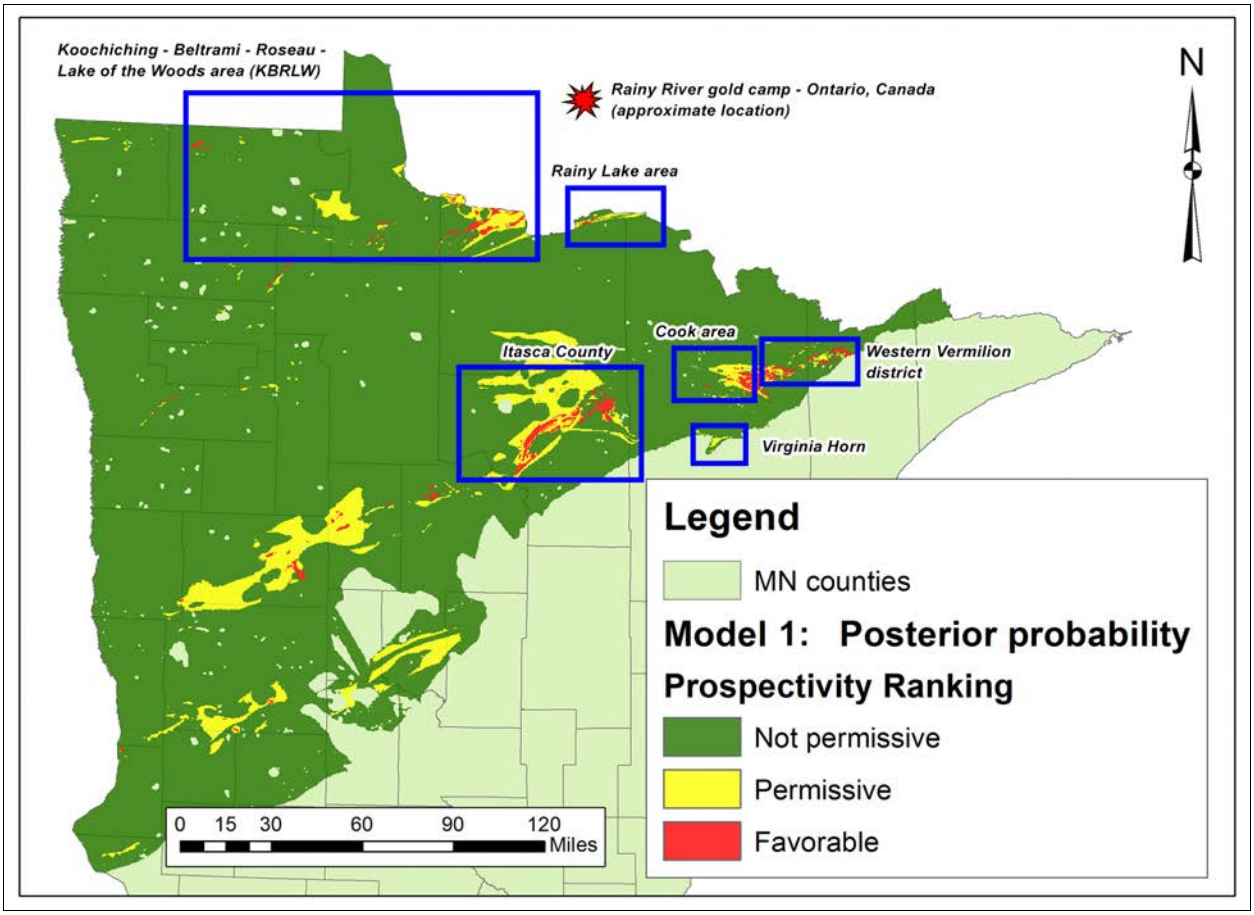


Figure 22: *Result of Model 1, and the six areas of gold exploration focus in Minnesota, offered by Severson (2011). All six contain unit cells ranked as Permissive for gold occurrences, and five contain unit cells ranked as Favorable.*

Chapter Six: Summary and Discussion

6.1 Project Summary

This project applied the data-driven weights-of-evidence method to gold exploration in northern Minnesota's Archean terranes. The study area represented the Wabigoon, Quetico, and Wawa subprovinces of the Canadian Shield where they underlie Minnesota.

