Mathematical Geology, Vol. 35, No. 3, April 2003 (© 2003)

Indicator Pattern Combination for Mineral Resource Potential Mapping With the General C-F Model[1]

Yong-Liang Chen[2,3]

As an uncertain reasoning model, the general C-F model was originally developed for processing the uncertainties of rule-based knowledge in the field of artificial intelligence. In this model, certainty factors and combined certainty factors are defined and used for expressing the strengths of knowledge rules and knowledge rule combinations, respectively. The certainty factor can reflect the believable degree of inferring hypothesis on the basis of a proof. Similarly, the combined certainty factor can reflect the believable degree of inferring hypothesis on the basis of the proof combination. It is a function of the related certainty factors and can be determined through combining the certainty factors via the combining rule of the general C-F model. In this paper, the general C-F model has been successfully applied to mineral resource potential mapping. We call this model as the applied form of the general C-F model. In this applied form, the certainty factor is applied to expressing the believable degree of inferring mineral occurrence on the basis of one of the map pattern states associated with the mineral occurrence. Correspondingly, the combined certainty factor is applied to expressing the believable degree of inferring the mineral occurrence on the basis of the map pattern state association. And it is also applied to expressing mineral resource potentials in the mineral resource potential mapping. In the current form, the first step in implementing the general C-F model is to estimate a pair of certainty factors for each map pattern under combination. The next step is to determine the combined certainty factor for the map pattern states coexisting in each locality of the mapping area. The last step is to generate the combined-certainty-factor raster map or the combined-certainty-factor contour map in order to select mineral resource targets. The applied form of the general C-F model is demonstrated on a case study to select mineral resource targets. The experimental results manifest that the model can be compared with the weights-of-evidence model in the effectiveness of mineral resource target selection.

KEY WORDS: uncertain reasoning, knowledge rule, certainty factor, combined certainty factor, mineral resource, target selection.

[1]Received 29 November 2001; accepted 11 December 2002.

[2]Mathematical Geology Research Institute, Jilin University, Geology Palace Building, 6 Ximinzhu Street, Changchun 130026, People's Republic of China; e-mail: wfx_88@163.net.

[3]The Key Lab of Resource Environment and GIS of Beijing, Capital Normal University, Huayuan Bridge, Beijing, 100037, People's Republic of China.

0882-8121/03/0400-0301/1 © 2003 International Association for Mathematical Geology

INTRODUCTION

GISs, served as a kind of high new technological tools, have been widely applied to geo-data management, digital geological mapping, mineral exploration, and mineral resource assessment. During the last decade, many researchers have been developing applied GIS-technologies and polygenetic geo-data digital models for geological applications of GISs. For example, the researchers from the BGS have developed a standard digitized model of geo-map for supporting digital geo-map production in the UK; some Australian researchers have been carrying on the study of digitized geo-data standards since 1991 and have been supported by many mining organizations in Australia; and some Chinese researchers have been developing GIS-based mineral prediction systems since 1995. To date, the majority of geo-data, used for mineral resource assessment, have been digitized. With the help of GIS-technologies, current mineral resource assessment has become faster, timelier, and more accurate than ever before. But some problems, such as polygenetic geo-data representation, geo-feature extraction, and geo-information synthesis, still need further investigation. In the author's opinion, developing statistical models, which can be easily implemented in practice and conveniently realized in GISs, is much more significant to GIS-based mineral resource assessment. In the last three decades, many statistical models have been introduced to quantify the relations between explanatory map patterns and target feature(s) in order to implement mineral resource assessment, such as logistic regression model (Agterberg, 1974; Agterberg, 1989), weights-of-evidence model (Agterberg, 1990; Agterberg, 1992; Agterberg, Bonham-Carter, and Wright, 1990), canonical favorability model (Pan, 1993a), indicator favorability model (Pan, 1993b), and extended weights-of-evidence model (Pan, 1996). A common characteristic of these models is to estimate a function, which best represents the relations between explanatory variables and target variable(s). Some of the above-mentioned models have been widely used in GIS-based mineral resource potential mapping.

In this paper, a new model for mineral resource potential mapping is introduced. The model is derived from the general C-F model (Liu, 2000; Wang, 2000; Wu and Liu, 1995) used in artificial intelligence. Different from the above-mentioned statistical models, it estimates not a statistical function but the certainties of associations between explanatory variables and target variable(s). The model is based on the following fundamental thought.

In an area under study, there exist a series of binary map patterns associated with binary mineral occurrence. The common way of mineral resource assessment is to regard map patterns and mineral occurrence as explanatory variables and target variable(s), respectively, and then estimate a statistical function to represent associations between the two types of variables. We might express associations between explanatory variables and target variable(s) by knowledge rules. In this consideration, we lay out a scheme for mineral resource potential mapping like this: for each map pattern, we may regard its one state and the mineral occurrence as the evidence and hypothesis of a production rule, respectively, and the certainty of association between them then can be measured by a certainty factor. Similarly, the certainty of association between the other state and the mineral occurrence can be measured by another certainty factor. These two certainty factors can reflect the believable degrees of the inferring mineral occurrence on the basis of each state of the map pattern, respectively. The association between a map pattern group and the mineral occurrence may be regarded as the organic combination of associations between each map pattern state constituting the map pattern group and the mineral occurrence. And the certainty of the association can be measured by a combined certainty factor, which is a function of all the certainty factors for the map pattern states constituting the map pattern group. This combined certainty factor reflects the believable degree of inferring the mineral occurrence on the basis of the map pattern group. So it can be served as a kind of mineral resource potential indicators. If only the all certainty factors for the all map pattern states constituting a map pattern group are known, the combining rule used in the general C-F model can be applied to determining the combined certainty factor for the map pattern group on the basis of the known certainty factors.

In mineral resource assessment, the mineral resource potential indicator for each locality in the mapping area needs to be estimated. Applying the general C-F model to mineral resource potential mapping, we may, first, determine two certainty factors for each map pattern under combination. The combined certainty factor for the map pattern group existing in each locality in the mapping area then can be determined on the basis of these certainty factors. Finally, a combined-certainty-factor raster map or a combined-certainty-factor contour map can be generated in GISs.

