71-41

P. G. KILLEEN
GEOPHYSICS

GEOLOGICAL SURVEY OF CANADA

DEPARTMENT OF ENERGY, MINES AND RESOURCES

PAPER 71-41

This document was produced by scanning the original publication.

Ce document est le produit d'une numérisation par balayage de la publication originale.

GEOMATHEMATICAL EVALUATION OF COPPER AND ZINC POTENTIAL OF THE ABITIBI AREA, ONTARIO AND QUEBEC

F. P. Agterberg, C. F. Chung, A. G. Fabbri,
A. M. Kelly and J. S. Springer

Price, $2.00

1972

ERRATA

Page No.	Line No.	Delete	Insert
3	36	Botol	Botbol
10	41	commonly	common,
14	26-27	multivariable	multivariate
21	4 and 8	E(Y X)	E (Y|X)
21	11	Hence, Σ(YX)	Hence, $\Sigma XY = 18$, $\Sigma X = 120$, $\Sigma Y = 27$, $\Sigma X^2 = 120$. When N = 644, we obtain
21	13	$\hat{B}_1$	$\hat{B}_o$
21	15	$(B_o + B_1)$	$(B_o + B_1X)$
28	Fig. 13c	$R^2 = 0.40$	$R^2 = 0.140$
30	25	correlation squared	correlation coefficient squared
38	5	heads is eight	heads is 0.5. If we toss sixteen coins, the expected number of heads is eight.
42	9	Paper 2	Special Paper No. 2

GEOLOGICAL SURVEY

OF CANADA

CANADA

PAPER 71-41

GEOMATHEMATICAL EVALUATION OF COPPER AND ZINC POTENTIAL OF THE ABITIBI AREA, ONTARIO AND QUEBEC

F. P. Agterberg, C. F. Chung, A. G. Fabbri,
A. M. Kelly and J. S. Springer

DEPARTMENT OF ENERGY, MINES AND RESOURCES

© Crown Copyrights reserved
Available by mail from *Information Canada*, Ottawa

from the Geological Survey of Canada
601 Booth St., Ottawa

and

Information Canada bookshops in

HALIFAX - 1735 Barrington Street
MONTREAL - 1182 St. Catherine Street West
OTTAWA - 171 Slater Street
TORONTO - 221 Yonge Street
WINNIPEG - 499 Portage Avenue
VANCOUVER - 657 Granville Street

or through your bookseller

Price: $2.00 Catalogue No. M44-71-41

Price subject to change without notice

Information Canada
Ottawa
1972

FOREWORD

This report represents the first quantitative estimates – quantitative in terms of statistical probability – by the Geological Survey of Canada of the undiscovered mineral resources of a region based on a statistical appraisal of certain geological and geophysical parameters. The method employed is still under development and is restricted by the very limited types of geoscience information that are available in consistent and quantifiable form for the whole of the region being assessed. The main value of these early results lies in the objective multivariate integration of certain geological and geophysical features and their systematic correlation with the occurrence of mineral deposits to yield tentative estimates of undiscovered resources.

The results reported are only one preliminary step in the use of geology to appraise the mineral potential of a region. Their use should be governed by full recognition of the limited scope of the geoscientific observations on which they are based. Ideally, they provide a quantified input to a much broader geological assessment of potential resources which would also incorporate consideration of essential local geological, geophysical and geochemical observations, large scale features such as geosynclines, plate tectonics, paleoclimates etc.; and appropriate conceptual models of ore genesis.

September 9, 1971.

S.C. Robinson,
Chief,
Economic Geology and
Geochemistry Division.

CONTENTS

Page

Foreword iii
Abstract/Résumé vii
Introduction 1
History of method 2
General discussion of method used 3
Geology and mineral deposits in the study area 6
The data base 9
    Geological data 9
    Mineral deposits 12
Methodology 13
    Preliminary statistical analysis 14
    Calculation of probability index 19
    Other probability maps for copper 26
    Relationship of method to methods of multivariate classification 32
Limitations 34
Use of results 37
References 41

Table 1. Totals of production and reserves for study area 8
2. Ten basic variables coded for Archean rocks 11
3. Frequency of occurrence of 10 x 10 km cells 18
4. Probability that the unit cell of 40 x 40 km contains x smaller 10 x 10 km cells with metal when m is given 38
5. Probability that a 10 x 10 km cell with metal contains a specific amount of metal 39

Appendix 1: Data for 57 copper and/or zinc deposits in Abitibi area 45
Appendix 2: Variables used to calculate values of Appendix 3 for copper potential 51
Appendix 3: Calculated values of copper potential for 644 cells; Outline of control area; control cells marked by solid circles 53
Appendix 4: Glossary of terms 57

Illustrations

Figure 1: Copper potential map ..... in pocket
2: Schematic block diagram for copper potential 5
3: Copper and zinc deposits in Abitibi area ..... in pocket
4: Zinc potential map ..... in pocket
5: Gold occurrences in Abitibi ..... in pocket
6: Size-frequency distributions for large copper and gold deposits 15
7: Histograms for grades of large copper and gold deposits 16
8: Average size of large gold deposits (1,000 oz.) plotted against total number of gold occurrences for 8 x 8 mile cells of a regular grid 17

Page

Figure 9: Average ratio of number of large deposits versus total number of deposits (large and small) plotted against total number of deposits per 8 x 8 mile cell 17
10: Cells used for coding, based on UTM grid 18
11: Copper potential maps corresponding to Figure 2 25
12: Copper potential maps computed by using fewer variables 27
13: Seven copper potential maps computed by progressively adding more geological variables to multiple regression equation 28-29
14: Copper potential map based on same variables as Figure 1 but with 44 additional variables to consider location of cells with respect to one another 32
15: Octic polynomials for copper and zinc, expressing systematic regional variations not explained by the 55 geological and geophysical variables used for Figure 1 ... 33

ABSTRACT

A number of geological features and geophysical parameters were coded systematically by superimposing a network of cells over the Abitibi area. A parameter expressing amount of metal per cell is calculated for cells (control cells) known to contain one or more large copper or zinc orebodies.

This parameter for known mineralization is related to the coded values of the geological and geophysical measurements inside and outside the control cells; the resulting relationship is extrapolated to all other cells in the area. The result is a set of predicted values which are scaled, summed for larger unit cells and contoured to give a probability index.

The metal potential of an area is then equal to the difference between total value as predicted by the probability index and total known value for proven reserves within the area. The predictions are subject to certain limitations in the method of approach and have wide confidence limits which can be calculated.

RÉSUMÉ

Les auteurs de l'étude ont effectué un codage systématique de plusieurs paramètres géologiques et géophysiques en surimposant un réseau de cellules sur la région de l'Abitibi. La méthode consiste à établir, pour chaque cellule de contrôle reconnue contenir un ou plusieurs importants massifs de minerai de cuivre et de zinc, un paramètre qui exprime la quantité de métal par cellule.

Ce paramètre est mis en regard des valeurs codées des mesures géologiques et géophysiques prises à l'intérieur et à l'extérieur des cellules de contrôle; le rapport qui résulte de cette comparaison est extrapolé pour être appliqué à toutes les autres cellules de la région. On obtient alors une série de prévisions quantitatives qui sont ramenées à une échelle commune, additionnées dans le cas des unités plus grandes, puis les courbes isoparamétriques sont établies pour obtenir un indice de probabilité.

Le potentiel en métaux d'un secteur donné équivaut donc à la différence entre la valeur totale connue établie d'après les réserves connues. Les prévisions sont sujettes à certaines limites dans la méthode d'approche, et leur exactitude, bien que très variable, peut être calculée.

GEOMATHEMATICAL EVALUATION OF COPPER AND ZINC POTENTIAL OF THE ABITIBI AREA, ONTARIO AND QUEBEC

INTRODUCTION

In 1967, the Geological Survey initiated studies designed to provide quantitative estimates of the mineral potential of a region from the application of statistical methods to the analysis of systematically mapped geological information. This paper presents the first results of these studies comprising maps displaying contours of the mineral probability index for copper and zinc in the Ontario – Quebec Gold Mining Community Region, approximately 33,000 square miles in area. These results are only one stage in a comprehensive evaluation of mineral potential of this region and are subject to modification by specialists in metallogeny and regional geology.

Traditionally, mineral potential has been assessed in qualitative terms in which formations, structures and other geological features comprising favourable environments for various types of mineral deposit, are identified. Such assessments have taken the form of metallogenic maps and reports on economic geology. More recently, research studies on the geology of deposits of economically important mineral commodities (copper, nickel, iron, etc.) have resulted in identification of geologically distinct types of deposits, definition of the geological environment favourable to their occurrence, development of concepts of processes leading to their deposition and, by comparison with the geology of a region, delineation of favourable areas.

An important further step consists of the application of statistics to combine individual opinions of several experts on the mineral potential of a region to reach consensus through the Delphi technique. This step can also provide a quantitative assessment as described by Barry and Freyman (1970) for the mineral endowment of northern British Columbia and the Yukon.

In the present study, statistical assessment of certain systematically mapped geological and geophysical variables and their relation to occurrence of known mineral deposits is used to assess the probability of occurrence of mineral deposits of various commodities throughout the region. A grid of cells of equal area is superimposed over the region, and some of the cells, containing mineral deposits of known size and grade, are selected as control cells. Using multivariate statistical techniques, the geology of each cell being assessed is compared with the geology of the control cells; the degree of geological similarity is used as an index of the probability of occurrence of mineral deposits of each commodity in the cells being assessed. The distribution of known mineral deposits throughout the region is also a factor.

The resulting maps are, in some respects, a quantitative equivalent to metallogenic maps. Together with these, they should be used as a basis for further appraisal by economic geologists, metallogenists, regional geologists and other specialists, using conceptual models, local geological

Original manuscript submitted: 29 November, 1971
Final version approved: 5 January, 1972

knowledge, broader geological features (e.g. tectonic plates, geosynclines, paleoclimates) and other variables not used for statistical analysis, to establish a more precise estimate of mineral potential. The Delphi method would be appropriate as a means of achieving consensus amongst these specialists.

The value of this method lies in the systematic measuring of geological variables, the objectivity of their correlation with occurrences of mineral deposits and the use of multivariate analysis to ascertain the relative importance of the variables, singly and in combination, to the occurrence of mineral deposits.

The obvious limitations of the method include the paucity of variables for which consistent data are available throughout the region, the assumptions that the applied geological variables are equally significant in different parts of the region and to different types of deposit, the possibility of undiscovered deposits in or near the control cells, the fact that coincidence of several local features in time and space may seem to be more significant than areally favourable features, and others that are discussed later.

Tests made of results, however, do indicate a significant correlation between the variables chosen and the occurrence of copper and zinc in the Abitibi region.

HISTORY OF METHOD

Quantification of geological features and subsequent use of mathematics, especially statistics, to treat the numerical data for classification, correlation and extrapolation are now widely applied in various branches of geology. Although the basic foundations for the framework of application were laid earlier, large-scale processing of data became possible with the advent of the computer. A review of computer usage in the various fields of geology has been given by different authors in Merriam (1969). The initial applications to economic geology were mainly devoted to the delineation of orebodies. For a review of developments in this field, reference is made to a Symposium volume edited by Weiss (1969) and a recent book by Koch and Link (1971).

An early major application of mathematical reasoning to exploration covering a larger region was made by Allais (1957). After examining certain statistics of the sizes and spatial distribution of orebodies, Allais proposed that both the size and the place of occurrence of an orebody were random events. From this, he developed a model with which the amount of metal to be found per unit area could be predicted. The approach was probabilistic in that the uncertainty of the predicted amount was assessed by the use of confidence limits. Since Allais gave no consideration to geology, the amounts predicted for any area depended only on the size of the area.

Griffiths (1966) refined the Allais model by allowing for the possibility of the clustering of orebodies and using a different function for the spatial distribution of deposits.

Harris (1965) made a significant advance over previous work by taking account of the regional geology of the areas under study. He quantified the geology by making measurements of areas and lengths of various geological features from maps. As a result, Harris was able to make predictions of the total value in dollars of metals to be found in various areas which depended not only on the sizes of the areas but also on their geology. The mathematical constraints of the methods used required that relatively large (400 square miles) units of area, be considered.

Many geological features can be treated as "presence/absence" variables if measured on small areas; for example, a mineral deposit either occurs in a cell or it does not occur. Again, a small area will not contain all rock types and other features present in the region. The presence or absence of a specific feature in an area may be a significant indication of the presence of some kind of mineralization. Recent work by Botbol (1971) and Vyshemirsky et al. (1971) is based on the statistical analysis of data of this type. An important result of these studies is that the presence or absence of combinations of geological features are usually better indicators of certain types of mineralization than the individual features taken alone.

GENERAL DISCUSSION OF METHOD USED

In the present report, we draw on several of the models discussed above. A more complete discussion will be found in the sections on methodology; the following is a brief account of what was done.

In a manner similar to most earlier studies, we subdivide the region under consideration into small (10 x 10 km) cells. We assume that the Allais model is essentially correct for these cells and also for (40 x 40 km) areas consisting of groups of cells, but that the mineral potential varies from cell to cell, being dependent on the geology. It is not required that the spatial distribution or sizes of deposits fit the specific distributions proposed by Allais or Griffiths; rather, we consider that the occurrence of one or more orebodies in the same cell is a single "event" (in a probabilistic sense). This leads to a situation where the number of new "events" in an area consisting of a group of cells is independent of the initial distribution of deposits throughout the individual cells.

Rather than subdivide the region into a "control area" and a "study area" as others have done (e.g. Harris, 1965; Kelly and Sheriff, 1969), we select a number of cells within the region to be used as control cells - these are not, in general, contiguous, the chief criterion being the existence of substantial known mineralization of a specific commodity.

As Harris has done, we have quantified a number of geological variables for all cells by making measurements from geological maps.

The methods of analysis, however, do not have the same constraints as those Harris used, so that we have been able to make the cells less than one-tenth the area of his.

We have also drawn on the work of Botol and Vyshemirsky et al. by forming new variables from pairs of the original set of variables which were measured.

Simply stated, the method consists of making a quantitative comparison between the geology of each cell where potential is to be predicted and the aggregate of control cells, taking into consideration the distribution of known mineral occurrences in the area.

The degree of similarity is used as a measure of the probability of the existence of undiscovered mineralization of stated amounts for the commodities being considered in the cells being assessed. The results are presented as probability contour maps for the number of 10 x 10 km cells with significant mineralization expected per group of sixteen cells covering a 40 x 40 km area.

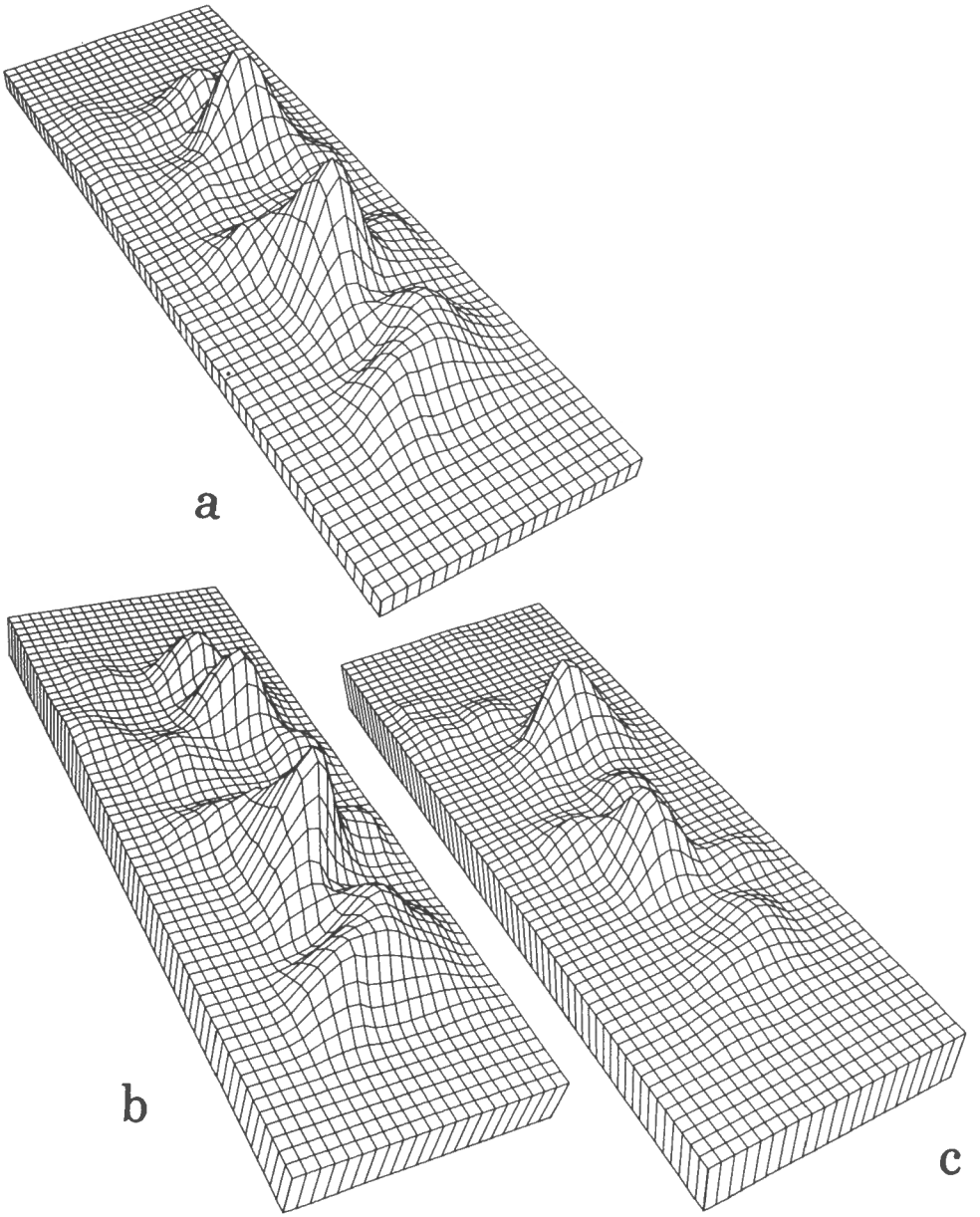
A brief explanation of the method is given on the copper potential map (Fig. 1). A glossary of terms with a specific meaning in this paper is given in Appendix 4.

The basic assumption for constructing a probability map such as Figure 1 is that the probability of occurrence of mineral deposits is systematically determined by composition and structure of the geological framework.

The contours apply to both discovered and undiscovered deposits of a commodity. The potential of any specific area is the difference between the total amount of mineralization predicted for that area (based on the contour values) and the total known mineralization in that area. The reader is presented with two separate patterns on a map (one for probabilities and one for known data) instead of a single pattern for the difference (potential). We do this so that the reader can evaluate the contours in a qualitative manner on the basis of their degree of matching in places where deposits were discovered in the past. Furthermore, the probability index has been scaled in such a manner that, on the average, the predicted amount of mineralization is equal to known amount in parts of the area which are considered to be well-explored and where mineralization is known to be. Consequently, for contouring, we have assumed that the potential is zero in the richest parts of the area where most mines occur. Known mineralization is quantified for statistical analysis as the sum of past production for mines and reserves as estimated in 1968. These reserves are "measured" and, in part, "indicated". It is likely, however, that in the control cells and in cells adjacent to the control cells, there exist more orebodies which have not yet been discovered or for which estimates of reserves have not yet been published for lack of development. During evaluation, the reader may rescale the contours on the map by multiplication by a constant factor to account for "possible" reserves in the mining camps. It would not be possible to rescale the contours by a simple operation if the map applies to the difference between contours and known values.

