PERGAMON

Computers & Geosciences 29 (2003) 3–13

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

Data-driven fuzzy analysis in quantitative mineral resource assessment

X. Luo[a], R. Dimitrakopoulos[b,*]

[a] WH Bryan Mining Geology Research Centre, The University of Queensland, Brisbane, Qld. 4072, Australia
[b] Department of Mining and Metallurgical Engineering, McGill University, Montreal, Canada

Received 14 February 2000; received in revised form 25 March 2002; accepted 3 April 2002

Abstract

The integration of geo-information from multiple sources and of diverse nature in developing mineral favourability indexes (MFIs) is a well-known problem in mineral exploration and mineral resource assessment. Fuzzy set theory provides a convenient framework to combine and analyse qualitative and quantitative data independently of their source or characteristics.

A novel, data-driven formulation for calculating MFIs based on fuzzy analysis is developed in this paper. Different geo-variables are considered fuzzy sets and their appropriate membership functions are defined and modelled. A new weighted average-type aggregation operator is then introduced to generate a new fuzzy set representing mineral favourability. The membership grades of the new fuzzy set are considered as the MFI. The weights for the aggregation operation combine the individual membership functions of the geo-variables, and are derived using information from training areas and L_1 regression.

The technique is demonstrated in a case study of skarn tin deposits and is used to integrate geological, geochemical and magnetic data. The study area covers a total of $22.5\,km^2$ and is divided into 349 cells, which include nine control cells. Nine geo-variables are considered in this study. Depending on the nature of the various geo-variables, four different types of membership functions are used to model the fuzzy membership of the geo-variables involved.

Keywords: Fuzzy sets; Data-driven methods; Mineral favourability index; Resource assessment; Exploration

1. Introduction

Advances in computer technologies play an increasingly important role in the exploration and assessment of mineral deposits. Geographic Information Systems are frequently used to evaluate mineral potential in exploration districts (Bonham-Carter, 1994) and provide tools to deal with multiple data sets, or layers, of diverse character from various sources. The combination and evaluation of these layers may be either visual/qualitative or quantitative. Quantitative integration of diverse exploration data and evaluation of results is an intricate task where mathematical/statistical approaches are employed to: (a) maximize the extraction of information from the data; (b) effectively combine diverse information; (c) provide tools to quantify inherent uncertainties; (d) rank potential targets; and (e) reduce data processing and evaluation time. Quantitative integration of diverse multi-source geo-information, including geological, geochemical, geophysical and remote-sensing data has been attempted within several mathematical or statistical frameworks. Notable approaches include regression-based methods (Sinclair and Woodsworth, 1970; Agterberg et al., 1972; McCammon, 1973; Chung and Agterberg, 1980), characteristic analysis (Botbol et al., 1978; Sinding-Larsen and Strand, 1981; McCammon et al., 1983), canonical correlation analysis (Pan and

*Corresponding author.
E-mail address: roussos@uq.edu.au (R. Dimitrakopoulos).

0098-3004/02/$ - see front matter
PII: S0098-3004(02)00078-X

Harris, 1992; Pan, 1993a, b) and Bayesian frameworks (Duda et al., 1978; Harris, 1984; Bonham-Carter et al., 1988; Singer and Kouda, 1988). Recent trends in modelling geo-information are based on weights of evidence (Bonham-Carter et al., 1988; Agterberg, 1989, 1992), the Dampster–Shafer belief and plausibility functions (Shafer, 1976; Moon, 1990; An et al., 1994), fuzzy sets (Zadeh, 1965; An et al., 1991; Chung and Fabbri, 1993), and neural networks (Singer and Kouda, 1996, 1997a, b). Furthermore, comprehensive presentation of existing methods for quantitative data integration in mineral exploration and mineral favourability analysis is available in textbooks (Bonham-Carter, 1994; Pan and Harris, 2000).

Previous approaches were based on the spatial mapping of mineral potential or favourability of mineral occurrence over an exploration play. Although not always explicitly defined, these approaches are formulated based on the notion of a mineral favourability index, MFI, $y(p)$, of an exploration cell p defined as an aggregation, $F(\cdot)$, of geo-variables x_i, $i = 1, \ldots, n$, and their transformations $g_{pi}(x_i)$ such that $y(p) = F(g_{p1}(x_i), \ldots, g_{pn}(x_n))$. This formulation requires a specific definition of $y(x)$, the transformation of the available data to a suitable form, and the definition of aggregation function F. In predicting favourability, the various techniques attempt to merge diverse data and may be challenged in their effectiveness and degree of integration. For example, most of the statistical approaches based on regression and characteristics or canonical correlation analysis transform the available data to binary or ternary form. Such techniques may be limited in the type of information that can be quantified. Bayesian methods and weights of evidence approaches may be limited by conditional independence requirements for the data used. Belief and plausibility functions may, in practice, be difficult to interpret and assess. A limitation of the above approaches is that they can be somewhat inflexible in expressing the different degrees of favourability of the mineral occurrence for each of the individual geo-variables considered.

Fuzzy sets provide an alternative framework that could improve upon some of the limitations in previous techniques. A fuzzy set-based formulation is used by An et al. (1991) to integrate geological and geophysical data from the Farley Lake area, Canada. The study uses the algebraic-sum and γ aggregation operators to outline favourable areas for base metal and iron deposits, based on an approach also advocated by Chung and Fabbri (1993). Although a notable development, the approach used to combine geo-information does not take into account the individual relative importance of geo-variables in quantifying favourability for mineral occurrence. In addition, the approach is 'knowledge-driven' and as such does not include objective 'data-driven' criteria in the combination of the geo-information used. A recent study (Cheng and Agterberg, 1999) presents a fuzzy set extension of the weight of evidence approach, which, differing from the approach presented herein, retains similar limitations to its progenitor.

The present study contributes a new data-driven approach in deriving a fuzzy mineral favourability index (FMFI) of mineral occurrence suitable for mineral exploration and resource assessment.

