Journal of
Applied Sciences

Regional Hydrothermal Copper Prediction Using Weight of Evidence Model in Northwestern Yunnan

Jacob Kporfor and Cheng Jiaguo

Institute of Mathematical Geology and Remote Sensing, China University of Geosciences,
Wuhan 430074, People's Republic of China

Abstract: Increasing cost of mineral exploration ventures over the time necessitates considerations of other innovations especially in delineating potential zones at the reconnaissance stage of mineral assessment. This study demonstrated the application of Weight of evidence model in delineating a very large and inaccessible area of 69052 sq.km in the northwestern part of Yunnan Province. The geological data sets compiled for the study includes, digital database of the following: geochemical stream sediment sampling results, lineament extracted from four landsat images, geological periods of the lithologies and mineral occurrences. The following preprocessing were meted to the data before weighting. Factor analysis was applied to the geochemical data, briefly illustrated, the geological period were reclassified and the lineaments were put into grid and the lineament density obtained and contoured before weighting. The selected three suitable diagnostic criteria; lineament, lithological ages and geochemical data sets were integrated using Weight-of-evidence model on 27 known hydrothermal deposits point to prepare the favorability map. The result of the prediction indicated a successful implementation of the model using the three recognition criteria above and also suggested new areas for a follow up exploration ventures.

Key words: Weight of evidence model, geographic information system, mineral exploration, conditional probability, favourability map

INTRODUCTION

The purpose of this study was to demonstrate the use of weight of evidence model as a choice in delineating potential areas in a very vast area of 69052.1 sq.km in the northwestern Yunnan Province. The main objectives were to evaluate effectiveness of the model in predicting regional hydrothermal copper deposit in the study area; to determine and evaluate the spatial associations between the copper bearing occurrences and variety of regional-scale geoscientific data; to determine the first order diagnostic recognition criteria for recognition the mentioned deposit type. The area is renowned in its mineral endowment, but the complexity of the topography makes it difficult to access. The use of quantitative mathematical methods is becoming very popular in solving these problems[1,2] The technique originally developed in a non-spatial context for medical[2] diagnosis and has been modified by Bonham-Cater *et al.*[1] to deal with spatial prediction-diagnosing mineral deposits using the symptoms of various geological phenomena.

Since the early 1990's computer-based Geographic Information Systems (GIS) applications has become an integral part of many mineral resource exploration programs[3]. A GIS is an integrated system of hardware, software and methodologies for the management of spatial (georeferenced) data. It facilitates data compilation and synthesis and permits exploratory analysis and modeling Evaluation of geosciences data with GIS can provide support for various geological investigations and aid decision-making processes, such as determining successful and cost-effective exploration or management strategies.

Three geological data set have been integrated using the above model to predict potential areas for hydrothermal copper prospects. The study reviews briefly on the regional geological setting and mineralization, applied factor analysis to the geochemical data to select only area associated with hydrothermal deposit through the minerals associations to avoid including area of different genetic types of copper mineralization. Flavorability has been produced indicating new areas for further prospecting.

Corresponding Author: Jacob Kporfor, Institute of Mathematical Geology and Remote Sensing,
China University of Geosciences, Wuhan 430074, People's Republic of China

WEIGHT OF EVIDENCE MODEL

The model is based on a probability framework on a set of theory[2] in which case the idea of posterior probability is emphasized. Considering mineral resource assessment, lets assume that a hypothesis (H) about the probability of mineral deposits of a series type D of a study area is related to a set of evidences $X(X=\{ X_1, X_2...X_n\})$; and $X_1, X_2...Xn$ which are conditionally independent with each other.

The probability of occurrence of mineral of type D without the consideration of any known evidence information is the prior probability (P{D}). The probability of occurrence of mineral deposit of type D given the condition of existing X in the study area is the posterior probability (P{D/X}).

