

○ VALIDATION AND SENSITIVITY ANALYSIS OF A MINERAL POTENTIAL MODEL USING FAVOURABILITY FUNCTIONS

Tsebaie Woldai, Department of Earth System Analysis, International Institute for Geoinformation Sciences and Earth Observation (ITC), Enschede, The Netherlands

Alberto Pistocchi, GECOSistema srl, Cesena, Italy

Sharad Master, Economic Geology Research Institute, School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa

An area in the Magondi Belt, Zimbabwe, has been chosen for mineral potential mapping using the favourability functions approach. The available datasets comprising of an old geological map, a detailed airborne total magnetic field survey, and geochemical samples at the nodes of an exploration grid, have been integrated using seven different inference techniques through the joint probability function under the conditional independence hypothesis. A geological conceptual model has been adopted for the representation of the mineralization occurrence, in order to select appropriate geospatial indicators of mineralization, while existing mines in the area have been used as a training set for the model. Among the different integration techniques which have been tested, some have proven to be robust in correctly predicting all the known mine sites. It has been thus possible to draw favourability maps using existing data, which indicate the most promising areas for exploration and detailed mapping efforts for mineral exploitation. Even more important, using sensitivity analysis of the favourability functions allowed to evaluate the most important factors controlling mineralization occurrence, and thus worth additional future investigation.

So far, a completely data driven approach has been supported because limitations due to lack of geological knowledge limited the possibility of using expert knowledge-based modifications of the conditional probabilities.

Improvements in prediction can be achieved through a more detailed geological description of the area based e.g. on the interpretation of remotely sensed data.

INTRODUCTION

Many authors (Duda et al, 1976; Chung and Fabbri, 1993; Bonham-Carter, 1994; Harris and Pan, 1999; Cheng and Agterberg, 1999) have shown the use of numerical techniques to link the occurrence of a phenomenon of interest to the local value of some attributes considered “symptoms” or causal factors for the phenomenon itself. The attributes are considered as evidence factors of the event, in the sense that the presence of each relevant attribute corresponds to a degree of “probability”, “possibility” or “likelihood” to find the event (Chung and Fabbri, 1993). The simultaneous verification of different attributes strengthens the favourability for occurrence of the phenomenon.

The use of these techniques has been increasingly receiving attention over the years in such disparate fields as medicine (Shortliffe and Buchanan, 1975), mineral exploration (e.g. Hecker-mann, 1986; Bonham-Carter, 1994) and geo-environmental hazards (e.g. Chung et al., 1995).

These techniques allow systematic exploration of relations between variables, which is particularly helpful when no clear expert judgement can be formulated, and a data driven approach, consisting in the derivation of statistical relations among variables, must be preferred.

Chung and Fabbri (1993) posed the problem under the framework concept of the “favourability functions” and showed how the latter can be used as a guide to model building ranging from purely data driven, statistical models to ones where favourability is defined based on expert knowledge.

Our paper shows a geographic information system (GIS)-based application in mineral potential mapping through this favourability functions approach. A particular emphasis is put on the issue of validating the prediction using a priori evidence of mineralisation corresponding to existing mines. An area in the Magondi Belt has been chosen to test the capability of an existing database to represent the actual knowledge of the area for mineralised deposit prediction. The idea is to incorporate as much of prior geological knowledge as possible, and to express it as weights in map overlay depicting the relative favourability of the terrain model for hosting the occurrence of mineral deposits.

GENERAL OUTLINE OF THE AREA AND GEOLOGICAL CONCEPTUAL MODEL OF THE MINERALISATION PROCESS

The study area is bounded by the longitudes 29° 50'E to 30° 15'E and 16° 15'S to 17° 30'S and is situated in the northern part of Zimbabwe (Figure 1). It stretches for 50 km from east to west and for 170 km from north to south covering an area of 8500 km². It is easily accessed through the town of Chinhoyi, which is approximately 130 km west of Harare. The two operating base metal mines, Alaska and Mhangura, are accessible from Chinhoyi. Alaska mine is approximately 12 km west of Chinhoyi and Mhangura mine is 70 km north of Chinhoyi.

Rocks of the early Proterozoic Magondi Supergroup outcrop on the northwestern flank of the exposed portion of the Archaean Zimbabwe craton. These rocks were deposited in the Magondi Basin (Figure 1), which extends under younger cover towards Botswana. The Magondi Supergroup is divided, from bottom to top, into the Deweras, Lomagundi, and Piriwiri Groups which were, according to Master (1991), deposited between c.2.16 and 2.0 Ga (Pb-Pb, Rb-Sr). The lowermost Dewera Group, which rests unconformably on Archean and earliest Proterozoic Granite-greenstone belt rocks and gneisses, is a mainly sedimentary sequence of up to 1.3 km thick, with subordinate mafic volcanics and pyroclastic rocks. The metasedimentary rocks consist of clastic terrigenous continental redbeds comprised of meta-arenites, rudites, and pelites, and minor carbonates and evaporites (chemical sediments) which were deposited in rift-related continental alluvial fan, braided stream, playa flat and playa lake environments. The Dewera Group contains several stratabound sediment-hosted copper-silver deposits. The largest of these deposits, at the Mangula and Norah mines, are situated at Mhangura (formerly Mangula, see Figure 1).

The Lomagundi Group unconformably overlies the Dewera Group. In places, the unconformity has been destroyed by intense tectonism during the Magondi orogeny. The Lomagundi Group, which is extensively distributed from north to south in the present study area, comprises of dolomite, phyllites, pockmarked quartzite, argillites, sandstones, intermediate volcanics and stripped slate in close association with arkose rocks of the Deweras Group. The lithological sequence of the Lomagundi Group is characteristic of marginal marine and shallow shelf settings.