Training sites were selected from a set of drill holes that intercepted anomalous gold, and from samples of glacial till that contained high counts of pristine gold grains. Proximity to major fault and shear systems, the presence of aeromagnetic anomalies, generalized rock types, and gold geochemistry in regional soil samples were all considered as evidence.

Spatial relationships between the training sites and the evidence were evaluated using weights tables. Currently-accepted gold exploration criteria were confirmed by these relationships. Patterns for each piece of evidence, representing favorable gold mineralizing conditions, were created. These patterns were integrated in two weights-of-evidence models.

The model results were assessed based on their ability to efficiently predict the training sites, and to efficiently predict the locations of historical gold-targeting exploratory drill holes in a blind test. Model 1, selected as best overall, had a success rate of 96.6 percent in classifying the training sites, and a prediction rate of 65.8 percent in the blind test.

Based on the model results, the study area was classified into three ranks, representing low to high potential for gold occurrences: not permissive, permissive, and favorable. Favorable areas have the best gold potential based on the presence of aeromagnetic anomalies, proximity to major fault and shear systems, and the presence of elevated or anomalous gold in soil samples, averaged across the major watersheds within the study area.

There are six areas of past and present gold exploration in Minnesota, as identified by Severson (2011). Model 1 classified tracts within five of these as Favorable for hosting gold occurrences, and all six as Permissive. Based on this comparison, the weights-of-evidence method

appears to be useful for generating new gold exploration targets within the study area. Many of the tracts ranked as favorable do not appear to have ever been drilled, so the bedrock in these areas has yet to be evaluated.

6.2 Limitations and Future Research Directions

While every effort was made to create meaningful evidence and gather quality training sites, no model is ever perfect. There are several ways that the models could potentially be improved, mainly by the addition and refinement of training sites and blind test sites. Counter-intuitively, doing this might actually lower the efficiency of classification but nonetheless increase the efficiency of prediction, by taking into account a larger variety of mineralization styles. However, even without these suggested improvements, both models performed very well with the available data. The model results are significant enough to warrant field testing of any favorable areas, provided they are first evaluated with respect to the local geology.

With a 96.6 percent efficiency of classification, the best way to improve the model would be to add more training sites. This would involve logging and sampling drill core at the MDNR core library in Hibbing, Minnesota, after evaluating historical exploration records in the MDNR archive[14].

Improving the evidence along with the training sites may also be possible. The soil geochemistry was reduced to three classes based on expert decisions prior to analyses, and this step may have been unnecessary. A test of the original geochemistry raster with cumulative-descending weights generally agreed with the background gold limit that was derived from the box-plot statistics. The bedrock geology raster was also reclassified using expert decisions. A test of the original map using categorical weights, and a reclassification based on the results, was very similar to the map classified by expert decisions. So here too, the expert decisions appear to have been unnecessary, but likely did not significantly alter the model results.

Also, while it was not available for this project, there is a statewide “C-horizon” soil sample

14 MDNR non-ferrous minerals exploration archive: http://minarchive.dnr.state.mn.us/

dataset available from the Natural Resources Research Institute (NRRI) in Duluth, Minnesota[15]. The C-horizon is the lowest part of the soil, just above the bedrock. This survey processed samples for gold grains much like the local-scale MDNR open-file projects 392 and 379 that were used here, and might contribute to a better modeling result. Finally, source datasets from the seminal PhD dissertation by Peterson (2001) on Archaen gold in Minnesota would be helpful; unfortunately, these datasets were unavailable for this thesis.