The contours given indicate whether or not a geological environment within a 40 x 40 km cell is favourable. The probability that mineralization will occur can be contoured but, as yet, we have not devised a method for contouring the known mineralization. If the difference is calculated for 40 x 40 km cells, this type of potential map would not correspond to a smooth surface but would show many local discontinuities coinciding with the artificial boundaries between cells of the grid overlain on the region.

The probability index is calculated by first comparing the pattern of control cells with the patterns of all geological (and geophysical) variables that were coded. The variable with the best match is selected. Next, the pattern of control cells is compared with all possible combinations of the variable selected and each of the remaining variables. This gives the best-matching combination of two variables. This procedure is repeated and, during each consecutive step, a new geological variable is added. The total effect of adding new variables decreases gradually when more and more variables are added to the list. At a given point, it makes little difference whether or not more variables are included. At this cut-off value, an equation is calculated for the relationship between known mineralization and those geological variables selected as "relevant". By means of this equation, a value is computed for every 10 x 10 km cell, which is, in turn, converted to a probability measure expressing the chances that a cell contains significant mineralization or not.



(a) Values of probability index before contouring;

(b) Ditto, by using 19 control cells in western Quebec only;

(c) Ditto by using 8 control cells in east-central Ontario only.

Figure 2. Schematic block diagram for copper potential; lines represent UTM-grid with spacing of 10 kms; boundaries as for Figure 10.

These probabilities are added for groups of sixteen cells at a time. The resulting sums give the probable number of mineralized 10 x 10 km cells per larger 40 x 40 km unit area. For copper, it is shown in the three-dimensional block diagram of Figure 2(a). Each grid-line intersection in Figure 2(a) represents one value of the probability index computed for the enclosing 40 x 40 km cell. The probability index can be contoured in the same manner as elevations above sea level are contoured on a topographical map.

The present method has been tested rather extensively as will be shown later. In Figure 2, we present the result of one of these tests. Figure 2(a) was calculated from data on copper deposits in twenty-seven control cells. Eight of these lie in the Ontario part of the Abitibi area and nineteen in Quebec. Suppose that no copper deposits had been found in Ontario. We may then use the nineteen control cells in Quebec to make a prediction for Ontario. The result is shown in Figure 2(b). The resemblance between Figures 2(a) and (b) indicates that the probability contours are not changed significantly by making use of the copper discoveries in Ontario.

The pattern changes, however, when the objective is to predict copper mineralization in Quebec by using the eight control cells in Ontario. This evaluation is shown in Figure 2(c). However, the discovery of the mining area near Timmins, Ontario would have led to the discovery of that near Noranda, Quebec, because of strong similarities in geology between these two areas.

The preceding remarks illustrated that our evaluation of mineral potential is largely "two-dimensional". By measuring features shown on maps, all cells in the larger region are compared with one another and classified on the basis of occurrence of mappable features coexisting with orebodies in the control cells.

The method can be summarized by saying that the pattern of known deposits is contoured in terms of variables which can be measured everywhere in the region. This is done in the hope that we are contouring not only known deposits, but also the undiscovered deposits.

On the whole, the Abitibi area is well-explored. When mineralization is significantly correlated to geological factors, one would expect the contoured probability index to coincide more or less with the pattern of known deposits if the region is well-explored everywhere. This is the case in Figure 1. However, in several sub-areas indicated by a pattern in Figure 3, the contours indicate copper potential in places where few or no deposits were discovered in the past.

Our predicted potential for such areas should be evaluated further by experts. For example, the amount of glacial debris in a favourable area may be such that local exploration by geophysical methods is hindered; or, on the basis of other factors, not used during the statistical analysis, a metallogenist or regional geologists may be able to suggest that a specific area is not likely to be mineralized although its measurable features are similar to those in the control cells.

GEOLOGY AND MINERAL DEPOSITS IN THE STUDY AREA

The Abitibi area studied in this paper covers the southern parts of the so-called Abitibi belt of the Superior Province of the Canadian Shield. In total, the area is 86,200 square kilometers (or 33,282 square miles).

It surrounds the mining camps of Timmins and Kirkland Lake in east-central Ontario and Noranda and Val-d'Or in western Quebec. The boundaries (see Fig. 1) are longitudes 76° and 84°, latitudes 47°30' and 49° in Ontario; and latitudes 48° and 49° in Quebec. This represents close to one per cent of Canada's total land area or six per cent of the Superior Province. Alternatively, it covers about five per cent of the combined areas of the Superior and Southern geological provinces.

A brief geological description has been given by McGlynn (1970, pp. 60-63) and an account of base metal deposits in the area by Lang (pp. 185-194).

The Abitibi area (Fig. 1) consists of ninety per cent Archean rocks, and of ten per cent Proterozoic rocks of the Cobalt Group. The oldest Archean volcanic and sedimentary rocks in the area form easterly trending belts separated by rocks of granitic composition. They were repeatedly intruded by mafic and ultramafic intrusions which are older than the acid intrusive stocks in the area. Locally, they are unconformably overlain by the predominantly sedimentary rocks of the Timiskaming Group.

The oldest rocks were repeatedly subjected to deformation and metamorphism during Archean time. Most folding and faulting took place during the Kenoran orogeny at which time many granitic stocks intruded. The latter range in age from 2,300 to 2,600 million years averaging 2,490 m.y. for the entire Superior Province (Stockwell, 1964).

In the southern parts of the area, this basement complex is overlain unconformably by Proterozoic sediments of the Cobalt Group. During Proterozoic time, the Archean rocks were intruded by at least three swarms of diabase dykes (Fahrig et al., 1965, and Fahrig, personal communication). These are the north-striking Matachewan swarm (about 2,485 m.y.) the southwest-directed Preissac swarm (approx. 2,000 m.y.) and the west-southwest-directed Abitibi swarm (approx. 1,230 m.y.).

Several major faults remained active after the Kenoran orogeny which closed Archean time. Several younger alkaline-syenite-carbonatite masses occur in the area (Carlson and Donovan, 1965). Paleozoic sediments of Ordovician and Silurian age occur south of Kirkland Lake in a northwesterly trending belt.

The area is, in many places, covered by glacial debris which often reaches considerable thickness, representing one of the major obstacles to exploration and the systematic mapping of bedrock geology.

Statistical data on large copper and zinc deposits are given in Appendix 1 and Figure 3. The areal distribution of copper, zinc and gold deposits is shown in Figures 1, 4 and 5. Most production of metal in the area was for the elements gold, silver, copper and zinc. The total production and reserve figures for these four elements are given in Table 1.

Most large base metal deposits in the area are lenticular, massive to disseminated, stratabound deposits enclosed by volcanic and sedimentary rocks of Archean age. These deposits contain mixtures, in various proportions, of pyrite, pyrrhotite, sphalerite, chalcopyrite and, in some instances, galena. They also may contain significant amounts of gold and silver. Although these deposits are essentially all of the same type, the relative amounts of metal recovered from them varies significantly, as shown in Figure 3. Zinc is not recovered from a number of deposits and significant copper may be absent in others.

Table 1

Totals of production and reserves for study area

Area Studied	Commodity or metal	Production (oz. or tons) x1000	Reserves (oz. or tons) x1000	P+R (oz. or tons) x1000	Percentages Reserves
Ontario area	Au	86, 941	4, 286	91, 227	
	Ag^1 (Au-Ag+ base metal)	35, 717	258, 114	293, 831	
	Ag^2 (Ag-Co)	85, 411	127	85, 538	
	$Ag^1 + Ag^2$	121, 128	258, 241	379, 370	
	Cu	207	801	1, 008	
	Zn	532	3, 878	4, 410	
Quebec area	Au	31, 061	4, 374	35, 435	
	Ag	79, 168	18, 741	97, 908	
	Cu	2, 226	234	2, 459	
	Zn	1, 682	513	2, 196	
Total area	Au	118, 002	8, 660	126, 662	6.86%
	Ag	200, 296	276, 982	477, 278	58.03%
	Cu	2, 432	1, 035	3, 467	29.85%
	Zn	2, 214	4, 391	6, 606	66.47%

In the western Quebec part of the area, pyrite-sphalerite bodies, from which zinc is recovered, tend to lie above pyrrhotite-chalcopyrite bodies with copper as the principal commodity.

Several authors believe that these deposits are related genetically to the volcanism. If this is so, they would be of Archean age and only slightly younger than the volcanic and sedimentary rocks in their vicinity. The deposits and their enclosing rocks were subsequently folded and tilted to their present attitudes during the Kenoran orogeny.

Goodwin and Ridler (1970) take this view and point out that the deposits occur in the vicinity of the sites of ancient volcanic centres. One of the characteristic features of these volcanic centres is the relative abundance of acid volcanic rocks (rhyolites and tuffs) in the upper parts of volcanic sequences, otherwise mainly consisting of mafic volcanics (basalts and andesites).

Kutina and Fabbri (1971) have done an extensive study on the relationship between copper and gold deposits and structural lineaments in the area. They found that structural control can be important, both locally and regionally.

Although it was mentioned above that the large copper-zinc deposits are essentially of one type, there are a few notable exceptions to this rule for the area. At the McIntyre gold mine near Timmins, Ontario, an important copper deposit was also found (No. 6, Appendix 1 and Fig. 5). At this place, the copper occurs as a dissemination of chalcopyrite in the Pearl Lake quartz-feldspar porphyry stock.

About half of the copper production from the Canadian Shield comes from nickel-copper mines. All important nickel deposits in the Canadian Shield are located in or near magmatic rocks of mafic or ultramafic composition, which were their source. This type of deposit is rare in the Abitibi area. In Appendix 1, the only large deposit of this type with significant copper reserves is No. 56, where the nickel and copper are associated with a mafic intrusion.

For the statistical analysis carried out to construct the contours of Figures 1 and 3, only large copper and zinc deposits were selected containing at least 1,000 tons of copper and/or zinc, respectively. They were treated as if they were of a single metallogenic type and were related to features of Archean bedrock only. Cells falling in the parts of the area covered by Proterozoic rocks were not used for this statistical comparison.

Large silver-cobalt orebodies are associated with mafic sills of Proterozoic age in the Cobalt Subprovince to the south.

Large, low-grade, banded cherty iron deposits with magnetite and/or hematite, of the Algoma type, are common in the sedimentary parts of the Archean volcanic-sedimentary piles, e.g. at Kirkland Lake. Their place of occurrence in the area is one of the factors considered during the statistical analysis.

Various types of geophysical expressions of the Archean bedrock also contribute information to the geomathematical classification of the 10 x 10 km cells. For example, Roscoe (1965) has pointed out that various types of volcanic rocks have a different signature on regional aeromagnetic maps. Bouguer anomalies provide approximate information on the thickness of the volcanic pile in the earth's crust that overlies rocks of granitic or granodioritic composition. This is because mafic volcanic rocks have a higher specific gravity than other types of Archean rocks in the area. Since the volcanic rocks were subjected to folding and erosion, the Bouguer anomaly does not provide a measure of the original thickness of the Archean volcanic-sedimentary complexes. Moorhouse (1965) has estimated that, for these rocks, the stratigraphic column reached a maximum thickness of 59,000 feet in the Quebec part of the area.

THE DATA BASE

Geological Data

In studies of this nature, geological data are usually quantified by superimposing a network of equal-area cells on maps of the area of interest and measuring areas, lengths, etc., in each cell. We have followed this procedure. Canadian topographic maps at scales of 1:250,000 and larger, published by the Surveys and Mapping Branch of the Department of Energy, Mines and Resources, have overprinted on them the Universal Transverse Mercator (UTM) grid system. In this system, the surface of the earth

is divided into 60 zones, each zone being 6° of longitude in width. A rectangular co-ordinate system is constructed in each zone; the origin of each zone being on the equator at a point 500,000 metres west of the central meridian of the zone for points in the northern hemisphere (The origin for each zone is in fact west of the zone; this convention is adopted so that there will be no negative co-ordinates). The position of a point is then specified by giving the zone number, the easting (i.e., the x-co-ordinate) and the northing (y-co-ordinate). For points in the southern hemisphere, a "false northing" of 10,000,000 metres is added, also to avoid negative co-ordinates.

The grid lines printed on Canadian maps at a scale of 1:250,000 are at 10 km intervals. This produces a convenient precise grid system for our purposes. Although most of the study area is in UTM zone 17, the easternmost 2° lies in zone 18. To preserve the equal-area cell structure therefore, it was necessary to extend the grid zone 17 eastwards into zone 18. In a small extension such as this, the errors introduced are negligible. Figure 3 shows the grid used, in relation to the geographic co-ordinate system.

Most geological maps used for data collection were at a scale of 1:253,440 (1 inch = 4 miles). The grids were fitted by photographic reduction of the 1:250,000 grids, matching the geographic co-ordinates of selected reference points (for one entire NTS lettered quadrangle or part of it) by trial and error to compensate for such factors as paper shrinkage. An additional complication was that some of the geological maps were based on a Lambert Conformal projection, whereas the grids were based on the Transverse Mercator projection. As a result, exact fits could not be obtained. However, in all cases, the errors were less than a distance equivalent to 0.25 km on the ground (i.e., 2.5%).

Each grid square was subdivided into 400 (1/2 x 1/2 km) sub-cells. Using a graticule, the rock type present at the central point of each sub-cell was noted, resulting in 400 "yes/no" type data for each cell.

Most "point-counting" was done using Ontario Department of Mines maps 2046 and 2116 and Map 1600-V accompanying Special Paper No. 2 of the Quebec Department of Natural Resources. In some areas more recent larger scale maps were used, notably for the area north of Timmins, Ontario, where use was made of a map compiled by J. L. Kirwan (Appended to Ph. D thesis, University of London, England, 1968).

Contacts between rocks of different types are often so irregularly shaped in nature that they must be generalized for representation on geological maps, the degree of generalization depending on the scale of the map. Furthermore, although rock types are shown as being mutually exclusive on maps, the ground truth is that interbedding of different rock types is commonly leading to another kind of generalization. Consequently, point count measurements made from maps may not be individually meaningful. However, totals for larger cells become increasingly representative of the mapped data as a whole.

In order to produce meaningful results with minimum bias, it is necessary to use as quantified variables only those geological features which are mapped consistently over the whole area. This imposes severe restrictions on the amount of information which can be derived from geological maps. Local areas in the neighbourhood of large operating mines or well-developed prospects tend to be more thoroughly explored and mapped in greater detail than less productive areas. Consequently, whereas one map may simply show a particular unit as "granitic rocks", an adjacent sheet may show its continuation subdivided into other units such as granodiorite, syenite

and quartz porphyry. The question of scale also comes in: for a given constant "graininess" of map, the detail will be greater on larger scale maps.

Another complication is that parts of the area are covered by very thick overburden and the bedrock character is unknown. These cells could not be used for analysis. Where there are adequate drillhole data, of course, this problem is obviated.

These considerations, plus the fact that all of the deposits included in the data base occur in rocks of Archean age, led to a decision to use a cell for analysis only if 25 per cent or more of its area was known to be covered by Archean rocks. The areas were recalculated to percentage values for Archean rocks by the elimination of points covered by rocks of other ages, thick overburden, water, or unmapped terrain. The result was seven generalized variables for Archean rocks, which are mapped consistently in the whole area. They are shown in Table 2.

For stratigraphic control of the sedimentary rocks, the approximate total length of banded cherty iron-formations per cell was included. These data were derived from maps compiled by G.A. Gross, of the Geological Survey of Canada, at a scale of 1:1,000,000 (G.S.C. Preliminary Series, 1963: Nos. 13, 14, 16, and 17).

Table 2

Ten basic variables coded for Archean rocks
(percentage values for first seven variables)

1. Granitic rocks (acid intrusions and gneisses)
2. Mafic intrusions (gabbros and diorites)
3. Ultramafics
4. Acid volcanics (rhyolites and pyroclastics)
5. Mafic volcanics (basalts and andesites)
6. Archean sedimentary rocks
7. Metamorphic rocks of sedimentary origin
8. Approximate total length of layered iron formations per cell
9. Average Bouguer anomaly per cell
10. Regional aeromagnetic anomaly at cell centre.

Many geophysical parameters are known to be good indicators of potential mineralization. In the study area, data were available from aeromagnetic and gravity surveys.

Aeromagnetic anomaly values for the centre of each cell were derived from unpublished 1:500,000 scale maps supplied by the Exploration Geophysics Division of the Geological Survey of Canada. The anomalies plotted were the residual values of total magnetism after subtraction of the earth's main magnetic field.

Bouguer anomaly values for all gravity measurement stations in the area were obtained from the Gravity Division, Earth Physics Branch, Department of Energy, Mines and Resources, Ottawa. Values for two or more stations in the same cell were averaged – values for cells with no stations in them were computed by interpolation from surrounding cells.

The result is ten basic measurement variables, for 644 cells, as shown in Table 2. In Figure 10, to be discussed in more detail later, the unused cells are indicated by minus signs.

Associations of basic variables can be used to provide important new variables. Forty-five such associations were derived from the measured variables, taking the products of all possible pairs. (There are 45 combinations of 10 items taken 2 at a time.)

The main reason for using artificially derived variables such as these is to allow for some of the many possible interactions found in geology. Some examples are: contact between two rock types or the proximity of an intrusion to a nearby rock type may be more important factors in the development of concentrated mineralization than the rock types themselves; the association of sedimentary rocks and iron formations defines a different facies than the sediments taken alone.

Mineral Deposits

The primary data on mineral deposits used for this study are represented in Figures 1, 3, 4 and 5 and in Appendix 1. These data were obtained from such publications as the Financial Post Survey of Mines, the Canadian Mines Handbook and from unpublished data files of the Mineral Resources Branch, Department of Energy, Mines and Resources, Ottawa, the Ontario Department of Mines, Toronto, and the Quebec Department of Natural Resources, Quebec City.

The main emphasis was on the large copper and zinc deposits shown in Figure 3. Figures 1 and 4 illustrate the sizes of these deposits. Appendix 1 gives data on the sizes, grades and main enclosing rocks. In most cases, the estimates of tonnages and grades are as of the end of 1968.

The definition of terms such as deposit, orebody etc., usually derives from economic, rather than purely scientific considerations. For example, many mines contain more than one distinct orebody. Conversely, more than one mine may have been established in a single mineralized zone. Consequently, it is often difficult to sort out the available data.

To overcome this difficulty, the entity we have used for statistical analysis is the aggregate of all mines, orebodies, etc., in each cell.

Only those deposits whose total production and reserves exceeded 1,000 tons of copper and/or zinc were used for analysis. Although this cut-off point is arbitrary, deposits of this size account for practically all known production and reserves in the area.