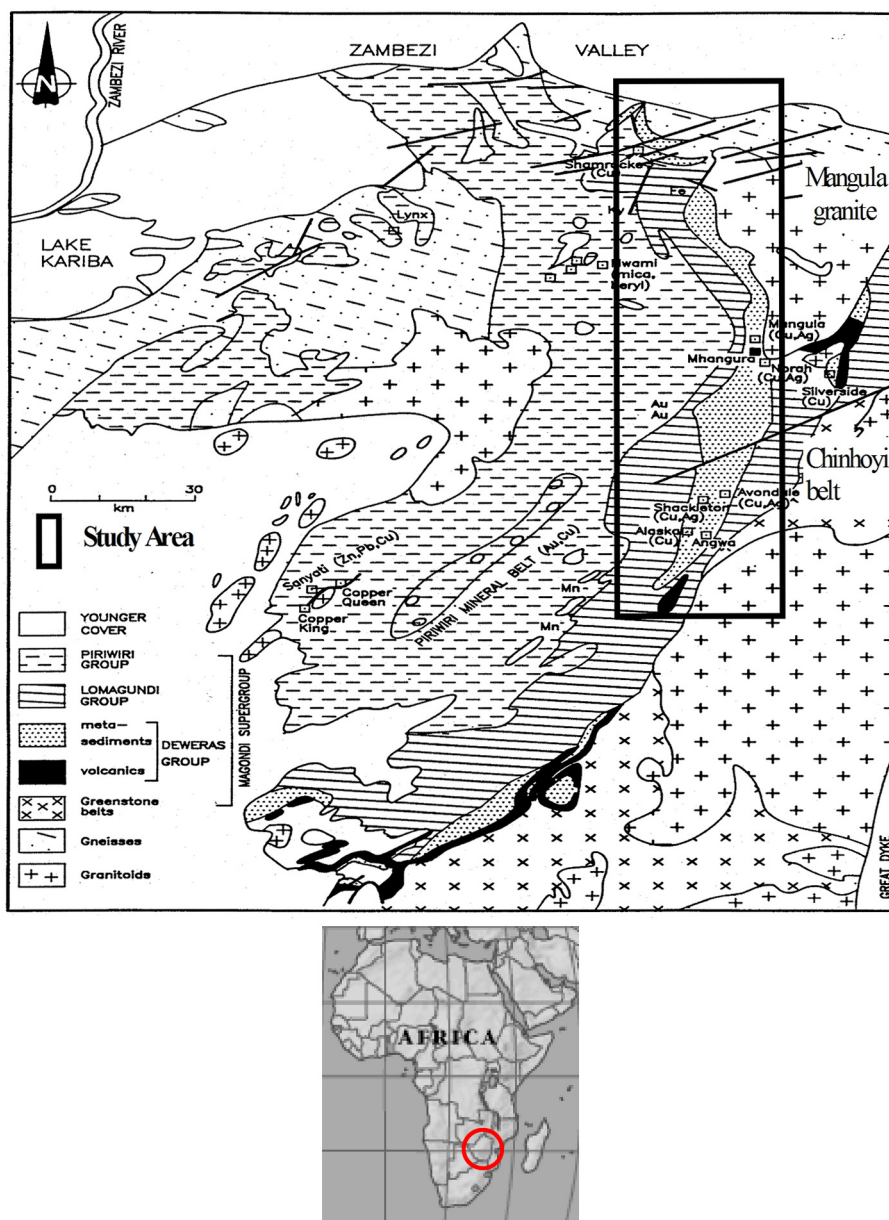


Figure 1 Regional geological map showing the basement rocks of the Magondi basin (modified after Master, 1991). Study area boundary: 29° 50'E to 30° 15'E and 16° 15'S to 17° 30'S.

The overlying Piriwiri Group, which can be considered the contemporaneous distal facies equivalent of the Lomagundi Group, comprises of alluvial fan and unworked sediments such as conglomerate, arkose, basal graphitic and pyritiferous slates with narrow bands of cherty man-ganiferous quartzite, greywackes and pelitic schist partly associated with chemical sedimentary

rocks. It shows the structure of repeated sedimentation, and the rocks are distributed in the central part of the study area successively from north to south. The Piriwiri Group was deposited in deep euxinic waters, possibly in continental slope, submarine fan and abyssal plain environments. Volcanogenic hydrothermal emanation on the sea floor may have given rise to the abundant cherts, manganese-rich beds, and minor iron-formations (Master, 1991).

The Magondi Supergroup is underlain by a basement complex consisting of Archaean granite-greenstone terrain (including a calc-alkaline magmatic arc succession) and gneisses of the Zimbabwean craton, and the earliest Proterozoic Great Dyke and related complexes that are intrusive into the Archaean rocks. The tectonic setting of the Magondi Supergroup was in rift-related continental back-arc basins which formed behind a magmatic arc produced by eastward subduction of oceanic lithosphere under the Zimbabwe craton. The Magondi Supergroup was deformed into a thin and thick-skinned fold-thrust belt and metamorphosed during the c.2.0-1.8 Ga (Rb-Sr, K-Ar) Magondi Orogeny (which resulted from collision of the magmatic arc and closure of the back-arc basin) and also by the Irumide and Pan-African zambesi orogenies (Master, 1991).

The Magondi Belt is a known source of base metal mineral potential with two major copper producing mines: Alaska and Mhangura. Some of the defunct mines in the area include: the Old Alaska, Hans and Shackleton (including the Avondale ore deposits) mines in the south, Mangula and Nora mines in the centre and Shamrocke Mine to the north of the study area. The economically valuable mineralisation in these mines is the strata bound sediment-hosted Cu-Ag- (Au-Pt-Pd-U) mineralization in the Deweras Group (Mangula, Norah, Shackleton/Avondale and Angwe mines), and the Cu-Ag mineralization of the Shamrocke Mine in the Lomagundi Group which was formed diagenetically by saline basin brines (Master, 1991). The massive sulphide deposits in the Piriwiri Group were probably “sedimentary exhalative” deposits which formed diagenetically during basin evolution. Minor Cu-Au occurrences of the “Piriwiri mineral belt” are, according to Master (idem), related to the intermediate volcanics and pyroclastic rocks in the Piriwiri Group.

From previous geological analyses (Kambewa, 1998; Lopez, 1998) a geological conceptual model of mineralisation has been assumed such that:

- I. The typical host rocks for mineralization are the Deweras group Arkoses; these are porous/permeable sedimentary rocks that allow the mineralized solutions to penetrate through. With respect to this, the overlaying dolomite layer may also be looked at as a favourable host rock type.
- II. The unconformity between the sedimentary cover layer of the area and the underlying igneous Archean basement is the main metal source and the area where most of the mineralization is likely to occur.
- III. The anticlinal axes and faults trending north to south, and faults trending NW-SE are to be regarded as preferential pathways for mineralization.
- IV. Cu and Ag anomalies are to be regarded as strong indicators of mineralization, although some copper-enriched leakage may have occurred close to streams eroding mine waste disposals. In Kambewa (1998), a Principal Components Analysis has been shown for metal concentrations all over the study area, and it had been pointed out how principal component 4 can be relating with the occurrence of a copper mineralization.
- V. The basic igneous rocks interlayered or intrusive into the Deweras Arkoses mark on the other side unfavourable lithologies for mineralization.

THE DATA INTEGRATION METHODOLOGY

We summarize in this section the concepts illustrated by Chung and Fabbri (1993) and Chung and Fabbri (1999), related to the favourability functions approach.

Suppose that A is the domain on which the analysis is performed, and F is the phenomenon of which we check the occurrence. If r data layers are available, and for the generic k -th layer ($k=1, \dots, r$), m_k classes of attributes, $k=1, \dots, m_k$, are individuated, we can define for each data layer a partition function:

$$V_k: A \Rightarrow \{1, 2, \dots, m_k\}$$

that assigns each pixel in A one of the classes in layer k . Furthermore, we can define for each layer another function ("value function"):

$$R_k: \{1, 2, \dots, m_k\} \Rightarrow [a, b]$$

which maps the occurrence of each layer in a value falling inside the interval $[a, b]$, where a and b depend on further assumptions made by the analyst (as it will be shown later). This value represents the degree of favourability, which is a measure of how reliable is the assumption that the phenomenon occurs once a particular class of the attribute is met.

The favourability function can then be defined as the functional composition of V and R for each data layer, i.e.:

$$F_k = R_k \circ V_k$$

The function R_k can be defined based on statistical relations emerging from data, or expert judgment.

In the present application, we used seven types of favourability functions based on data, namely Bayes' theorem, certainty factors, and five fuzzy logic algorithms, for which a synthetic description and references are provided in the following. It must be noticed that, in addition to the described ones, many other different techniques can be used, as the weights of evidence (e.g. Cheng and Agterberg, 1999; Bonham-Carter, 1994; Chung and Fabbri, 1993), the belief functions (Shafer, 1976), and linear regression over probabilities (Chung et al., 1995; Chung and Fabbri, 1999). Chung and Fabbri (1993) also suggest a technique to cope with lacking information on evidence.

The interval extremes a , b must be assumed by the analyst according to his interpretation of the 'reliability': if it is considered that reliability is the same as 'probability', then $a=0$, $b=1$. If the measure of reliability is set equal to the certainty factor (Shortliffe and Buchanan, 1975; Heckermann, 1986), then $a=-1$, $b=1$. If a different technique is chosen, other values might be necessary. The results of these calculations are numbers expressing indexes of favourability with respect to the considered phenomenon; they are not to be seen as direct measures of probability in the classical sense.