As mentioned in several chapters of this thesis, the method used to extract the major fault and shear structures from the geologic map was done quickly, and without expert knowledge of Minnesota's Precambrian geology. It might be possible, with expert consultation, to extract structures based on recognized deformation events, some of which may be known to be more or less associated with gold mineralization.

Conditional independence was not a factor in the two models developed, because the results were translated into a relative favorability ranking. However, several "test" models (results of which are not presented here), indicated that the geologic map, in large part interpreted from aeromagnetic data, was a major source of conditional dependence in this project. So if necessary, the geologic map could be omitted from future models, following the example of Nykänen and Salmirinne (2007).

Notwithstanding the possible improvements listed above, the models developed in this thesis appear valid and useful for gold prospecting: the areas ranked as favorable should be investigated in the field. Follow-up mapping and perhaps basal till sampling is recommended.

6.3 Dissemination of Results

By objective, data-driven methods, this research has reconfirmed six areas already prospected for gold in northern Minnesota's greenstone belts, and also suggested a new area to the southwest, in Becker and Hubbard counties (Figure 15, enlarged in Appendix). To encourage further exploration in all of these areas, parts of this thesis will be prepared as a short paper, for submission

15 NRRI presentation, statewilde C-horizon soil survey: http://www.nrri.umn.edu/egg/presentations.html

to applicable geoscience journals, mining companies with interests in gold, and key personnel at the MDNR. Finally, the posterior probability results of Model 1 are available for download (as a three-class raster in esri GRID format) from http://www.bhartley.com/projects/mngold.html.

Whether or not these models, or some improved future models, lead to the discovery of new gold occurrences is the ultimate test of their accuracy. In the words of Francis Pettijohn, who contributed a great deal to the understanding of Precambrian geology in the Lake Superior region, "The rocks are the final court of appeal."

References

Agterberg, F.P., and Q. Cheng. 2002. Conditional Independence Tests for Weights-Of-Evidence Modeling. *Natural Resources Research* 11 (4): 249-255.

Bonham-Carter, G.F., F.P. Agterberg, and D.F. Wright. 1989. Weights of Evidence Modelling: A New Approach to Mapping Mineral Potential. in *Statistical Applications in the Earth Sciences,* ed. F.P. Agterberg and G.F. Bonham-Carter. *Geological Survey of Canada, Paper 89-9* : 171-183.

———. 1988. Integration of Geological Datasets for Gold Exploration in Nova Scotia. *Photogrammetric Engineering and Remote Sensing* 54 (11): 1585-1592.

Bonham-Carter, G.F. 1994. *Geographic Information Systems for Geoscientists.* Oxford, UK: Elsevier.

Card, K.D. 1990. A Review of the Superior Province of the Canadian Shield, a Product of Archean Accretion. *Precambrian Research* 48: 99-156.

Carranza, J.M. 2009a. *Geochemical Anomaly and Mineral Prospectivity Mapping in GIS.* Ed. Martin Hale. Amsterdam, Netherlands: Elsevier.

———. 2009b. Objective Selection of Suitable Unit Cell Size in a Data-Driven Model of Mineral Prospectivity. *Computers and Geosciences* 35: 2032-2046.

Chung, C.F., and A.G. Fabbri. 2008. Blind Tests and Spatial Prediction Models. *Natural Resources Research* 17 (2): 107-118.

———. 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. *Natural Hazards* 30: 451-472.

Guiterrez, M., V.M.R. Gomez, T.A. Hererra, and D.M. Lopez. 2012. Exploratory Analysis of Sediment Geochemistry to Determine the Source and Dispersion of Ba, Fe, Mn, Pb, and Cu and in Chihuahua, Northern Mexico. *Journal of Geography and Geology* 4, (4): 26-39.

Hardie, C., D. Runnels, P. Live, S.E. Daniel, D.G. Ritchie, A. Coulson, G. Cole, D. El-Rassi, and D. Tolfree. 2013. *Feasibility Study of the Rainy River Gold Project.* NI 43-101 Technical Report. Prepared for New Gold, Inc (July 31, 2013). Montreal, QC: BBA, Inc.