The odd probability (O{D}) is the ratio of the probability that an evidence occur P{D}; to the probability that it does not occur, such as P{ $\overline{D}$ }. Expressed as:

$$O\{D\} = \frac{P\{D\}}{P\{\overline{D}\}} \qquad (1)$$

Based on the hypothesis H on D with given X, its odd is:

$$O\{D/X\} = O\{D\}\frac{P\{D/X\}}{P\{\overline{D}/X\}} \qquad (2)$$

In ordinary weights of evidence model X_1 in the X need be divided into subsets, which can be expressed in terms of set of theory.
A (presence or true) and $\overline{A}$ (absence or false). The two subsets have following purposes.

$$A\cup \overline{A} = X_i \text{ (A and } \overline{A} \text{ make up } X_i) \quad i =1,2,.....n \qquad (3)$$
$$A\cap \overline{A} = 0 \qquad \text{(A and } \overline{A} \text{ are mutually exclusive)} \qquad (4)$$

With the definitions and the properties of binary pattern outlined above, the following equations can be deduced:

$$P\{D/X\} = P\{D\}\frac{\{X/D\}}{P\{X\}}$$
$$= P\{D\}\frac{P\{X_1/D\}}{P\{X_1\}}\frac{P\{X_2/D\}}{P\{X_2\}}............\frac{P\{Xn/D\}}{P\{Xn\}} \qquad (5)$$

$$O\{D/X\} = O\{D\}\frac{P\{X/D\}}{P\{X\overline{D}\}}$$
$$= O\{D\}\frac{P\{X_1/D\}}{P\{X_1\overline{D}\}}\frac{P\{X_2/D\}}{P\{X_2/\overline{D}\}}............\frac{P\{X_n/D\}}{P\{X_n/\overline{D}\}} \qquad (6)$$

$$Log\{O\{D/X\}\} = Log\{O\{D\}\}+Log\frac{P\{X/D\}}{X/\overline{D}} = Log\{O\{D\}\}+$$
$$Log\{\frac{P\{X_1/D\}}{P\{X_1/\overline{D}\}}\}+Log\{\frac{P\{X_2/D\}}{P\{X_2/\overline{D}\}}\}+.....Log\{\frac{P\{X_n/D\}}{P\{X_n/\overline{D}\}}\} \qquad (7)$$

If X_i is present:

$$\frac{P\{X_i/D\}}{P\{X_i/\overline{D}\}}$$

is called sufficient ratio (LS). Otherwise, it is called necessity ratio (LN). Corresponding to the natural logarithm of LS is the weight of positive weight of evidence(W^+) and the natural logarithm of LN is the negative weight of evidence (W^-). For convenience let W_o denote the natural logarithm of O{D} .

Thus the Eq. (6) and (7) can be written as

$$O\{D/X\} =O\{D\}LS_1 \text{ (or } LN_1)\, LS_2(\text{or } LN_2)....LSn \text{ (or } LNn) \qquad (8)$$

$$Log\{O\{D/X\}\} = W_0+W_1{}^++(\text{or } W_1{}^-)+W_2{}^++(\text{or } W_2{}^-)...\,....+ W_n{}^+$$
$$(\text{or } W_n{}^-) = \sum_0^n W_i \qquad (9)$$

The posterior probability P{D/X} can be converted from the equation above as follows:

$$P\{D/X\} = \frac{\exp(\text{logit}\{O\{D/X\}\})}{1+\exp(\text{logit}\{O\{D/X\}\})} \qquad (10)$$

The values of P{D/X} calculated with using Eq. (7) are identical to those calculated directly using Eq. (5).

The advantage of using the conditional probability expressions (5 or 8) is that the weights are easier to be interpreted than the probability factors.

In most applications of ordinary Weight of evidence model, the contrast C is defined as Eq. (1.11) and used to select the cut-off to divide continuous variables into binary patterns

$$C = W_i{}^+\text{-}W_i{}^-, I = 1,2...n \qquad (11)$$

In general, absolute weights values between 0 and 0.5 are mildly predictive; values 0.5 and 1 are moderately predictive ;values between 1 and 2 are strongly predictive and greater than 2 are extremely predictive

Geological setting: Yunnan Province is located in the Himalayan fold belt. It is characterized by complex geological structures, perfectly developed strata, frequent magmatic activities and various degree of metamorphism. The province consists of dominant basaltic lavas and subordinate pyroclastic rocks an example the Permian Emeishan Basalt Formation that is one of Earth's LIPs (large igneous provinces). The Emeishan volcanic successions unconformably overlie the late Middle