BAYES' THEOREM

If the favourability is assumed to coincide with the probability that a certain phenomenon F occurs given the occurrence of a set of attribute classes E_1, \dots, E_n , then according to Bayes' theorem, under the hypothesis that E_1, \dots, E_n are conditionally independent, we can write (Chung and Fabbri, 1999):

$$\text{Probability}\{F \text{ occurs given } E_1, \dots, E_n\} = \\ = P\{F / E_1 \dots E_n\} = (pps_1 * \dots * pps_n) * (ppa_1 * \dots * ppa_n) / (psF)^{n-1} * pps_{1 \text{ to } n}$$

where:

- $pps_i, i=1, \dots, n$ is the prior probability that a certain attribute class occurs, and can be estimated by the percentage of the total area where an attribute class occurs
- $pps_{1 \text{ to } n}$ is the prior joint probability of the attribute classes considered. This can be assumed as the percentage of the total area where all the classes occur together
- $ppa_i, i=1, \dots, n$ is the probability of finding F given the occurrence of the attribute class E_i . This can be computed according to the formula (Chung and Fabbri, 1993):

$$ppa_i = 1 - (1 - (area_i - 1)nb(I))$$

where $area_i$ is the area where class i is met, and $nb(I)$ is the area of class I where F is also met
- psF is the prior probability of F all over the area, and can be evaluated as the percentage of the total area where F is met.

In the case of Bayes' theorem, then, the value function R_k for each attribute class is given by ppa_i , and the functional composition by the above expression for $P\{F / E_1 \dots E_n\}$.

According to this rule, a map can be computed for each combination of the attribute classes occurring. This can be done through a routine cross operation within the raster GIS used.

It is worth stressing that prior probability psF needs to be estimated for the computation of certainty factors described below, but its use in absolute terms is meaningless since it is practically impossible to know the probability of future occurrences of phenomena. The prediction score must then be taken as an indicator of favourability to the phenomenon in general terms, and not as a numerical estimate of probability. In this sense, psF is simply a scaling factor to ensure numerical consistency.

CERTAINTY FACTORS

If certainty factors (Shortliffe and Buchanan, 1975) are used, then rules change according to the following:

1) the certainty factor for an attribute class can be defined as:

$$CF_i = [P\{F/E_i\} - P\{F\}] / [P\{F/E_i\}(1 - P\{F\})] \text{ if } P\{F/E_i\} > P\{F\}$$

$$CF_i = [P\{F/E_i\} - P\{F\}] / [(1 - P\{F/E_i\}) P\{F\}] \text{ if } P\{F/E_i\} < P\{F\}$$

with $i=1, \dots$, number of thematic data classes of the causal factors

2) for 2 classes, the certainty factor is computed according to the following rules:

$$CF_{1+2} = CF_1 + CF_2 - (CF_1 * CF_2), \text{ if both } CF_1 \text{ and } CF_2 \text{ are nonnegative}$$

$$CF_{1+2} = CF_1 + CF_2 / [1 - \min(|CF_1|, |CF_2|)], \text{ if } CF_1 \text{ and } CF_2 \text{ have opposite sign}$$

$$CF_{1+2} = CF_1 + CF_2 + (CF_1 * CF_2), \text{ if both } CF_1 \text{ and } CF_2 \text{ are negative}$$

3) the procedure applies iteratively for more maps by computing first the $CF_{1+2}=CF_{12}$, then $CF_{13}=CF_{12}+3$, as shown above, and so on.

In this case the value function is $P(F/E_l)$, which is computed according to the formulas presented for Bayes' theorem; the composition is given by $CF1_n$ where n is the number of attributes.

FUZZY LOGIC OPERATORS

As the last method we used, fuzzy sets theory (Zadeh, 1965) was applied by calculating the 'fuzzy sum', 'fuzzy product', 'fuzzy and', 'fuzzy or' and 'fuzzy gamma function'. The membership functions were assumed to be equal to the estimates of the probability of finding F given the class E_l (Zadeh, 1968), i.e. ppa_l , being:

- 1) 'fuzzy and' = $\min(ppa_l)$, $l=1, \dots, n$
- 2) 'fuzzy or' = $\max(ppa_l)$, $l=1, \dots, n$
- 3) 'fuzzy product' = $\prod(ppa_l)$, $l=1, \dots, n$
- 4) 'fuzzy sum' = $1 - \prod(1 - ppa_l)$, $l=1, \dots, n$
- 5) 'fuzzy gamma operation' = $(\text{fuzzy sum})^\gamma (\text{fuzzy product})^{1-\gamma}$

In this way, rules for map overlaying are defined so that the analyst can evaluate the influence of the different occurrences in attribute data all over the study area, in order to detect the favourability of sites for the occurrence of further phenomena. Again, the value function for each attribute class is given by ppa_l and the functional composition provided by the above rules.

DATA SELECTION AND ANALYSIS

As pointed out in the conceptual geological model, three types of factors are deemed to be important for mineralization:

- the lithological unit
- the distance from the rock basement, which is considered to be a source of copper minerals
- the presence of geological lineaments, and particularly faults and joints, through which mineralized solutions migrate.

For each factor, an appropriate representation must be found and used for subsequent analysis.

As far as the first and third factors are to be described, the existing geological maps (Kambewa, 1998; Master, 1991; Lopez, 1998) have been digitised (Figure 2). Qualitative classes have been used to represent the function V_k for lithological units. For the case of the lineaments, a raster distance map has been computed. This represents a continuous field, which has been sliced by considering the distribution of the five existing mines. Three of them were located within 500 m from the known lineaments, while the others were within 3000 m. This has brought to choose a map slicing using three classes, one with distances <500 m, another with distances between 500 and 3000 m, and the third with distances >3000 m. Using this slicing prevents the occurrence of spurious zonation in favourability scores, due to insufficient training sets. For instance, if the distance range between 500 and 3000 m would have been sliced into more than one class, some of them would have shown zero conditional frequency of the training evidence set, thus yielding the senseless statement that distances between 500 and 3000 m might be less favourable than distances of 3000 m.