The tonnage/grade estimates are reasonably accurate for mines and developed prospects. It is more difficult to classify the smaller occurrences by size, since they are not usually explored in three dimensions. As guidelines, we used classifications made by Dugas et al. (1967), Ferguson et al. (1971) and Shklanka et al. (1969) for data obtained from provincial departments. Gold occurrences (Fig. 5) were defined on the basis of grade – e.g., if one or more assays of a quartz vein indicated high gold content, it would be classified as a gold occurrence. Copper occurrences were more difficult to classify, since small amounts of high-grade chalcopyrite mineralization are common in the area. For zinc, the problem is less severe as sphalerite mineralization in the area is less widely dispersed than chalcopyrite. The larger copper and zinc occurrences are indicated by triangles in Figures 1 and 4. Some of the symbols represent several mineralized zones

within the same claim (1/4 x 1/4 mile) or group of adjacent claims. Usually these are less than 0.5 km apart. The spatial distribution of occurrences was such that no strict mapping rules could be followed at the scale shown; consequently, the patterns shown should be regarded as generalizations.

The clusters of deposits shown on the 1:500,000 maps do not provide a representative picture of the bedrock geology, as the local areas surrounding known large deposits have been, in general, more thoroughly explored. This may lead to an artificial intensification of the occurrence density in such areas.

In considering mineral deposits data, one encounters an immediate classification problem. If the data used are of a very general nature, (e.g., simply the number and sizes of all deposits regardless of type), there can be only one class. The resolving power of the prediction models is then low and the possibility of identifying those geological features strongly associated with a particular type of deposit does not exist.

On the other hand, if the parameters of each deposit are specified in great detail, the number of classes rises to the point where almost every deposit is unique; in this case, statistical analysis may have little or no meaning.

The problem, then, is to evolve a classification scheme in which a sufficient number of deposits will fall into each class to allow meaningful statistical analysis, while having enough classes to permit adequate resolution of the correlations between each class and its geological setting. The lack of adequate computer-processable mineral deposits files has prevented proper study of this problem in the past; however, this situation is now being rectified and work is in progress on the problem.

In this study, a fairly general classification scheme, which groups deposits by commodity, amount of contained metal and geological age, has been used. Although it is probably sub-optimal, it has proved adequate for our purposes.

METHODOLOGY

No attempt is made in this section to give a detailed, mathematical exposition of the techniques employed. Most of them are elementary and will be found in standard textbooks or reference works on mathematical statistics. Instead, the methods are discussed very briefly for the mathematically inclined reader. Those not so inclined may skip the first two parts of this section without loss of appreciation of the remainder of the paper. The section consists of three parts. In the first, the data on occurrences of copper and gold in the area are subjected to a preliminary statistical analysis without consideration of geological factors.

In the second part, the methods used to calculate a probability index for copper and zinc in the Abitibi area are discussed in more detail. When the method is applied (see third part), one is faced with having to take a number of decisions with regard to the types of variables to be related to one another, size of study area, spacing of grid overlain on the area, selection of control cells and size of control area. These factors will influence the strength of the relationship between mineralization and coded variables to be used for prediction.

The choice of these variables is very important for the end result. Initially, however, several decisions must be taken in a manner which may not be better than arbitrary. However, during analysis, by trial

and error, one may be able to improve the results considerably by modifying the factors initially selected. This does not mean that the data base (see previous section) is changed; rather that it is treated in a slightly different manner to bring out a relationship between mineralization and coded variables which is statistically significant and from which distorting factors have been eliminated as far as possible.

For example, in one preliminary experiment, we tried to relate the place of occurrence of some 718 copper deposits in the area to the coded geological variables. The result was little better than random because the smaller copper deposits shown by triangles in Figure 1 are not significantly correlated to their geological setting as coded for 10 x 10 km cells by the method discussed in the previous section. On the other hand, if the control cells are selected on the basis of large copper deposits, a statistically significant relationship emerges rapidly.

In other experiments, it was found that a single analysis for a large region may give better results than a number of similar analyses for smaller, separate parts of the same region. In other work based on smaller areas and fewer variables, it became apparent that the cell size of 10 x 10 km was too small and that better results could be obtained by also considering geological data from cells surrounding the control cells. The contours presented in Figures 1 and 4 of this paper will be tested for possible adverse effects of this nature in the third part of this section.

Another problem to be assessed here is which geological variables singly and in combination are most strongly correlated to copper mineralization. This problem was studied in detail. However, at this time, we can only provide a partial answer which cannot be separated from multivariable methodology. For example, it is well-known that in the Abitibi area, the large copper-zinc deposits tend to occur in places where acid volcanic rocks are relatively abundant. This would appear to be confirmed by the statistical analysis used to construct the contours of Figures 1 and 4 since acid volcanic is part of the first set of variables selected as most relevant out of the total set of 55 variables. However, we also did the statistical analysis with a modified data base where acid and mafic volcanic were lumped together as volcanics. The interesting result of the experiment was that the final contoured probability index has a pattern (Fig. 12(b), see later) which is not markedly different from that shown in Figure 1. However, more geological variables were needed before the same type of pattern was obtained. This does not mean that acid volcanics are not correlated to the occurrence of mineralization in the area. Rather, the fact that it is not required as a variable to do a geomathematical evaluation illustrates that the geological and geophysical variables that were used are strongly intercorrelated. From a statistical point of view, a combination of several geological variables may be as strongly correlated to a specific type of mineralization as another variable which, taken singly, is a more direct indicator.

Preliminary Statistical Analysis

Elementary statistical methods, including the use of histograms and the chi-squared goodness-of-fit test, were used to examine the distribution of deposit sizes, grades and the spatial distribution patterns. Figures 6 through 9, inclusive, illustrate these.

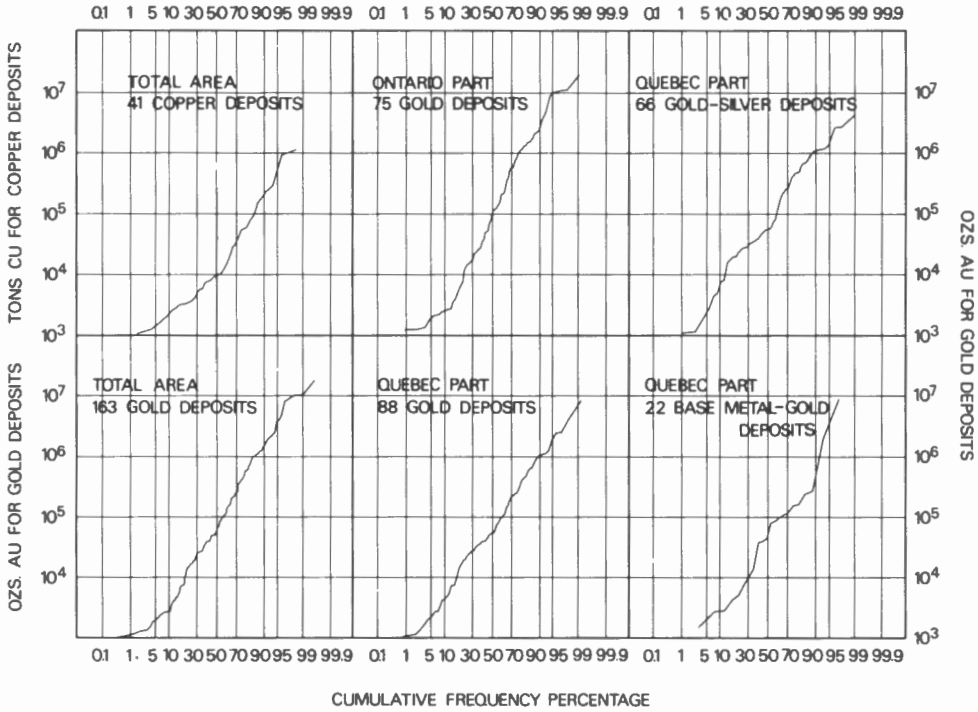


Figure 6. Size-frequency distributions for large copper and gold deposits plotted on logarithmic probability paper.

Note that the plots shown in Figure 6 are logarithmic. The shape of the curve indicates that the sizes of both copper and gold deposits are log-normally distributed, although Figure 6A suggests that the larger copper deposits may not satisfy this distribution. Chi-squared tests confirm these suggestions.

The histograms in Figure 7 indicate that the variability of average grade for large deposits is small in comparison with the variations in total size.

An interesting test was performed as part of some earlier, preliminary work on the same area. The ratio of the number of large deposits of gold and copper per unit area to the total number of known occurrences of these same commodities per unit area was computed. ("Large" in this context is determined by the criteria mentioned earlier, i.e., production plus reserves greater than 1,000 oz. of gold or 1,000 tons of copper.) At the time these calculations were made, 8 x 8-mile cells were being used, so that the unit of area was 64 square miles.

This ratio, R, tended to be constant in areas where the density of occurrences (of all sizes) exceeds a threshold value of 4 or 5 per cell (Fig. 9). For gold, the threshold value of 5 occurrences per 8 x 8-mile cell is also the approximate lower limit of density in gold-producing areas (see Fig. 8). A similar result for copper could not be obtained either for lack of sufficient data or because larger and smaller copper deposits are not as strongly related to one another as large and small gold deposits, or, more probably, because both preceding reasons are valid to some extent.

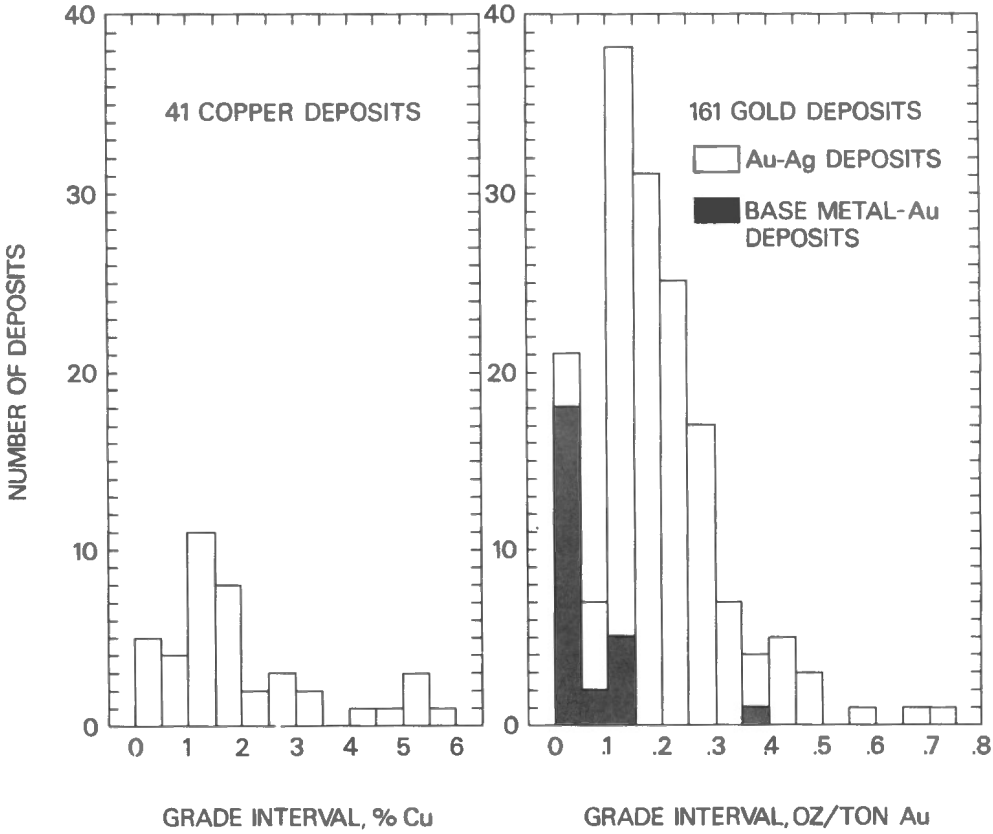


Figure 7. Histograms for grades of large copper and gold deposits.

Although Figure 9 shows the threshold value to be the same in both the Ontario and Quebec parts of the region the average ratio, $\overline{R}$, is smaller in Ontario than Quebec, for both gold and copper. These measures may be useful, in that both R and $\overline{R}$ are probably not very sensitive to variation in the intensity of exploration.

The significance of these results is not immediately obvious. We do not know to what extent they reflect natural populations of mineral deposits, since the data are no doubt biased, to an unknown degree, by economic effects.

It is probably safe to assume that the threshold value and lower tail of the size-frequency distributions are almost completely determined by economic factors.

The spatial distribution patterns for copper and gold occurrences have been studied in more detail by Agterberg and Fabbri (in press) who used methods of harmonic analysis.

Apart from these complications of an economic nature, the results of the preliminary analysis indicate that the spatial abundances of large and small deposits are related to one another. In general, the deposits form clusters. If the density increases, the average size of the deposits also increases. This relationship was more clearly established for gold than for copper.

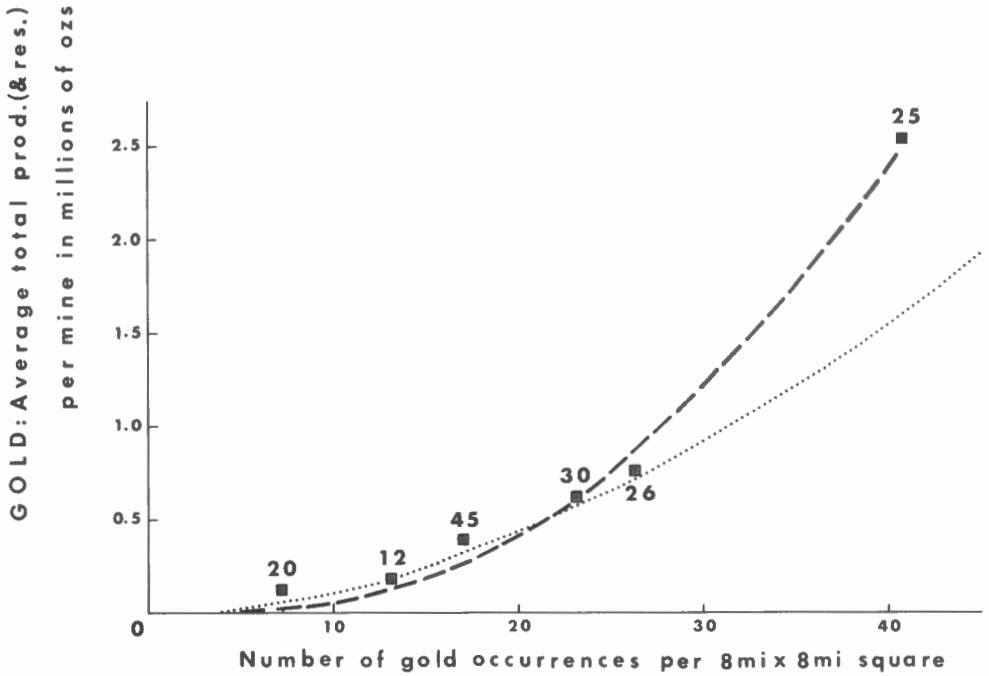


Figure 8. Average size of large gold deposits (1,000 oz.) plotted against total number of gold occurrences for 8 x 8 mile cells of a regular grid; Broken line is best-fitting curve for cells with more than 5 occurrences; solid line is same with corrections for some 8 x 8 mile cells which were relocated by moving them across important shear-zones on the geological map.

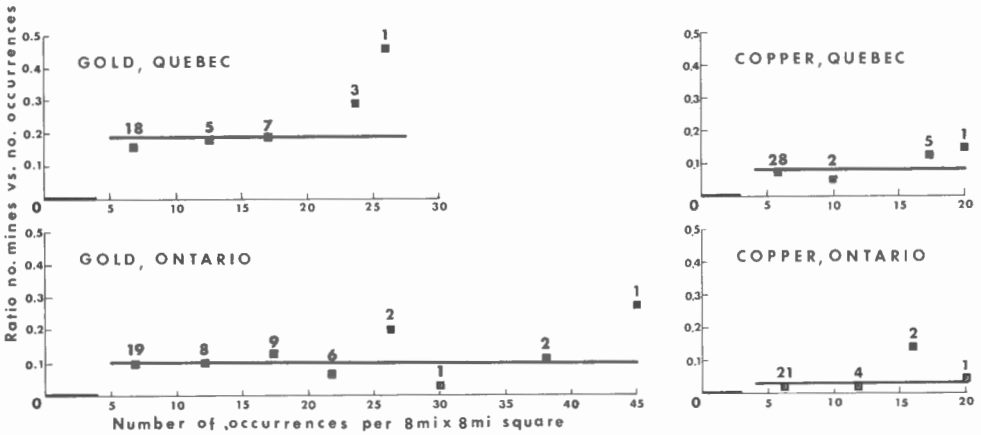
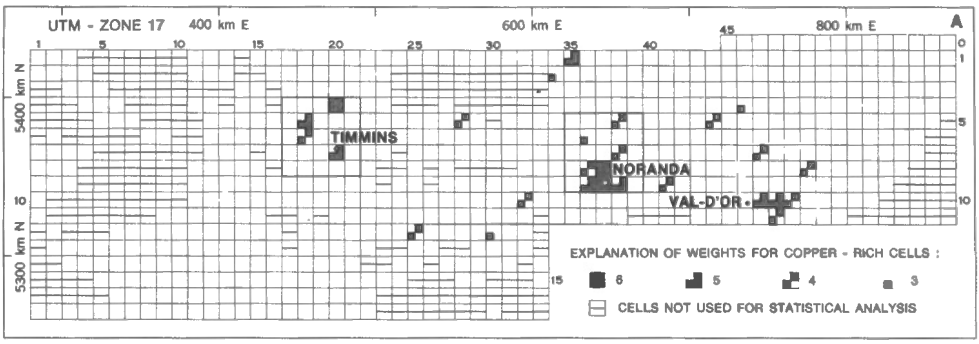
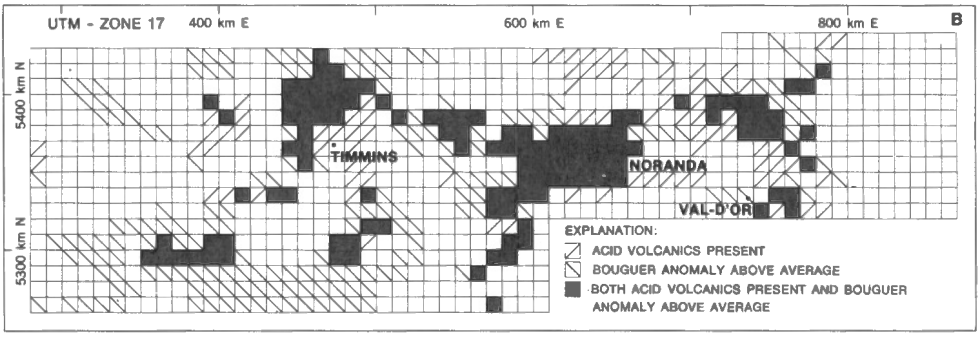


Figure 9. Average ratio of number of large deposits versus total number of deposits (large and small) plotted against total number of deposits per 8 x 8 mile cell. The best "fits" are horizontal lines for values based on cells with more than 5 gold or 4 copper occurrences.





(a) Control cells for copper and smaller control area consisting of two 50 x 50 km cells;

(b) Simplified example of presence-absence data for two variables and their combination.