2. Fuzzy sets

Geological information and data interpretations used in mineral exploration are inherently ambiguous. The quantitative precision of expressions like "relatively high", "high", "fair", "low", and "relatively low" or "fairly favourable/unfavourable for the mineral occurrence", as well as the grey areas between these expressions, is difficult to define. Fuzzy set theory (Zadeh, 1965; Zimmermann, 1991) provides a mathematical framework to represent the linguistic and data ambiguities frequently encountered in mineral exploration, geological information analysis and interpretation. The theory formally associates any statement with a quantifiable measure indicating the degree of possibility of the statement.

2.1. Definition and examples

If X is a collection of objects denoted by x, fuzzy set A in X is the set of ordered pairs

$$A = \{(x, \mu_A(x)) | x \in X\}, \tag{1}$$

where $\mu_A(x)_A$ is termed the membership function or membership grade of x in A (Zadeh, 1965). $\mu_A(x)$ maps X to the membership space M. When M contains only the two points 0 and 1, A is a non-fuzzy set and $\mu_A(x)$ is identical to the characteristic function of a regular, non-fuzzy, set. The range of $\mu_A(x)$ is $[0, 1]$, where 0 expresses non-membership and 1 expresses full membership.

A simple example may be the fuzzy set A, "the number of faults is favourable for mineral occurrence", described as

$$A = \{(1, 0.3), (2, 0.7), (3, 0.9), (4, 0.8), (5, 0.6), (6, 0.4)\}.$$

The most favourable situation for mineral occurrence is when the membership function is the highest. Here, the membership function is 0.9 and corresponds to the presence of three faults in an exploration cell. As another example, the fuzzy set "distance to a fault is favourable for mineral occurrence" is defined by Eq. (1) and a membership function $\mu_A(x) = 1/[1 + (x/500)^2]$. The membership function indicates that the closer a cell is to a favourable fault, the higher the favourability for mineral occurrence.

2.2. Aggregation operations

Aggregation operations on fuzzy sets are used to combine them to a single set (Dubois and Prade, 1985; Klir and Folger, 1988; Zimmermann, 1991). An aggregation operation F is defined as the mapping

$$F:[0,1]^n \to [0,1] \tag{2}$$

for $n \geqslant 2$. When applied to n fuzzy sets $A_1, A_2, \ldots, A_n$ defined on $X_1, X_2, \ldots, X_n$, the aggregation operation F produces an aggregate fuzzy set A defined on $X = (X_1, X_2, \ldots X_n)$, by operating on the membership functions of each $x = (x_1, x_2, \ldots, x_n) \in X$ in the aggregated set

$$\mu_A(x) = F(\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots, \mu_{A_n}(x_n) | x_i \in X_i, i = 1, \ldots, n) \tag{3}$$

for each $x \in X$. Permissible aggregation operations must satisfy the boundary conditions

$$F(0,0,\ldots,0) = 0 \quad \text{and} \quad F(1,1,\ldots,1) = 1 \tag{4}$$

and be monotonic and non-decreasing in all their arguments, i.e.

$$a_i, b_i \in, [0,1], \text{ if } a_i \leqslant b_i, i = 1 \text{ to } n, \text{ then } F(a_1, a_2, \ldots, a_n) \leqslant F(b_1, b_2, \ldots, b_n). \tag{5}$$

Aggregation operations often used in mineral exploration are outlined in Appendix A.

2.3. Aggregation operations and mineral favourability

Different aggregation operators may reflect different geological criteria in mineral exploration. For the criterion "if the measurements of geo-variable X_1 AND geo-variable X_2 in a cell p are favourable for mineral occurrence, then p is favourable for mineral occurrence", the corresponding operator can be selected either as a Cartesian product or an algebraic product; for the criterion "if the measurement of geo-variable X_1 OR the measurement of X_2 in cell p is favourable for mineral occurrence, then p is favourable for mineral occurrence", the corresponding operator can be selected as an algebraic sum or a bounded sum. If neither "AND" nor "OR" can represent the favourability of the mineral occurrence in terms of X_1 and X_2 (for example, for the criterion "the favourability of mineral occurrence is determined by the favourabilities of X_1 and X_2") a γ-operator may be selected as the corresponding aggregation operator.

3. A fuzzy mineral favourability index

3.1. Definition and estimation

In mineral exploration, an exploration terrain can be represented by a grid of cells and each cell assessed as a potential exploration target. The mineral favourability index of an exploration cell is assessed from the available geo-variables and characterized by proposition A: "Cell is favourable for the mineral occurrence." In fuzzy set theory, proposition A can be considered as a fuzzy set with a membership function which can quantify how good a member the cell is to fuzzy set A. In the proposition A above, specific values of the membership function quantify the degree or grade that a given cell is favourable for mineral occurrence.

The FMFI of cell p is defined as $y_f(p) = \mu_A(x(p))$, the membership grade of the fuzzy set A: "mineral favourability" in p. Furthermore, we may consider the fuzzy set A_i: "Measurement of geo-variable X_i in a given cell $p, x_i(p)$, is favourable for mineral occurrence" with a membership function $\mu_{A_i}(x_i(p))$. According to Eq. (3), the FMFI can then be represented as an aggregation operation on $\mu_{A_i}(x_i(p)|x_i \in X_i, i = 1, \ldots, n)$

$$y_f(p) = \mu_A(x(p)) = F(\mu_{A_1}(x_1(p)), \mu_{A_2}(x_2(p)), \ldots, \mu_{A_n}(x_n(p))). \tag{6}$$