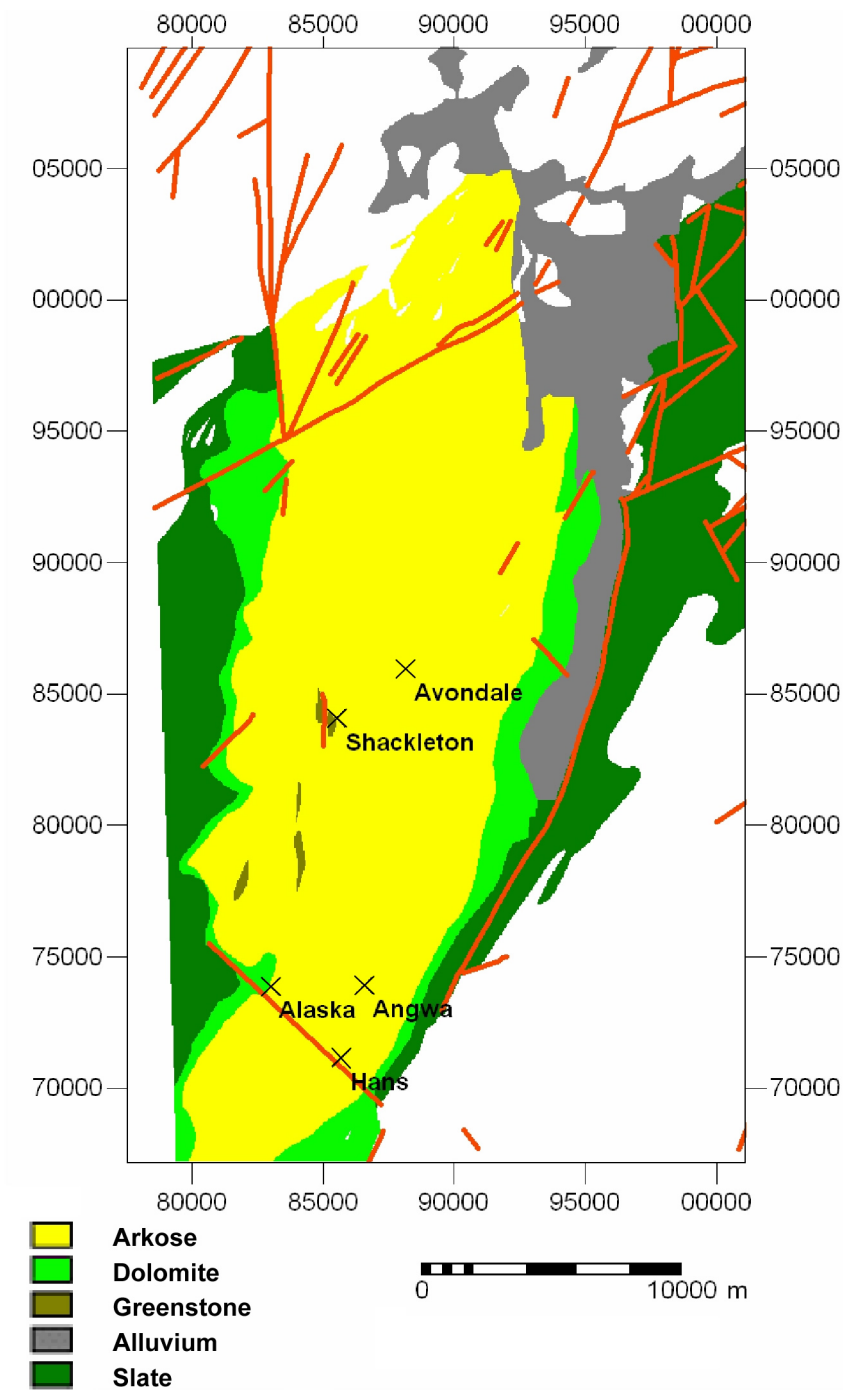


Figure 2 Geology of the study area: lithological units and main lineaments (red lines); existing mines are shown

The distance from the rock basement (second factor) is well represented by the total magnetic field obtained from airborne survey (Figure 3).

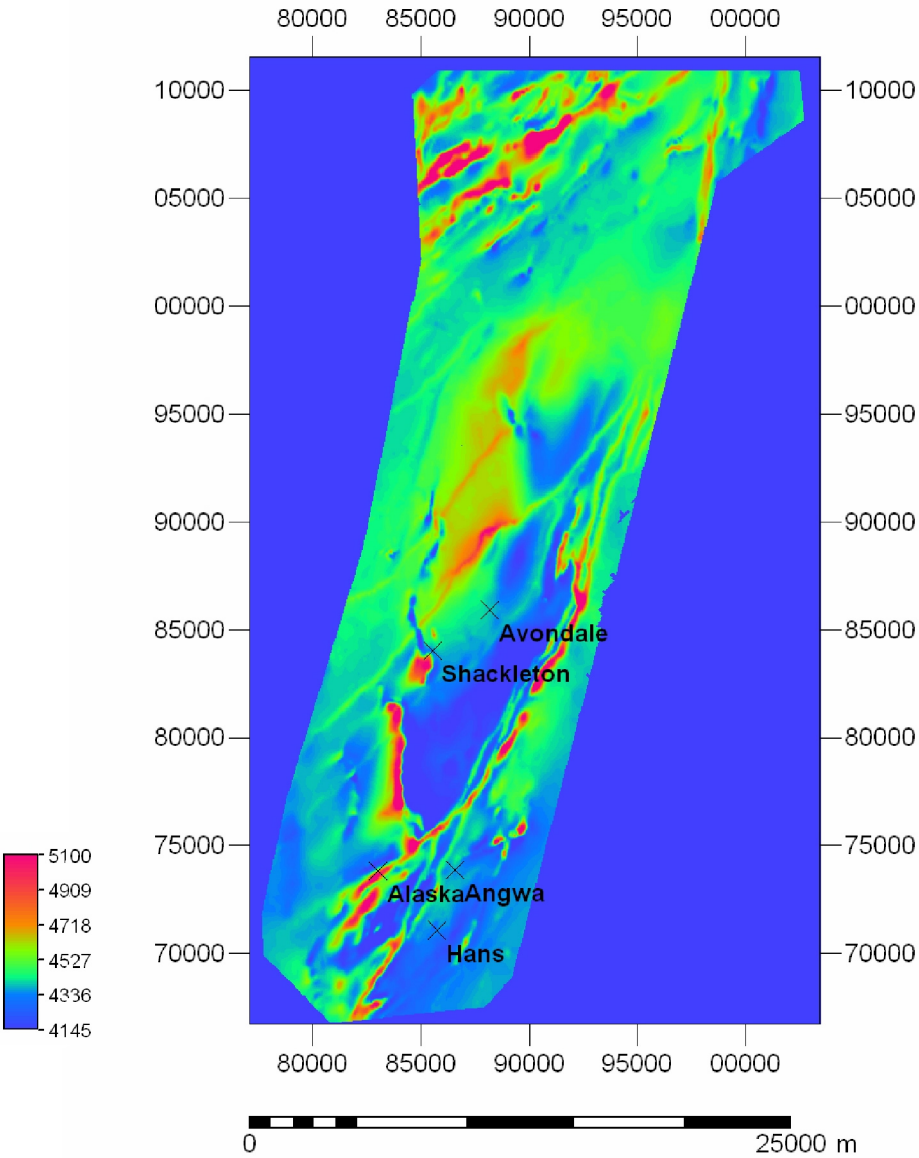


Figure 3 Aeromagnetic map, total magnetic field (nT).

Since the total magnetic field came as a continuous field, it has been sliced into 4 separate value classes; slicing provides a quick and conceptually simple way to bring these themes to discrete classes. This operational simplicity is achieved at the price of potential bias in the analyses, due

to considering an arbitrarily chosen slicing instead of the whole information carried by the continuous field distribution.

In addition to the supporting factors related to the geological structure of the area, an extensive database from geochemical exploration, consisting of measures of concentration of nine elements (Cu, Zn, Pb, Co, Fe, Ni, Ag, Au, As) taken from cores sampled approximately 4 m below terrain surface has been retrieved. The database has been used to map geochemical concentrations, potentially relevant as supporting factors for the occurrence of mineralisation.

A correlation analysis of the data shows that the element concentration occurrence does not follow a clear association pattern. Table 1 shows the computed correlation matrix, which makes it possible to suppose that all elements are almost uncorrelated. From exploratory interpolation it has been possible to state that Pb, Ni and Co don't show any particular spatial pattern, their highest and lowest values being scattered randomly over the study area. For this reason, these elements are not relevant as spatial indicators and have been discarded in subsequent analyses. In addition, Ag, Au and As show very often extremely low concentrations, and don't have any physical relation with the occurrence of the copper deposits considered here, although Kambewa (1998) recognised Ag as a strong evidence factor. This has led to not considering them in modeling. Eventually, only Cu, Fe and Zn have been retained as somehow meaningful geological indicators. Figure 4 shows a nearest neighbour interpolation of the data for these variables. A variogram surface analysis has been done to check for the isotropy of the values. All elements showed an approximately isotropic spatial correlation.