Figure 10. Cells used for coding, based on UTM grid; those with less than 25% coverage of Archean rock are indicated.

Table 3

Frequency of occurrence of 10 x 10 km cells with one or more large copper and/or zinc deposits

Number of large deposits per cell	Number of cells with Copper deposits	Number of cells with Zinc deposits
0	617	627
1	19	13
2 (one or more)	6	3
3 (one or more)	1	0
7 (one or more)	1	1
one or more	27	17

In practical exploration, relationships of this type are well-known and have led to the rule that more exploration is warranted in places where discoveries have been made in the past. It is likely that this rule remains valid for the Abitibi area, in that undiscovered deposits are likely to occur in the vicinity of known deposits, particularly at some depth below the surface, where they are not easily found.

The spread in average grade values for the copper and zinc deposits of Appendix 1 (also see Fig. 7) is relatively small compared with the size values expressed by total amount of metal present. However, the size-frequency histogram becomes more or less symmetrical when a logarithmic scale is used for amount of metal. In the statistical analysis of the next section, total amount of copper and/or zinc per cell will be weighed by using the logarithm of total amount of metal present. An analogous procedure was used for the representation of individual large deposits in Figures 1, 4 and 5.

Calculation of Probability Index

As mentioned previously, data for two or more deposits falling in the same cell were combined. The 41 large copper deposits in the area fall in 27 control cells; the 26 zinc deposits occur in 17 control cells. The frequency distribution of number of copper and zinc deposits per control cell is shown in Table 3. Most control cells have one deposit, a few two or three, but one (cell no. 37, 8) near Noranda has seven large copper and zinc deposits.

Statistical analysis was done on the basis of values which are the common logarithm (base 10) of combined size for large deposits falling in the same cell. These copper "weights" are represented graphically in Figure 10(a). They constitute a pattern which was systematically correlated to the 55 geological and geophysical variables and to combinations of these variables.

In order to illustrate the method of calculation, a simplified example from Agterberg and Robinson (in press) is discussed first. The coded values for the 644 cells of the data base are numerical values. For variables 1-8 in Table 2, they can be reduced to binary data by replacing the numerical value by a one for cells where a feature is present and maintaining the zero when the feature is absent. The same was done for variables 9 and 10 by subtracting the mean for all 644 cells from cell values and replacing positive numbers by ones and negative numbers by zeros. The meaning of the 45 additional variables for pairs of basic variables now is as follows. A new variable for a pair is one when both variables constituting the pair are one (since $1 \times 1 = 1$). If one or both of the original variables are zero, the new variable is also zero ($1 \times 0 = 0 \times 1 = 0 \times 0 = 0$). Those familiar with set theory will recognize that the additional variables are intersections of two sets of cells each characterized by the presence of one or two basic variables.

The copper weights now can be connected to each of the 55 simplified variables. It was done by computing the 55 ordinary correlation coefficients. In general, one variable out of 55 has the largest correlation coefficient. This variable is selected for further analysis. For the copper in Abitibi, it is the combination of acid volcanics present and Bouguer anomaly above average. The latter two variables are represented both singly and in combination in Figure 10(b).

For further illustration, the copper mineralization pattern can also be reduced to binary form. For this the numerical value for every control cell is set equal to one. By graphical comparison of Figure 10(a) to 10(b), we count that 18 out of 27 control cells coincide with 18 out of 120 cells where acid volcanics are present and the Bouguer anomaly is above average. On the other hand, 9 out of 27 control cells fall in 644 - 120 = 524 other (non-black) cells in Figure 10(b). Suppose that one black cell is selected at random from Figure 10(b), then the probability that it is a control cell is 18/120 = 15.0 per cent. For a non-black cell in Figure 10(b), we have a probability of 9/524 = 1.7 per cent that it is a control cell. Figure 10(b) can now be replaced by a probability map by assigning a probability of 15.0 per cent to all black cells and of 1.7 per cent to other cells. This model would be appropriate if all large copper deposits had been discovered. Suppose, however, that instead of 27 cells with at least one major deposit, there are X such cells in the region, X minus 27 of which remain to be discovered. It would be reasonable to assume that out of these (X - 27) undiscovered deposits, as before, 18/27 = 67% will fall in black cells of Figure 10(b), and 33% in non-black cells.

Of course, we do not know how many cells with significant undiscovered mineralization occur in the region. This problem is solved by defining a control area to be used for scaling. For Figure 10(a), we assumed that 25 cells near Timmins and 25 cells near Noranda are well-explored and that all copper-rich cells have been discovered here. In total, 12 out of 50 cells in the control area are control cells. On the other hand, 32 black cells of Figure 10(b) fall in the control area.

Suppose that as before, the probability, is 15.0% that a black cell contains significant copper mineralization. The sum of one 15.0% probability for each of the 32 black cells in the control area amounts to 32 x 15.0% = 4.80. To this, 18 x 1.7% = 0.31 is added for the remaining (non-black) cells in the control area. Our estimated total becomes 5.11. It is the total number of copper-rich cells initially predicted to exist in the control area, and is less than the number actually known to exist in the control area that number being 12.

For this reason, we correct all probabilities by the factor 12/5.11 = 2.35. This gives a modified probability of 2.35 x 15.0% = 35.2% for black cells in Figure 10(b) and 4.0% for other cells. The final set of probabilities is such that the quantity "predicted-minus-known" is exactly zero in the control area.

We can now compute the potential for any block of cells. For the entire area, the probabilities add up to X = 120 x 35.2% + 524 x 4.0% = 63.2. Subtraction of the total number of control cells would give 63 - 27 = 36 cells for total potential of the Abitibi region or 36/27 = 1.33 times as much as has been already found. This calculation was carried through to its conclusion for illustration only. A better estimate of potential will be obtained later by adding other variables. One objection to this initial estimate is that all non-black cells in Figure 10(b) would have a 4% probability of being rich in copper. However, many of these cells are underlain by acid intrusions and gneisses for which the calculated probability should be less, since none of the known large deposits in the Abitibi area occur in a cell underlain by granitic Archean rocks only.

The previous results can also be obtained by performing a linear regression analysis to relate control cells in Figure 10(a) to black cells in Figure 10(b). The regression model is of the form

$$E(Y\ X) = B_o + B_1 X$$

When B_o and B_1 are estimated by the method of least squares,

$$\hat{B}_1 = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{N\Sigma X^2 - (\Sigma X)^2}$$

and

$$\hat{B}_o = \frac{\Sigma Y - \hat{B}_1 \Sigma X}{N}$$

The expression E(Y X) denotes the expected value of Y when X is known; Y is one for control cells and zero otherwise; X is one for black cells in Figure 10(b) and zero otherwise.

Hence, $\Sigma(Y\ X)$

$$\hat{B}_1 = \frac{644 \times 18 - 120 \times 27}{644 \times 120 - 120 \times 120} = 0.132824$$

and

$$\hat{B}_1 = \frac{27 - 0.132824 \times 120}{644} = 0.017176$$

We see that $\hat{B}_o = 1.7\%$ denotes the previously calculated probability for non-black cells whereas $(B_o + B_1)$ is equal to 15.0% for black cells (with X = 1). The advantage of using regression analysis is that instead of binary data, the original numerical values for the variables can be used. Also, by doing a multiple regression analysis, the "dependent" variable Y can simultaneously be related to two or more variables.

A conventional measure for the strength of the relationship between two variables is the square of the correlation coefficient $R^2_{y.x}$

with

$$R^2_{y.x} = \frac{\left[N\Sigma XY - (\Sigma X)(\Sigma Y)\right]^2}{\left[N\Sigma X^2 - (\Sigma X)^2\right]\left[N\Sigma Y^2 - (\Sigma Y)^2\right]}$$

For the previous example, $R^2_{y.x} = 0.067$

In most practical applications of regression analysis, such a low value would be regarded as indication that there is only a weak correlation between the variables X and Y. This is contradicted by a visual comparison of Figures 10(a) and (b) which suggests that the two variables are significantly correlated. The reason for a low value of $R^2_{y.x}$ is that the present result depends strongly upon the size of cells of the grid overlain on the area.

Suppose that the cells are made smaller by decreasing the spacing between grid lines. This has little effect upon the terms ΣXY, ΣY, ΣY^2 and $(\Sigma Y)^2$. For example, ΣY cannot exceed 41 which is the total number of large copper deposits in the area. On the other hand, the terms N, ΣX, ΣX^2, and $(\Sigma X)^2$ will increase rapidly when cell size decreases. The effect of this is that $R^2_{y.x}$ also will decrease. The explanation is that the variables X and Y have different geological characteristics. Y is related to the presence of mineral deposits which are relatively small bodies plotting as points on a small-scale map. On the other hand, the X's correspond to features which are more widespread in the area.

When multiple regression is used, $R^2_{y.x}$ is replaced by the so-called squared multiple correlation coefficient $R^2_{y.x_1\,x_2..}$ providing a measure of the strength of the multiple correlation. When the cell size is kept constant, this parameter can still be used to compare different solutions obtained by multiple regression. The statistical model used for multivariate analysis was

$$E(Y|X_1, X_2, X_3, \ldots) = B_o + B_1X_1 + B_2X_2 + B_3X_3 + \ldots$$

Y represents the logarithm (base 10) of total tonnage copper in control cells and is zero elsewhere. X_1, X_2, X_3, ... represent the geological variables that were coded. In total, there are 55 geological variables which are strongly interrelated with one another.

Multiple regression can be used to estimate the coefficients B_o, B_1, B_2, ... B_{55}. However, since the X_i's are, in part, percentage values which add up to 100% per cell, it is not possible to do a full regression on all 55 variables. In fact, there may exist a number of perfect or practically perfect linear relationships between the geological variables. In general, this has the effect that computing of the B's becomes difficult because of propagation of the truncation errors caused by the limited precision of the digital computer.

We used stepwise regression to obtain a solution. This method is fully described in Draper and Smith (1968). A computer algorithm for the basic operations was developed by Efroymson (1960). Miesch and Connor (1968) used the same method for their more elaborate algorithm and gave geological applications.

Stepwise regression essentially consists of the following procedure:

1. Select, from the simple correlation coefficients, the geological variable most highly correlated with metal value.
2. Using partial correlation coefficients, select the next most highly correlated variable. Perform a multiple regression.

2a. Having included two variables, repeat step 2 once, then go to step 3.

3. Perform a backward elimination; this is done by treating each variable, except the last one entered, as if it were the last, and calculating its contribution to the regression. If its contribution is significant, it is retained; if not, it is eliminated.
4. Repeat steps 2 and 3 until either all variables have been used or a pre-determined "Q-value" (explained below) has been reached.

The Q-values represent the probability that a given step is not significant in terms of the conventional F-distribution. This is a standard

method of determining the "goodness-of-fit" of a regression; it is explained in Draper and Smith (op. cit.) and most other standard works.

The stepwise regression that gave the contours in Figure 1 is shown in Appendix 2. The rank indicates the order in which the variables entered the regression. The product variable for acid volcanics and mafic volcanics was selected first. It has the lowest Q-value. Remember that we were operating on the data base itself without making the simplification leading to Figure 10(b) where Acid volcanics x Bouguer anomaly was selected first.

For this run, we specified that the multiple regression would be stopped at a predetermined Q-value equal to 0.5. The variable included at step 16 was later deleted at step 19, but this was the only backward pass made for $Q<0.5$. In total, 26 variables (out of 55) were selected. Their regression equation is of the form

$$\hat{Y} = -0.14863 + 0.00045\, X_1 + 0.00147\, X_2 \ldots\ldots\ldots$$

where $\hat{Y}$ is the calculated value for logarithm of tonnage copper per cell; X_1 is Acid volcanics x Mafic volcanics; X_2 is Granitic rocks x Acid volcanics; etc. The final regression coefficients are listed in Appendix 2.

By means of this equation, a value was calculated for each of the 644 cells. These are shown in Appendix 3 (They were multiplied by a factor of 10 and rounded, in order to made direct use of a computer print-out).

Calculation of moving averages on the basis of these values combined in blocks of sixteen at a time, and contouring the result gives the pattern shown in Figure 1.

In order to assign values to contours in Figure 1, we extended the argument used for the simplified example of Figure 10.

Within the study area, a control area was selected to contain mining camps of Timmins, Kirkland Lake, Noranda and Val d'Or. It consists of the 145 cells outlined by a frame in Appendix 3.

The simple method that we used in practice was as follows:

1. Take the sum of all calculated values in the control area.
2. Divide this sum by the number of control cells in the control area. This gives a scaling factor.
3. Divide calculated values for all cells in the area by the scaling factor.
4. Sum the scaled values for overlapping blocks of 16 cells.
5. Contour the results by computer program.

For the data of Appendix 3 the scaling factor amounts to 3.20 When all values are divided by 32 (accounting for the multiplication by 10 used for the print-out) they become, by approximation, probabilities. The probability for a single cell then indicates the chances that this cell contains a significant amount of copper. Addition, for blocks of 16 cells, gives the probable number of 10 x 10 km cells with metal per surrounding 40 x 40 km cell. From here on, we can use Allais' method for computing total amount of metal per area, as will be discussed in the section on use of the probability maps. The preceding method of assigning values to the contours, which can be interpreted readily, is approximate. The validity of the approximation can be tested, however.

The dependent variable Y used for regression is a so-called compound random variable.

It means that a calculated value $\hat{Y}$, obtained by regression reflects one or both of two variables expressing different properties of copper deposits in the region:
If $\hat{Y}$ is high, this means (1) that there is a high probability that the cell is rich in copper; or (2) that, if copper deposits are present, they have a high probability of being large. These two explanations are quite different; in reality, both may be valid to some extent. For convenience, we assumed that the first factor related to place of occurrence (1) is more important and that the second factor related to size of deposits (2) can be rejected. In practice, $\hat{Y}$ is then rewritten as the product $\hat{P} \times \bar{Y}$ where $\hat{P}$ is the estimated probability that the cell is mineralized and $\bar{Y}$ the average amount of metal per cell. If this explanation is valid, $\hat{Y}$ should be uncorrelated to Y for control cells.

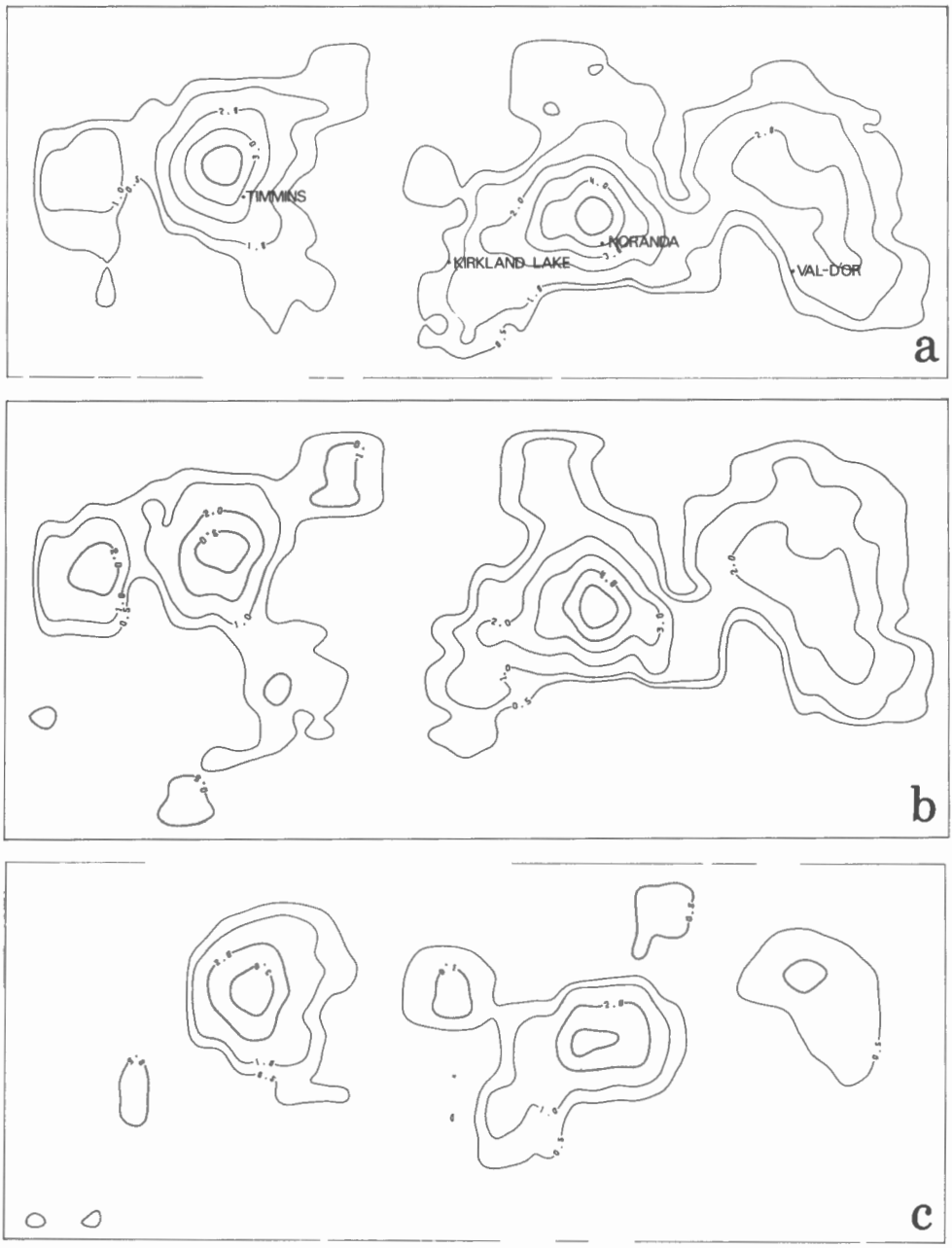
For copper, there are 27 control cells with Y-values (weights shown graphically in Figure 10(a) whose distribution is normal by close approximation. The conventional correlation coefficient between Y and $\hat{Y}$ amounts to 0.31. It can be tested for significance by consulting standard statistical tables for the product-moment correlation coefficient. Since there are only 27 data, the value 0.31 is not significant at the 95 per cent level. For zinc (17 control cells), this coefficient amounts to 0.11 which is also not statistically significant.

For the 644 cells on Figure 10a and Appendix 3, an initial set of probabilities $\hat{P}$ therefore is obtained by dividing all calculated values $\hat{Y}$ by $\bar{Y} = 4.31$, representing the average copper weight assigned to control cells. As before, these probabilities will underestimate the true probabilities. This effect is compensated for by summing $\hat{P}$ for the control area (145 cells in Appendix 3) and comparing the sum with the number of control cells in the control area, this being 24 for copper. The ratio of these two sums amounts to 1.35, and all initial estimates of $\hat{P}$ must be multiplied by it. This procedure gives exactly the same result as the simpler method described above, when the values $\hat{Y}$ were divided by 3.20 being the sum of $\hat{Y}$ in the control area divided by number of control cells (= 24) in this area.

The method is approximate and it is possible that some of the corrected probabilities for individual cells fall outside the interval $[0, 1]$ where they should lie. In Appendix 3, several of the calculated values are negative, which would correspond to negative probabilities for these cells. The effect of these values is largely eliminated when sums are taken for 16 values at a time.