The definition of the aggregation operation $F(\cdot)$is critical for derivation of the FMFI. All of the aforementioned aggregation operations have the following conversion property:

$$F(\mu_{A_1}(x_1), \mu_{A_2}(x_2)) = F(\mu_{A_2}(x_2), \mu_{A_1}(x_1)),$$

where $F(\cdot)$ denotes any of these aggregation operations. As a result, the relative importance of the individual fuzzy subset may be neglected; for example, in copper deposits occurring in gabbros, mineral occurrence is mostly associated with a gravity anomaly, X_1, and also occasionally associated with a geochemical copper anomaly in soil, X_2. If for a given cell p, due to a possible deep-seated copper deposit, $x_1(p)$ is favourable but $x_2(p)$ is unfavourable for mineral occurrence, then $\mu_{A_1}(x_1(p)) = 1, \mu_{A_2}(x_2(p)) = 0$ and

$$y_f(p) = \mu_A(x(p)) = F(\mu_{A_1}(x_1(p)), \mu_{A_2}(x_2(p))) = F(1,0)$$

while cell p could be highly favourable for mineral occurrence. On the other hand, if for a cell $q, x_1(q)$ is unfavourable (e.g. a local absence of gabbros and gravity anomalies) while $x_2(q)$ is favourable for mineral occurrence, then $\mu_{A_1}(x_1(q)) = 0, \mu_{A_2}(x_2(q)) = 1$ and

$$y_f(q) = \mu_A(x(q)) = F(\mu_{A_1}(x_1(q)), \mu_{A_2}(x_2(q))) = F(0,1)$$

while q is unfavourable for mineral occurrence.

This can be expressed as

$$\mu_A(x(p)) = F(1,0) > F(0,1) = \mu_A(x(q)),$$

stressing the importance of the relative contribution of the information used; it is not generally satisfied by the aggregation operations previously mentioned.

To overcome the limits of common aggregation operations and express the degree of importance and relevance of individual geo-variables to mineral occurrence, the following weighted average is suggested as the suitable aggregation operation for an FMFI:

$$y_f(p) = \mu_A(x(p)) = \sum_{i=1}^{n} \lambda_i \mu_{A_i}(x_i(p)), \tag{7}$$

where λ_i is the weight indicating the relative importance of the fuzzy set A_i for A. Conditions in Eq. (4) lead to the unbiasness condition

$$F(1, 1, \ldots, 1) = \sum_{i=1}^{n} \lambda_i = 1.$$

In addition, for any set of values $\{\mu_{A_i}(x_i(p)), i = 1, \ldots, n\}$, it is required that

$$F(\mu_{A_1}(x_1(p)), \mu_{A_2}(x_2(p)), \ldots, \mu_{A_n}(x_n(p))) \geqslant 0$$

which leads to non-negative weights $\lambda_i \geqslant 0, i = 1, \ldots, n$. The required aggregation operation is therefore defined as

$$\mu_A(x(p) = \sum_{i=1}^{n} \lambda_i \mu_{A_i}(x_i(p)),$$
$$\sum_{i=1}^{n} \lambda_i = 1 \quad \forall \lambda_i \geqslant 0. \tag{8}$$

3.2. An example

Consider the fuzzy set A: "Cell p is favourable for the occurrence of a gabbros hosted copper deposit." Cell p is to be assessed from geological information grouped in the form of three relevant fuzzy sets:

A_1 = the gravity survey in p is favourable for mineral occurrence;

A_2 = the elemental copper in soil in p is favourable for mineral occurrence;

A_3 = the distance between p and a fault system is favourable for mineral occurrence.

If, for example, the related weights are derived from Eq. (8) and found to be $\{\lambda_1, \lambda_2, \lambda_3\} = \{0.4, 0.2, 0.4\}$, then the FMFI is

$$\mu_A(x(p)) = 0.4\,\mu_{A_1}(x_1(p)) + 0.2\,\mu_{A_2}(x_2(p)) + 0.4\,\mu_{A_3}(x_3(p)).$$

For cells p and q with measurements it is

$$p : \{x_1(p), x_2(p), x_3(p)\} = \{45\,\text{mgal}, 15\,\text{ppm}, 0.4\,\text{km}\}$$

and

$$\{\mu_{A_1}(45\,\text{mgal}), \mu_{A_2}(15\,\text{ppm}), \mu_{A_3}(0.4\,\text{km})\} = \{0.9, 0.1, 0.8\},$$

$$q : \{x_1(q), x_2(q), x_3(q)\} = \{5\,\text{mgal}, 65\,\text{ppm}, 0.3\,\text{km}\}$$

and

$$\{\mu_{A_1}(5\,\text{mgal}), \mu_{A_2}(65\,\text{ppm}), \mu_{A_3}(0.3\,\text{km})\} = \{0.1, 0.95, 0.85\}.$$

Then, the FMFI is $\mu_A(x(p)) = 0.7$ and $\mu_A(x(q)) = 0.57$, indicating p is more favourable for the mineral occurrence than q.

The definition of membership functions $\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots, \mu_{A_n}(x_n)\}$ and the derivation of weights $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ are both crucial for the effectiveness of the FMFI. The membership functions $\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots, \mu_{A_n}(x_n)\}$ can be defined objectively based on the characteristics of a geo-variable using the normal practice in exploration, whereas the weights $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ can be derived through a variety of optimal methods. All methods must be constrained by the conditions of unbiased and non-negative weights (Eq. (8)).

4. Application in exploring skarn tin deposits

4.1. Geology and control cells

The study area is located in southwestern China near the contact zones of the south margin of the Lugu granitic body (Yang et al., 1986) as shown in Fig. 1, and includes several skarn tin deposits. The exposed rocks in the study area include Presinian slightly metamorphosed sandstone and carbonate rocks, Triassic/Jurassic sandstones, and Quaternary sandstones. Tin deposits occur exclusively in Presinian rocks.

The Lugu granites are the major magmatic rock type in the study area, with their outcrop occupying about $180\,\text{km}^2$. The Lugu granites in the study area are related to skarn tin deposits both in space and in time. Petrochemically the granites contain high contents of Sn (58 ppm), Be (8.8 ppm), and B (1084 ppm) and represent an abundant resource of tin mineralization in the study area. Tin deposits occur in all but the external contact zone of the Lugu granitic body. The width of the external contact zone is approximately 2 km.