Harris, J.R., L. Wilkinson, and M. Bernier. 2001. Analysis of Geochemical Data for Mineral Exploration using a GIS – A Case Study from the Swayze Greenstone Belt, Northern Ontario, Canada. *Geological Society of London, Special Publications* 185: 165-200.

Kerrich, R., and K.F. Cassidy. 1994. Temporal Relationships of Lode Gold Mineralization to Accretion, Magmatism, Metamorphism and Deformation – Archean to Present: A Review. *Ore Geology Reviews* 9: 263-310.

Kotz, S., N. Balakrishnan, and N.L. Johnson. 2000. *Continuous Multivariate Distributions.* Hoboken, NJ: Wiley Publishing.

Marjoribanks, R. 1997. *Geological Methods in Mineral Exploration and Mining.* London, UK: Chapman & Hall.

Minnesota Department of Natural Resources. *International Falls Drill Core Descriptions and Chemistry, Koochiching County, Minnesota: Project 378,* by Frey, B. Open-file report, Lands and Minerals Division, Mineral Potential Evaluation Section. Saint Paul, MN 2012.

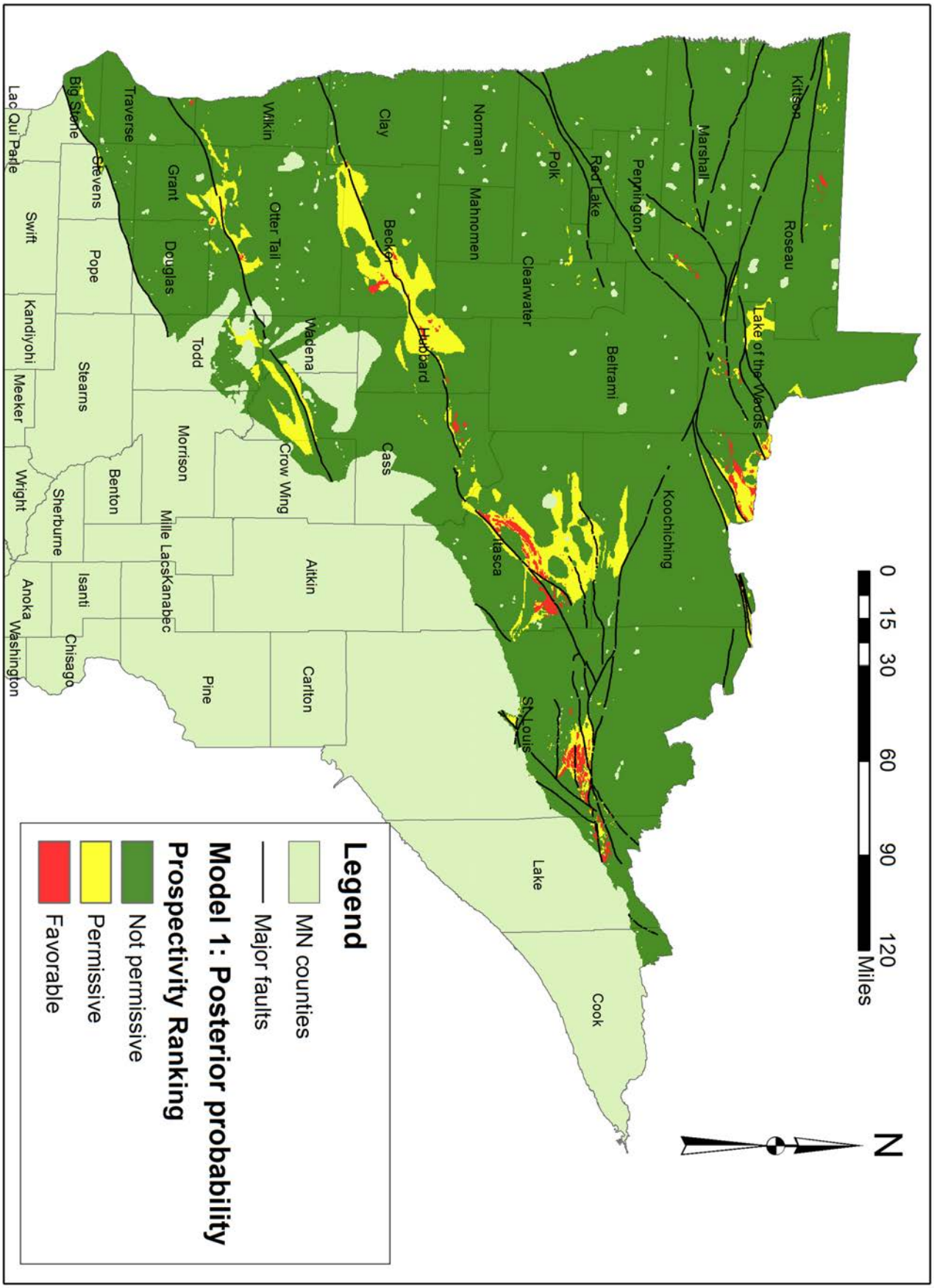
Minnesota Department of Natural Resources. *Drill Core Evaluation of Vermilion Greenstone Belt Gold Mineralization, Northeastern Minnesota: Project 373.* Frey, B. and A. Hanson. Open-file report, Lands and Minerals Division, Mineral Potential Evaluation Section. Saint