In the contouring program, negative values for sums of 16 values are automatically replaced by zeros. It is also possible that probabilities computed this way exceed 100%, although this did not happen here. The largest value in Appendix 3 is 28 giving a probability of 28/32 = 88%.

The values for cells in Appendix 3 show a number of trends. Their local pattern of variation is irregular because they reflect local discontinuities (mainly contacts between rock types) in the geological framework. By enlarging the cell of reference, such local effects tend to disappear. The "moving average" becomes increasingly smooth and can be contoured for 40 x 40 km cells. A disadvantage of smoothing is that some resolution is lost. If the geology is favourable at a specific place, the probability index for 40 x 40 km cells will reflect this at points up to 20 km away, both in the



(a) Probability index scaled with respect to control area of 50 cells shown in Figure 10A;

(b) Ditto, computed from 19 control cells in Quebec; control area consists of 25 cells near Noranda;

(c) Ditto, from 8 control cells in Ontario; control area consists of 25 cells near Timmins.

Figure 11. Copper potential maps corresponding to Figure 2.

north-south and the east-west directions. Nevertheless, regional interpretation is greatly facilitated by using such contours and the probability index is meaningful if related to all terrain within the surrounding 40 x 40 km cell.

We have conducted a series of experiments to test the pattern of contours in Figure 1 by changing size of the control area, number of control cells, number of geological variables included for computing, sizes of grid-cells and by considering systematic variations in the pattern of copper mineralization perhaps not adequately expressed by the coded variables.

Some of these experiments are discussed in the third part of this section.

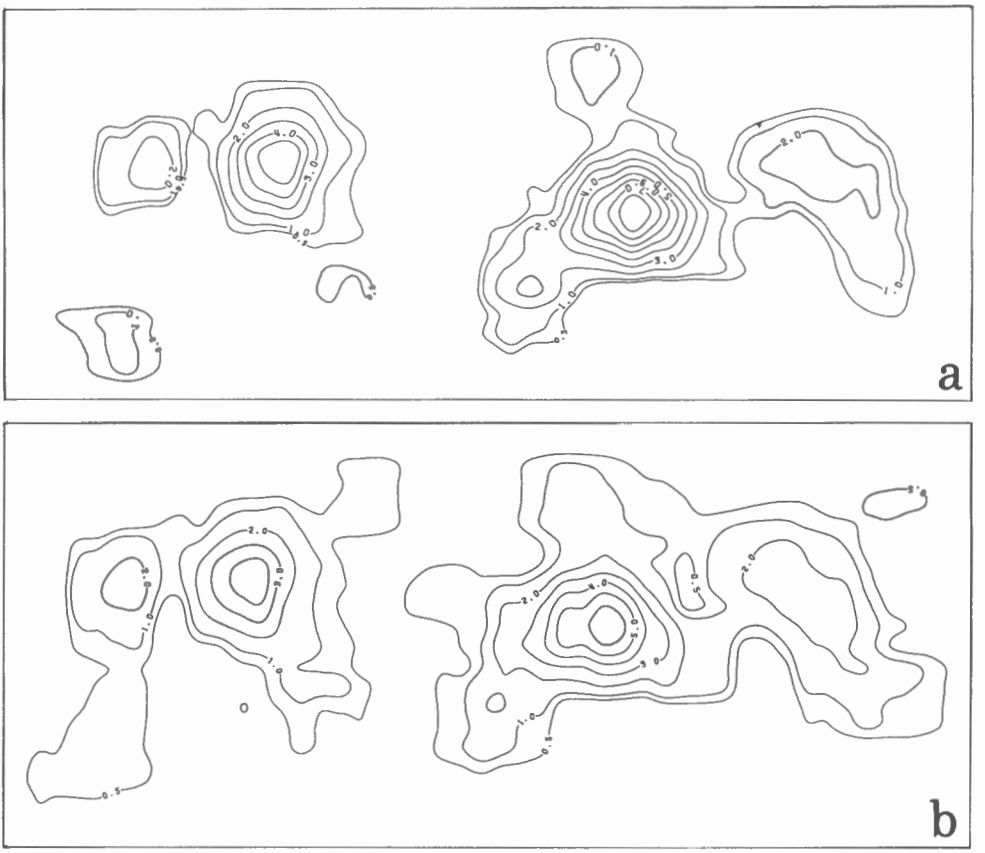
Other Probability Maps for Copper

The first experiment was discussed in part in the section on summary of method. Two 50 x 50 km squares near Timmins and Noranda are depicted in Figure 10(a). Together they can be used as a control area consisting of 50 cells, or about one third of the control area of 145 cells shown in Appendix 3. The contoured probability index for 55 geological variables, based on this new control area, is shown in Figure 11(a). Being similar, the results indicate that the exact shape and size of the control area is not a critical factor in this type of application.

The statistical analysis was repeated by omitting the 8 and 19 control cells falling in Ontario and Quebec, respectively. For the prediction based on Quebec (also see Fig. 2B), the 25 cells near Noranda were used as a control area. Those near Timmins were used for prediction from the 8 control cells in Ontario. The results are shown as contoured probability maps in Figures 11(b) and (c).

Mathematically, it is of interest that the 644 calculated values underlying Figure 11(a) are equal to the sum of the calculated values underlying Figures 11(b) and (c), respectively. It indicates that we can divide the pattern of probabilities for a larger region into components each of which is based on part of the control cells. For Figure 11, this sub-classification of control cells was based on geographical location. In other situations it can be based on types of deposit, if the control cells are defined for deposits which belong to different classes from a metallogenetic point of view. Several experiments were done to check the extent to which the contour map of Figure 1 is influenced by specific geological variables or groups of variables. The results obtained for two runs are shown in Figure 12. The statistical method used was the same as that for constructing Figure 1. The only difference was that fewer basic variables were used (see Table 2). In the run leading to Figure 12(a), variables 8, 9 and 10 of Table 2 were not used and variables 6 and 7 were combined into a single one by addition of percentage values for each cell. This left six basic variables from which 15 additional variables were constructed by combining the basic variables pairwise.

The result (Fig. 12(a)), is approximately equivalent to what can be coded using 4-mile geological maps and disregarding geophysical data, etc., obtained from other sources. On the whole, Figure 12(a) resembles the contoured pattern of Figure 1. Locally, however, there are discrepancies. For example, the peak near Noranda is higher in Figure 12(a) and there is a new area of high copper potential indicated to the south of the Timmins area.



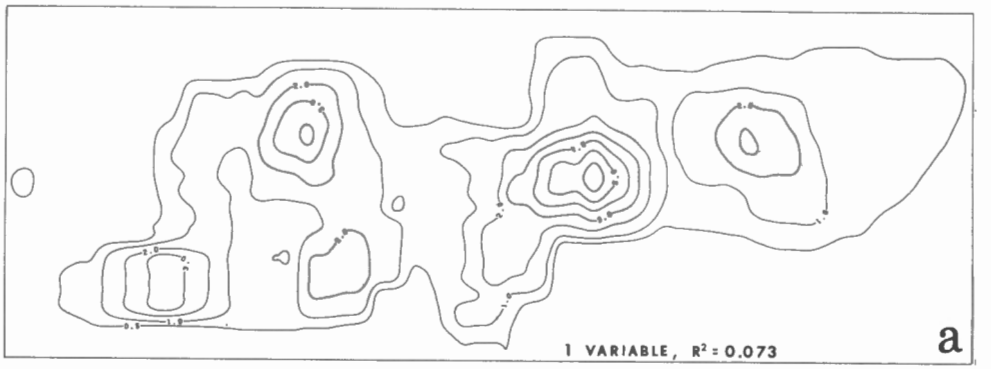
(a) Basic variables 8 – 10 (Table 2) deleted and 6 and 7 added to give a single new variable;

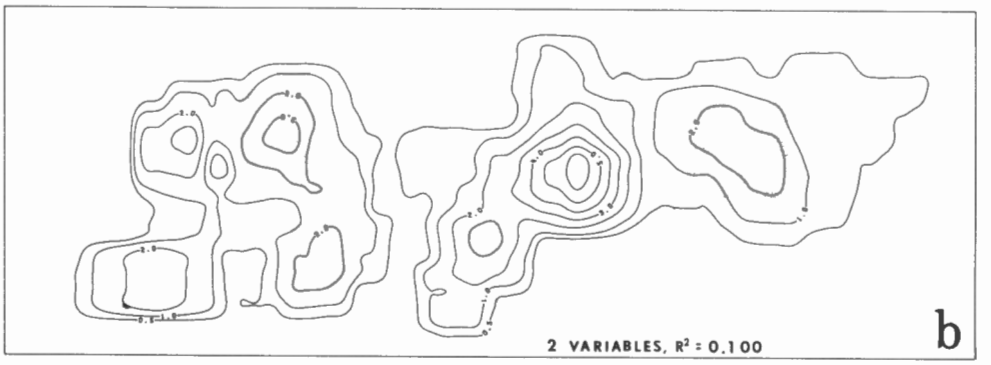
(b) Basic variables 4 and 5 were added to give undifferentiated volcanics as a single new variable.

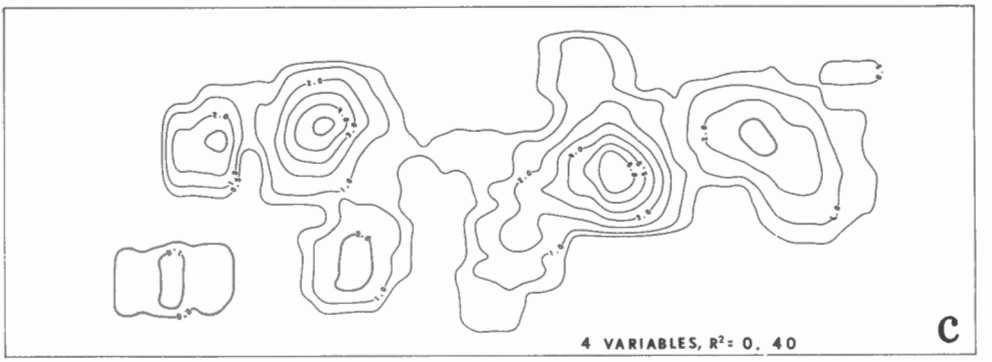
Figure 12. Copper potential maps computed by using fewer variables.

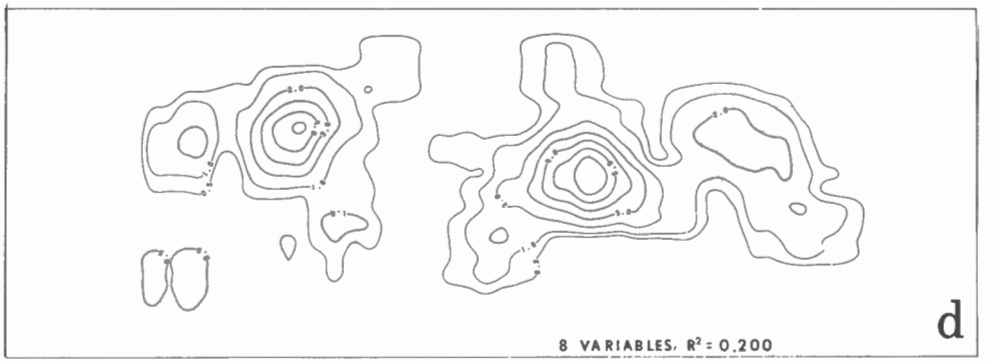
We recall that in Appendix 2, the first five variables that entered the stepwise repression for copper, contain acid volcanics as a factor, a factor which is singly most strongly correlated to occurrence of large copper (and zinc) deposits. We combined basic variables 4 and 5 in Table 2 into a single variable (undifferentiated volcanics) and redid the analysis. This gave 9 basic variables and 45 geological variables in total. The resulting potential map (Fig. 12(b)) is very similar to that shown in Figure 1.

The results of Figure 12 confirm that the geological variables in the data base are strongly intercorrelated. By working with many variables as was done for Figure 1, we admit considerable redundancy in the data base. However, during the multiple regression, this redundancy is automatically eliminated. For this type of multivariate statistical analysis, is may be best to start off with as many variables as possible and to let the elimination of redundancies be done by computer. However, the method is


1 VARIABLE, R^2 = 0.073
a


2 VARIABLES, R^2 = 0.100
b


4 VARIABLES, R^2 = 0. 40
c


8 VARIABLES, R^2 = 0.200
d

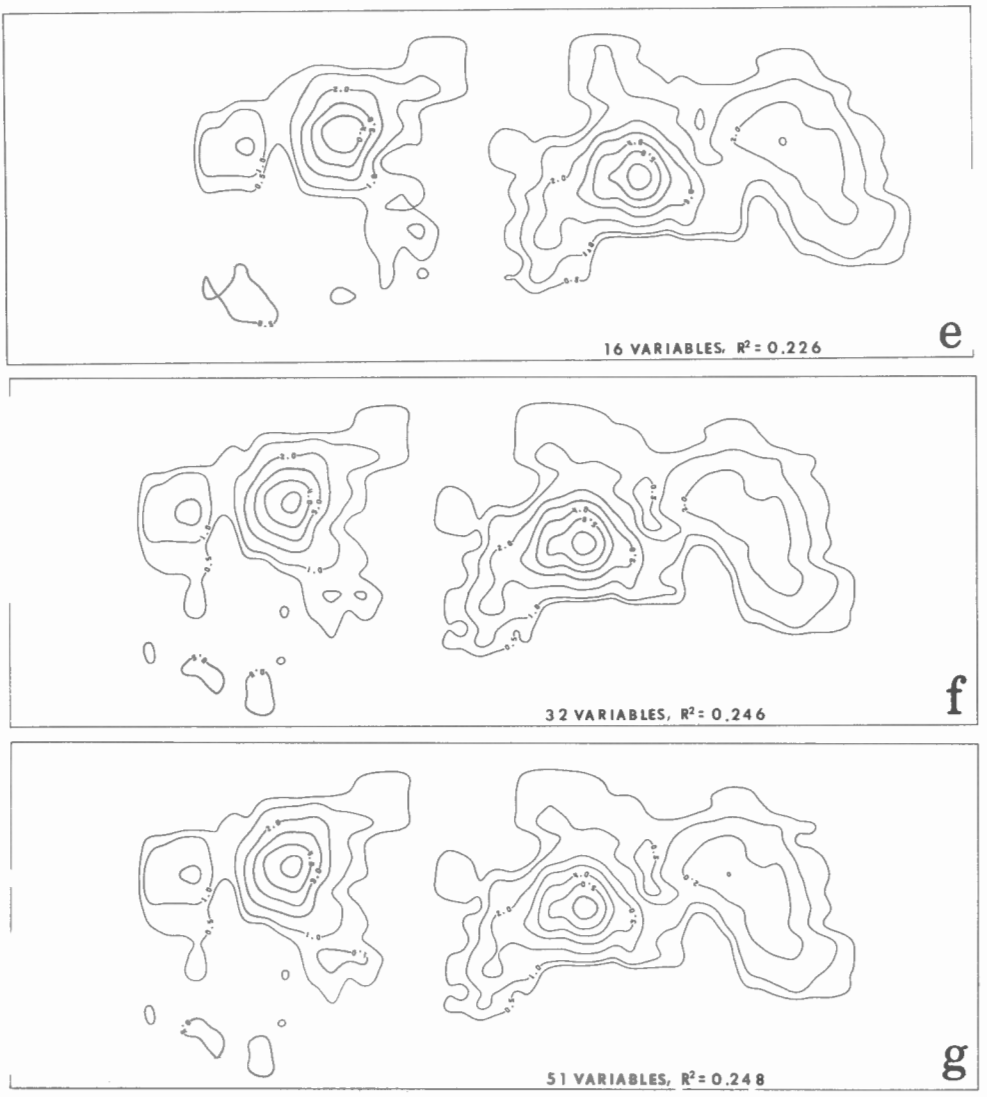


Figure 13. Seven copper potential maps computed by progressively adding more geological variables to multiple regression equation. Method is "forward selection" (as applied by Agterberg, 1964); The contours of Figure 1 are based on the refined method of "stepwise regression" with $Q < 0.5$ using 26 variables (as applied by Miesch and Connor, 1968). Example shows that addition of new variables rapidly becomes redundant.

well-suited to solve specific problems such as how relevant it is to map one or more variables in a region if the object it to establish the mineral potential of that region.

The ranks and weights of the variables as shown in Appendix 2 individually, are not meaningful indicators of relevance since they are also determined by the strong interrelationships between the geological variables. In fact, these effects increase very rapidly when more than one or two geological variables are used for multivariate analysis. Some of these problems were discussed in more detail by Agterberg (1971) and Agterberg and Kelly (1971). They are among the main reasons for considering, simultaneously, data for all cells in an area when a geomathematical evaluation is done, both when the multivariable relationship is being established and when individual cells are being classified.

The next experiment to be discussed was done to test for possible effects related to changing the predetermined Q-value of stepwise regression which, until now, was kept at 0.5. A run was done whereby the Q-value was set equal to 1.0 implying that all variables may enter the regression equation. Out of 55 variables, 51 variables did actually enter. Four were left out because there are several perfect or nearly perfect linear relationships between them and other variables. For example, basic variables 1-7 of Table 2 add up to 100% for each cell, preventing a solution of the regression coefficients when these seven variables are used at the same time. Seven full runs were done with the results shown in Figure 13(a)-(g). The stepwise regression for these runs was terminated after 1, 2, 4, 8, 16, 32 and 51 steps, respectively. The multiple correlation squared (R^2) for each run is shown in Figure 13. It increases from 0.073 to 0.248. However, it had reached 0.246 after 32 variables were used; the remaining variables had practically no effect on this measure of the strength of the relationship between copper mineralization and geological variables. The final pattern was reached much earlier, after 8 variables were used (Fig. 13(d)). Later changes achieved by adding more variables were minimal.

Suppose that in Appendix 2, we set the Q-value at 0.10. The cut-off then occurs after selection of approximately 9 variables. (This number is "approximate" because backward passes may slightly change the outcome.

The previous results indicate that it is not very important where we set the Q-level provided that it is at least 0.1, or when at least eight variables are used. We normally use Q = 0.5, admitting considerable redundancy among the variables that are used. A reason for doing this is that the computed Q-value may vary in an irregular manner at first, but its rate of change usually stabilizes at a point somewhere near Q = 0.3 where backward passes also begin to occur. Our experience with other runs for fewer variables was that setting Q = 0.1 as the cut-off level may result in a premature termination of the regression, where a higher cut-off value of Q would have led to significantly better results.

The dependent variable Y used during these runs is definitely not normally distributed because of abundant zero-values. This is one reason why the problem of which is the best solution cannot be decided exclusively by rigid statistical inference. Several results shown in Figures 11-13 could be slightly better than the one selected for Figure 1. The significance of slight differences in pattern could not be assessed by us. It was pointed out before (Fig. 10(b)) that taking a single variable for computing of copper potential already provides useful results. However, it also led to unrealistically

high probabilities for cells underlain by granites and gneisses. The same objection can be raised against Figures 13(a) and (b), where there are too many high probabilities for single cells. This is indicated by the shape of the 0.5 contour in these diagrams, which extends into barren terrain toward the east. This contour is also affected by the presence or absence of cells (– signs in Fig. 10(a) with less than 25 per cent of mappable Archean bedrock, since these cells were counted as zeros when the probability index for 40 x 40 km cells was computed. In the last experiment to be discussed here, the possibility was considered that the spacing of the UTM grid used for counting and analysis is too narrow and that a better map could have been obtained by relating the known mineralization to more regional changes in the constitution of the geological framework. Methods to deal with possibilities of this type are not elementary and a model that is satisfactory remains to be developed. A two-stage least-squares model approach was discussed by Agterberg and Cabilio (1969), and Agterberg (1970). For large copper and zinc deposits in Abitibi, the problem is not serious, since for these the geology of 10 x 10 km cells provide good control. Nevertheless, the following results are interesting.