N-E trending faults control the distribution of the ore deposits, and their ages determine the shape and size of ore bodies. The tin deposits in the study area are usually associated with magnetite mineralization; therefore, the presence of magnetite mineralization can be regarded as a favourable direct indicator for the occurrence of tin deposits. Since a magnetic anomaly may also indicate the presence of magnetite mineralization, it is viewed as an indirect index for mineral occurrence. Soil geochemical surveys show that a combination of anomalous tin, copper, lead, and zinc may indicate tin mineralization near the earth's surface.

The study area covers a total of $22.5\,\text{km}^2$ and is divided into 349 cells with a unit cell size of $0.64\,\text{km}^2$.

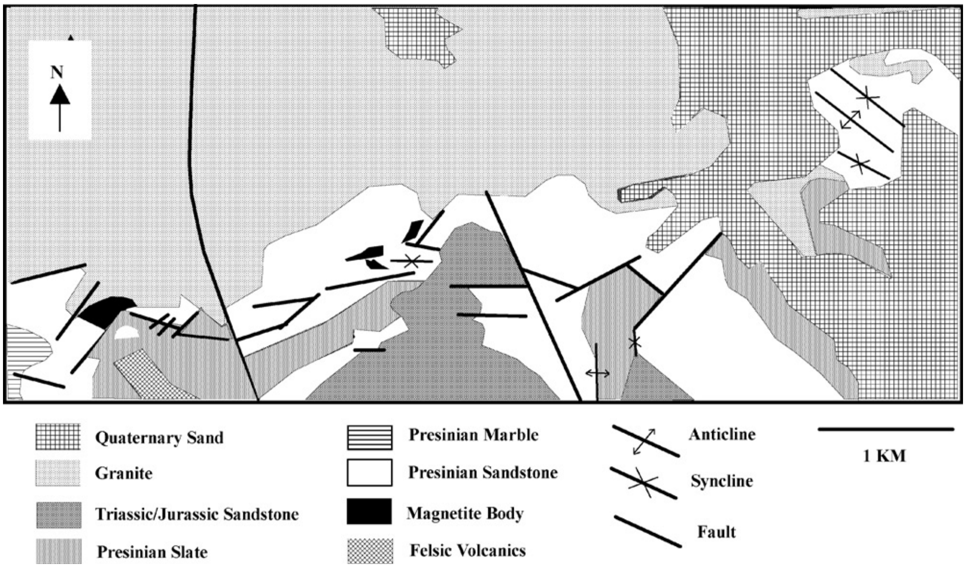

Fig. 1. Geological map of study area. Southern margin of Lugu granite deposit is shown in northwest part of area.

There are nine control cells while tin mineralization occurs in all control cells.

4.2. Geo-variables and their membership functions

Variables $X_1 - X_9$ are selected in the study area based on the discussion above. Table 1 shows the selected geo-variables and their characteristics. The nine geo-variables are as follows: X_1: The average Sn concentration of soil in a given cell. X_2: The average soil Cu content in a cell. X_3: The average Pb concentrations in soil at a given cell. X_4: The average Zn concentrations in soil at a given cell. X_5: The distance between the centre of a given cell and the boundary of the granite. The variable is positive if the cell is in the external contact zone or negative if the cell is in the internal contact zone. X_6: The ratio of area occupying Presinian sandstone-slate at a given cell. X_7: Coded as 1 or 0 based on the presence or absence of N-E trending faults in a given cell. X_8: Coded as 1 or 0 based on the presence or absence of magnetite mineralization in a given cell. X_9: Coded as 1 or 0 based on the presence or absence of magnetic anomalies in a given cell.

For each geo-variable X_i, the corresponding fuzzy set is proposed as A_i: "Measurement of X_i in a given cell p, $x_i(p)$, is favourable for the occurrence of skarn tin deposits." Four types of membership functions are defined as follows:

(i) *Geochemical anomalies* in the concentrations of elements Sn, Cu, Pb and Zn are favourable for the occurrence of tin deposits. They are recognized in terms of their average m and standard deviation s, while the membership function of the corresponding fuzzy set can

Table 1
Geo-variables used in this study

Geo-variable	Definition of geo-variable	Type of measurement
X_1	Sn	Concentration in soil (ppm)
X_2	Cu	Concentration in soil (ppm)
X_3	Pb	Concentration in soil (ppm)
X_4	Zn	Concentration in soil (ppm)
X_5	Contact zone of granite	Distance to the boundary (m)
X_6	Presinian sandstone-slate	Ratio of area covered
X_7	N-E trending faults	Presence
X_8	Magnetite mineralization	Presence
X_9	Magnetic anomalies	Presence

be defined from the membership function

$$\mu_A(X) = \begin{cases} 1 - \dfrac{bs}{x - am + bs}, & x > am, \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

where constants $a, b \geqslant 0$. The properties of this model are

$\mu_A(x) = 0$ when $x \leq \mathrm{am}$
$\mu_A(x) = 0.5$ when $x = am + bs$
$\mu_A(x) \to 1$ when $x \to +\infty$
$\mu_A(x)$ is a monotonic increasing function.

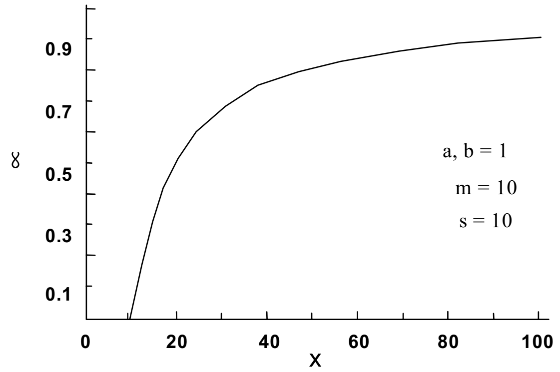

Fig. 2. First type of membership function defined in Eq. (9).