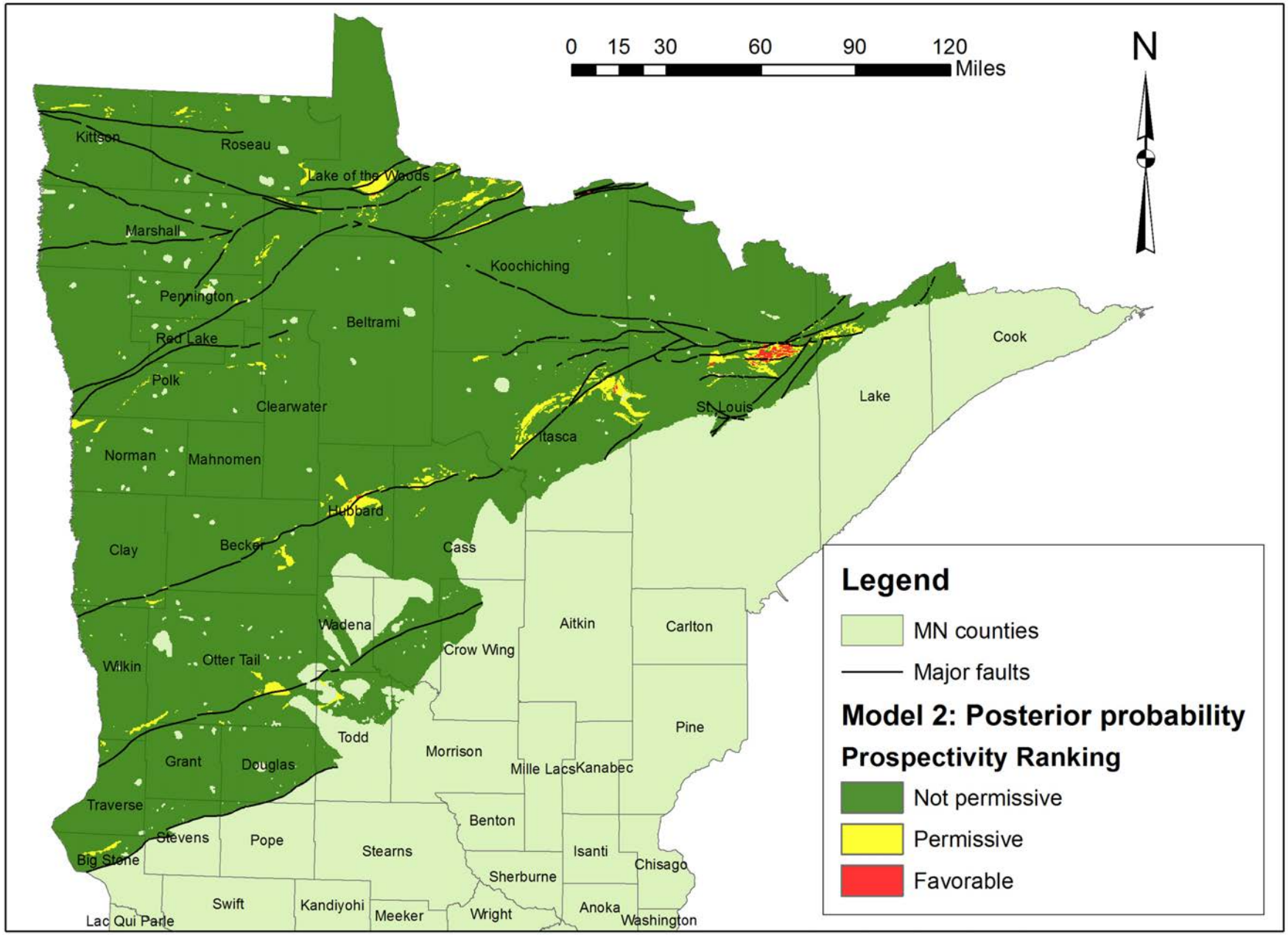
Paul, MN 2010.
Minnesota Minerals Coordinating Committee. *Upgrade of Aeromagnetic Databases and Processing Systems at the Minnesota Geological Survey,* by Chandler, V.W. Open-file report, Minnesota Geological Survey. Saint Paul, MN 2007.
Moon, C.J., M.K.G. Whately, and A.M. Evans. 2006. *Introduction to Mineral Exploration.* Oxford, UK: Blackwell Publishing.
Muir, T.L., B.R. Schnieders, and M.C. Smyk (Compilers and Editors). 1995. *Geology and Gold Deposits of the Hemlo Area.* Geological Association of Canada – Toronto '91, Hemlo Field Trip Guidebook*,* 120p.
Natural Resources Research Institute. *The History of Gold Exploration in Minnesota.* Severson, M.J. Technical report 2011/43. Duluth, MN 2011.
Nykänen, V., H. Salmirinne. 2007. Prospectivity Analysis of Gold Using Regional Geophysical and Geochemical Data from the Central Lapland Greenstone Belt, Finland. *Gold in the Central Lapland Belt, Finland.* Ed. V. Juhani Ojala. In *Geological Survey of Finland, Special Paper 44:* 251-269.
Ojakangas, R.W. 2009. *Roadside Geology of Minnesota.* Missoula, MT: Mountain Press Publishing Company
Peterson, D. 2001. Development of Archean Lode-Gold and Massive Sulfide Deposit Exploration Models Using Geographic Information System Applications: Targeting Mineral Exploration in Northeastern Minnesota from Analysis of Analog Canadian Mining Camps. PhD diss., University of Minnesota-Duluth.
Poulsen, K.H., K.D. Card, and J.M. Franklin. 1992. Archean Tectonic and Metallogenic Evolution of the Superior Province of the Canadian Shield. *Precambrian Research* 58: 25-54.