CU	1								
PB	0.217514	1							
ZN	0.127237	0.230512	1						
FE	<u>0.399333</u>	<u>0.422935</u>	0.205321	1					
CO	<u>0.33941</u>	<u>0.328598</u>	0.188015	<u>0.703861</u>	1				
NI	0.088643	0.159756	0.050928	0.175653	<u>0.374099</u>	1			
AS	0.054031	0.023898	0.297403	0.110737	0.101505	0.05664	1		
AU	0.105329	0.020772	0.009198	0.020203	0.040742	0.042179	-0.00843	1	
AG	0.118025	0.045308	0.025826	0.060796	0.003236	0.002897	0.01292	0.071024	1
	CU	PB	ZN	FE	CO	NI	AS	AU	AG

Table 1 Correlation matrix of geochemical data (correlations >0.3 highlighted)

Their actual meaning for the analysis has been tested during model calibration and validation as discussed below. An additional test on geochemical data has been done by computing the auto-krigeability and codispersion coefficients (Wackernagel, 1997, pp. 143-150) of copper with respect to the other elements. The graphs showing these coefficients for different lag spacing values are shown in Figure 5.

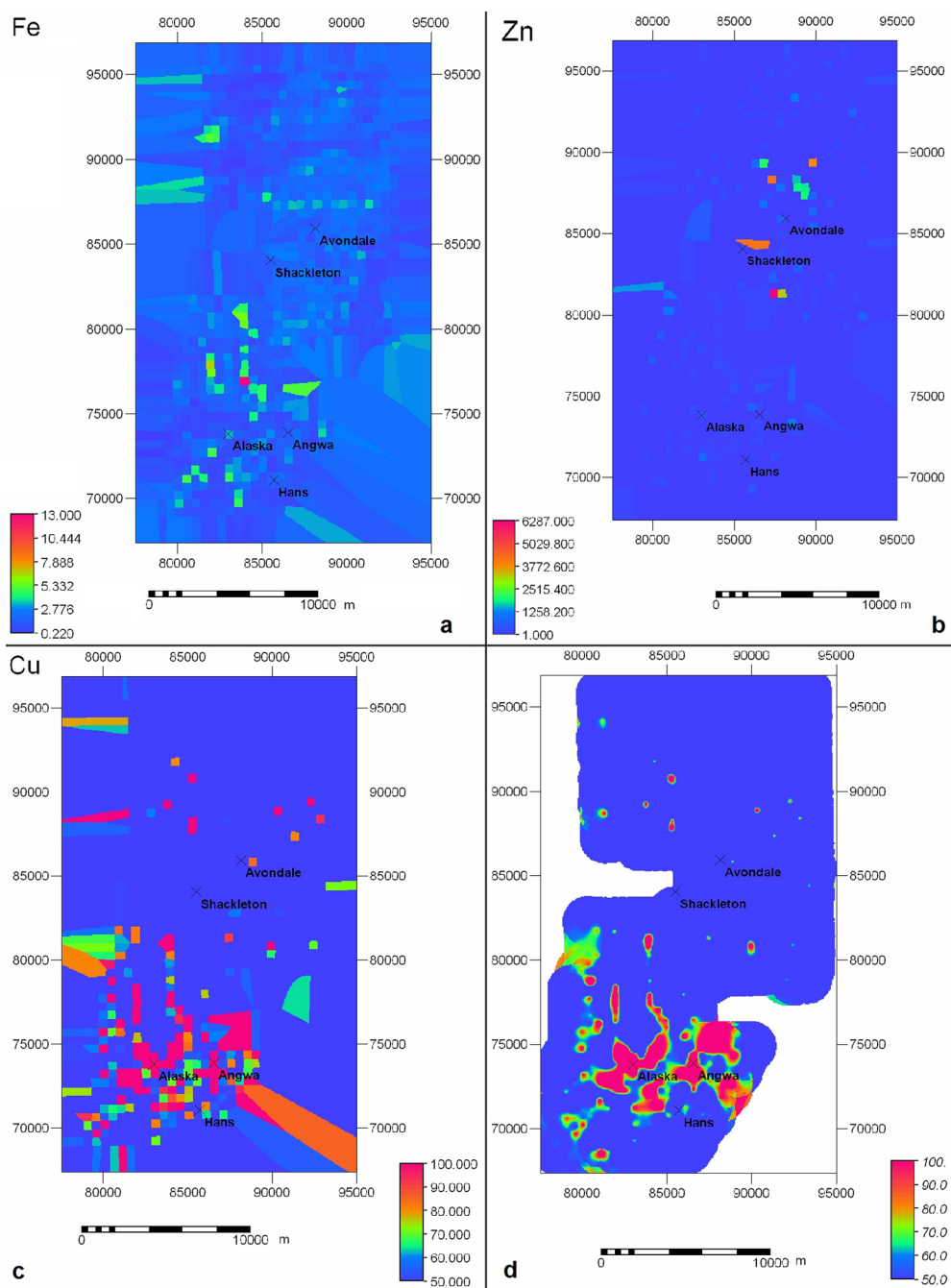


Figure 4 Nearest neighbour exploratory interpolations of iron (a), zinc (b) and copper (c) anomalies. (d) represents the refined interpolation of Cu using kriging.