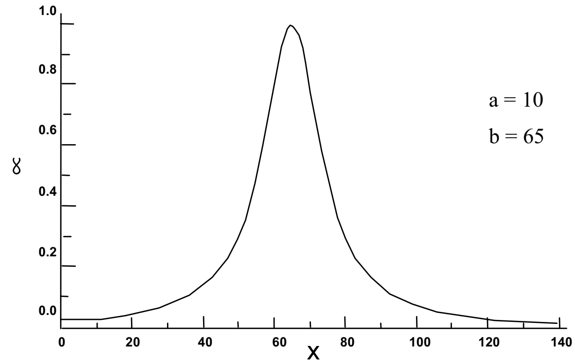

Fig. 3. Second type of membership function defined in Eq. (10).

The selection of coefficients a and b depends on the definition of anomalies. It is highly likely that no anomaly is present when the concentration is less than the mean, implying that constant a is 1. The unclear boundary between "anomaly" and "background" is around the mean plus one standard deviation, implying that b is also 1. In the present study, a and b are selected as 1, which is sufficient for the scope of this analysis. The function model (Eq. (9)) becomes

$$\mu_A(X) = \begin{cases} 1 - \dfrac{s}{x - m + s}, & x > m, \\ 0, & \text{otherwise}, \end{cases}$$

where m and s are the mean and standard deviation of the geochemical measurements in a cell. This model is shown in Fig. 2.

(ii) *The distance* to the centre of external contact zones is recognized as an indicator for possible mineralization. The smaller the distance, the higher the favourability for mineral occurrence. The following membership function can define the corresponding fuzzy set:

$$\mu_A(X) = \frac{1}{1 + \left((x - b)/a\right)^2}, \tag{10}$$

where $a, b > 0$, a denotes the range and b expresses the centre (see Fig. 3).

The properties of the model in Eq. (10) are

$\mu_A(x)$ is symmetric to the centre b,
$\mu_A(x) = 0.5$ when $x = b \pm a$,
$\mu_A(x) = 1$ when $x = b$,
$\mu_A(x) \to 0$ when $x \to \pm\infty$.

In the present study, the centre and range of the external contact zone are all approximately 1 km. Thus, constants a and b are set to 1 and the membership function in Eq. (11) becomes

$$\mu_A(X) = \frac{1}{1 + (x - 1)^2}.$$

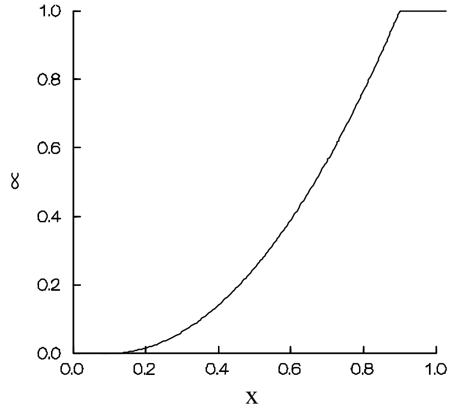

Fig. 4. Third type of membership function defined in Eq. (11).

(iii) *The ratio* of the area occupying Presinian sandstone-slate is regarded as favourable for mineralization when it is over 90%. The degree of favourability decreases as the ratio decreases and is regarded as unfavourable when it drops to less than 10%. Accordingly, the membership function of the corresponding fuzzy set can be defined from

$$\mu_A(X) = \begin{cases} 1, & 1x \geqslant 0.9, \\ (x - 0.1)^2/(0.9 - 0.1)^2, & 0.1 \leqslant x < 0.9, \\ 0, & x < 0.1 \end{cases} \tag{11}$$

as shown in Fig. 4.

(iv) *The presence* of N-E trending faults, magnetite mineralization, and magnetic anomalies can be expressed through membership functions which correspond to a ternary model. This model expresses favourability as a function of existence of the corresponding geo-variable and it is

$$\mu_A(X) = \begin{cases} 1, & x \text{ exist}, \\ 0.5, & x\text{'s existance unknown}, \\ 0, & \text{does not exist}. \end{cases} \tag{12}$$

4.3. Control cells

The aggregation operation defined by the system of Eq. (8) is preferably derived using a set of control cells in the exploration area under consideration. Control cells provide key information about the local relative importance of geo-variables to mineralization and are selected on mainly two criteria: (a) they must be well explored; and (b) they must be available to quantification by all relevant geo-variables.

Control cells in the study area with examples of membership function values for tin, magnetite mineralization, and N-E trending faults are shown in Figs. 5–7, respectively. The control cells in the study area have the following characteristics:

(a) All control cells are distributed along the contact zone of the south margin of the Lugu granitic body.
(b) Most control cells have high membership values of the fuzzy set corresponding to the Sn anomaly, indicating the presence of Sn content in these cells.
(c) Most control cells exhibit magnetite mineralization.
(d) Most control cells contain N-E trending faults.

4.4. Deriving fuzzy mineral favourability indices

When a set of control cells is well defined and meets the criteria previously described, the weights of the aggregation operation in Eq. (8) can be obtained by

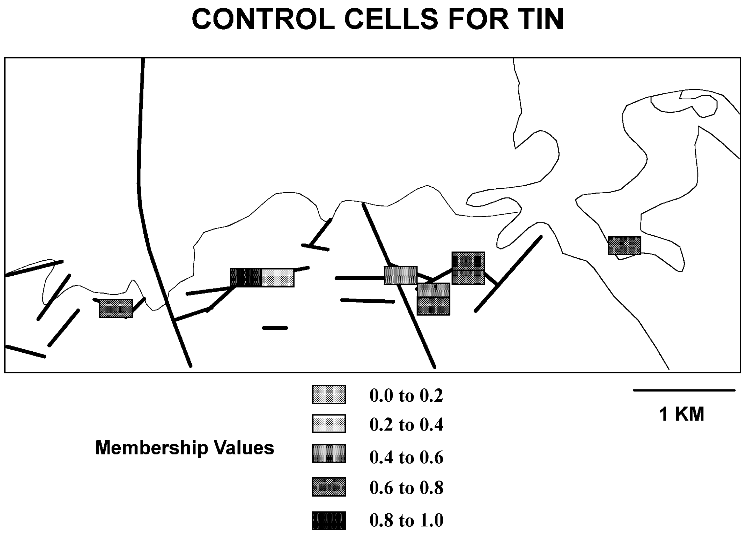

Fig. 5. Control cells with membership values for tin.