For each locality in the mapping area, we can determine the map pattern group existing in the locality through overlapping all map layers associated with the mineral occurrence in GISs. According to this map pattern group, we can choose all of the certainty factors for the map pattern states constituting the map pattern group, and then combine all these certainty factors into the combined certainty factor for the map pattern group existing in the locality.

REVIEW ON THE GENERAL C-F MODEL

The general C-F model is the extended form of the C-F model. These two models are derived from the theory of certainty factors, in which the uncertainties of rule-based knowledge are measured by certainty factors. Both models have been successfully applied to quantitative blood disease diagnosis.

Representation of Uncertain Knowledge

In the theory of certainty factors, an uncertain knowledge can be expressed by the following production rule:

$$\text{If} \quad E \quad \text{then} \quad H \quad \text{with} \quad \text{CF}(H, E)$$

here, E and H are the evidence and hypothesis of the knowledge rule, respectively. CF(H, E), called certainty factor or rule strength of the knowledge rule, is the believable degree of inferring hypothesis H on the basis of evidence E. The evidence E may be either simple evidence or a combination of some evidences. The hypothesis H may be either a single hypothesis or a series of hypotheses.

Definition of Certainty Factor

In the C-F model, the strength of a knowledge rule is defined as

$$\text{CF}(H, E) = \text{MB}(H, E) - \text{MD}(H, E)$$

here, CF(H, E) is the certainty factor expressing the rule strength of the above production rule, MB(H, E) and MD(H, E) are called measure belief and measure disbelief, respectively. MB(H, E) represents the belief increase of hypothesis H when evidence E is true; MD(H, E) represents the disbelief increase of hypothesis H when evidence E is true.

Let $p(H)$ and $p(H \mid E)$ denote the prior probability and conditional probability of hypothesis H, respectively. Then MB(H, E) and MD(H, E) can be defined as follows:

$$\text{MB}(H, E) = \begin{cases} 1, & \text{if} \quad p(H) = 1 \\ \dfrac{\max\{p(H|E), p(H)\} - p(H)}{1 - p(H)}, & \text{otherwise} \end{cases}$$

$$\text{MD}(H, E) = \begin{cases} 1, & \text{if} \quad p(H) = 0 \\ \dfrac{\min\{p(H \mid E), p(H)\} - p(H)}{-p(H)}, & \text{otherwise} \end{cases}$$

According to the above-mentioned definitions to the certainty factor, measure belief, and measure disbelief, CF(H, E) can be expressed by

$$\mathrm{CF}(H,E)=\begin{cases}\dfrac{p(H\mid E)-p(H)}{1-p(H)} & \text{if } p(H,E)\geq p(H)\\[2ex] \dfrac{p(H\mid E)-p(H)}{p(H)} & \text{if } p(H,E)<p(H)\end{cases} \tag{1}$$

This form of the certainty factor is used in the C-F model. The form used in the general C-F model is an extended form of this certainty factor. It can be expressed by

$$\mathrm{CF}'(H,E)=\begin{cases}\dfrac{p(H\mid E)-p(H)}{[1-p(H)]p(H\mid E)} & \text{if } p(H,E)\geq p(H)\\[2ex] \dfrac{p(H\mid E)-p(H)}{p(H)[1-p(H\mid E)]} & \text{if } p(H,E)<p(H)\end{cases} \tag{2}$$

The certainty factor $\mathrm{CF}'(H,E)$ can reflect the strength of association between hypothesis H and evidence E. Its range is $[-1,+1]$. When the certainty factor $\mathrm{CF}'(H,E)>0$, it indicates the evidence E supporting the hypothesis H to a certain believable degree; when the certainty factor $\mathrm{CF}'(H,E)<0$, it indicates the evidence E opposing the hypothesis H to a certain believable degree; when the certainty factor $\mathrm{CF}'(H,E)=0$, it indicates the evidence E neither supporting nor opposing the hypothesis H. For simplicity, $\mathrm{CF}'(H,E)$ is still expressed as $\mathrm{CF}(H,E)$ in the following paragraphs.

Combination of Certainty Factors

If there exist several evidences associated with one hypothesis, we may infer the hypothesis on the basis of each of the evidences and determine the certainty factor corresponding to each inference. When all the related certainty factors have been estimated, the knowledge–uncertainty combining rule given in the general C-F model will be applied to combining these certainty factors into a combined certainty factor, which expresses the believable degree of inferring the hypothesis on the basis of all the evidences associated with the hypothesis. The combining rule used in the general C-F model is derived from the counterpart of the C-F model. The following paragraphs will illustrate it:

Suppose that there exist m evidences, $E_1, E_2, \ldots, E_m$, associated with the same hypothesis H. Let us consider the case of $m=2$. If inferring hypothesis H on the basis of E_1 and E_2, we may establish the following production rule:

$$\text{If}\quad E_1\wedge E_2\quad\text{then}\quad H\quad\text{with}\quad \mathrm{CF}(H,E_1\wedge E_2)$$

This production rule is a combination of the following two production rules:

$$\text{If}\quad E_1\quad\text{then}\quad H\quad\text{with}\quad \mathrm{CF}(H,E_1)$$

$$\text{If}\quad E_2\quad\text{then}\quad H\quad\text{with}\quad \mathrm{CF}(H,E_2)$$

here, $\mathrm{CF}(H,E_1)$ and $\mathrm{CF}(H,E_2)$ can be calculated using Eq. (2). $\mathrm{CF}(H,E_1 \wedge E_2)$ is a combined certainty factor, which is a function of $\mathrm{CF}(H,E_1)$ and $\mathrm{CF}(H,E_2)$. According to the knowledge uncertainty combining rule used in the general C-F model, it can be expressed by

$$\mathrm{CF}(H, E_1 \wedge E_2) = \begin{cases} \mathrm{CF}(H, E_1) + \mathrm{CF}(H, E_2) - \mathrm{CF}(H, E_1)\mathrm{CF}(H, E_2), \\ \quad \text{if } \mathrm{CF}(H, E_1) \geq 0 \text{ and } \mathrm{CF}(H, E_2) \geq 0 \\ \dfrac{\mathrm{CF}(H, E_1) + \mathrm{CF}(H, E_2)}{1 - \min(|\,\mathrm{CF}(H, E_1)\,|, |\,\mathrm{CF}(H, E_2)\,|)}, \\ \quad \text{if } \mathrm{CF}(H, E_1) \cdot \mathrm{CF}(H, E_2) < 0 \\ \mathrm{CF}(H, E_1) + \mathrm{CF}(H, E_2) + \mathrm{CF}(H, E_1)\mathrm{CF}(H, E_2) \\ \quad \text{if } \mathrm{CF}(H, E_1) < 0 \text{ and } \mathrm{CF}(H, E_2) < 0 \end{cases} \tag{3}$$