Pemian carbonate formation (i.e. the Maokou limestone) and are in turn covered by the upper Triassic or Jurrassic sediment in the central part. The early Permian thick carbonate sequence began to deposit as response to extensive transgression and basin subsidence. The Yangtse-craton was separated by an unconformity which formed as a response of the Dongwu movement. It consists of Proterozoic metamorphic rocks overlain by a thick Cambrian-Triassic carbonate and shale sequence. In the Devonian, the SW margin of the craton underwent extensional faulting, deposition of a thick sequence of Devonian-Triassic marine sedimentary rocks and in a few places eruption of basalts. These rocks were uplifted and folded during the Early Jurassic to Latest Cretaceous (190-65 Ma) Yanshanian orogeny, which formed as a response to subduction of the Pacific Plate under the Eurasian plate movement. The process of splitting and amalgamation took place from time to time in this region, which resulted in the formation of three main ore-bearing strata. Three main metallogenetic models can be described as three types of deposits, magmatic, metamorphic and tectonic mobile zone.

Mineralization: The known copper, gold, lead-zinc, silver and molybdenum occurrence in the study area occur in several forms, hydrothermal, epithermal, hypothermal, mesothermal, magmatic, volcano-sedimentary, strata-bound, skarn and saline lake deposits. The host rocks in most cases are shale, carbonates, sandstone, conglomerate and partially altered andesine. The principal alteration processes are decalcification, silification, sericitization, sulfurdation, epidotization and chloritization.

Porphyry-type the most common type in the area occurs when large amounts of hot water carrying small amounts of metals pass through permeable rocks and deposits the metals. They are located mainly in two settings within the area, the orogenic belts; at island arcs and at continental margins. The plutonic porphyry types occur in batholithic settings with mineralization principally occurring in one or more phases of plutonic host rock. Volcanic types occur in the roots of volcanoes, with mineralization both in the volcanic rocks and in associated comagmatic plutons. Mineralization occurring around dykes and granitic rocks with porphyritic textures, breccia zones with angular or locally rounded fragments with sulphides between fragments or in fragments, epidote and chlorite, quartz and sericite and secondary biotite alteration-with fractures coated by sulphides, or quartz veins with sulphides. Original sulphide minerals in these deposits are pyrite, chalcopyrite, bornite and molybdenite. Most of the sulphides occur in veins or plastered on fractures; most are intergrown with quartz or sericite. In many cases, the deposits have a central very low grade zone enclosed by 'shells' dominated by bornite, then chalcopyrite and finally pyrite, which may be up to 15% of the rock[4]. Molybdenite distribution is variable, radial fracture zones outside the pyrite halo may contain lead-zinc veins with gold and silver mineralization.

Recognition criteria: Hydrothermal copper deposits are scattered throughout the northwestern area, which are mainly porphyry type. Their genetic characteristic recognition criteria appear to be; age of host-rock lithology; association with volcanic or magmatic-plutonic rocks; hydrothermal alteration zones, geological structures with high tectonic settings and fractures. The criteria based on the heat sources rock and faults are not considered further as a result of the reworking of the area within the geological period and could only be appropriate when working on local bases

PREDICTION OUTLINE

The GIS-based mineral potential mapping process can be broken down into five main steps[5]:

- Building a spatial digital database
- Extracting predictive evidence for a particular deposit type, based on exploration model
- Calculating weights for each predictive map (evidential theme)
- Combining the evidential themes to predict mineral potential
- Validation of the model.

Spatial datasets: The datasets for this study are compiled from the Yunnan Geological survey. These are:

- The digital based mineral deposit data consists of 42 established mines on the study area out of these 27 are identified with hydrothermal mineralization.
- Database of the Lithology
- Lineament digitized from 1:200.000 geological maps derived from shaded relief total intensity magnetic field image
- Database of geochemical mainly from stream sediment sampling covering almost entire study area.