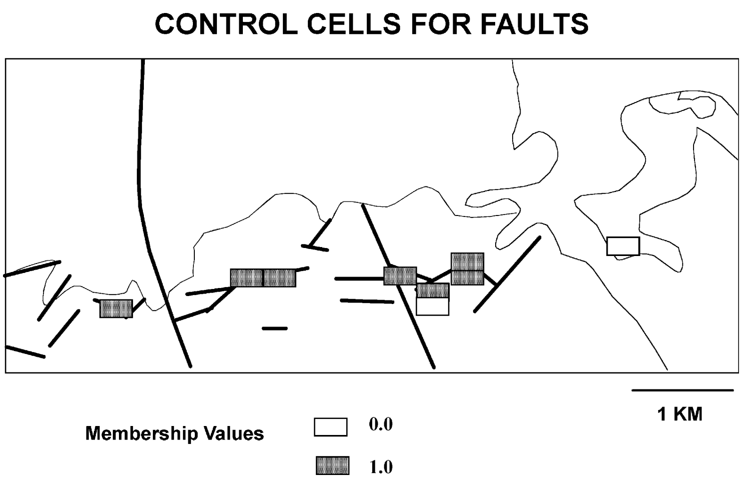

Fig. 6. Control cells with membership values for presence of faults.

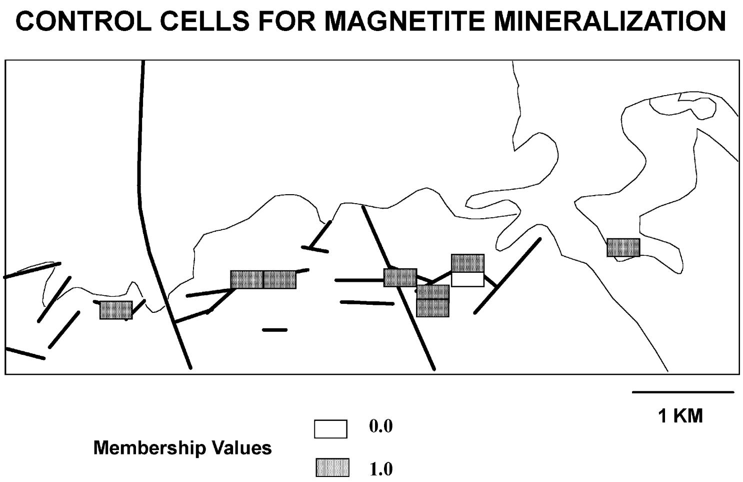

Fig. 7. Control cells with membership values for magnetite mineralization.

Table 2
Weights of fuzzy sets derived from control cells in study area

Fuzzy set	A_1 (Sn)	A_2 (Cu)	A_3 (Pb)	A_4 (Zn)	A_5 (contact zone)	A_6 (Presinian rocks)	A_7 (faults)	A_8 (magnetite mineralization)	A_9 (magnetic anomaly)
Weight	0.21	0.04	0.03	0.03	0.08	0.12	0.18	0.20	0.11

restricted L_1 regression. Consider the following optimization formulation:

$$\begin{array}{l} \text{minimize } \sum_{\mathrm{k}=1}^{\mathrm{K}} |\mu_A(x(p_k)) - \sum_{i=1}^{n} \lambda'_i \mu_{A_i}(x_i(P_k))| \\ \qquad \text{subject to } 0 \leqslant \lambda'_i 1 \ \forall \lambda'_i, \end{array} \tag{13}$$

where K denotes the number of control cells and p_k denotes a control cell ($k = 1, \ldots, K$). The minimization problem in Eq. (13) can be solved by extending the techniques of interval linear programming to include interval constraints on weights (Armstrong and Hultz, 1977; Zanakis and Rustagi, 1982). The desired weights in Eq. (8) are obtained from

$$\lambda_i = \lambda'_i / \sum_{i=1}^{n} \lambda'_i. \tag{14}$$

Table 2 shows the weights from restricted L_1 regression in the study area using the available control cells. The weights of the fuzzy sets corresponding to anomalies in concentrations for the elements Cu, Pb and Zn are relatively small, implying a weak, distant relationship between these geo-variables and mineralization. The weights of the fuzzy sets corresponding to the presence of Sn anomalies, N-E trending faults, and magnetite mineralization are relatively large: $\lambda_1 + \lambda_7 + \lambda_8 > 0.5$, implying that mineral occurrence is dominated by these three geo-variables. This result is consistent with previous exploration experience in the study area.

The recognition of the relative importance and significance of geo-variables to mineral occurrence is critical to mineral prediction, as stressed earlier. As a result, the weighted average operator in Eq. (8) is superior to other aggregation operators. In the present study, when measurements of the elements copper, lead and zinc in a given cell p are favourable and measurements of tin, magnetite mineralization and N-E trending faults are unfavourable for mineral occurrence, then the presence of mineralization in p is highly unfavourable. On the other hand, when measurements of copper, lead and zinc are unfavourable but measurements of tin, magnetite mineralization, and N-E trending faults are favourable, then p is highly favourable for mineral occurrence. In short, geological consistency is preserved by the weights shown in Table 2.