Similarly, we can establish a production rule as follows:

$$\text{If} \quad E_1 \wedge E_2 \wedge \ldots \wedge E_m \quad \text{then} H \quad \text{with} \quad \mathrm{CF}(H, E_1 \wedge E_2 \wedge \ldots \wedge E_m)$$

here, the combined certainty factor $\mathrm{CF}(H, E_1 \wedge E_2 \wedge \ldots \wedge E_m)$ is a function of the following m certainty factors: $\mathrm{CF}(H,E_1)$, $\mathrm{CF}(H,E_2)$, ..., and $\mathrm{CF}(H,E_m)$. To determine it, we need to make similar $m - 1$ combinations. In each combination, two (combined) certainty factors are combined into a combined certainty factor using Eq. (3). The whole combining procedure is like this: first, $\mathrm{CF}(H, E_1)$ and $\mathrm{CF}(H, E_2)$ are combined into $\mathrm{CF}(H, E_1 \wedge E_2)$; and then $\mathrm{CF}(H, E_1 \wedge E_2)$ and $\mathrm{CF}(H, E_3)$ are combined into $\mathrm{CF}(H, E_1 \wedge E_2 \wedge E_3)$; ......; finally, $\mathrm{CF}(H, E_1 \wedge E_2 \wedge \ldots \wedge E_{m-1})$ and $\mathrm{CF}(H, E_m)$ are combined into $\mathrm{CF}(H, E_1 \wedge E_2 \wedge \ldots \wedge E_m)$.

APPLIED FORM OF THE GENERAL C-F MODEL

Suppose that we have m explanatory binary map patterns to be integrated: Z_j, $j = 1, 2, \ldots, m$. Let Y be the binary target variable to be assessed in the mapping area. Y usually is a mineral descriptor, e.g., mineral occurrence. The following forms are common:

$$Z_j(x) = \begin{cases} Z_j^+, & \text{if feature } j \text{ is present at location } x \\ Z_j^-, & \text{otherwise} \end{cases}$$

$$Y(x) = \begin{cases} Y^+, & \text{if the target variable exists at location } x \\ Y^-, & \text{if the target variable does not exist at location } x \end{cases}$$

Obviously, map pattern Z_j has two states: Z_j^+ and Z_j^-. So the m binary map patterns under combination have $2m$ map pattern states in all. In the mapping area, all of these $2m$ map pattern states are associated with mineral occurrence Y^+. Because Z_j^+ and Z_j^- cannot coexist in the same locality, there must coexist m map pattern states in each locality in the mapping area. For one locality, we need to infer mineral occurrence Y^+ on the basis of the m map pattern states coexisting in the locality. The believable degree of this inference is measured by a combined certainty factor, which can be determined through combining the m certainty factors to the m map pattern states coexisting in the locality. In this applied form of the general C-F model, the combined certainty factor with respect to each locality in the mapping area is defined as the mineral resource potential indicator of the corresponding locality.

Certainty Factor Calculation

Consider map pattern $Z_j (j = 1, 2, \ldots, m)$. According to the spatial distribution of map pattern Z_j, the two types of subareas constitute the mapping area: the first denotes those subareas where map pattern Z_j exists; the second indicates those subareas where map pattern Z_j is absent or unvalued. As the two states of map pattern Z_j,Z_j^+ and Z_j^-, appear in the above two types of subareas, respectively. In the first type of subareas, we can infer mineral occurrence Y^+ on the basis of map pattern state Z_j^+; similarly, in the second type of subareas, we can infer mineral occurrence Y^+ on the basis of map pattern state Z_j^-. The believable degrees of these two inferences can be measured by two certainty factors, which can be calculated using Eq. (2). Replacing E and H in Eq. (2) with Z_j and Y^+, respectively, we can obtain the following applied form of Eq. (2):

$$\mathrm{CF}(Y^+, Z_j) = \begin{cases} \dfrac{p(Y^+ \mid Z_j) - p(Y^+)}{[1 - p(Y^+)]p(Y^+ \mid Z_j)} & \text{if } p(Y^+ \mid Z_j) \geq p(Y^+) \\ \dfrac{p(Y^+ \mid Z_j) - p(Y^+)}{p(Y^+)[1 - p(Y^+ \mid Z_j)]} & \text{if } p(Y^+ \mid Z_j) < p(Y^+) \end{cases} \quad (j = 1, 2, \ldots, m) \tag{4}$$

here, $\mathrm{CF}(Y^+, Z_j)$ is a certainty factor; $p(Y^+)$ and $p(Y^+ \mid Z_j)$ are the prior probability and conditional probability of mineral occurrence Y^+, respectively. In Eq. (4), Z_j may be equal to Z_j^+ with respect to the first type of subareas or Z_j^- with respect to the second type of subareas. Therefore, conditional probability $p(Y^+ \mid Z_j)$ can be written more explicitly as $p(Y^+ \mid Z_j^+)$ for the first type of subareas or $p(Y^+ \mid Z_j^-)$ for the second type of subareas, and the corresponding certainty factor $\mathrm{CF}(Y^+, Z_j)$ can be written more explicitly as $\mathrm{CF}(Y^+, Z_j^+)$ or $\mathrm{CF}(Y^+, Z_j^-)$.

In practice, $p(Y^+)$, $p(Y^+ \mid Z_j^+)$, and $p(Y^+ \mid Z_j^-)$ can be estimated from the samples of the studied area. Substituting expressions $p(Y^+)$ and $p(Y^+ \mid Z_j^+)$ into Eq. (4), we can obtain one certainty factor $\mathrm{CF}(Y^+, Z_j^+)$; and substituting expressions $p(Y^+)$ and $p(Y^+ \mid Z_j^-)$ into Eq. (4), we can obtain another certainty factor $\mathrm{CF}(Y^+, Z_j^-)$. Likewise, each pair of certainty factors for each of the m map patterns under combination can be obtained.