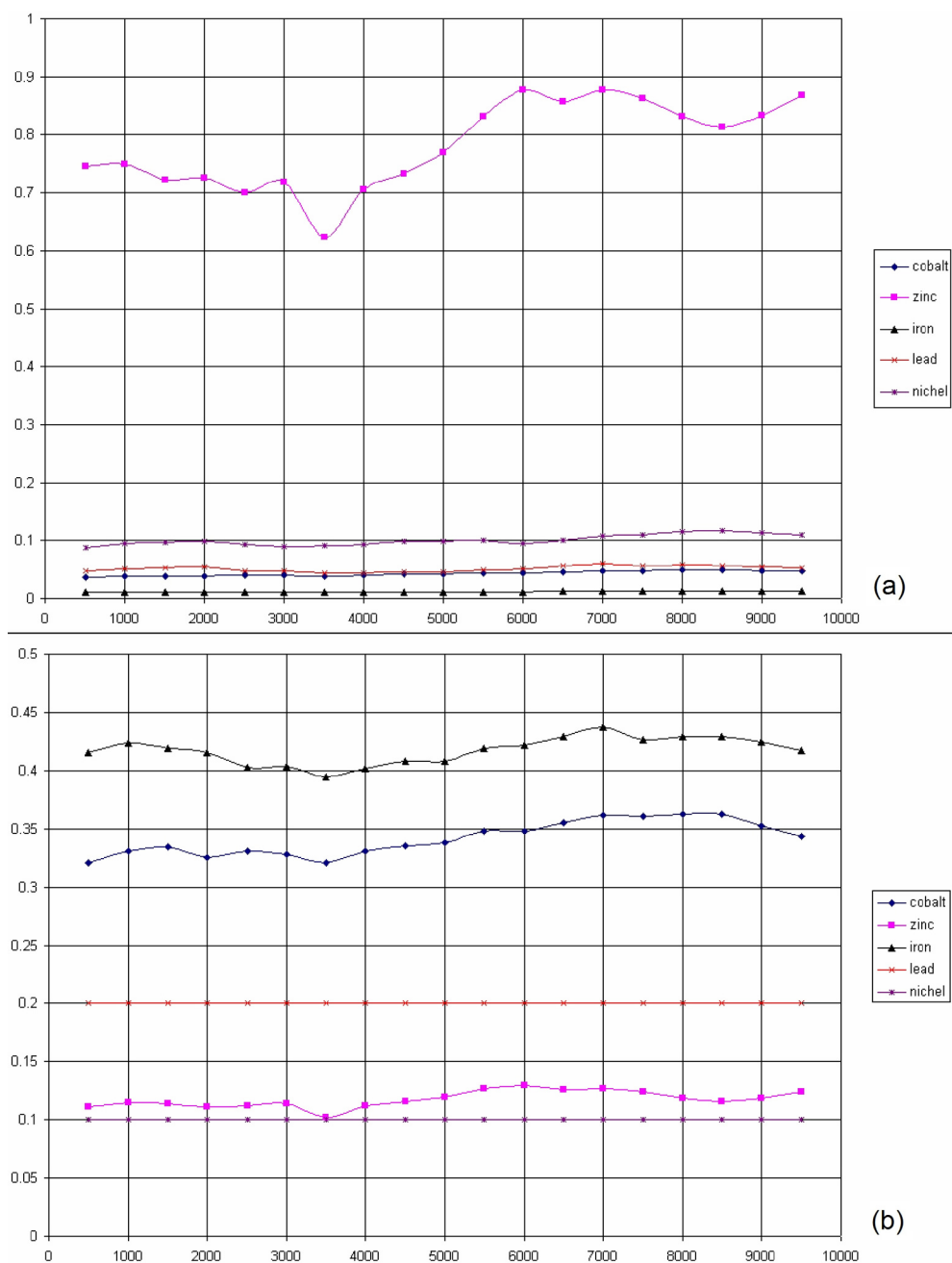


Figure 5 Autokrigability (a) and codispersion coefficients (b) of copper with respect to the other elements.

A constant value of the auto-krigeability coefficient (AC) of a regionalized random variable with respect to another, means that kriging that variable and co-kriging it using the other as co-variable, are not significantly different. As one can see, all AC's don't have any particular trend and can be considered stationary for the purpose of the test, although a certain variation is seen in the case of zinc. Thus, in first approximation, copper can be considered as an autonomous variable, as already highlighted from the correlation analysis.

Codispersion coefficients (CC's) are practically constant for all spatial scales, which supports the intrinsic correlation hypothesis, i.e. a similar correlation matrix holds at all spatial scales (Wackernagel, 1997). Moreover, correlation coefficients between the variables are generally rather low. For this reason, each variable can be considered as autonomous and there seems to be no exception to this at any spatial scale. It has then been chosen to interpolate Cu, Fe and Zn values independently using nearest neighbour, and to use the element concentration maps as separate causal factors for the modeling. Also, a more refined representation of chemical concentrations has been attempted for Cu, which is expected to be the only representative element for mineralization. Kriging was performed on the log-transform of the data, using an exponential variogram model with range 1000 m and sill 0.8, nested with a nugget effect of 0.2. The result is also shown in Figure 4.

Eventually, six causal factors have been selected: airborne total magnetic field, copper concentration, zinc concentration, iron concentration, geological units, and distance from geological lineaments. The data have been successively processed for the modeling analysis.

Together with causal factors, integration modeling requires definition of a training set of mineralized areas, which have been identified in the location of five known existing mines (Avondale, Schackleton, Alaska, Angwa, and Hans, as reported in the figures). In order to have an appropriate representation of the mines, since the shape and dimension of the mineralized deposits was not known, a square window of $N \times N$ pixels around each mine location point has been used. The pixels in the window received the value 1 (presence of mineralization) while the other pixels received 0 (absence of mineralization). This allows production of a binary map with five 1-valued windows and the rest equal to 0. The number of pixels, N , was chosen by subsequent trials, seeking for the representation which maximizes the prediction performance according to the criteria explained below. It was found that using $N=5$ proved satisfactory.

RESULTS

Once the data have been prepared, a modeling study has been performed using the favourability functions approach. The latter has been shown to require estimation of the frequencies of simultaneous occurrence of a certain combination of the causal factors in known mineralised sites. As a starting point, the five known mines were used. Following this criterion, favourability maps can be produced that, in principle, define the most promising areas for exploration and detailed mapping efforts for mineral exploitation, on the example of Figure 6. This map, presented for illustration purposes only, is not sufficiently robust for practical use because the number of known mineralised sites was too limited. Rather, we focus on the use of the favourability functions approach for the validation of predictions that can be obtained by combining the six causal factors described in the previous section.

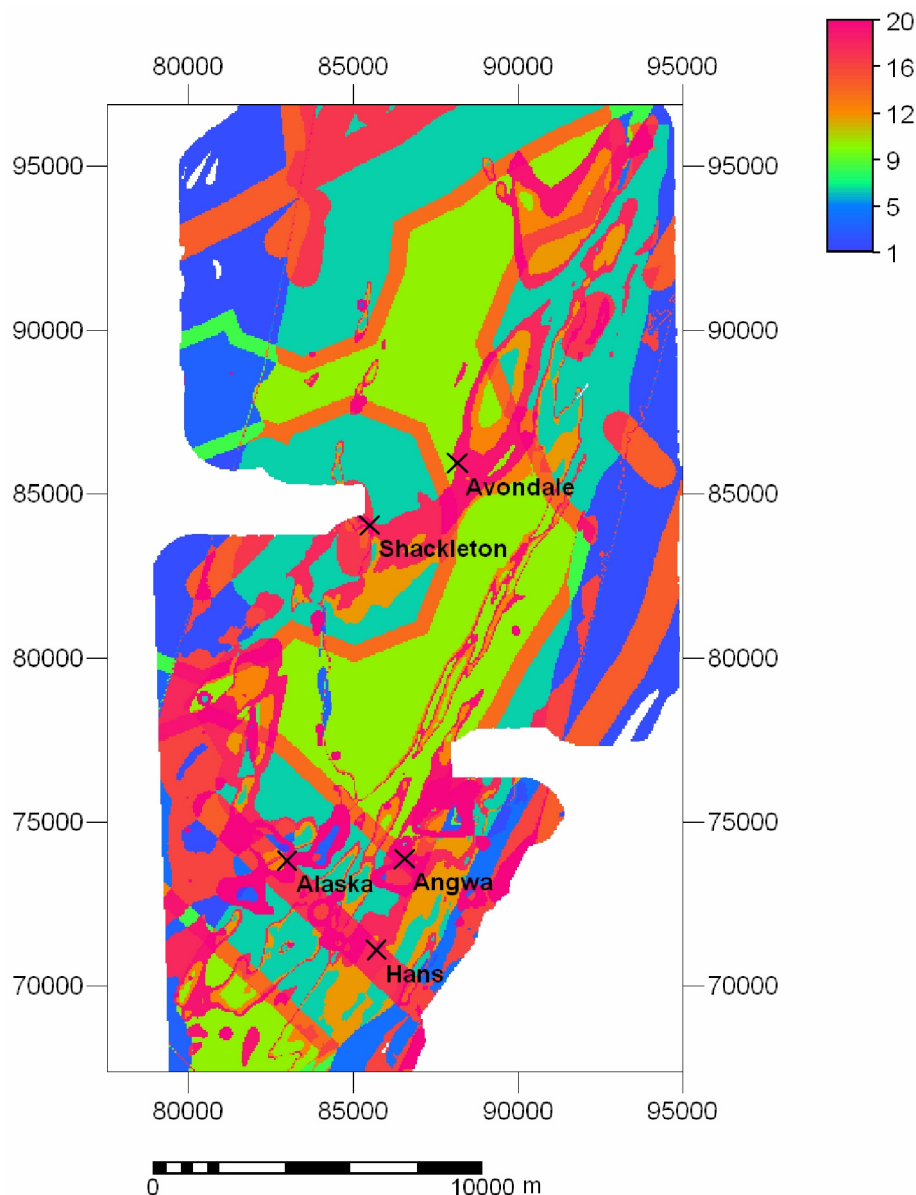


Figure 6 Example of prediction map: fuzzy-sum (the 20 classes are shown according to the legend).