To the 55 variables in the data base, we added 44 new variables, each of which is a function of one or both of the two geographical coordinates describing the location of a cell within the region. Together, these 44 variables make a polynomial of the eighth degree (octic polynomial). Otherwise, the 99 variables were treated as the 55 variables in previous runs ($Q<0.5$). The result is shown in Figure 14.

The part in the final regression equation containing the selected part of the 44 new variables can be represented separately. This gives a continuous polynomial surface of the eighth degree. It is shown for both copper and zinc in Figure 15. If the 55 original variables provide an adequate representation of the relationship between mineralization and geological framework, there should be little difference between the patterns of Figures 1 and 14. Also, the patterns of Figure 15 should be close to zero everywhere in the study area. These conditions are approximately satisfied for copper and zinc in the Abitibi area. Possible exceptions are the east-west trending copper peak west of Val d'Or and a zinc peak in this same area extending farther northward.

A comparison of the contours to occurrence of known copper deposits in Figure 1 suggests that in the vicinity of Val d'Or, the contours underestimate the amount of copper actually found in this area. In fact, this is the maximum negative discrepancy between contours and known data in the region. For a 40 x 40 km cell known to contain 5 control cells, the probability index is only 2. This would give a negative copper potential which, of course, is not possible. However, the probability index for 40 x 40 km cells can be regarded as the expected value of a binomial frequency distribution (see later). For $m = 2$, there is a 4 per cent probability that this actual number is 5 or 6. Consequently, the discrepancy near Val d'Or could be explained as a statistical fluctuation within the model used for Figure 1.

On the other hand, the discrepancy does not exist in the contours of Figure 14. If Figure 15 is compared with geological and metallogenetic maps for the region, it is immediately suggested that the unexplained copper peak is genetically related to the Cadillac Break, a major east-west stratigraphic discontinuity, accompanied by shear zones.

If this explanation is accepted, we can ask why structural lineaments such as the Cadillac Break were not coded and used in the data base, particularly since several large copper and zinc deposits elsewhere in the

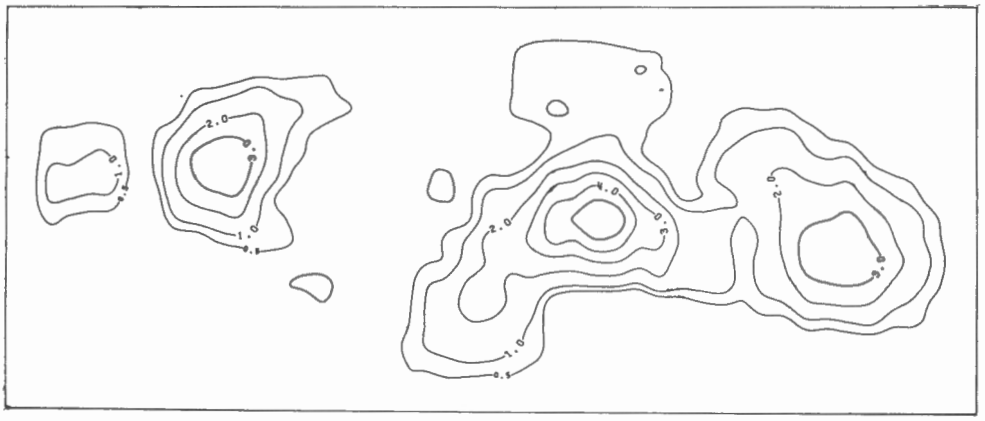


Figure 14. Copper potential map based on same variables as Figure 1 but with 44 additional variables to consider location of cells with respect to one another; Note that the "fit" in the Val d'Or area is improved.

area are also genetically related to geological structures. The reason is that it is difficult to describe in a systematic manner, structures in highly deformed, metamorphic terrain which is also extensively covered by glacial debris of various types. Locally, structural control is an important variable. By the construction of structural "trajectories" (Kutina and Fabbri, 1971) and relating patterns derived by harmonic analysis of occurrences to known lineaments (Agterberg and Fabbri, in press) some regional extrapolations can be made on the basis of structures. However, this topic is still in the field of research and has not led to simple practical methods useful for evaluating mineral potential on the Canadian Shield.

Sinclair and Woodsworth (1970), in a study of estimating exploration potential in an area near Terrace, B.C., have used distance of mineral deposits from a fault as a variable for multiple regression. This method can be used in areas where there is a well-described fault. In the Abitibi Belt there occur several important shear zones. The method employed for deriving Figures 14 and 15 accounts for the possibility that one or more of these shear zones are relevant factors since the relative location of all cells with respect to one another and implicitly to the shear zones is considered. Thus it is likely that the Cadillac Break is of both regional and local importance for copper mineralization but other faults and shear zones such as the Destor-Porcupine Fault may be of local importance only.

Relationships of Method to Methods of Multivariate Classification

Although the preceding calculation of the probability index was explained in terms of a linear least squares model, the procedure is closely related to several "classification" techniques such as cluster analysis, factor analysis (both R- and Q- modes) and, notably, discriminatory analysis. For those familiar with classification techniques, we formulate our problem in slightly different terms.

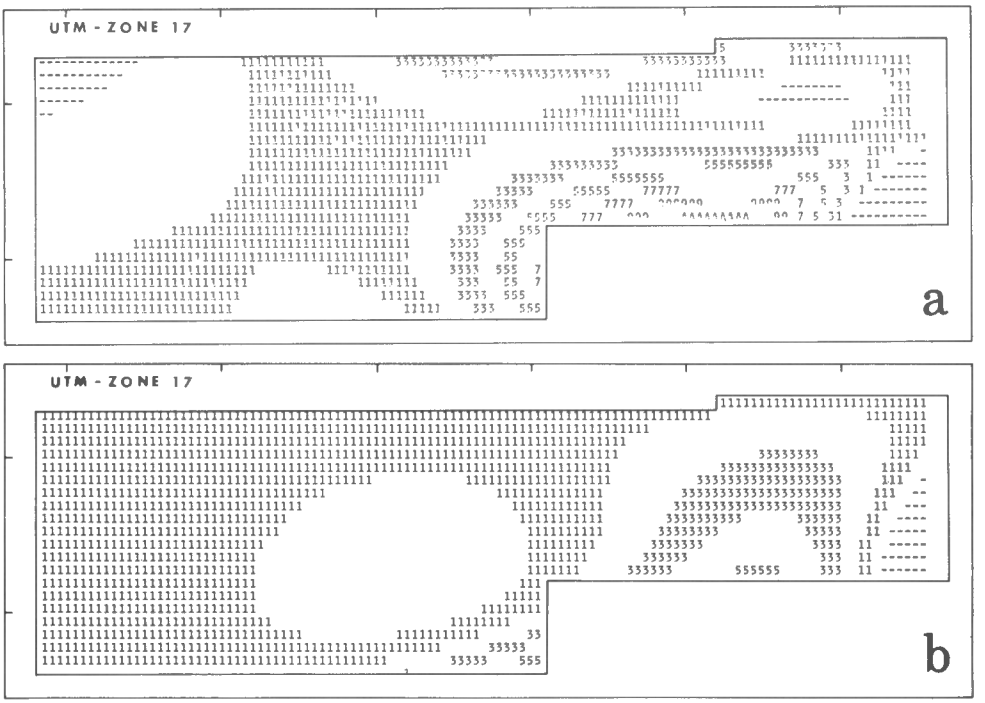


Figure 15. Octic polynomials for copper and zinc, expressing systematic regional variations not explained by the 55 geological and geophysical variables used for Figure 1.

There are 644 cells in the area and for each cell, p variables were considered. This gives a geological data matrix X of (p + 1) columns and 644 rows. (The first column of X consists of 644 ones). Initially, we had p = 55, but because of elimination of some variables p = 26 was used for Figure 1. At most, we had p = 51 for Figure 13(g).

Suppose that we form the square matrix X^TX with (p+1) columns and rows. This matrix is closely related to that usually taken as a starting point for cluster analysis and factor analysis (R-mode). In fact, in practice, we convert X^TX to the correlation matrix R for convenience. Essentially, we did not do a classification of geological data on the basis of X^TX (or R) but formed the square matrix $X(X^TX)^{-1}X^T$ with 644 columns and rows. Each column of $X(X^TX)^{-1}X^T$ corresponds to one cell in the area. Imagine that we made 644 different maps, one for each cell. Every map would have 644 values represented by one column of $X(X^TX)^{-1}X^T$. A single map would express the degree of similarity of all cells with one particular cell, which may be a control cell containing one or more large orebodies.

For copper, the 27 columns of $X(X^TX)^{-1}X^T$ for control cells can be added after weighing them according to values ranging from 3 to 6 (see Fig. 10(a)). The resulting 644 sums are shown in Appendix 3. Of course, these sums can be divided into parts corresponding to specific types of control cells. This was done by taking separate sums for the 19 and 8 control cells in western Quebec and east-central Ontario, respectively (see Figs. 2 and 11).

Stated in these terms the problem can be regarded as a classification problem. However, the classification of cells is done on the basis of similarity with control cells for one or more variables Y expressing mineralization. The data reduction is not done on X^TX or a related matrix (as in cluster analysis and R-mode factor analysis). It is also not done on XX^T or a related matrix (as in Q-mode factor analysis) but on $X(X^TX)^{-1}X^T$ with respect to cells where $Y > 0$. Those familiar with discrimatory analysis will recognize that the present method is, from a computational point of view, closely related to that of calculating a discriminant function to separate two populations from one another.

LIMITATIONS

The formation of a mineral deposit at a given place was, in general, determined by many physical and chemical factors and conditions most of which are unknown. In order to reconstruct the processes, which resulted in a specific deposit, the geologist is forced to work with limited data for a complex three-dimensional assemblage of rocks preserved after the metallogenic processes had ended. The changes that took place in the course of time cannot be studied directly.

It is possible to map three-dimensionally the geological setting of large orebodies which are being mined out. The surrounding areas of mines also may be well-known because of extensive drilling. By a detailed interpretation of the setting of deposits, much can be said about provenance, processes and favourable conditions. This has led to guidelines for prospecting, where the objective is to locate features elsewhere in the country indicating that favourable conditions may have existed. Features to look for may include certain types of faults and folds, other structures, rock types with specific properties such as porosity and permeability, contacts or other associations between two or more different types of rock, etc.

The method used in this paper is different. We could not make use of the many detailed relationships derived from the setting of specific deposits, mainly because most features listed above are rather specialized and unknown almost everywhere over a large region. Instead of this, our data base contains very limited facts available for most 10 x 10 km cells in the region. Consequently, the statistical results will not allow us to assess the probability that a specific type of deposit was formed at a given place which could, for example, be a good target for drilling.

Another set of factors providing useful indicators of possible occurrences of hidden orebodies are those more directly related to the physical (and chemical) presence of deposits. They include certain types of geophysical anomaly (e.g. those related to susceptibility, conductivity or specific gravity of the ore), geochemical halos, and mineralogical indicators (presence of specific ore minerals in the country rock, various types of alteration).

A mineral deposit qualifies as an orebody only if certain economic conditions, such as the anticipated price of copper and zinc, are satisfied. This leads to numerous other factors to be considered such as size, grade and shape of orebody, accessibility by roads and railways, amount of overburden or depth below the surface, applicability of specific techniques of mining and beneficiation, etc.

These factors should be used to improve upon our evaluation of regional potential. Most of them represent local conditions indicating the sites where drilling and/or mining is feasible.

For this paper, we have coded the regional geology and used some geophysical variables. Limitations are that the size of various features resulting from processes such as geosynclinal subsidence, orogeny or plate tectonics is larger than the study area and they cannot be adequately treated. The present method does not evaluate features such as paleoclimates. Also, if coding is restricted to features from one region, no use is made of relationships between mineralization and geological framework for deposits in other parts of the world. The significance of specific factors is rarely expressed in a clear-cut manner in nature and their relevance may become obvious only after a comparative metallogenic study of many similar deposits from different areas.

Regional geochemical surveys involving the systematic collection of chips of rock, samples of water, mud, etc., and trace element determination can provide useful information at the regional level. If the region was at one time glaciated, the composition of tills, eskers, etc., may be useful to trace transportation routes of specific types of boulders or heavy mineral grains back to sources where deposits may be hidden.

Other types of data which could be quantified systematically but were not used include chemical analyses of rocks, age determinations, shapes of certain intrusions, and petrographic modal analyses, e.g. for the study of regional metamorphic patterns and processes.

The data used by us are believed to be representative of a limited set of features. The main criterion for measuring them is their widespread availability. They are not screened before the statistical analysis but during this processing. We allowed the presence of both irrelevant and redundant facts in the data base which should be filtered out statistically. An advantage of this automatic coding is its objectivity. Initially, all facts in the data base are equally important in that they will all be correlated to one another and to the occurrence of deposits.

Problems of multivariate statistical analysis include the fact that homogeneous multivariate populations can only rarely be defined for data from large regions, because the coded variables tend to vary from place to place within the area. During some initial experiments an area of the Superior Province was divided into parts and eigenvectors were extracted from the multivariate correlation matrices for these parts. No patterns could be established by experiments of this type. The apparent lack of consistency in overall composition of the geological framework makes it difficult to develop a relationship between occurrence of deposits in one part of an area and to extrapolate the results to other parts. This problem was solved by proceeding as discussed in the previous section. Another problem of statistical analysis is the fact that data for the (dependent) variable to be predicted are incomplete. The amount of information on abundance of deposits in the upper parts of the earth's crust varies according to what is truly present and intensity of exploration. Since exploration depends on many other factors (see below), we do not really know whether information on the dependent variable is complete or not at any given place.

Finally, most multivariate techniques in statistical textbooks apply to situations where the variables fulfill specific auxiliary conditions such as normality. Many geological data are of the presence-absence type. If they are present, their amount can be expressed by a number (e.g. a percentage value). When they are absent, this number becomes zero. Unless the cells are made very large so that each cell contains all coded variables,

zero values will occur, giving the frequency distribution of the variables a distinctly dichotomous nature which prevents transformation into Gaussian curves.

The methods used in this paper to avoid these problems are not necessarily optimal. More research is needed, especially for the multivariate statistical analysis of presence-absence data since this type of information will increase when more specific variables are coded, cells are made smaller and more distinct types of mineral deposits are treated separately.

During the coding and presentation of results on maps, differences in cartographic projection between maps at different scales presented a problem. For example, the contours in Figure 1 were constructed on the basis of the UTM-grid, which becomes slightly deformed when a Lambert conformal projections is used for representation. In the present work, the resulting discrepancies are very small and can be neglected, but cartographic problems of this nature will become more important for larger regions.

Finally, one of the more severe limitations of the approach is the uncertainty regarding the completeness of information on the occurrence of mineral deposits. For copper, the data consisted of total production of all mines to the end of 1968 and reserves estimated at that time. Total reserves (see Table 1) for copper amount to thirty per cent of total estimated tonnage. They are largely for proven reserves. It is likely that these numbers would increase significantly when possible reserves could be considered. Estimates of possible reserves can best be attained in well-explored areas, usually in the vicinity of known deposits, where the geology is well-documented. Fair estimates could perhaps be gathered through the Delphi method.

In order to construct the contours of Figure 1, we assumed that copper potential is zero in about 23 per cent of total area (145 cells in Appendix 3) enclosing the mining towns of Timmins, Kirkland Lake, Noranda and Val d'Or where about 90 per cent of the control cells are located. The areas with copper potential by Figure 1, therefore, must fall largely outside this control area.

For example, in Figure 3, an area near Timmins of about 3,240 square miles (or 10% of total study area) was analyzed for copper potential using the contours of Figure 1. This indicated that five cells with significant copper probably remain to be found as compared with four cells known to exist. This estimate of potential can be converted to tonnage of copper and confidence limits can be computed. However, when we assume that there remains undiscovered copper potential in the control area, the estimate of five new cells must be revised. Suppose that, in terms of control cells, the copper content of the control area is twice what is known to exist. The potential for the control area then increases from zero to 24 new cells (or the total number of cells with copper increases from 24 to 48). This causes the potential for the Timmins area to increase from 5 to 15 new cells.

If the total potential for the control area is revised, estimates for all other areas must also change, but not in direct proportion to the change in the control area. As pointed out earlier, this is one of the reasons for presenting contours in Figures 1 and 4 which are not corrected for known deposits. They can then be revised more readily than contours reflecting the difference between the known and unknown.

The two problems of lack of completeness of data on large copper deposits and variations in intensity of exploration are closely interrelated. Intensity of exploration is the result of factors such as abundance and types of

outcrops, type and emphasis of exploration activity and amount of interpretation on the basis of local geology and geophysical and/or geochemical surveys. The latter factor includes number and types of drillholes per cell. Abundant information is in existence for these factors but it is widely scattered over many different sources and not of a standard form. Moreover, much information used for exploration in the past was lost or is still kept confidential. These factors were not considered in a direct manner by us. Use of the Delphi method has already been suggested as one method to obtain estimates of possible reserves in or near the control cells. Another possibility is historical studies of the type undertaken by Harris (1965) who plotted annual production against time for mining areas, estimating final totals on the assumption that trends of the past can be extrapolated into the future.

A limitation is that the models can only predict the distribution of mineralization in two dimensions. The third dimension, depth, can be considered in the well-explored parts of the area. It is then necessary to perform a detailed study of the geometrical configuration in three dimensions of all large deposits and the geology of their immediate environment and to attempt systematic statistical extrapolations in three-dimensional space for the control cells and cells adjacent to them. This would improve the results presented in this report which are based on generalized and restricted measures for total amount of copper and zinc per control cell.

As mentioned in the introduction, the contoured probability index is not a sufficient basis for evaluating the mineral potential for a commodity in a region. The factors listed in this section, and others, should be considered by experts to revise the initial crude predictions based on multivariate analysis. At this time, we are far from a situation where all relevant facts can be measured and included in a single comprehensive statistical model. Moreover, experienced exploration geologists may be able to frame conceptual models based on relatively few facts. Such predictions will vary from person to person but, again, a consensus might be reached through the Delphi method as discussed in the introduction.

USE OF RESULTS

The following method of evaluating the probability index and converting it into tonnage of copper or zinc expected in an area, was selected for its simplicity. It is based on the assumption that model of Allais (1957) is applicable within each area selected for predicting undiscovered resources.