Generation of the Combined Certainty Factor Raster Map

Consider the case for $m = 2$. Not losing generality, we might consider two map patterns Z_j and Z_k ($j \neq k$, j, $k = 1, 2, \ldots, m$). In this case, four types of subareas constitute the studied area. These four types of subareas sequentially correspond to the four types of map pattern state combinations: $Z_j^+ \wedge Z_k^+$, $Z_j^+ \wedge Z_k^-$, $Z_j^- \wedge Z_k^+$, and $Z_j^- \wedge Z_k^-$. In each of these four types of subareas, we can infer mineral occurrence Y^+ on the basis of the corresponding map pattern state combination. The believable degrees of these four inferences can be measured respectively by the four combined certainty factors, $\mathrm{CF}(Y^+, Z_j^+ \wedge Z_k^+)$, $\mathrm{CF}(Y^+, Z_j^+ \wedge Z_k^-)$, $\mathrm{CF}(Y^+, Z_j^- \wedge Z_k^+)$, and $\mathrm{CF}(Y^+, Z_j^- \wedge Z_k^-)$.

The combined certainty factor to a two-map-pattern-state combination can be determined by combining the two certainty factors to the two combined map pattern states via the combining rule expressed in Eq. (3). Replacing E_1,E_2, and H in Eq. (3) with Z_j,Z_k, and Y^+, respectively, we can transform Eq. (3) into the following applied form:

$$\mathrm{CF}(Y^+, Z_j^\wedge Z_k) = \begin{cases} \mathrm{CF}(Y^+, Z_j) + \mathrm{CF}(Y^+, Z_k) - \mathrm{CF}(Y^+, Z_j)\mathrm{CF}(Y^+, Z_k), \\ \quad \text{if } \mathrm{CF}(Y^+, Z_j) \geq 0 \text{ and } \mathrm{CF}(Y^+, Z_k) \geq 0 \\ \dfrac{\mathrm{CF}(Y^+, Z_j) + \mathrm{CF}(Y^+, Z_k)}{1 - \min(|\,\mathrm{CF}(Y^+, Z_j)\,|, |\,\mathrm{CF}(Y^+, Z_k)\,|)}, \\ \quad \text{if } \mathrm{CF}(Y^+, Z_j) \bullet \mathrm{CF}(Y^+, Z_k) < 0 \\ \mathrm{CF}(Y^+, Z_j) + \mathrm{CF}(Y^+, Z_k) + \mathrm{CF}(Y^+, Z_j)\mathrm{CF}(Y^+, Z_k) \\ \quad \text{if } \mathrm{CF}(Y^+, Z_j) < 0 \text{ and } \mathrm{CF}(Y^+, Z_k) < 0 \end{cases} \tag{5}$$

$$(j \neq k;\, j, k = 1, 2, \ldots, m)$$

here, $\mathrm{CF}(Y^+, Z_j \wedge Z_k)$ is a combined certainty factor; $\mathrm{CF}(Y^+, Z_j)$ and $\mathrm{CF}(Y^+,Z_k)$ are the two certainty factors to the two map pattern states under combination. In Eq. (5), Z_j may be equal to Z_j^+or Z_j^+, and Z_k may be equal to Z_k^+ or Z_k^+. Therefore, $\mathrm{CF}(Y^+, Z_j)$ can be written more explicitly as $\mathrm{CF}(Y^+, Z_j^+)$ or $\mathrm{CF}(Y^+, Z_j^-)$; $\mathrm{CF}(Y^+, Z_k)$ can be written more explicitly as $\mathrm{CF}(Y^+, Z_k^+)$ or $\mathrm{CF}(Y^+, Z_k^+)$; and the corresponding combined certainty factor $\mathrm{CF}(Y^+, Z_j \wedge Z_k)$ can be written more explicitly as CF(Y^+, $Z_j^+ \wedge Z_k^+$), CF(Y^+, $Z_j^+ \wedge Z_k^-$), CF(Y^+, $Z_j^- \wedge Z_k^+$), or CF(Y^+, $Z_j^- \wedge Z_k^-$). These four combined certainty factors correspond to the four types of map pattern state combinations, respectively.

When there exist m map patterns under combination, theoretically, 2^m possible types of subareas may constitute the studied area. Our task is to determine the combined certainty factor for each type of subareas on the basis of the m certainty factors to the m map pattern states coexisting in each type of subaresas. In most cases, this is not easy because it is difficult to determine the boundaries and areas of each type of subareas. However, the following two-step procedure for calculating all combined certainty factors is practical and effective:

a. Divide the map area into small regular-grid cells, whose size is sufficiently small so that one cell usually contains only one mineral deposit. The prior probability and m pairs of conditional probabilities of mineral occurrence Y^+ can be estimated from this cell set and then m pairs of certainty factors can be calculated using Eq. (4).
b. Determine the map pattern states coexisting in the first grid cell, and then combine the m certainty factors for these m map pattern states into a combined certainty factor using Eq. (5). Do the same with all the other grid cells; finally the combined certainty factor for each grid cell can be obtained.

The above two-step procedure can effectively avoid determining the boundaries and areas of polygons, which are generated by combining all map patterns associated with the mineral occurrence. What's more, it can generate the grid cells, which constitute the sample set used for the estimation of the prior probability and conditional probabilities of the mineral occurrence.

Estimation of the Prior Probability and Conditional Probabilities

Suppose that there are n small-grid cells in the mapping area, and that among them $n(Y^+)$ cells contain mineral deposits. Then the prior probability of mineral occurrence Y^+ can be estimated by

$$p(Y^+) = \frac{n(Y^+)}{n} \tag{6}$$

Form a two-way contingency table for Z_j ($j = 1, 2, \ldots, m$) and Y. The elements in the table are the frequencies of joint occurrence for different state combinations of Z_j ($j = 1, 2, \ldots, m$) and Y (Zhou and Xia, 1993). The conditional probabilities of mineral occurrence Y^+ can then be estimated:

$$p(Y^+ \mid Z_j^+) = \frac{n(Z_j^+ Y^+)}{n(Z_j^+)} \quad j = 1, 2, \ldots, m \tag{7a}$$

$$p(Y^+ \mid Z_j^-) = \frac{n(Z_j^- Y^+)}{n(Z_j^-)} \quad j = 1,\ 2, \ldots, m \tag{7b}$$

here, $n(Z_j^+ Y^+)$ is the frequencies that Z_j^+ and Y^+ occur concurrently; $n(Z_j^- Y^+)$ is the frequencies that Z_j^- and Y^+ occur concurrently; $n(Z_j^+)$ is the frequencies that Z_j^+ occurs; and $n(Z_j^-)$ is the frequencies that Z_j^- occurs.