4.5. Results for favourability assessment and target identification

The FMFIs of all 349 cells in the study area are generated from the aggregation operation in Eqs. (7) and (8) and use the derivations in Eqs. (13) and (14). Fig. 8 shows the 20 cells with the highest FMFIs including the nine control cells. In comparison to Figs. 5–7, those cells showing the presence of a tin anomaly, N-E trending faults and magnetite mineralization are associated with high FMFIs ($y_f(p) \geqslant 0.63$),

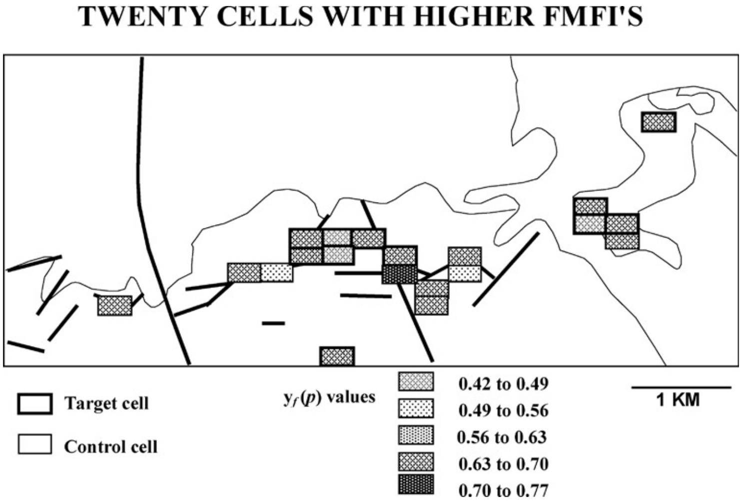

Fig. 8. Twenty cells with higher FMFIs, including control cells.

reflecting the dominant role of these geo-variables in the possible presence of tin mineralization. The FMFIs in the remaining cells not shown in Fig. 8 are substantially lower: ($y_f(p) < 0.3$) in their FMFIs and the cells show a marked, natural split from the cells shown in Fig. 8. Note that the interpretation of the FMFIs is made in the context of identifying a sequence of possible exploration targets in the study area. In general, FMFI values should not be interpreted in a strict sense; that is, the chances of tin deposits in cell with FMFI of 0.70 are higher than those of a cell with FMFI of 0.65; both cells are highly likely targets.

In the western section of the study area, there is no cell with high FMFIs except a control cell, suggesting that the potential mineral occurrence is very limited. In the eastern section, there are three cells with relatively high FMFIs ($0.49 < y_f(p) < 0.56$), clustered mostly around a control cell, showing a high potential for tin occurrence. The central part of the study area exhibits a large cluster of high FMFI cells ($y_f(p) \geqslant 0.56$) about the control cells. N-E trending faults are well developed in this section, suggesting a relatively high potential for Sn mineral occurrences. An additional possible target area is indicated in the southern central section of the study area.

5. Summary and conclusions

The fuzzy set framework provides the analytical tools to deal with qualitative and quantitative, multi-source and multi-character geological information, information which may be linguistically ambiguous, imprecise, and incomplete. The present study defines the degree of mineral favourability in terms of a fuzzy mineral favourability index. The proposed FMFI combines favourabilities from individual geo-variables rather than direct measurements of these variables and characterizes the degrees of these favourabilities with membership functions. The character of any pertinent geo-variable is used to define physically meaningful membership functions for the available observation, as shown in the case study.

The key element of the proposed approach is the use of an unbiased, weighted average aggregation operator to define the favourability grade of an exploration cell. The data-driven determination of weights for the contribution of individual geo-variables, that is their membership functions, is based on supervised training using control cells and restricted L_1 regression. The use of weights emphasizes the discrimination of the importance of each geo-variable in the mineral favourability of a cell, as well as its geological appropriateness.

The case study shows the practical aspects of the proposed fuzzy mineral favourability index in identifying exploration targets for skarn tin deposits in the vicinity of the Lugu granite body in southwestern China. The proposed method identifies eleven geologically consistent targets for further detailed exploration.

Several improvements of the proposed approach may be considered in the future. When deriving weights for Eq. (7), additional optimality criteria may be developed, whereas in the situation of large numbers of control cells, the fuzzy approach may be extended to include spatial dependencies in geo-variables. An additional consideration could be the introduction of a relative confidence measure associated with the FMFI through the definition of a companion fuzzy set "information confidence in cell p".

Acknowledgements

Funding was provided by the National Science and Engineering Research Council of Canada grant OGP0105803 to R. Dimitrakopoulos. Thanks are in order to D. Singer, M. Scott and D. Hedger and the two journal reviewers for their comments on earlier drafts of the manuscript.

Appendix A. Common aggregation operations on fuzzy sets

Some common aggregation operations on fuzzy sets used in mineral exploration are given below. Additional information for aggregation operations used in mineral favourability analysis may be found in Bonham-Carter (1994) and Pan and Harris (2000):

(a) The Cartesian product of a series of n fuzzy sets $A = A_1 \times A_2 \times \cdots \times A_n$ has a membership function defined by

$$\mu_A(x) = \min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots, \mu_{A_n}(x_n)\}.$$

The aggregated membership value $\mu_A(x)$ is the minimum value and not affected by others.

(b) The algebraic sum $A = A_1 + A_2 + \cdots + A_n$ has a membership function defined by

$$\mu_A(x) = \prod_{i=1}^{n}(1 - (1 - \mu_{A_i}(x_i)).$$

The aggregated membership value $\mu_A(x)$ is not smaller than $\max\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots, \mu_{A_n}(x_n)\}$ and not greater than $\min\{1, \sum_{i=1}^{n} \mu_{A_i}(x_i)\}$.

(c) The algebraic product $A = A_1 \cdot A_2 \cdot \ldots \cdot A_n$ has a membership function defined by

$$\mu_A(x) = \prod_{i=1}^{n} \mu_{A_i}(x_i).$$

The aggregated membership value $\mu_A(x)$ is not greater than the minimum value and tends to be very small due to the effect of multiplying values less than 1; this operator decreases more rapidly than the Cartesian product.

(d) The bounded sum $A = A_1 \oplus A_2 \oplus \cdots \oplus A_n$ has a membership function defined by

$$\mu_A(x) = \min\left\{1, \sum_{i=1}^{n} \mu_{A_i}(x_i)\right\}.$$

This operator increases more rapidly than the algebraic sum.