Raines, G.L. 1999. Evaluation of Weights of Evidence to Predict Epithermal-gold Deposits in the Great Basin of the Western United States. *Natural Resources Research* 8 (4): 257-276.
Robb, L.J. 2005. *Introduction to Ore-Forming Processes.* Oxford, UK: Blackwell Publishing.
Sawatzky, D.L., G.L. Raines, G.F. Bonham-Carter, and C.G. Looney. 2010a. *Spatial Data Modeler (SDM): ArcGIS 10 Geoprocessing Tools for Spatial Data Modeling Using Weights of Evidence, Logistic Regression, Fuzzy Logic, and Neural Networks.* http://www.ige.unicamp.br/sdm/default_e.htm (Last accessed October 2013).
Sawatzky, D.L., G.L. Raines, G.F. Bonham-Carter, and C.G. Looney 2010b. *Spatial Data Modeler (SDM) for ArcGIS 10*: Help files.
Stevens, S.S. 1946. On the Theory of Scales Measurement. *Science* 103 (2684): 677-680.
Taylor, S.R., and S.M. McLennan. 1995. The Geochemical Evolution of the Continental Crust. *Reviews of Geophysics* 33 (2): 241-265.
U.S. Department of the Interior. *Resource Materials for a GIS Spatial Analysis Course – Revision of Lectures,* by Raines, G.L. Open-file report 01-221, U.S. Geological Survey. Reno, NV 2006.
U.S. Department of the Interior. *The National Geochemical Survey – Database and Documentation.* Open-file report 2004-1001*,* U.S. Geological Survey. Reston, VA 2004. http://tin.er.usgs.gov/geochem/doc/home.htm (Last accessed October 2013)
U.S. Department of the Interior. *Descriptive and Grade-Tonnage Models of Archean Low-Sulfide Au-Quartz Veins and a Revised Grade-Tonnage Model of Homestake Au,* by Klein, T.L. and W.C. Day. Open-file report 94-250, U.S. Geological Survey. Reston, VA 1994.