CASE STUDY

A case study is given in this section to demonstrate the use of the applied form of the general C-F model described above. The method was applied to a Pb–Zn polymetallographic province in northern Xinjiang Uygur Autonomous Region, China. In this metallographic province, several dozens of Pb–Zn polymetallic deposits have been found. In order to delineate Pb–Zn polymetallic potential targets for the further mineral exploration in this metallographic province, both the weights-of-evidence model and the applied form of the general C-F model are applied to the Pb–Zn polymetallic potential target selection. The two mineral resource potential maps generated by the two models are compared with each other.

The Regional Ore-Controlling Factors

In the studied area, regional ore-controlling factors mainly include basement complexes, intrusive rocks, the Devonian system, the Carboniferous system, regional structures, geochemical anomalies, and regional mineralized rocks.

The basement complexes, inferred on the basis of the magnetic and gravitational data, are widely distributed in the studied area. In the depth, they are continuously distributed; but on the surface, they are locally distributed here and there because other geological bodies cover on them. They control the distributions of both those geological bodies and most Pb–Zn polymetallic deposit clusters. They provide good metallogenetic surroundings for the regional Pb–Zn poly-metallization in the studied area.

Intrusive rocks are widely distributed in the studied area. They control the distribution of the most known Pb–Zn polymetallic deposits. Some of them constitute the ore-forming host rocks and provide much mineral substance for the Pb–Zn poly-metallization. The others are not obviously related to the Pb–Zn poly-metallization, but they may provide thermal energy that can accelerate the migration of mineralizing elements because most geochemical anomalies are distributed around them. According to the statistical data, intersections between intrusive rocks and regional structures are usually the most favorable localities for the Pb–Zn poly-metallization in the studied area.

The Devonian system is distributed in the Paleozoic volcanic–sedimentary basins. The rocks of the Devonian period can be classified into two types: one is sedimentary rock and the other is volcanic sedimentary rock. The former is little related to Pb–Zn poly-metallization, but the latter is closely related to the Pb–Zn poly-metallization. Some of the volcanic sedimentary rocks are the adjacent rocks of Pb–Zn polymetallic deposits. They can be served as the indicative stratum for the Pb–Zn polymetallic deposit prospecting in the studied area.

Compared to the Devonian system, the distribution of the Carboniferous system is more limited. The rocks of the Carboniferous period are mainly sedimentary rocks, which are somewhat related to regional Pb–Zn poly-metallization. They control the distribution of some small Pb–Zn polymetallic deposits in the studied area.

Regional structures are widely distributed in the studied area. They control the distribution of most geological bodies, especially intrusive bodies. Some of them directly control the distribution of Pb–Zn polymetallic deposits. They provide good tectonic surroundings for regional Pb–Zn poly-metallization.

Geochemical anomalies are widely distributed in the studied area. Anomalous elements mainly include Au, Ag, Cu, Pb, and Zn. The geochemical anomalies of these elements usually constitute the anomalous ring surrounding some large intrusive bodies. They can directly reflect the regional Pb–Zn poly-metallization around the intrusive bodies. Nearly all the known Pb–Zn polymetallic deposits are located in the geochemical anomalies in the studied area.

Besides the above-mentioned ore-controlling factors, mineralized rock is also an important regional ore-controlling factor in the studied area. They are widely distributed along the margins of large intrusive bodies in the studied area. It reflects that Pb–Zn poly-metallization is closely related to regional magmatism.

Map Pattern Preparation

On the basis of the above-mentioned regional ore-controlling factors, the following map patterns for Pb–Zn polymetallic target selection are predrawn in MapInfo platform:

a. The distributive map of the minimum river basins with gold anomalies (Fig. 1);
b. The distributive map of the minimum river basins with silver anomalies (Fig. 2);
c. The distributive map of the minimum river basins with copper anomalies (Fig. 3);

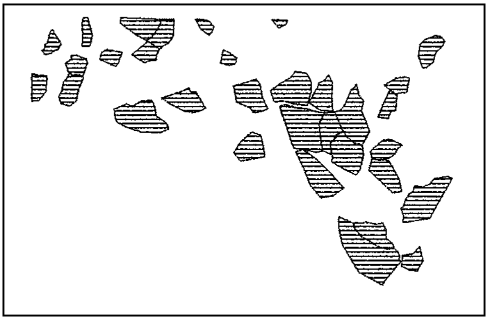

Figure 1. The distributive map of the minimum river basins with gold anomalies.

d. The distributive map of the minimum river basins with lead anomalies (Fig. 4);
e. The distributive map of the minimum river basins with zinc anomalies (Fig. 5);
f. The distributive map of regional Pb–Zn poly-metallization (Fig. 6);
g. The distributive map of regional linear structures (Fig. 7);
h. The distributive map of regional granites (Fig. 8);
i. The distributive map of regional diorites (Fig. 9);
j. The distributive map of the early Devonian system (Fig. 10);
k. The distributive map of the middle Devonian system (Fig. 11);
l. The distributive map of the early Carboniferous system (Fig. 12);
m. The distributive map of the known Pb–Zn polymetallic deposits (Fig. 13).

Besides the above map patterns (layers), a special map layer called statistical unit map layer should be generated in the GIS. It can be formed through

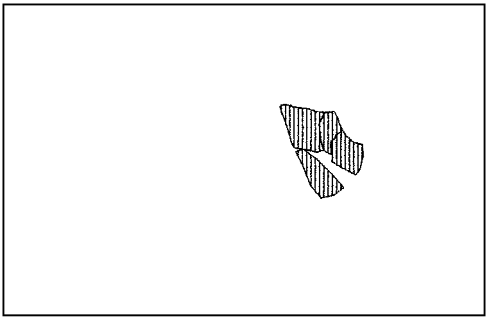

Figure 2. The distributive map of the minimum river basins with silver anomalies.

