Data-driven based logistic function and prediction-area plot for mineral prospectivity mapping: a case study from the eastern margin of Qinling orogenic belt, central China

Hongyang Bai
China University of Mining and Technology
Yuan Cao
Henan Bureau Of Geology And Mineral Exploration And Development
Heng Zhang
Henan Bureau Of Geology And Mineral Exploration And Development
Wenfeng Wang (wangwenfeng@cumt.edu.cn)
China University of Mining and Technology https://orcid.org/0000-0002-2200-0250
Chaojun Jiang
Henan Bureau Of Geology And Mineral Exploration And Development
Yongguo Yang
China University of Mining and Technology

Research Article

Keywords: Data-driven, Logistic function, Prediction-area plot, Mineral prospectivity mapping, Qinling 0-orogenic belt.

Posted Date: June 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-306330/v1

Data-driven based logistic function and prediction-area plot for mineral prospectivity mapping: a case study from the eastern margin of Qinling orogenic belt, central China

Hongyang Bai [a,c], Yuan Cao [b], Heng Zhang [b], Wenfeng Wang [a,c,d,*], Chaojun Jiang [b], Yongguo Yang [a,c]

[a] *School of Resources and Geosciences, China University of Mining and Technology, Xuzhou 221116, China*

[b] *The Second Geology Prospecting Institute of Henan Bureau of Geology and Mineral Exploration and Development, Zhengzhou 451464, China*

[c] *Key Laboratory of Coalbed Methane Resources and Reservoir Formation Process of the Ministry of Education, China University of Mining and Technology, Xuzhou 221000, China*

[d] *School of Geology and Mining Engineering, Xinjiang University, Urumqi 830047, China*

Abstract

The present work combines data-driven based logistic function with prediction-area plot for delineating target areas of orogenic gold deposits in eastern margin of Qinling metallogenic belt, central China. Firstly, the values of geological and geochemical information layer were transformed into a series of fuzzy numbers with a range of 0-1 through a data-driven based logistic function on the basis of mineralization theory of the orogenic gold deposits. Secondly, the prediction-area(P-A) plot was performed on the above evidence layers and their corresponding fuzzy overlay layers to pick out a proper prediction scheme for mineral prospectivity mapping(MPM) based on the known gold occurrences. What's more, to further prove the advantages of this method, we also used a knowledge-driven approach for comparison purpose. Finally, with the concentration-area(C-A) fractal model, the fractal thresholds were determined and a mineral prospecting map was generated. The result, five of the six known gold deposits are located in high and moderate potential areas (accounts for 18.6 % of the study area), one in low potential area (accounts for 38.4 % of the study area) and none in weak potential area (accounts for

* Corresponding author at: School of Resources and Geosciences, China University of Mining and Technology, Xuzhou 221116, China.
E-mail address: wenfwang@vip.163.com.

43 % of the study area), confirmed the joint application of data-driven based logistic function and prediction-area plot a simple, effective and low-cost method for mineral prospectivity mapping, which can be a guidance for further work in the research area.

Keywords: Data-driven; Logistic function; Prediction-area plot; Mineral prospectivity mapping; Qinling orogenic belt.

Declarations

Funding: This study was supported by the National Natural Science Foundation of China (Nos. U1903207, 41972176 and 41872248) as well as A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Conflicts of interest: No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication.

Availability of data and material: All data generated or analyzed during this study are included in this article.

Code availability: Not applicable.

1. Introduction

Mineral prospectivity mapping is a comprehensive research, during which geological engineers depict metallogenic target areas with known information and data in the study area to guide further exploration. It is essentially a classification technique (Yousefi and Carranza, 2015a), by which the study area could be divided into areas with high, moderate and low favorability of mineralization, respectively (Knox-Robinso, 2000; Abedi, 2012). Simply, its objective is to portray the smallest area where usually contains the most mineral deposits in the study area. In the above process, however, two key difficulties remain yet. One is to convert evidence layers with different orders of magnitude values into a same space and integrate them (Yousefi and Nykänen, 2016), the other is to determine a group of reasonable thresholds to demarcate the study area(Knox-Robinso, 2000).