The validation procedure consisted in using four out of the five existing mines as the evidence layer to compute frequency of occurrence of mineralisation. In this way, a favourability map can be build for any of the seven techniques outlined above, that reflects the frequency of occurrence of the causal factors at the four mines. Then the fifth existing mine was used for testing the prediction by recording the favourability score assigned to the pixels containing the discarded mine.

It has been argued that if the prediction model works, then the “validation mines” should always be found in the highest scoring percentage of the area.

For simplicity, the prediction maps were divided into 20 classes of favourability, each covering approximately 5% of the study area sorted in increasing favourability order. Thus, for instance, class 20 covers the 5% area with the highest prediction score, class 19 the 5% with second highest score and so on. According to this criterion, the scores of the area of the “validation mines” should always keep in a range of high favourability scores. For this case study, it has been chosen to consider acceptable all the predictions that assigned to all the “validation mines” a score class greater than or equal to 15, i.e. according to which all the “validation mines” were in the 25% highest favourability score area. Figure 7 shows the results obtained with the different techniques and two options for the causal factors, namely: (1) using lithological classes, distance from lineaments, aeromagnetic field, and the Cu concentration shown in Figure 4 (d); this option has been denoted by “4 factors, refined Cu; (2) using six causal factors, including Zn and Fe and the previous ones with nearest neighbour interpolation of all chemicals.

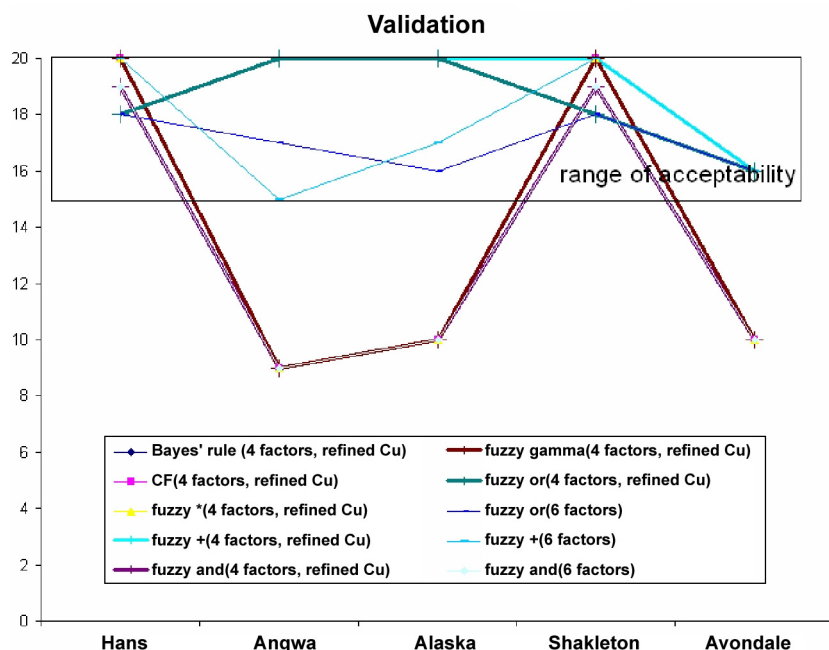


Figure 7 Validation of the prediction using four and six causal factors.

As expected, different predictors behaved differently, and only two of the seven made an acceptable prediction of all the “validation mines”, namely the fuzzy or operation, and the fuzzy sum operation. It is worth noting that adding Zn and Fe produced worse results with these two predictors, than using only four causal factors.

DISCUSSION

Making a prediction using in turn only five of the six causal factors allowed us to test the sensitivity of the predictors to each of them. It is of relevance here to discuss the point for the two predictors that behaved acceptably in all cases, i.e. the fuzzy or, and fuzzy sum algorithms. For reference, also a sensitivity analysis for fuzzy product, which failed the validation test, is reported. Figure 8, Figure 9 and Figure 10 show the results obtained.

The well-performing predictors were sensitive to the most complete and detailed data, i.e. airborne total magnetic field, and to the distance to lineaments. The fuzzy or predictor proves to be also sensitive to lithological units in one case. Other factors, such as zinc and iron, in the case of the fuzzy or, produce better results; however, it is worth noting that neglecting just one of them produces a much worse prediction than using both together, which indicates they introduce noise. Moreover, no physical nor geological link has been established between these elements and copper mineralization; this further supports the idea that these factors should be neglected. For the fuzzy product, an improvement when neglecting either distance to lineaments or lithological units is observed. The same behaviour is observed in all other predictors that fail to pass validation. This further supports the indication that the rules of combination of the factors are in this case inadequate, except for fuzzy sum and fuzzy or.

Using a refined interpolation of Cu (Figure 4 (d)) has limited effect on the calculation of the frequencies of the causal factors at existing mines, although it tends to improve the predictions. However, when considering mapping it can provide a better prediction pattern reflecting the kriged surface instead of the patchy one obtained from nearest neighbor interpolation.

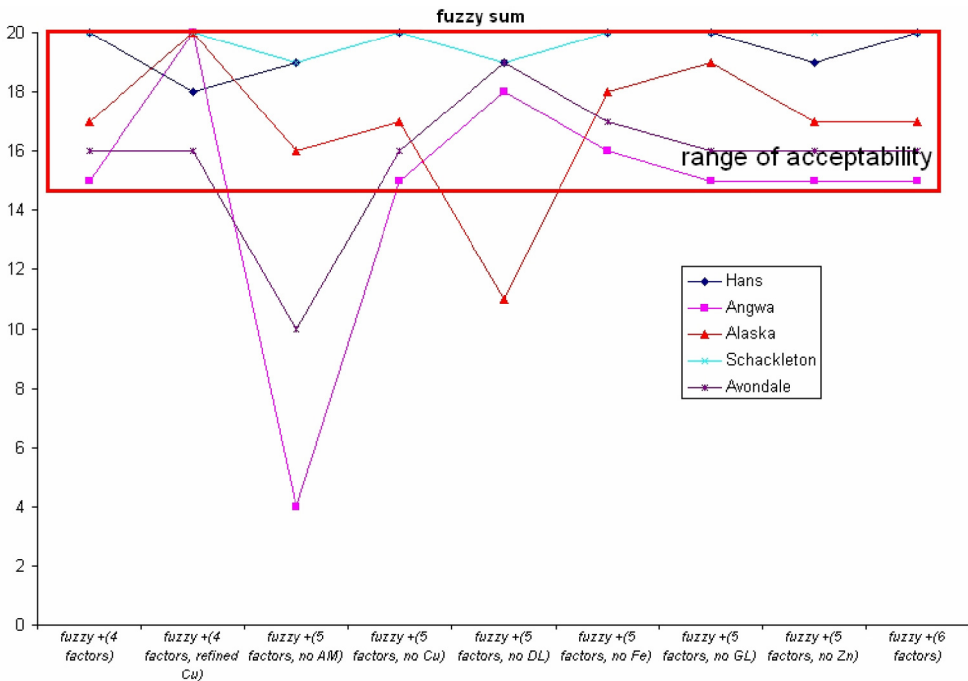


Figure 8 Sensitivity analysis for fuzzy sum.