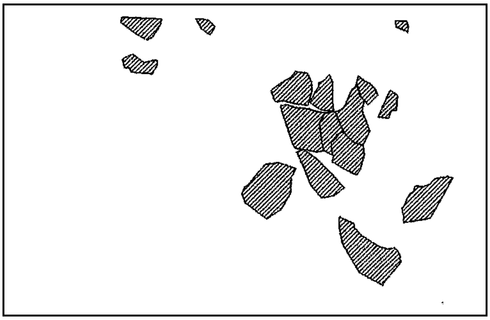

Figure 3. The distributive map of the minimum river basins with copper anomalies.

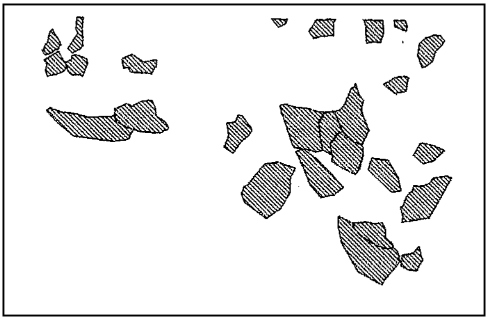

Figure 4. The distributive map of the minimum river basins with lead anomalies.

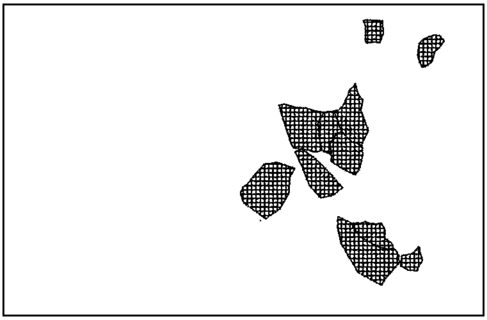

Figure 5. The distributive map of the minimum river basins with zinc anomalies.

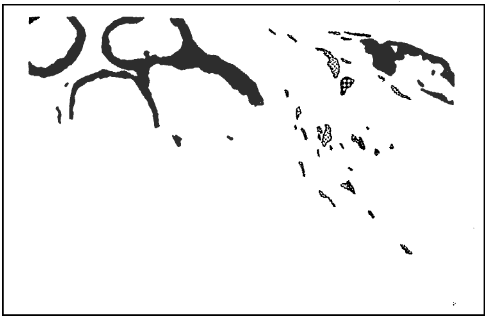

Figure 6. The distributive map of regional Pb–Zn poly-metallization.

Figure 7. The distributive map of regional linear structures.

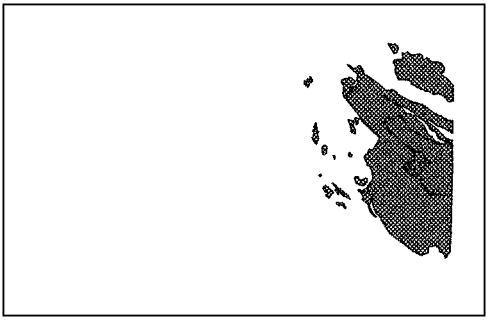

Figure 8. The distributive map of regional granites.

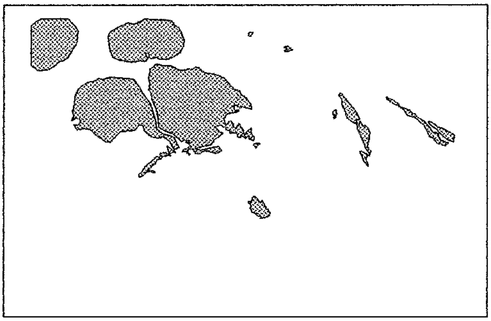

Figure 9. The distributive map of regional diorites.

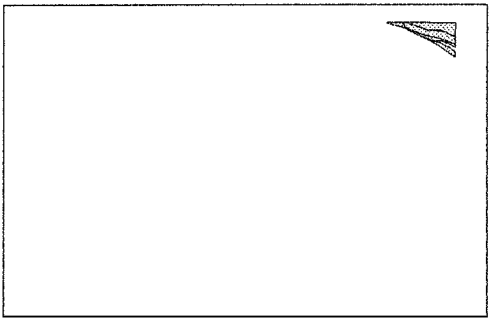

Figure 10. The distributive map of the early Devonian system.

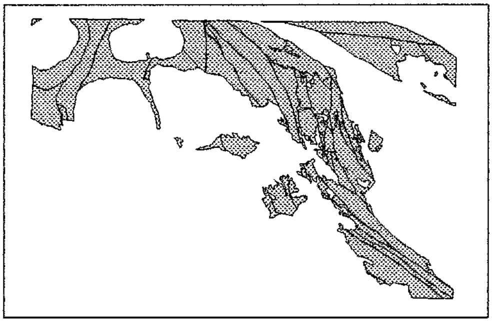

Figure 11. The distributive map of the middle Devonian system.

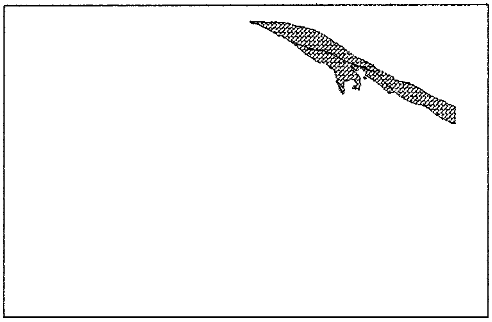

Figure 12. The distributive map of the early Carboniferous system.

automatically generating a certain number of small grid cells in the GIS. In this case study, 3800 small grid cells are generated. They are distributed in 50 scanning lines, and 76 small grid cells are located in each scanning line. The grid-cell (statistical unit) distributive map of the studied area is shown in Figure 14.