48 Recent decades have witnessed researches on foregoing issues by many scholars. All of them could
49 be grouped into three categories(Yousefi and Nykänen, 2016; Yousefi and Carranza ,2016; Du et al.,
50 2016). One is the knowledge-driven method, which mainly assess mineralization evidences based on the
51 knowledge of geological experts and assigns different weights to each mineralization factor. It is
52 generally suitable for study areas with low exploration levels(Carranza, 2010), in which a large amount
53 of mineralization data is absent. Among them, fuzzy sets and fuzzy logistic, proposed by Zadeh (1965),
54 has been widely applied and proved of great value in MPM by lots of geologists (An, 1991; Cheng and
55 Agterberg, 1999; Knox-Robinson, 2000; Carranza and Hale, 2001; Luo and Dimitrakopoulos, 2003;
56 Nykänen, 2008; Lusty, 2012; Abedi, 2013; Ford et al., 2016; Du et al., 2016). However, the method has
57 a weakness of suffering the experts' bias, thus, leading to the result that different experts often have
58 different opinions(Yousefi and Carranza, 2015a). Another is the traditional data-driven method, which
59 can better establish the relationship between known mines and various evidence layers for metallogenic
60 prediction(Porwal et al. 2003a, b; Carranza, 2008; Carranza et al., 2008a; Zuo, 2011; Liu et al. 2014,
61 2015). Nevertheless, its deficiency, requires a batch of known mines training to establish this connection,
62 leading it infeasible in areas with low exploration levels. As a result, it would be deviation from the fact
63 as unknown deposits have a lesser chance to participate in such training. Besides, there are also
64 approaches(or system) of integration (Yousefi et al., 2019) or hybridization of the two aforementioned
65 methods(Chung and Fabbri, 1993; Billa et al., 2004; Porwal et al., 2004, 2006; Roy et al., 2006; Carranza
66 et al., 2008b; Cassard et al., 2008).Whereas, most of whom still cannot get out of the above dilemma or
67 just in development (eg. exploration information system promoted by Yousefi et al.)
68 In order to solve these problems, many geologists attempted not to use training points to assign
69 weights to evidence layers(Luo, 1990; Chung and Fabbri, 1993; Carranza and Hale, 2002; Luo and

70 Dimitrakopoulos, 2003; Carranza, 2010).However, their choice of experience functions are also depend
71 on expert judgment, leading them to suffer the same limitations as knowledge-driven or data-driven
72 methods.

73 Recently, a data-driven approach based on logistic functions proposed by Yousefi et al.(2012, 2013,
74 2014) and Yousefi and Carranza (2015a) could well overcome the defects of the above two methods. It
75 assigns weights to the evidence layers in a data-driven way without experts' bias and without using the
76 known mines. In this paper, the logistic function is applied to convert the evidence values into fuzzy
77 values with the range of 0-1. Whereafter, the P-A plot advocated by Yousefi and Carranza(2015a, 2015b,
78 2016)was applied to evaluate the prediction ability and fuzzy overlay was used to integrate these fuzzy
79 layers to generate a prospective map. Consequently, three prospective maps were obtained and assembled
80 and the best one was identified. Finally, a prospectivity map was obtained by means of concentration –
81 area (C-A) model classification. With 5 out of 6 known Au occurences predicted in 18.6 percent of the
82 study area, a satisfactory prospectivity map was obtained. It should be noted that all the aforementioned
83 spatial analyses were based on the raster format in geographical information system (GIS).

84 **2. Geological Setting**

85 The study area lies on the southwest of Henan province, central China, covering an area about
86 1001km^2. It is located tectonically in the junction between the southern margin of the north China plate
87 and the northern margin of the Yangtze Plate (the central and eastern part of the Qinling structural belt),
88 with Jingziguan-Shigang complex syncline, Xixia fault depression basin generated (Meng and Zhang,
89 2000; Bureau of geology and mineral resources of Henan province, 1989). Crossing four metallogenic
90 zones (north Qinling orogenic belt, south Qinling fold belt, Xixia-Tongbai polymetallic metallogenic belt
91 and the Danjiangkou gold-silver-vanadium-antimony metallogenic zone), the study area is characterized

by complex geological conditions, frequent magmatic activities, and significant structural deformation-(regional) metamorphic effects. Consequently, the metallogenic geological conditions are superior, with vanadium and gold as mainly metals as well as marbles and limestones as main non-metallic mineral, enabling it an area with rich reserves and prospecting potential(Bureau of geology and mineral resources of Henan province, 1989). Location and brief geological maps of the study area are shown in Fig.1.