DATA PROCESSING AND ANALYSIS

Geochemical data: The geochemical data covers almost the entire study area. This is stream sediment data consisting of 16 elements shown in the Table 1. In the area different genetic type of copper can be found and to differentiate these factor analysis[6] had been carried out

Table 1: Component matrix

|	Component				
	1	2	3	4	5
ag	0.085	0.641	-0.256	0.264	-0.089
as	0.079	0.615	0.147	0.196	0.278
au	0.116	0.344	-0.210	0.428	-0.348
ba	-0.093	0.262	0.266	-0.267	-0.335
co	0.941	-0.055	0.070	-0.025	-0.017
cr	0.827	-0.146	-0.022	-0.027	0.109
cu	0.716	0.168	0.111	0.053	-0.131
mo	0.107	0.320	0.581	-0.072	-0.272
ni	0.835	-0.136	-0.019	-0.016	0.111
pb	0.043	0.686	-0.258	-0.469	0.095
sb	0.065	0.525	-0.192	0.457	-0.092
sn	-0.03	0.212	0.283	0.245	0.750
u	-0.216	0.204	0.605	-0.047	-0.060
v	0.865	-0.044	0.086	-0.052	-0.052
w	-0.019	0.163	0.494	0.071	-0.015
zn	0.077	0.536	-0.211	-0.625	0.130
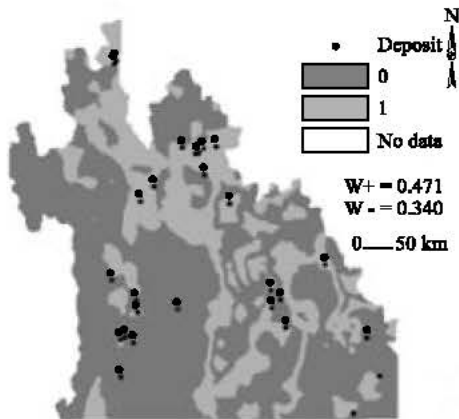
Fig. 1: Binary Map of geochemical data crossed with Cu deposits showing favorable areas of Cu deposits
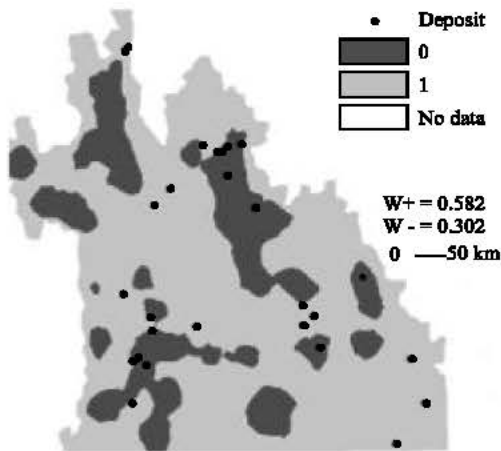

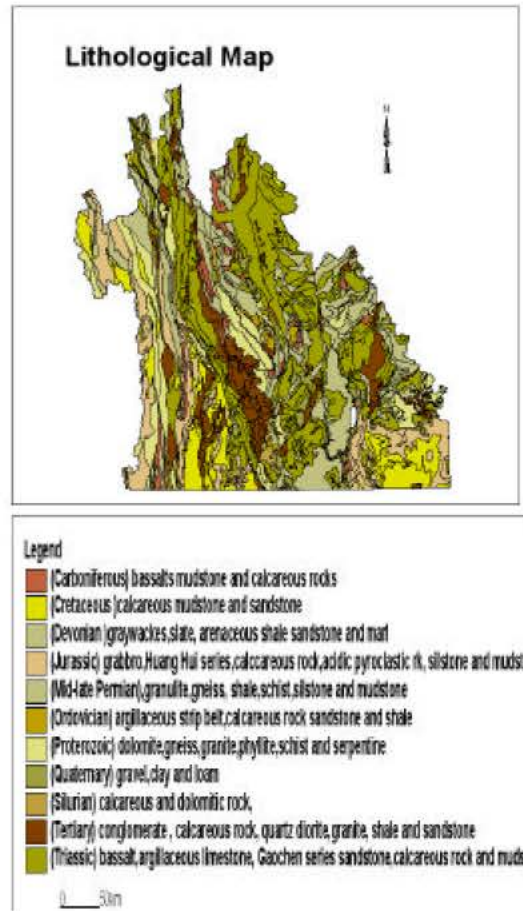Fig. 2: Binary layer of lineament overlaid with Cu deposits