Attributive Table Construction and Certainty Factor Calculation

The attributive table of the statistical unit map layer can be generated as follows. First, we may define an attributive table with one column, in which grid cell numbers are saved. Then, one column is added into the table and the statistical unit map layer is overlapped with the mineral deposit map layer; the result for this overlapping is saved in the increased column. Next, 12 columns are added into the table and the statistical unit map layer is overlapped with each explanatory map layer sequentially; the result for each overlapping is saved in each of the

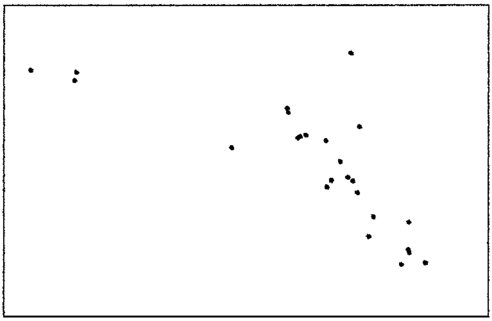

Figure 13. The distributive map of the known Pb–Zn polymetallic deposits.

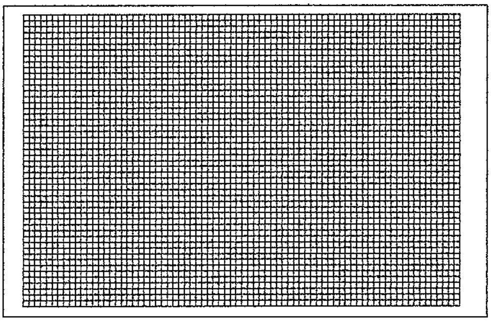

Figure 14. The distributive map of the grid cells generated in MapInfo.

12 increased columns. In this case study, the number of grid cells is 3800, so an attributive table with 3800 records and 14 columns is generated in the GIS.

In order to express the spatial distributive information of the map patterns under combination in the attributive table, value 1 or 0 needs to designate the corresponding column of the attributive table. The method for this value designation is as follows. The statistical unit map layer is overlapped with the mineral deposit map layer. For each grid cell, if there exists any known Pb–Zn polymetallic deposit in it, the second column of the corresponding record of the attributive table is valued by 1, otherwise by 0. Similarly, the statistical unit map layer is overlapped with each explanatory map layer. For each grid cell, if its center is located in one entity of the overlapped map layer, the corresponding column of the corresponding record of the attributive table is valued by 1, otherwise by 0.

According to the data in the attributive table, we can calculate both the prior probability of mineral occurrence Y^+ using Eq. (6) and the conditional probabilities of mineral occurrence Y^+ using Eqs. (7a) or (7b). Then the certainty factor for each map pattern state can be calculated via Eq. (4). Similarly, the two weights, W_j^+ and W_j^-, can be calculated on the basis of the prior probability and conditional probabilities of mineral occurrences Y^+ and Y^-. Both the two weights and two certainty factors for each map pattern are listed in Table 1.

Certainty Factor Combination and Mineral Target Selection

One grid cell in the statistical unit map layer corresponds to one record of the attributive table. According to the values of each record of the attributive table, we can determine the map pattern states coexisting in the corresponding grid cell, and then using Eq. (5) we can combine all of the certainty factors for these map pattern states into a combined certainty factor, which is served as the mineral resource potential indicator of the grid cell. Such being the case, the combined certainty

Table 1. Estimated Two Certainty Factors and Two Weights for Each Map Pattern

Regional ore-controlling factors	CF^+	CF^-	W^+	W^-
Gold anomaly river basin	0.736747	−0.489215	1.334639	−0.671806
Silver anomaly river basin	0.902568	−0.283089	2.328602	−0.332804
Copper anomaly river basin	0.761901	−0.369514	1.435069	−0.461264
Lead anomaly river basin	0.749863	−0.493625	1.385745	−0.680478
Zinc anomaly river basin	0.813662	−0.384954	1.680191	−0.486058
Regional poly-metallization	0.707705	−0.129671	1.229990	−0.138884
Regional linear structure	0.657635	−0.297144	1.071879	−0.352603
Regional granite	−0.219870	2.684166E-02	−0.248295	2.720848E-02
Regional diorite	−0.646844	8.325511E-02	−1.040846	8.692604E-02
The early Devonian system	−1.000000	7.678051E-03	0.000000	0.000000
The middle Devonian system	0.713794	−0.772237	1.251044	−1.479448
The middle Carbonic System	0.354514	−1.585873E-02	0.437752	−1.598582E-02

factor for every grid cell in the statistical unit map layer can be determined. When the combined certainty factors for all the grid cells in the statistical unit map layer have been determined, a combined-certainty-factor raster map can be generated, and the corresponding combined-certainty-factor contour map can be drawn. For this case study, the combined-certainty-factor contour map of the studied area is shown in Figure 15, and the ore-forming posterior probability contour map is shown in Figure 16. In these two contour maps, the corridors between each pair of adjacent contour lines are filed with different black colors from gray–white to absolutely black. The corridors with bigger combined certainty factors are filled with relatively black colors, so those subareas with the absolutely black color in the two contour maps are the most favorite target areas for Pb–Zn polymetallic

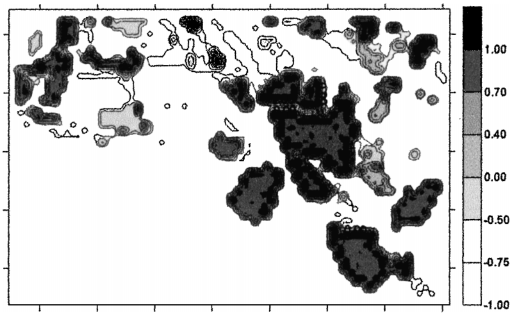

Figure 15. Combined-certainty-factor contour map. (The small lighty shaded circles within the dark areas are not deposits but the grid cells with comparitively lower combined certainty factor values)

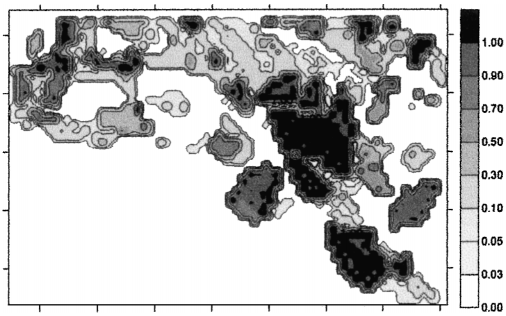

Figure 16. Ore-forming posterior probability contour map. (The small lighty shaded circles within the dark areas are not deposits but the grid cells with comparitively lower posterior values).

deposit prospecting. From these two contour maps, we can obtain the following results:

- The contour lines in these two maps have similar shapes. And the subareas with the absolutely black color in one contour map correspond to the counterparts in the other map.
- Both of the two contour maps can show that there exist several subareas with relatively black colors in the northeastern corner of the studied area. To date, Pb–Zn polymetallic deposits have not been found in these subareas, so these subareas are selected as Pb–Zn polymetallic potential targets. In future mineral prospecting, we should pay more attention to these subareas.
- Imposing the known mineral deposit map layer onto the two contour maps, we can see that nearly all of the known Pb–Zn polymetallic deposits in the studied area are located in the subareas with relatively black colors. Therefore, the two statistical models established in the studied area are reasonable and the Pb–Zn polymetallic potential targets selected by the two models are believable.