The results of the calculations by computer are shown for copper in Appendix 3. There are hundreds of numbers for individual cells based on tens of variables. A further data reduction is required for better understanding of the trends in the region. By dividing the calculated values of Appendix 3 by 32.0, we obtain a probability for each cell. These probabilities can be grouped into classes and represented on maps by different symbols. The resulting patterns are informative (see e.g. Agterberg and Kelly, 1971). Nevertheless, the patterns remain complex and each of the values is subject to an uncertainty which cannot be estimated by simple methods, unless it is assumed that the auxiliary condition of normality is satisfied. However, because of the abundant zero values, we are dealing with a strongly bimodal distribution for metal and, therefore, cannot rely on the normality assumption which is being violated.

The probability that a single cell contains one or more large deposits also can be seen as an expected value. The expected value for any group of cells is equal to the sum of the expectations for the individual cells. This concept can be illustrated as follows. If we toss a coin, the probability of having heads is eight. This sum of the individual probabilities is an expected value. Actually, the number of times that heads turns up is a random variable described by the binomial frequency distribution.

For copper, we took sums of probabilities for groups of sixteen cells at the time. This sum is meaningful as an expected value (m). We can assume that the number of smaller cells with metal (x) satisfies a binomial frequency distribution. The general equation for this distribution then is

$$P(x) = \binom{n}{x} (m/n)^x (1-m/n)^{n-x}$$

where P(x) denotes the probability that exactly x cells will have metal given that the expected number is m. In Figures 1 and 4, n is equal to sixteen and the unit cell of reference is 40 x 40 km. Thus, each value of m defines a sequence of probabilities P(x). Some of these sequences are shown in Table 4. They were derived from the above equation of P(x).

Table 4

Probability (in per cent) that the unit cell of 40 x 40 km contains x smaller 10 x 10 km cells with metal when m is given

(x=)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	(=x)
m															
0.5	60	31	8	1											
1.0	36	38	19	6	1										
2.0	12	27	29	19	9	3	1								
4.0	1	5	13	21	23	18	11	5	2	1					
8.0				1	3	7	12	17	20	17	12	7	3	1	

We argued that m is a meaningful number as the sum of sixteen different expected values for 10 x 10 km cells. However, in order to evaluate m statistically, we assumed that the probability that a single cell has copper is m/16 for the sixteen cells per larger 40 x 40 km cell. This is equivalent to replacing probabilities for 10 x 10 km cells by sixteen-point two-dimensional moving average values.

Alternatively, it can be said that a modification of the Allais model is adopted for 40 x 40 km cells. The mineralized 10 x 10 km cells can fall anywhere within a 40 x 40 km cell. Allais assumed that the number of deposits expected per equal-area cell in a region is the same for all cells. We assumed that the number of smaller cells with metal per larger equal-area cell in a region differs from place to place in the area and is given by the probability index.

The second basic asumption of Allais (1957) was that each deposit has a specific probability that its size falls within a given range. The size-frequency distribution can be that for deposits in the same area or for deposits from another area.

The total amount of copper per control cell in Abitibi ranges from 1,084 to 1,179,900 tons of copper. On the average, it is about 140,000 tons. This represents a wide range, as also illustrated by the groupings in Table 5.

Table 5

Probability that a 10 x 10 km cell with metal contains a specific amount of metal

Class (tons of metal)	Probability (in per cent) Copper	Zinc
1,000 - 20,000	59	29
20,000 - 200,000	22	41
>200,000	19	29

In total, there were twenty-seven control cells for copper and seventeen for zinc. The zinc average amounts to about 300,000 tons of zinc.

The expected tonnage of copper per 40 x 40 km cell is obtained by multiplication of m by 140,000. The area of the unit cell is 40 x 40 km (25 x 25 miles, approx.) or 618 square miles. Hence, the probable tonnage of copper per square mile (M) amounts to (140,000 x m)/618 = 277 m.

Confidence limits for probable tonnage in an area are even wider than those of m which were exemplified in Table 4. They cannot be calculated by using a simple equation, because total tonnage of copper per area, according to the Allais model, is a compound random variable.

In our application, this compound random variable is determined by (1) a binomial distribution for number of cells with metal, and (2) the size-frequency distribution for amount of copper per control cell. The second distribution is approximately lognormal. Simple analytical expressions for compound distributions of this type have never been developed. However, the problem can be solved by Monte Carlo simulation. For this, we assumed that tonnage copper per 10 x 10 km cell has a probability of 1/27 of being equal to any one of the twenty-seven tonnage copper values of control cells.

The uncertainty in estimates of probable total tonnage copper per area will be illustrated by two examples.

If the contoured index of Figure 1 amounts to 4, the probable tonnage of copper in the surrounding 40 x 40 km cell is 4 x 140 = 560 thousand. From Table 4, we know that the estimated value 4 is reasonably precise. For example, x = 2, 3, 4, 5, or 6 with a total probability of 86 per cent.

Approximate estimation for the uncertainty in the estimate of 560,000 tons Cu, indicated that the median is 300,000 tons. This means that the probability is 50 per cent of having more than 300,000 tons, and also 50 per cent of less than 300,000 tons. The lower and upper quartiles are

50,000 and 850,000 tons, respectively. This means that with 50 per cent probability, the actual amount will fall between 50,000 and 850,000 tons. Further, there is a 2.5 per cent probability that the amount is less than 3,500 and a 2.5 per cent probability that it exceeds 2.2 million tons. The great width of these confidence limits is comparable to that arising from attempts to predict sizes of oil pools in barrels (cf, Kaufman, 1963). The uncertainty decreases somewhat when the size of area for evaluation is made larger.

A second example is given to illustrate the estimation of the total amount of copper in undiscovered large deposits for a larger region.

The Timmins area depicted in Figure 3 has significant copper potential. It consists of eighty-four cells or 8,400 km^2 (or about 3,240 square miles).

Addition of 84 probabilities for cells and rounding gives a total expected value of 9. Alternatively, the average contour value for the depicted area can be determined from the contour map (e.g. by methods commonly used to compute average elevation for an area on a topographical contour map). It gives approximately 1.76. This number then must be multiplied by the number of times (=5 1/4) the sub-area is greater than the unit cell of 40 x 40 km. Rounding off again gives nine as the answer.

From this, the number of control cells in the sub-area must be subtracted (=4, see Table 3 or Fig. 10(a)). The potential is equal to 9-4 = 5 cells with copper. Multiplication by 140,000 (average amount of copper per control cell) gives 700 thousand tons of copper potential. By Monte Carlo simulation, it was also estimated that the median is 400,000 tons. The quartiles fall at 100,000 and 1,150,000 tons, respectively.

The preceding calculations are useful. They should, however, be interpreted with moderation, because of the limitations of the method outlined in the previous section. The estimates of probable amounts of copper are likely to be conservative, mainly since we assumed that, at the end of 1968, the amount of mineable copper in the vicinity of known mines (control area of Appendix 3) was exhausted except for 30 per cent of total tonnage copper for the region in proven reserves (see Table 1). The copper potential in the entire region, as calculated from the contours in Figure 1, is about 1.3 million tons of copper, nearly all of which is concentrated in the three areas indicated by a pattern in Figure 3. If, for example, we assume that only 50 per cent of mineable copper truly present in the well-explored parts of the area was used for control, our estimate of the total copper potential in the Abitibi region increases to 6.5 million tons of copper.

The result of another use made of Figure 1 are the three anomalies shown in Figure 3. These are sub-areas where the probability index exceeds the density of control cells with known deposits in many places. The following simple operation was applied to Figure 1 in order to locate these three anomalies.

1. Contour values at centres of 10 x 10 km cells were rounded off to the nearest integer.
2. The number of control cells falling in the surrounding larger 40 x 40 km cell was subtracted from the values for (1).
3. The moving average for square blocks of sixteen values for the difference was taken, and
4. Values greater than 0.8 were retained and indicated by a pattern in Figure 3.

This operation is equivalent to filtering the difference between calculated probabilities and known data (0 or 1) for 10 x 10 km cells by means of a double triangular filter and using an arbitrary cut-off value.

The three areas selected by this method have in common that their geology (as coded for our data base) is similar to that at places where large copper deposits are known to exist. The potential of those sub-areas should be evaluated in more detail by considering other factors (see previous section). Because of the two-dimensional nature of this study, little can be said about true copper and zinc reserves in the region. Much of these may be established by further drilling in the vicinity of known deposits and this would result in an increase of estimates of potential based on Figure 1 and 4.

REFERENCES

Agterberg, F.P.
1970: Multivariate prediction equations in geology; J. Mathematical Geol., vol. 2, p. 319-324.

1971: A probability index for detecting favourable geological environments; Ninth Symposium on Decision-Making in the Mineral Industry; Can. Inst. Mining Met., Sp. Vol. 10, Montreal, 1970, p. 82-91.

Agterberg, F.P., and Cabilio, P.
1969: Two-stage least-squares model for the relationship between mappable geological variables; J. Mathematical Geol., vol. 1, p. 137-153.

Agterberg, F.P., and Fabbri, A.G.
in press: Harmonic analysis of copper and gold occurrences in the Abitibi area of the Canadian Shield; Proc. Tenth Symposium Application of Computers in Mineral Industry, Johnnesburg, South Africa, 1972 Proceedings.

Agterberg, F.P., and Kelly, A.M.
1971: Geomathematical methods for use in prospecting; Can. Mining J., vol. 92, no. 5, p. 61-72.

Agterberg, F.P., and Robinson, S.C.
in press: Mathematical problems of geology; Proc. 38th Sesion Internatl. Statist. Inst., Washington, D.C., 1971 (Preprint, 28 p.)

Allais, M.
1957: Method of appraising economic prospects of mining exploration over large territories; Management Sci., vol. 3, p. 285-347.

Barry, G.S., and Freyman, A.J.
1970: Mineral endowment of the Canadian Northwest - A subjective probability assessment; Can. Inst. Mining Met. Bull., vol. 63, p. 1031-1042.

Botbol, J.M.
1971: Characteristic analysis as applied to mineral exploration; Ninth Symposium on Decision-making in the Mineral Industry; Can. Inst. Mining Met., Sp. Vol. 10, Montreal, 1970.

Carlson, H.D., and Donovan, J.F.
1965: Geological Compilation Map, Chapleau-Foleyet; Ont. Dept. Mines, Map 2116.

Draper, N.R., and Smith, H.
1966: Applied regression analysis; John Wiley & Sons, Inc., New York.

Dugas, J., Latulippe, M., and Duquette, G.
1967: Annoted bibliography on metallic mineralization in the regions of Noranda, Matagami, Val d'Or, Chibougamau; Quebec Department of Natural Resources, Paper 2.

Efroymson, M.A.
1960: Multiple regression analysis; in Ralston, A., and Wilf, H.S. Eds., Mathematical methods for digital computers; John Wiley & sons, New York, p. 191-203.

Fahrig, W.F., Gaucher, E.H., and Larochelle, A.
1965: Paleomagnetism of diabase dykes of the Canadian Shield; Can. J. Earth Sci., vol. 2, p. 278-298.

Ferguson, S.A., Groen, H.A., and Haynes, R.
1971: Gold deposits of Ontario; Ontario Dept. Mines Northern Affairs, Mineral Resources Circ. No. 13.

Goodwin, A.M., and Ridler, R.H.
1970: The Abitibi Orogenic Belt; Geol. Surv. Can., Paper 70-40, p. 1-30.

Griffiths, J.C.
1966: Exploration for natural resources; J. Operations Res. Am., vol. 14, p. 189-209.

Harris, D.P.
1965: An application of multivariate statistical analysis to mineral exploration; unpubl. Ph. D. Thesis, The Pennsylvania State University, March 1965.

Kaufman, G.M.
1963: Statistical desision and related techniques in oil and gas exploration; Prentice-Hall, Inc., Englewood Cliffs, N.J.

Kelly, A.M., and Sheriff, W.J.
1969: A statistical examination of the metallic mineral resources of British Columbia; Proc. Symp. on Decision-Making in Min. Exploration, II, University British Columbia, Vancouver, p. 221-243.

Koch, G.S. Jr., and Link, R.F.
1971: Statistical analysis of geological data, volume II; John Wiley & Sons, Inc., New York.

Kutina, J., and Fabbri, A.G.
1971: Relationship of structural lineaments and mineral occurrences in the Abitibi area of the Canadian Shield; Geol. Surv. Can., Paper 71-9.

Lang, A.H.
1970: Base metals, Canadian Shield; in Geology and Economic Minerals of Canada; R.H.W. Douglas, Editor; Geol. Surv. Can., Econ. Geol. Rept., no. 1, p. 185-200.

McGlynn, J.C.
1970: Superior Province; in Geology and Economic Minerals of Canada; R.H. Douglas, Editor; Geol. Surv. Can., Econ. Geol. Rept., no. 1, p. 54-84.

Merriam, D.F., Editor
1970: Computer applications in the earth sciences; Plenum Press, New York; 281 p.

Miesch, A.T., and Connor, J.J.
1968: Step-wise regression and nonpolynomial models in trend analysis; Kansas Geol. Surv. Computer Contr., no. 27.

Moorhouse, W.W.
1965: Stratigraphic position of sulphides in the Archean; Trans. Can. Inst. Mining Met., vol. 58, p. 947-950.

Roscoe, S.M.
1965: Geochemical and isotopic studies, Noranda and Mattagami areas; Trans. Can. Inst. Mining Met., vol. 68, p. 279-285.

Shklanka, R., Editor
1969: Copper, nickel, lead and zinc deposits of Ontario; Ontario Dept. Mining Met., vol. 58, p. 947-950.

Sinclair, A.J., and Woodsworth, G.L.
1970: Multiple regression as a method of estimating exploration potential in an area near Terrace, B.C.; Econ. Geol., vol. 65, p. 998-1003.

Stockwell, C.H.
1964: Fourth report on structural provinces, orogenies and time-classification of rocks of the Canadian Precambrian Shield; Geol. Surv. Can., Paper 64-17.

Vyshemirsky, V.S., Dmitriev, A.N., and Trofimnk, A.A.
1971: Criteria for prediction of giant oil pools; World Petrol. Congr., Moscow, June 1971, Sp. Paper 8, 15 p.

Weiss, A., Editor:
1969: A decade of digital computing in the mineral industry; Special Volume, Am. Inst. Mining Met. Ptrol. Eng.; New York.

APPENDICES 1 TO 4

APPENDIX 1

Data for 57 copper and/or zinc deposits in Abitibi area

DEPOSIT NO.	NAME OF DEPOSIT AND/OR MINING COMPANY	Period	PRODUCTION Tons of metal '000 Tons	Grade %	Year	RESERVES Tons of metal '000 Tons	Grade %	TOTAL P+R Tons of metal '000 Tons	OTHER METALS PRESENT	MAIN ENCLOSING ROCKS	CELL NO. X, Y
1	Kam-Kotia Mine (Kam-Kotia Mines Ltd.)	Cu 1943-68	55	1.28	1968	26	1.13	81 Cu	Ag, Au	andesite, greenstone, tuff, rhyolite, diabase	18, 5
		Zn 1943-68	30	.86	1968	84	3.98	114 Zn			
2	Jameland Mines Ltd.	Cu			1968	10	1.71	10 Cu	Au	andesite, rhyolitic tuff, agglomerate, breccia, diorite	18, 5
		Zn			1968	12	7.21	12 Zn			
3	Canadian Jamieson Mines Ltd.	Cu 1965-69 March	12	2.60	1969 March	5	2.47	17 Cu	Ag, Au	rhyolite, andesite	18, 6
		Zn 1966-69 March	28	6.27	1969 March	8	4.03	35 Zn	Ag, Au		
4	Mordey Prospect (Genex Prospect, United Oblaski Mining Co. Ltd.)	Cu			1966	3	2.47	3 Cu	Au, Ag	rhyolite, prophyry, andesite	18, 6
		--						--			
5	Kidd Creek Mine (Ecstall Mining Ltd., Texas Gulf Sulphur Co. Inc.)	Cu 1965-68	108	1.56	1968	705	1.33	813 Cu	Pb, Ag, Au	felsitic tuff, breccia, mafic intrusive	20, 4
		Zn 1965-68	474	6.82	1968	3,752	7.08	4,227 Zn			
6	McIntyre Mine (McIntyre Porcupine Mines Ltd.)	Cu 1963-68	24	.76	1968	32	.79	56 Cu	Ag, Au, Mo, W	porphyry, andesite, basalt, tuff	20, 7
		--						--			
7	Alexo Mine (Noranda Mines Ltd.)	Cu 1913-19	.3	.50	1919	.1	.50?	.4 Cu	Ni	peridotite, basalt, andesite	
		--						--			
8	Ryan Lake Mine (Min-Ore Mines Ltd., Pax International Mines Ltd.)	Cu 1948-65	2	1.35	1965	5	1.00	7 Cu	Ag, Au	peridotite, porphyry, syenite	25, 12
		--						--			
9	Matarrow Mine (Floyd Property, Jarrow Mines Ltd.)	--						--	Ag, Pb	iron formation	
		Zn 1952-53	.5	1.15				.5 Zn			
10	Gates Mine (Miller Lake-O'Brien Mine, Siscoe Mines Ltd.)	Cu 1954-65	.04	.01				.04 Cu	Ag, Co, Bi, Pb, Zn	conglomerate, diabase	
		--						--			
11	Centre Hill Mine (Munro Copper Mines Ltd.)	Cu 1967-68	2	1.80	1968	12	2.00	14 Cu	Au	rhyolitic agglomerate, tuff	28, 5
		Zn			1968	10	1.65	10 Zn			
12	Ethel Copper Prospect (Ethel Copper Mines Ltd.)	Cu 1962-66	.2	1.19?				.22 Cu	Ag, Au	diabase, arkose	28, 15
		--						--			
13	Dane Copper Mine (Dane Mining Co. Ltd.)	Cu 1911-45	.1	?				.1 Cu	Au, Zn	basalt	
		--						--			