Fig. 3: The map illustrates the major lithologies and their geological periods

and five factors were obtained to show different forms of mineralization. The first factor illustrate Cu in association with ultrabasic minerals such Ni, V etc. signifying metamorphic mineralization in an orogenic environment, the second factor where copper occurs with Pb, Zn, Ag despite the low value of Cu can be classified as hydrothermal mineralization, the third mineral association shows acidic magmatic deposition. The fourth shows mineralization in a low temperature environment could be classified as epithermal and the fifth Cu with Sn, very high temperature mineraliation. The second factor has been selected and contoured using Kriging in ArcGIS and exported for the weighting. The data was overlain by the mineral occurrence to determine the favorable zone based on (W^+ and W^-) (Fig. 2) and the multiclass map reclassified to binary map. The cut-off was selected from $W^+=0.2$ to $W^+=0.65$ to represent favorable geochemical anomalous zone which covers approximately about 32% of the area. Fifteen out of the 27 mines fall in this region shown in Fig. 1.

Table 2: weight of evidence analysis of the selected recognition criteria

Geochemical domain

Class	A (Sq km)	No. Dep	W+	s (W+)	W−	s (W−)	C	s (C)	Stud (C)
-1.189	203.253	0							
-0.806	5705.992	1	-0.82	1.00	0.05	0.20	-0.8675	1.0191	-0.8512
-0.588	9955.567	1	-1.37	1.00	0.12	0.20	-1.4949	1.0191	-1.4668
-0.464	7419.928	2	-0.39	0.71	0.04	0.20	-0.4256	0.7349	-0.5791
-0.247	13419.660	6	0.12	0.41	-0.03	0.22	0.1508	0.4630	0.3257
0.136	11866.489	9	0.65	0.33	-0.21	0.24	0.8618	0.4084	2.1104
0.810	8245.406	4	0.20	0.50	-0.03	0.21	0.2320	0.5419	0.4281
1.994	9279.827	3	-0.21	0.58	0.03	0.20	-0.2341	0.6125	-0.3822
4.076	1933.133	1	0.27	1.00	-0.01	0.20	0.2740	1.0193	0.2688
Lineament domain									
549.312	20.670	0							
1288.664	203.121	0							
2283.803	884.060	0							
3623.221	4418.699	1	-0.56	1.00	0.03	0.20	-0.5890	1.0192	-0.5779
5426.024	15809.117	4	-0.45	0.50	0.10	0.21	-0.5516	0.5418	-1.0180
7852.525	29856.155	10	-0.17	0.32	0.11	0.24	-0.2810	0.3986	-0.7050
11118.500	16464.087	11	0.52	0.30	-0.25	0.25	0.7703	0.3918	1.9661
15514.375	525.750	1	1.57	1.00	-0.03	0.20	1.6008	1.0200	1.5694
Lithological age domain									
Carbonefereous (C-P)	1886.34	2	0.99	0.71	-0.05	0.2000	1.0436	0.7352	1.4194
Creataceous (K-E)	6146.36	3	0.22	0.58	-0.02	0.2042	0.2423	0.6125	0.3955
Devonian (D2-3)	4198.50	2	0.19	0.71	-0.01	0.2000	0.2077	0.7350	0.2826
Jurassic Huang Hui (J2-3)	8252.66	3	-0.08	0.58	0.01	0.2042	-0.0868	0.6125	-0.1417
Late-Permian (P3)	5472.24	4	0.62	0.50	-0.08	0.2086	0.6996	0.5419	1.2910
Mid-late Permian (P2)	5333.96	1	-0.27	1.00	0.01	0.1962	-0.2813	1.0192	-0.2760
Ordovician-Silurian (O-S)	403.23	0							
Proterozoic (Pt)	7303.90	0							
Quatenary (Qhall)	1805.18	0							
Tretiary (Tr)	7123.63	3	0.11	0.58	-0.01	0.2042	0.1245	0.6125	0.2033
Triassic (T2MS)	13528.98	0							
Triassic (T3BS)	4303.07	6	1.27	0.41	-0.19	0.2183	1.4554	0.4632	3.1423
Triassic (T1-2MS)	2947.79	2	0.55	0.71	-0.03	0.2000	0.5808	0.7351	0.7901

Table 3: Contingency table conditional independence

	Geochemical	Lithology
Lineament	3.12	0.00
Geochemical		0.30

Probability values depend on chi-squared values and number of degrees of freedom.