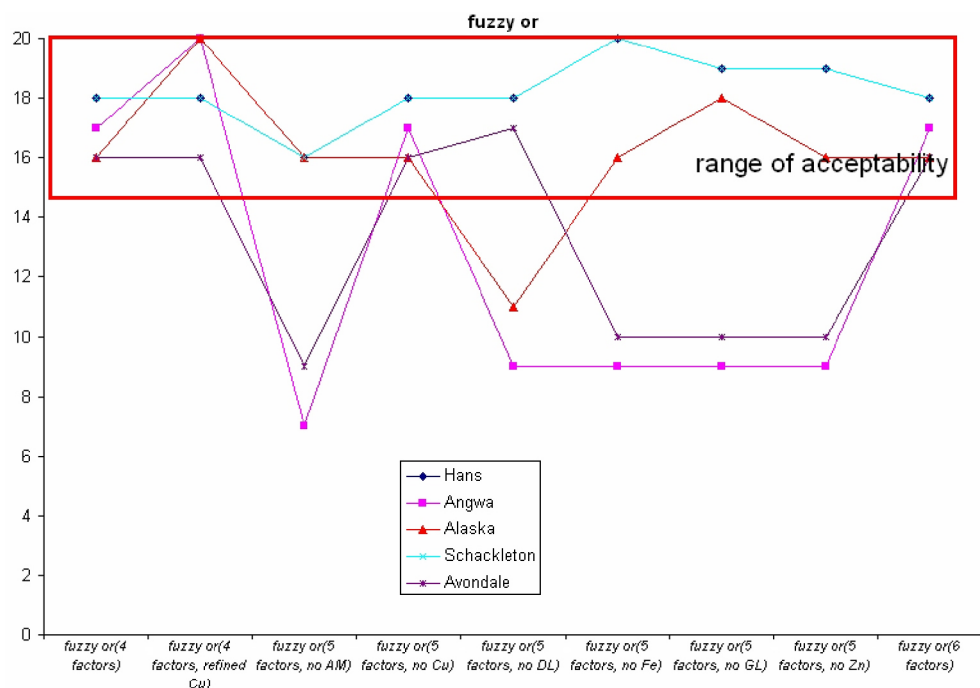


Figure 9 Sensitivity analysis for fuzzy or.

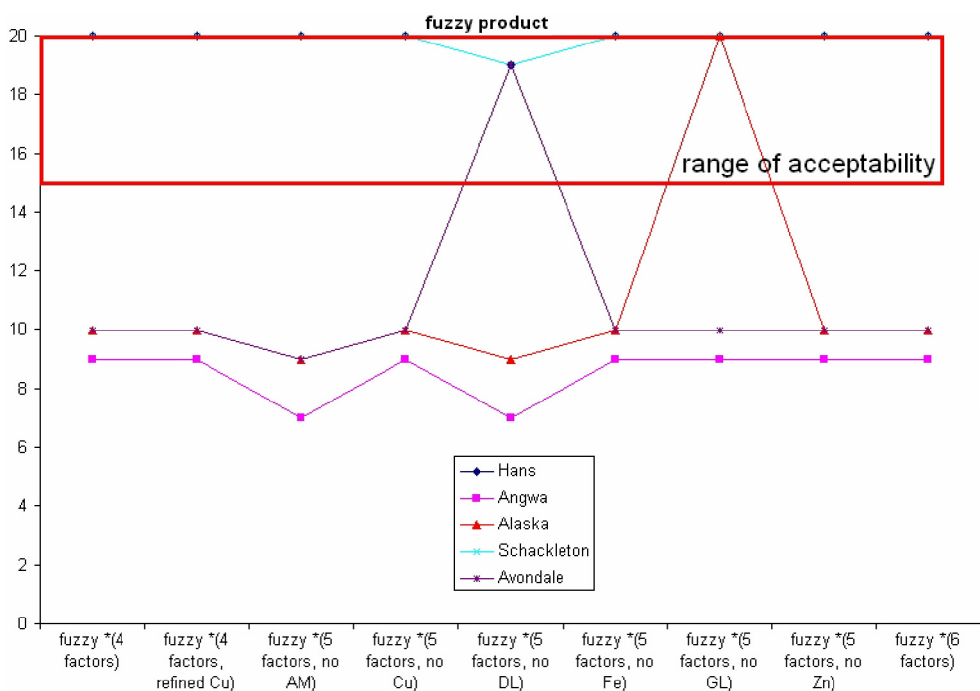


Figure 10 Sensitivity analysis for fuzzy product.

CONCLUSIONS

From the exploratory analysis and modeling here presented it has been possible to obtain a positive indication on the geological conceptual model adopted for mineralization, by stating that a prediction can be drawn by integrating the factors corresponding to the distance to lineaments, lithological unit, aeromagnetic field, and copper concentration in soil.

It has been possible to evaluate different predictors, highlighting that different integration rules have substantially different behaviour, and in some cases fail to predict mineralization.

So far, a completely data driven approach has been supported since limitations due to lack of geological knowledge limited the possibility of using expert's knowledge-based modifications of the conditional probabilities. Improvements in prediction can be achieved through a more detailed geological description of the area based e.g. on the interpretation of remotely sensed data, and through a better representation of the favourability functions, either via an improved survey of known mineralisation sites, or via the introduction of value functions based on expert's judgement.

REFERENCES

- Bonham-Carter, G. *GIS for Geoscientists: Modeling with GIS*. New York: Elsevier; 1994.
- Cheng, Q; Agterberg, F.P. Fuzzy weights of evidence method and its application in mineral potential mapping. *Natural Resources Research* 1999; 8 (1).
- Chung, C.F., Fabbri, A.G., van Westen, C.J. 'Multivariate regression analysis for landslide hazard zonation'. In: Carrara, A., Guzzetti, F., (eds.) *Geographical Information Systems in Assessing Natural Hazards*. Dordrecht: Kluwer Academic Publishers; 1995.
- Chung, C.F., Fabbri, A.G., *The representation of geoscience information for data integration*. Non Renewable Resources 2. 1993. pp. 122–139.
- Chung, C; Fabbri, A.G. 'Probabilistic prediction models for landslide hazard mapping'. *Photogrammetric Engineering and Remote Sensing* 1999; 65 (12): 1389–1399.
- Duda, R., Hart, P., Nilsson, N., 1976. 'Subjective bayesian methods for rule-based inference systems'. In: *Proceedings of the National Computer Conference*, Vol. 45, AFIPS.
- Harris, D; Pan, G. Mineral favorability mapping: a comparison of artificial neural networks, logistic regression, and discriminant analysis, *Natural Resources Research* 1999; 8 (2): 93–109.
- Heckermann, D. 'Probabilistic interpretation of the MYCIN's certainty factors'. In: Kanal, L.N., Lemmer, J.F. *Uncertainty in Artificial Intelligence*. Elsevier; New York: 1986.
- Kambewa, C. Use of GIS in exploration for sediment-hosted copper deposits in the Makonde Basin, Zimbabwe [Unpublished MSc Thesis]. ITC, Delft, The Netherlands. 1998. 102pp.
- Lopez, C., E., Integration of aeromagnetic, thematic mapper and geochemical data for mineral exploration in the Makonde district, Zimbabwe [Unpublished MSc Thesis]., ITC, Delft, The Netherlands. 1998. 70pp.
- Master, S., The origin and controls on the distribution of copper and precious metal mineralization at the Mhangura and Norah Mines, Mhangura, Zimbabwe [Unpublished PhD Thesis]. Witwatersrand, South Africa. 1991. 382pp.
- Shafer, G. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press; 1976. 220 pp.

Shortliffe, E; Buchanan, B.G. 'A modeling of inexact reasoning in medicine' *Mathematical Biosciences* 1975; 23: 351–379.

Wackernagel, H. *Multivariate Geostatistics*. Berlin: Springer; 1997. 200 pp.

Zadeh, L.A. 'Fuzzy sets'. *IEEE Informatics and Control* 1965; (8): 338–353.

Zadeh, L.A. 'Probability measures of fuzzy events'. *Journal of Mathematical Analysis and Application*, 1968; 10: 1999–2010.

Cite this article as: Woldai, Tsehaie; Pistocchi, Alberto; Master, Sharad. 2006. 'Validation and sensitivity analysis of a mineral potential model using favourability functions'. *Applied GIS* 2 (1). pp. 2.1–2.19. DOI: 10.2104/ag060002.