APPENDIX 1

Data for 57 copper and/or zinc deposits in Abitibi area

DEPOSIT NO.	NAME OF DEPOSIT AND/OR MINING COMPANY	Period	PRODUCTION Tons of metal '000 Tons	Grade %	Year	RESERVES Tons of metal '000 Tons	Grade %	TOTAL P+R Tons of metal '000 Tons	OTHER METALS PRESENT	MAIN ENCLOSING ROCKS	CELL NO. X, Y
14	Argonaut Mine (Upper Beaver Mines Ltd., Lake Beaverhouse Mines Ltd.)	Cu 1925-68	3	?	1968	1	1.19	4 Cu	Au, Ag, W, Mo	basalt, andesite, syenite	32, 10
		--						--			
15	Amity Mine (Amity Copper and Gold Mines Ltd.)	Cu 1928-32	.03	4.75	1932	1	6.00	1 Cu	Zn, Pb	intermediate tuff, iron formation	30, 12
		--						--			
16	Patterson Mine (Fidelity Mining and Investment Corp.)	Cu 1928-56	.2?	5.27?				.2 Cu	Ag	granitic felsite, acid tuff, iron formation	
		--						--			
17	Trethewey-Ossian Mine (Fidelity Mining and Investment Corp.)	Cu 1948-62	.1?	?				.1 Cu	Zn, Ag, Au	intermediate tuff, granitic felsite, acid tuff, iron formation	
		--						--			
18	Abana Mine (Normetal Mining Corp. Ltd.)	Cu 1937-68	213	2.33	1968	26	2.15	239 Cu	Ag, Au, Pb	rhyolite tuff, quartz porphyry	35, 1
		Zn 1937-68	408	4.47	1968	75	6.25	483 Zn			
19	Duvan Copper Co. Ltd.	Cu 1960	.2	10.04	1960	3	2.50	3 Cu	Ag	grewacke, tuff, basic tuff	34, 2
		--						--			
20	Hunter Mine (Beattie-Duquesne Mines Ltd.)	Cu 1956-57	1	1.00				1 Cu	Au, As, Mo, Ag	andesitic tuff, porphyry, sericite schist, andesite, sediments, schists	36, 6
		--						--			
21	Inmont Copper Mine (Robb-Montbray Mines Ltd.)	Cu 1934-35	.08	6.50				.08 Cu	Au, Te, Zn	rhyolite	
		--						--			
22	Halliwell Gold Mines Ltd.	Cu 1938	.04	1.50				.04 Cu	Au	andesite, aplite	
		--						--			
23	Aldermac Mine (Aldermac Copper Corp. Ltd., West Wasa Mines Ltd.)	Cu 1930-43	31	1.49	1943	12	2.00	43 Cu	Au, Ag	felsic tuff, agglomerates, rhyolite	36, 9
		--						--			
24	O'Leary-Malartic Mines Ltd.	Cu			1959	.4	1.25	.4 Cu		biotite gneiss, diabase	
		--						--			
25	Valray Explorations Ltd.	Cu			1953	.9	3.85	.9 Cu		grewacke	
		--						--			
26	New Norzone Mines Ltd	--						--	Pb, Ag, Au, Cu	grewacke	36, 10
		Zn			1948	7	8.50	7 Zn			
27	Vauze Mines Ltd. (B1 and B2)	Cu 1961-65	11	2.81	1964	.1	1.40	11 Cu	Au, Ag	andesite, rhyolite	37, 8
		Zn 1961-65	4	1.00	1964	.04?	.50?	4 Zn			

APPENDIX 1

Data for 57 copper and/or zinc deposits in Abitibi area

DEPOSIT NO.	NAME OF DEPOSIT AND/OR MINING COMPANY	Period	PRODUCTION Tons of metal '000 Tons	Grade %	Year	RESERVES Tons of metal '000 Tons	Grade %	TOTAL P+R Tons of metal '000 Tons	OTHER METALS PRESENT	MAIN ENCLOSING ROCKS	CELL NO. X, Y
28	Lake Dufault Mines Ltd.	Cu 1964-68	80	3.99	1968	15	1.60	95 Cu	Au, Ag	siliceous breccia, quartz-feldspar porphyry, rhyolite	37, 8
		Zn 1964-68	111	5.49	1968	13	2.10	124 Zn			
29	Old Waite Mine (Waite Amulet Mines Ltd., Noranda Mines Ltd.)	Cu 1928-49	59	4.70				59 Cu	Au, Ag	andesite, rhyolite	37, 8
		Zn 1928-49	37	2.97				37 Zn			
30	East Waite Mine (Waite Amulet Mines Ltd., Noranda Mines Ltd.)	Cu 1951-61	62	4.13	1961	1	2.57	63 Cu	Au, Ag	andesite, rhyolite breccia	37, 8
		Zn 1951-61	49	3.26				49 Zn			
31	New West Amulet Mine (Waite-Dufault Mines Ltd., Bedford Hill Mines Ltd., Waite Amulet Mines Ltd.)	Cu			1966	4	1.50	4 Cu	Zn	rhyolite, rhyolite breccia	36, 8
		--						--			
32	Amulet 'F' Mine (Waite Amulet Mines Ltd., Noranda Mines Ltd.)	Cu 1930-51	10	3.40				10 Cu	Au, Ag, Pb,	rhyolite breccia	37, 8
		Zn 1930-51	10	3.40				10 Zn			
33	Amulet 'C' Mine (Waite Amulet Mines Ltd., Noranda Mines Ltd.)	Cu 1930	13	5.12				13 Cu	Au, Ag	rhyolite breccia, andesite, feldspar porphyry	37, 8
		Zn 1930	51	8.50				51 Zn			
34	Amulet 'A' and 'lower A' (A: Waite Amulet Mines Ltd., lower A, Lake Dufault Mines Ltd.)	Cu 1937-62	271	5.12	1961	16	5.60	287 Cu	Au, Ag	andesite	37, 8
		Zn 1937-62	290	5.47	1961	6	8.70	296 Zn			
35	West McDonald Mines Ltd.	Cu 1955-59	.09	.82				.09 Cu	Au, Ag, Pb	rhyolite breccia, sericite schist	38, 8
		Zn 1955-59	30	2.85				28 Zn			
36	Joliet Québec Mines Ltd.	Cu 1964-68	25	1.20	1968	6	1.21	31 Cu	Ag, Au	rhyolite breccia, andesite tuff	37, 9
		--						--			
37	Lyndhurst Mining Co. Ltd.	Cu 1956-57	3	1.77	1967	1	2.10	4 Cu	Au, Ag, Pb, Zn	acid pyroclasics	38, 5
		--						--			
38	Mobrun Copper Ltd.	Cu			1956	19	.62	19 Cu	Au, Ag	rhyolite andesite	38, 7
		Zn			1956	71	2.32	71 Zn			
39	South Dufault Mines Ltd.	Cu			1960	.8	1.60	.8 Cu	Zn, Ag, Au	rhyolite, rhyolite, breccia, andesite	
		--						--			
40	Delbridge No. 2 Mine (Delbridge Mines Ltd., D'Eldona Gold Mines Ltd.)	Cu			1968	3	1.03	3 Cu	Au, Ag	rhyolite breccia, tuff	38, 9
		Zn			1968	39	11.97	39 Zn			
41	Quemont Mining Corp. Ltd.	Cu 1948-68	173	1.21	1968	5	.89	179 Cu	Ag, Au	rhyolite breccia, rhyolite dykes, quartz-feldspar-porphyry dykes, diabase, chlorite-sericite schist	38, 9
		Zn 1947-68	229	1.59	1968	13	2.22	242 Zn			

APPENDIX 1

Data for 57 copper and/or zinc deposits in Abitibi area

DEPOSIT NO.	NAME OF DEPOSIT AND/OR MINING COMPANY	Period	PRODUCTION Tons of metal '000 Tons	Grade %	Year	RESERVES Tons of metal '000 Tons	Grade %	TOTAL P+R Tons of metal '000 Tons	OTHER METALS PRESENT	MAIN ENCLOSING ROCKS	CELL NO. X, Y
42	Horne Mine (Noranda Mines Ltd.)	Cu 1927-68	1,081	1.90	1968	68	2.54	1,149 Cu	Ag, Au, Zn	diabase, syenite prophyry, quartz porphyry, rhyolite, rhyolitic explosives, chlorite schists	37, 9
		--						--			
43	New Mic-Mac Mines Ltd.	Cu 1942-47	2	.27				2 Cu	Au, Ag	andesite	41, 9
		--						--			
44	Dumagami Mines Ltd.	Cu			1965	3	.29	3 Cu	Au, Zn, Ag	tuffs, felsic agglomerate	41, 9
		--						--			
45	New Formaque Mines Ltd.	Cu			1959	8	1.80	8 Cu	Ag, Se, Bi	acid volcanics	44, 5
		Zn			1959	13	1.86	13 Zn			
46	Trinity Chibougamau Mines Ltd. (North Trinity Mining Corp. Ltd.)	Cu			1954	2	1.15	2 Cu	Au, Ag	acid agglomerate, chloritic volcanics	46, 4
		Zn			1954	1	.74	1 Zn			
47	Manitou-Barvue Mines Ltd. (Barraute Twp.)	--						--	Pb, Ag, Cu	rhyolite breccia	47, 5
		Zn 1952-57	125	2.24	1957	153	3.63	279 Zn			
48	Belfort Mines Ltd. (Roymont Mines Ltd.)	Cu			1964	.5	.21	.5 Cu	Ag, Au, Pb	acid tuffaceous rocks, sediments	47, 6
		Zn			1964	18	7.00	18 Zn			
49	Barvallée Mines Ltd.	Cu			1958	3	1.23	3 Cu	Ag	acid-intermediate flows and tuffs, cherty layers	47, 7
		Zn			1958	12	5.71	12 Zn			
50	Consolidated Mogador Mines Ltd. (Vendome Mines Ltd., Canadian Shield Mining Corp.)	Cu			1957	5	.47	5 Cu	Pb, Au, Ag, Ni	tuff, basic intrusives, sheared basic volcanics	47, 7
		Zn			1957	82	7.30	82 Zn			
51	Greene Stabell Mines Ltd.	Cu 1933-37	.2	.28				.2 Cu	Au, Ag, Te, Zn	andesite, granodiorite	
		--						--			
52	East Sullivan Mines Ltd.	Cu 1949-66	152	.92				152 Cu	Au, Ag, Pb	diorite, acidic to intermediate volcanics plus fragmentals, syenite porphyry	47, 10
		Zn 1949-66	78	.47				78 Zn			
53	Manitou-Barvue Mines Ltd. (Golden Manitou Mines Ltd.)	Cu 1942-68	34	.33	1968	.8	.72	34 Cu	Ag, Au, Pb	sericite schists	48, 10
		Zn 1942-68	261	2.54	1968	10	3.37	271 Zn			
54	Québec Manitou Mines Ltd.	Cu			1967	10	1.26	10 Cu	Au, Zn, Pb, Ag	granodiorite, sericite schist	48, 10
		--						--			

APPENDIX 1

Data for 57 copper and/or zinc deposits in Abitibi area

DEPOSIT NO.	NAME OF DEPOSIT AND/OR MINING COMPANY	Period	PRODUCTION Tons of metal '000 Tons	Grade %	Year	RESERVES Tons of metal '000 Tons	Grade %	TOTAL P+R Tons of metal '000 Tons	OTHER METALS PRESENT	MAIN ENCLOSING ROCKS	CELL NO. X, Y
55	Dunraine Mines Ltd. (Rainville Copper Mines Ltd.)	Cu 1956-68 --	4	1.39	1968	8	1.22	12 Cu --	Au, Ag	tuffaceous rocks, granodiorite	48, 11
56	Zulapa Mining Corp. Ltd.	Cu --			1964	8	.48	8 Cu --	Ni	diorite	50, 8
57	Abitibi Copper Mines Ltd. (Abitibi Metals Mines Ltd.)	Cu --			1966	8	.88	8 Cu --	Au, Zn	rhyolite, partly fragmental, granodiorite	49, 10

APPENDIX 2:

Variables used to calculate values of Appendix 3 for copper potential

Rank	Name of Variable	Q-value	Regression coefficient
0.	Constant term		-0.14863
1.	Acid volcanics x Mafic volcanics	0.36×10^{-9}	0.00045
2.	Granitic rocks x Acid volcanics	0.44×10^{-4}	0.00147
3.	Acid volcanics x Sediments	0.70×10^{-3}	0.00182
4.	Acid volcanics x Iron formations	0.45×10^{-4}	-0.00334
5.	Acid volcanics x Aeromag.	0.63×10^{-3}	-0.01958
6.	Mafic intrusions x Acid volcanics	0.14×10^{-3}	0.00187
7.	Metamorphic rocks x Iron formations	0.87×10^{-3}	0.00618
8.	Mafic intrusions x Metamorphic rocks	0.91×10^{-2}	0.00129
9.	Granitic rocks x Sediments	0.32×10^{-1}	0.00037
10.	Sediments x Bouguer anomaly	0.13	0.00027
11.	Mafic volcanics x Sediments	0.021	0.00017
12.	Acid volcanics	0.13	0.05068
13.	Sediments x Aeromagnetic anomaly	0.12	0.00560
14.	Sediments	0.14	-0.01726
15.	Ultramafics x Sediments	0.27	0.00049
16.	Granitic rocks x Iron formations	0.30	-------
17.	Iron formations x Bouguer anomaly	0.10	-0.00490
18.	Iron formations	0.19	-0.40777
19.	Mafic volcanics x Iron formations	0.09	0.00171
--	Backward pass, No. 16 deleted	0.65	-------
19.	Iron formations x Aeromag.	0.20	0.02133
20.	Ultramafics x Iron formations	0.29	-0.00332
21.	Sediments x Metamorphic rocks	0.36	-0.00025
22.	Acid volcanics x Metamorphic rocks	0.29	0.00123
23.	Mafic intrusions x Mafic volcanics	0.41	0.00026
24.	Granitic rocks	0.38	0.00128
25.	Granitic rocks x Mafic volcanics	0.44	0.00003
26.	Mafic intrusions x Sediments	0.50	-0.00030

APPENDIX 3

Calculated values of copper potential for 644 cells (numbered as in Figure 10a). All values were multiplied by ten and rounded. Outline of control area (145 cells). Control cells (in total 27) marked by solid circles.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	0	0								-2	-4	0		0	4	2	0	3	-1
2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	0	0	0							0	0	0			0	0	2	-1	6
3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
			0	0	0									0	0		6	-2	15	6
4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
				0	0	0						0		10	0		2	10	12	13
5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	+	●
					0	0	0					20	6	6	0	2	3	18	15	7
6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	+	+
		0						0	0			14	1	1	0	0	1	28	13	2
7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●
	1	0	0	0								14	0	0		0	3	9	6	11
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	3	0	0								0	0	0	0	0		-2	3	0	2
9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	0	0								0	1	0	0	0	0	0	0	2	2
10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0										0	3	7	2	4	-5	3	0	0	1
11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0								0	1	0	0	0	0	0			0		1
12	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
		0				0	0	0	0	0	0	0	0	0	0	0		0	0	0
13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
		0	0	0	0	0	0	0	1	0	0	4	0	0	0	0		0	0	10
14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
			0	0	0	0	0	3	8	-9	3	1	5	0	0	0	0	0	-1	0
15	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
					0	0	0	0	5	7	-2	1	1	0	0	0	0	0	0	-1
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
						0	0	0	0	0	0	0	5	1	0	1	1	5	4	0
17	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	1	0	0	0

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	+	+	+	+	+
	2	0	0	9	10	2	0	0	0	-9	-3	2	4	9	8	7	1	5	3	4
2	+	+	+	+	+	+	+	+	+	+	+	+	+	●	+	+	+	+	+	+
	-1	2	4											3	-1	0	0	-1	5	0
3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	11	4	2										0	0	1	0	0	2	1	0
4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	2	5	3				0							0	0	4	0	1	1	2
5	+	+	+	+	+	+	+	●	+	+	+	+	+	+	+	+	+	●	+	+
	1	-2	-3	5	0	-3	2	4	5		0	0	0	0	6	12	8	13	0	3
6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	+	+	+	+
	5	4	1	0	0	0	1	5	0	3	3	2	0	2	4	12	6	0	0	0
7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	+	+
	1	1	1		0	0	0	4	0	0	3	2	15	2	13	25	11	15	4	2
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	●	+	+	+
	2	-1			0		0					5	20	5	5	11	20	15	7	0
9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	●	●	●	+	+
	3	8		0	0		1	0	0	0	0	0	4	4	12	18	17	18	11	5
10	+	+	+	+	+	+	+	+	+	+	+	●	+	+	+	+	+	+	+	+
	1	8	0	2	3	0	0	0	1	5	12	10	7	0	5	0	-6	-4	0	0
11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	2	0	0	0	0	0	1	0	1	4	-2	6		0	0	0	0	0		
12	+	+	+	+	●	+	+	+	+	●	+	+	+	+	+	+	+	+	+	+
	0	1	8	1	5	0	1	0	0	13	0	12								
13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	-4	2			-2	0			0	0	1	4								
14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1			0	0		0	0	0	0	0		0							
15	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	-1	0	0	0				0	7				0							
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	-1																		
17	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	-3								-11										

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
					2	1	0	-1	-3	-4	2	0	0	0	1	0	0	0	0
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1	0	0	2	0	1	3	2	4	2	0	0	0	0	0	0	0	0	0
2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	0	1	2	7	3	3	3	0	1	3	2	0	0	0	0	0	0	0
3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1	0	0	1	1	1	1	3	3	-1	1	0	0	0	0	0	0	0	0
4	+	+	+	+	+	●	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	1	11	7	2	1	9	0	0	2	1	2	0	0	0	0	0	0	0
5	+	+	+	●	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	1	7	10	6	9	11	2	3	3	0	0	0	0	0	0	0	0	
6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	0	0	6	11	6	2	5	7	3	10	0	0	0	0	0	0	0		
7	+	+	+	+	+	+	●	+	+	+	+	+	+	+	+	+	+	+	+
	2	1	2	1	1	1	12	6	4	1	1	0	0	0	0	0	0		
8	+	+	+	+	+	+	+	+	+	●	+	+	+	+	+	+	+	+	+
	0	0	0	0	2	0	5	6	2	2	0	0	0	0	0	0			
9	●	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	22	0	1	1	0	1	8	4	2	1	7	0	0	0	0				
10	+	+	+	+	+	+	●	●	●	+	+	+	+	+	+	+	+	+	+
	9	0	2	0	0	2	5	8	6	2	-1	3	0	0					
11	+	+	+	+	+	+	+	●	+	+	+	+	+	+	+	+	+	+	+
			0	0	-2	-2	0	17	1	8	9	8							

APPENDIX 4

GLOSSARY OF TERMS

Cell (10 x 10-km cell or square): Unit area of original grid overlain on area of study for coding of geological features; Universal Transverse Mercator (UTM)-grid with 10-km spacing for Abitibi area (see Fig. 10).

Study area: Area consisting of all 10 x 10-km cells with data used for statistical analysis; includes control cells and control area. In Abitibi, study area consists of 644 cells each of which has more than 25 areal per cent Archean rocks on the geological map (see Fig. 10).

Control cells: Those 10 x 10-km cells with at least one large deposit (e.g. there are 27 control cells for copper in Abitibi, see Fig. 10).

Large copper (or zinc) deposit: Mine or developed prospect with total production and 1968 reserves of at least 1,000 tons of metal.

Calculated value (Predicted value): Value for a 10 x 10-km cell obtained by multivariate statistical analysis used to match copper (or zinc) values to all values for all variables (644 x 55 = 35,420 values for Abitibi) inserting zero copper (or zinc) values for cells without deposits (see Appendix 3).

Control area: Area used for scaling calculated values so that the predicted number of cells with at least one large copper (or zinc) deposit is equal to number of control cells in this area; measures 145 cells for Figures 1 and 4 (see Appendix 3).

n: number of 10 x 10-km cells per unit cell (n= 16 for copper and zinc).

Unit cell: Area of reference for probability index; consists of a block of n cells; measures 40 km on a side for copper and zinc in Abitibi, is oriented according to north-south direction, and can be superimposed anywhere on the contour map.

m: probability index, contoured pattern for sums of 16 calculated values for cells. Since these sums form a smooth pattern, they were contoured by machine after scaling; expresses number of smaller 10 x 10-km cells with one or more large deposits expected per unit cell.

x: exact number of smaller 10 x 10-km cells with one or more large deposits per unit cell.

P(x): Binomial probability that a unit cell contains exactly x smaller 10 x 10-km squares with one or more large deposits.

M: ($\approx$ 227m for copper): probability index converted to expected tonnage of copper per square mile. For zinc, M $\approx$ 630 m.