Table 4: Result of weight-of-evidence

Evidential lithology	Class field	W0	W1	Contrast*	Student (C)
Lithology	Value	-0.847	0.823	1.6698	3.961
Lineament	Value	-0.302	0.582	0.8838	2.281
Geochemistry	Value	-0.340	0.471	0.8106	2.104

Lineament data: Lineaments are extracted from the Landsat images using lines and circular features. Intersection of radial features and the lines are most favorable for mineralization. The lineament are then put into grid and the line density calculated using MAPGIS a software developed by CUG. The result is then imported in ARCGIS and contoured using kriking.

The layer has then been weighted and an evidence layer has been chosen. The cut-off value was selected with (W^+ = 0.52 to W^+ = 1.57). This represent the anomalous zone for lineament and this occupies about 24.6% of the area and 12 out of the 27 mines are within the area shown in Fig. 2.

Lithology: The database provides the various geological periods associated with most of the formation shown

in Fig. 3. The classification of the geological period is the one adopted by the Yunnan Geological Survey. The various time units were reclassified to the major geologic periods and reclassification is again done on the major units, which have mineral occurrence in order to avoid over prediction. The cut-off value range from (W^+ = 0.22 to W^+ = 2) which includes Silurian rocks, Triassic Gaoshen series (T3BS), Triassic (T1-2MS), Late-Permian formations, Creataceous and the Carboniferous formations shown on (Table 2). This occupies nearly 31% of the area and 18 of the mines are located in this region shown in Fig. 4.

TEST OF CONDITIONAL INDEPENCE

Conditional independence is done to check the maps that can be used in the final prediction. This is tested by formulating the null hypothesis of pairwise conditional independence of the predictor binary maps. The χ_c^2 tests is run to check if the hypothesis is rejected at 95% significance level.

According to Bonham-carten et al.[2], if two binary predictor maps B_1 and B_2 are conditionally independent with respect to a set of mineral occurrences points D:

Then, $N\{B_1 \cap B_2\} = N\{B_1 \cap D\}N\{B_2 \cap D\}/N\{D\}$
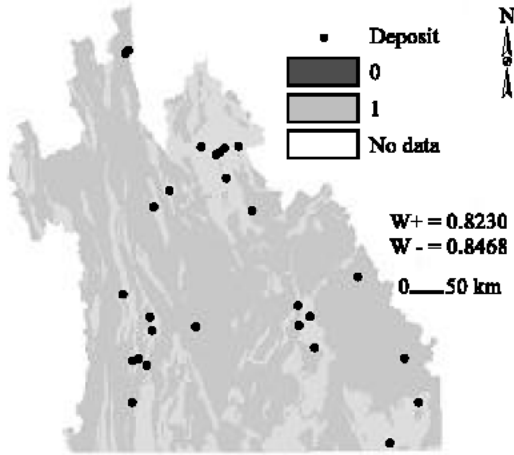
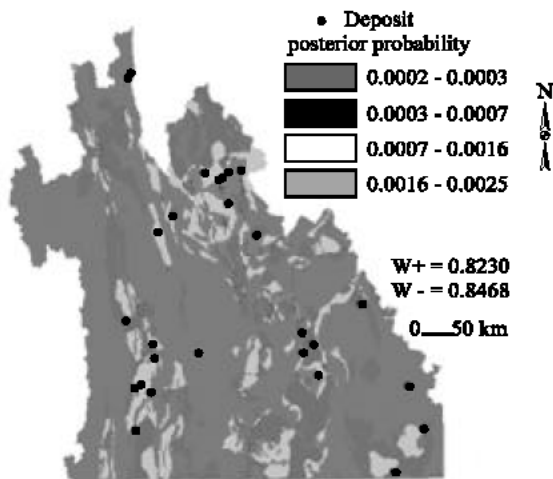Fig. 4: Binary map extracted from lithology and deposits

Fig. 5: Map of conditional probability of hydrothermal Cu occurrence in Northwestern Yunnan Fig. 5 (prior prob = 0.0004)

The mineral occurrence are considered as points, or single unit cells, the resulting values of χ^2 are unaffected by the units of the area measurement. The calculated values can then be compared with the standard values of χ^2 to verify if the conditional hypothesis holds at 95% probability with one degree of freedom. The standard value of χ^2 with one degree of freedom at 95% significance level is 3.84[7] all the values shown on Table 3 from result WOE are less than 3.84.