(e) The γ-operator defines a membership function

$$\mu_A(x) = \left(\prod_{i=1}^{n} \mu_{A_i}(x_i)^{1-\gamma}\right)\left(1 - \prod_{i=1}^{n}(1 - \mu_{A_i}(x_i))^{\gamma}\right),$$

where $\gamma \in [0, 1]$. The γ-operator is a combination of the algebraic product and the algebraic sum. When γ is 1, this operator is the same as the algebraic sum; and when γ is 0, this operator is the algebraic product. The choice of γ provides a balance between the strong effects of the algebraic sum and the weak effects of the algebraic product.

References

Agterberg, F.P., 1989. Computer programs for mineral exploration. Science 245, 76–81.

Agterberg, F.P., 1992. Combining indicator patterns in weights of evidence modeling for resource evaluation. Nonrenewable Resources 1 (1), 39–50.

Agterberg, F.P., Chung, C.F., Fabbri, A.G., Kelly, A.M., Springer, J.S., 1972. Geomathematical evaluation of copper and zinc potential of the Abitibi area, Ontario and Quebec. Geological Survey of Canada, Paper 71–74.

An, P., Moon, W.M., Rencz, A., 1991. Integration of geological, geophysical, and remote sensing data using fuzzy set theory. Canadian Journal of Exploration Geophysics 27 (1), 1–11.

An, P., Moon, W.M., Bonham-Carter, G.F., 1994. Uncertainty management in integration of exploration data using the belief function. Nonrenewable Resources 3 (1), 60–71.

Armstrong, R.D., Hultz, J.W., 1977. An algorithm for restricted discrete approximation problem in the L_1 norm. SIAM Journal on Numerical Analysis 14 (3), 555–565.

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists: Modelling with GIS. Pergamon, Oxford, 398pp.

Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1988. Integration of geological datasets for gold exploration in Nova Scotia. Photogrammetric Engineering and Remote Sensing 54 (11), 1585–1592.

Botbol, J.M., Sinding-Larsen, R., McCammon, R.B., Gott, G.B., 1978. A regionalized multivariate approach to target selection in geochemical exploration. Economic Geology 73, 534–546.

Chung, C.F., Agterberg, F.P., 1980. Regression models for estimating mineral resources from geological map data. Mathematical Geology 12 (5), 473–488.

Chung, C.F., Fabbri, A.G., 1993. The representation of geoscience information for data integration. Nonrenewable Resources 2 (2), 122–139.

Cheng, Q., Agterberg, F.P., 1999. Fuzzy weights of evidence method and its application in mineral potential mapping. Natural Resources Research 8 (1), 27–35.

Dubois, D., Prade, H., 1985. A review of fuzzy set aggregation connectives. Information Sciences 36, 85–121.

Duda, R.O., Heart, P.E., Barrette, P., Gasching, J.G., Konolige, K., Reboh, R., Slocum, J., 1978. Development of the Prospector consultation system for mineral exploration. Final Report, SRI Projects 5821 and 6415: Stanford Research Institute International, Menlo Park, CA, 202pp.

Harris, D.P., 1984. Mineral Resource Appraisal. Oxford University Press, New York, 445pp.

Klir, G.J., Folger, T.A., 1988. Fuzzy Sets, Uncertainty, and Information. Prentice-Hall, Englewood Cliffs, NJ, 355pp.

McCammon, R.B., 1973. Nonlinear regression for dependent variables. Mathematical Geology 5, 365–375.

McCammon, R.B., Botbol, J.M., Sinding-Larsen, R., Bowen, R.W., 1983. Characteristic analysis, 1981-final program and a possible discovery. Mathematical Geology 15 (1), 59–83.

Moon, W.M., 1990. Integration of geophysical and geological data using evidential belief function. IEEE Transactions on Geoscience and Remote Sensing 28 (4), 711–720.

Pan, G.C., 1993a. Regionalized favourability theory for information synthesis in mineral exploration. Mathematical Geology 25 (5), 603–631.

Pan, G.C., 1993b. Canonical favourability model for data integration and mineral potential mapping. Computers & Geosciences 19 (8), 1077–1100.

Pan, G.C., Harris, D.P., 1992. Estimating a favourability equation for the integration of geodata and selection of mineral exploration targets. Mathematical Geology 24 (2), 177–202.

Pan, G.C., Harris, D.P., 2000. Information Synthesis for Mineral Exploration. Oxford University Press, New York, 461pp.

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ, 297pp.

Sinclair, A.J., Woodsworth, G.J., 1970. Multiple regression as a method of estimating exploration potential in an area near Terrace, B.C. Economic Geology 65 (8), 998–1003.

Sinding-Larsen, R., Strand, G., 1981. Quantitative integration of mineral exploration data. Episodes 1981 (1), 9–12.

Singer, D.A., Kouda, R., 1988. Integrating spatial and frequency information in the search for Kuroko deposits of the Hokuroku District, Japan. Economic Geology 1988 (83), 18–29.

Singer, D.A., Kouda, R., 1996. Application of a feedforward neural network in the search for Kuroko deposits in the Hokuroku District, Japan. Mathematical Geology 28 (8), 1017–1023.

Singer, D.A., Kouda, R., 1997a. Classification of mineral deposits into types using mineralogy with a probabilistic neural network. Nonrenewable Resources 6 (1), 27–32.

Singer, D.A., Kouda, R., 1997b. Use of a neural network to integrate geoscience information in the classification of mineral deposits and occurrences. In: Gubins, A.G., (Ed.), Proceedings of Exploration '97, Toronto, Canada, pp. 127–134.

Yang, Z., Cheng, Y., Wang, H., 1986. The Geology of China. Clarendon Press, Oxford, 303pp.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353.

Zanakis, S.H., Rustagi, J.S., 1982. Optimization in Statistics. North-Holland Publishing Company, New York, 333pp.

Zimmermann, H.J., 1991. Fuzzy Set Theory and Its Applications. Kluwer Academic Publishers, Boston, 399pp.