CONCLUSION

In GIS-based mineral resource potential mapping, the common practice is to estimate a statistical function for representing the relation between explanatory map patterns and target feature(s). Before applying this function, it is usually needed to test the believable degree of the function in order to determine whether the function is significant or not. For example, a statistical test must be done before applying a regression equation to prediction. Besides this, some statistical models require explanatory variables satisfying certain statistical distribution.

For example, the majority of multivariate analysis methods, such as the cluster analysis, the discriminant analysis, and the factor analysis, require explanatory variables satisfying normal distribution, and both the weights-of-evidence model and the extended weights-of-evidence model require explanatory map patterns satisfying conditional independence. Therefore, before applying statistical models to mineral resource potential mapping, we must carry on some preprocessing work, such as statistical tests, considering application limitations of the method. But the new method described in this paper requires no restriction to explanatory map patterns and needs no statistical test before application. Applying this new model, we directly estimate the certainties of associations between explanatory map patterns and the mineral occurrence instead of estimating a statistical function. Compared with other statistical models, this new model can effectively decrease the preprocessing workload.

Compared with the weights-of-evidence model, this new model has the following similar characteristics:

The two certainty factors, CF_j^+ and CF_j^-, defined in the new model are much similar to the two weights, W_j^+ and W_j^-, defined in the weights-of-evidence model. From Table 1, we can see that the two certainty factors, CF_j^+ and CF_j^-, can be compared with the two weights, W_j^+ and W_j^-, respectively. If the positive weight W_j^+ is not equal to zero, CF_j^+ and W_j^+ always have the same sign; similarly, if the negative weight W_j^- is not equal to zero, CF_j^- and W_j^- always have the same sign.

The strength of association between an explanatory map pattern and the target variable can be measured by the contrast:

$$\mathrm{CC}_j = \mathrm{CF}_j^+ - \mathrm{CF}_j^- \tag{8}$$

Large contrast values imply strong associations between Z_j and Y; small contrast values indicate the opposite. CC_j can be positive or negative, which indicates positive or negative associations between Z_j and Y, respectively. CF_j^+ and CF_j^- always take opposite signs. If CF_j^+ is positive, the association between Z_j and Y is positive, and *vice versa*. The contrast CC_j defined here is similar to the contrast C_j defined by Agertberg, Bonham-Carter, and Wright (1990).

Estimating the two certainty factors for each map pattern is much similar to calculating the two weights defined in the weights-of-evidence model. Like the two weights, W_j^+ and W_j^-, the two certainty factors, CF_j^+ and CF_j^-, are the functions of the prior probability and conditional probabilities of the mineral occurrence. Therefore, physical significances of CF_j^+ and CF_j^- can be compared with the counterparts of W_j^+ and W_j^-, respectively.

Determining the combined certainty factors is a key step in implementing the new model described here. However, it is totally different from calculating the log posterior odds defined in the weights-of-evidence model. It is a step-by-step procedure, during which we need repeatedly use the combining rule given in the general C-F model to combine two (combined) certainty factors until all certainty factors under combination have been combined.

From the case study, we can see that the general C-F model and the weights-of-evidence model can be compared with each other in the effectiveness of mineral resource target selection.

ACKNOWLEDGMENTS

The author thanks Prof Suyu Wang for her constructive suggestions, which improved the paper. The author also thanks his tutor Prof Dayou Liu for his generous helps with his postdoctoral research. The research here is supported by the Open Laboratory of Research on Nontraditional Mineral Resources.

REFERENCES

Agterberg, F. P., 1974, Automatic contouring of geological maps to detect target areas for mineral exploration: Math. Geol., v. 6, no. 4, p. 373–395.

Agterberg, F. P., 1989, LOGDIA-FORTRAN 77 program for logistic regression with diagnostics: Comput. Geosci., v. 15, no. 4, p. 599–614.

Agterberg, F. P., 1990, Combining indicator patterns for mineral resource evaluation, *in* China University of Geosciences (eds.), Proceedings of international workshop on statistical prediction of mineral resources, Vol. 1: Wuhan, China, p. 1–15.

Agterberg, F. P., 1992, Combining indicator patterns in weights-of-evidence modeling for resource evaluation: Nonrenewable Res., v. 1, no. 1, p. 39–50.

Agterberg, F. P., Bonham-Carter, G. F., and Wright, D. F., 1990, Statistical pattern integration for mineral exploration, *in* Gaal G. and Merriam, D. F. (eds.), Computer applications for mineral exploration in resource exploration: Pergamon Press, Oxford, p. 1–21.

Liu, D. Y., 2000, Numerical methods for treating uncertainty and vagueness in knowledge-based systems: Jilin University Press, Changchun, 280 p. (In Chinese).

Pan, G. C., 1993a, Canonical favorability model for data integration and mineral potential mapping, Comput. Geosci., v. 19, no. 8, p. 1077–1100.

Pan, G. C., 1993b, Indicator favorability theory for mineral potential mapping, Nonrenewable Resour., v. 2, no. 4, p. 292–311.

Pan, G. C., 1996, Extended weights-of-evidence modeling for the pseudo-estimation of metal grades, Nonrenewable Resour., v. 5, no. 1, p. 53–76.

Wu, Q. Y., and Liu, J. N., 1995, Artificial intelligence and expert systems: University of National Defence Sciences and Technologies Press, Changsha, 260. p. (In Chinese).

Wang, W. S., 2000, Principles and applications of artificial intelligence: Publishing House of Electronic Industry, Beijing, 291. p. (In Chinese).

Zhou, G. Y., and Xia, L. X., 1993, Nonquantitative data analysis and its application: Science Press, Beijing, 295. p. (In Chinese).