World Gold Council. October 2013. *The Direct Economic Impact of Gold.* London, UK: Pricewaterhouse Coopers LLP

Wright, D.F., and G.F. Bonham-Carter. 1996. VHMS Favourability Mapping With GIS-Based Integration Models, Chisel Lake-Anderson Lake Area. In *EXTECH I: A Multidisciplinary Approach to Massive Sulphide Research in the Rusty Lake-Snow Lake Greenstone Belts, Manitoba.* Ed. G.F. Bonham-Carter, A.G. Galley, and G.E.M. Hall. *Geological Survey of Canada Bulletin 426:* 339-376, 387-401.

Appendix


Kittson
Roseau
Marshall
Pennington
Red Lake
Polk
Norman
Mahnomen
Clay
Becker
Wilkin
Otter Tail
Grant
Douglas
Traverse
Big Stone
Stevens
Pope
Swift
Lac Qui Parle
Kandiyohi
Meeker
Wright
Todd
Wadena
Hubbard
Clearwater
Lake of the Woods
Beltrami
Koochiching
Itasca
Cass
Crow Wing
Aitkin
Carlton
St. Louis
Lake
Cook
Morrison
Mille LacsKanabec
Pine
Benton
Stearns
Sherburne
Isanti
Chisago
Anoka
Washington
0 15 30 60 90 120
Miles
N
Legend
MN counties
Major faults
Model 1: Posterior probability
Prospectivity Ranking
Not permissive
Permissive
Favorable


0 15 30 60 90 120
Miles
N
Kittson
Roseau
Lake of the Woods
Marshall
Koochiching
Pennington
Beltrami
Red Lake
Cook
Polk
Clearwater
St Louis
Lake
Itasca
Norman
Mahnomen
Hubbard
Clay
Becker
Cass
Wadena
Aitkin
Carlton
Crow Wing
Wilkin
Otter Tail
Pine
Todd
Morrison
Grant
Douglas
Mille Lacs
Kanabec
Traverse
Benton
Stevens
Pope
Stearns
Isanti
Big Stone
Chisago
Sherburne
Swift
Kandiyohi
Meeker
Wright
Anoka
Lac Qui Parle
Washington
Legend
MN counties
Major faults
Model 2: Posterior probability
Prospectivity Ranking
Not permissive
Permissive
Favorable