PROSPECTIVITY MODEL

The predictive exploration model was generated by summing the weights of evidence of three binary maps representing the hydrothermal copper recognition criteria.

The various overlap combinations of the binary maps result in the highest cumulative weights for the locations where favorable conditions are present.

DISCUSSION

The low value of the prior probability is the result of the few number of mineral occurrences for this particular study. This affects the low value gotten from the posterior probability values.

Reines[8] Suggested the following test if the weights make sense; the deposits must be found in the high favorability area; use other methods to cross check if the results tally. In general the results of the prediction needs to be validated to see the predictive ability of the results but this is quite an expensive thing to do.

The prediction result demonstrates that most of the known deposits occur in the predictive region. They could be classified into three major group, high medium and low. With a prior probability of 0.0004, the blue color shown on the legend represent the background value, the green color represent low predictive zone which occupies about 21% of the area, yellow color the medium which occupies 12% and the red color the high predictive region occupying about 3% with 9 mines located or near it as shown clearly on Fig. 5.

Table 4 summaries the results of the weight, contrast and student (C) calculated from the weight evidence model. The three evidential maps have contrast value nearly or greater than 0.5 and their students (C) values greater that 2 and they are conditionally independent to each other at significant level 95% (Table 3) so the can be used in weight evidence model. The lithological age is the best recognition criteria according to the available data the with contrast value of 1.6698 and student (C) value of 3.961 (Table 2)

CONCLUSIONS

The results of the evaluation concluded the successful implementation of weight-of-evidence model in the Northwestern of Yunnan as a preliminary measure is delineating potential areas for further exploration. The results has shown new areas which need to be check as a way of validation.

Weight of evidence has its associated advantages and disadvantages. Since it is a data-driven, it does not allow for subjective decisions on the choice of evidences especially determining the cut-off weight. There is the possibility of loosing information during the conversion of multi-class map to binary maps. It's main advantage is the ease of carrying the method especially where there is the need to make a quick decision.

REFERENCES

1. Bonham-Cater, G.F., F.P. Agterberg and D. Wrigght, 1989. Weight-of-evidence Modeling: A New Approach to Mineral Potentials. In: Agterberg F.P. and G.F. Boham-Cater, (Eds.) Statistical Applicattions in the Earth Geological Survey Canada, Paper 89-9, pp: 171

2. Agterberg, F.P., G.F. Bonham-Carter and D.F. Wright, 1990. Statistical Pattern Integration for Mineral Exploration. In: Gaal, G. and D.F. Merriam, (Eds.), Computer Applications in Resource Estimation Prediction and Assessment of Metals and Petroleum. New York, Pergamon Press, p: 1-12.

3. Mark, J.M., 1990. Mineral Potential Modelling of gold and Silver Mineralization in the Nevada Great basin. A GIS Based approach Using Weight of Evidence USGS.

5. Arc-WofE User Guide June, 1998.

4. Qintong S.G., 1991. Lead-Zinc Deposit in Nujiang-Lanangjiang-Jinshajiang, pp: 148-158.

6. John W. and Sons, 1986. Statistics and Data Analysis in Geology. 2nd Edn., pp: 527-551.

7. Asadi, H.H., 1999a. Integrated analysis of aeromagnetic, Lansat TM and mineral occurrence data for epithermal gold exploration in northwestern Iran. Proceedings of 13th International Conference on Applied Geology Remote Sensing Vancouver, British Columbia, Canada 1-3 March, pp: 8.

8. Raines, G.L., 1999. Evaluttion of weight of evidence to predict epithermal-gold deposits in the great bain of Western United states. Natural Resources Res., 8: 257-276.