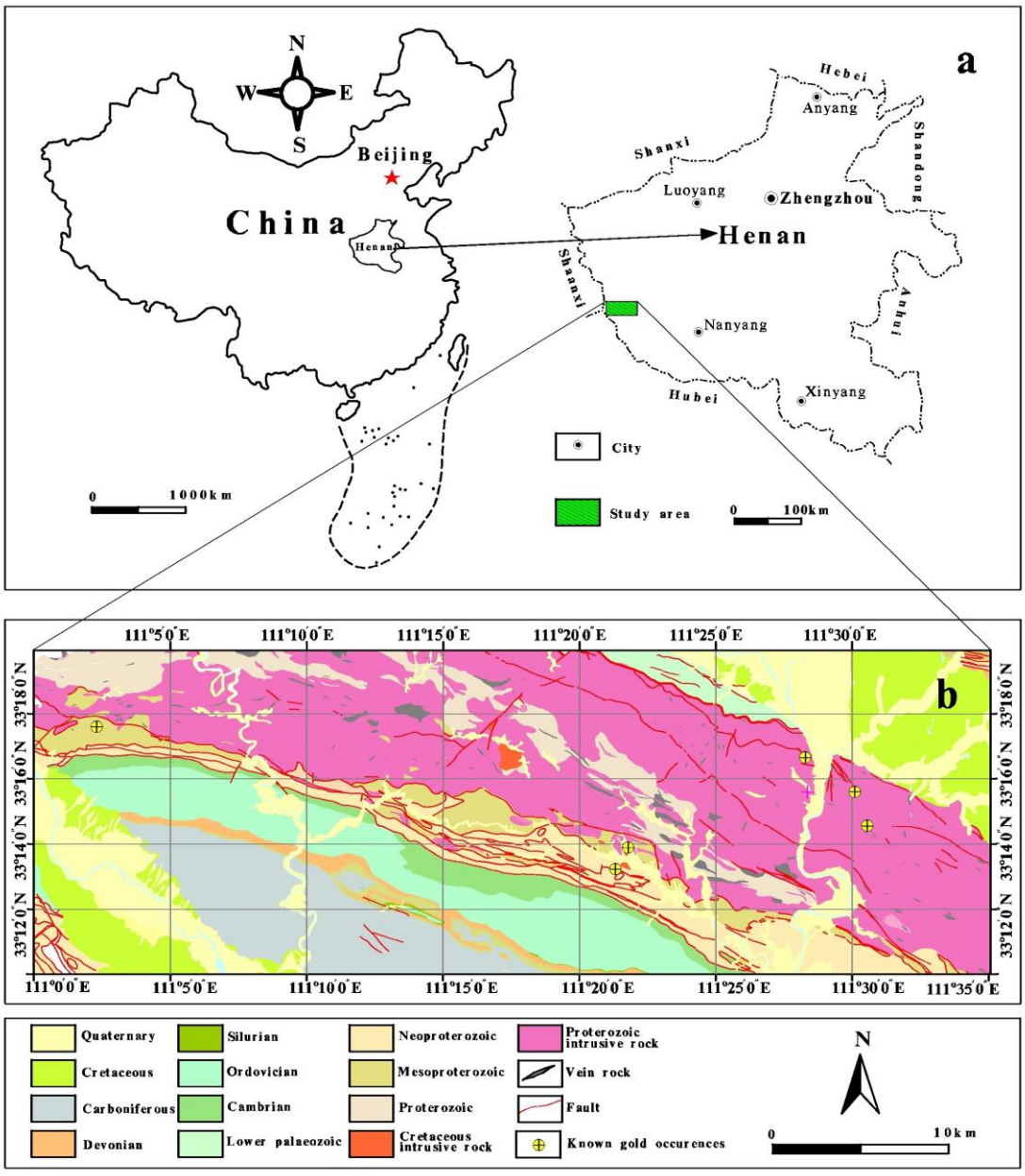


Fig.1. (a)Location of the study area and (b)brief geological map of the study area(modified from Bureau of geology and mineral resources of Henan province, 1989).

3. Data set

3.1 Data sources

In this paper, we used the 1: 50000 regional geological map and 1: 50000 stream sediment survey prepared by the second geology prospecting institute of Henan bureau of geology and mineral exploration and development. A total of 4036 samples were collected, processed and analyzed with a sampling density of 4-6 points per square kilometer in about 1000 square kilometers in the study area. Each of the original samples weighed more than 150g with a particle size less than 60 mesh (<216μm).

A multi-element (Au, Ag, As, Sb, Cu, Pb, Zn, Mo, W, Cd) was obtained with Graphite Furnace Atomic Absorption Spectrometry(Au), Atom Emission Spectrometry (Ag), Atomic Fluorescence Spectrometry (As, Sb) and Inductively Coupled Plasma Optical Emission Spectrometry(Cu, Pb, Zn, Mo, W, Cd). The detection limits were 0.3ppb for Au, 0.02ppm for Ag, 0.2ppm for As, 0.04ppm for Sb,1ppm for Cu, 1ppm for Pb,5ppm for Zn,0.3ppm for Mo, 0.3ppm for W, 0.04ppm for Cd.

Tight quality control was maintained at every stage of the process according to the National Standard of Geochemical Survey (DZ/T 0011-1991) and Geological and Mineral Laboratory Testing Quality Management (DZ01304-2006).

3.2 Data preprocessing

On the basis of Hengl's (2006) research, we determined a proper cell with 200m × 200m for all of the evidence maps. Subsequently, the geological data and geochemical data were processed to obtain evidence values in each cell. In this way, the study area contains a total of 25,024 cells.

For orogenic gold deposits, the metallogenic geological conditions mainly include ore-bearing rocks, ore-controlling structures, and magma activities (Kerrich and Wyman 1990; Groves et al., 1998; Kerrich

et al.， 2000; Goldfarb et al.， 2001). At the same time, as its ore-bearing rocks could not be more complex and diverse, a variety of rocks can generate ore-bearing rocks when other mineralization conditions are in place. As a result, only ore-controlling structures and magmatic activities are considered in actual research in this paper.

On the basis of 1:50,000 geology survey, the main geological information related to the mineralization, such as magmata and faults, were extracted as the evidence layers.Meanwhile, based on the measurement of 1:50,000 stream sediment survey, Stage Factor Analysis (SFA) was performed to carry out the multi-element closely related to the gold mineralization in the study area. Subsequently, the corresponding multi-element factor were grid-processed to get their value in each cell.

3.2.1 Geological evidences, mainly heat sources and faults.

Magma, as the main heat source, takes an important role in the formation of gold deposits. It can not only provide corresponding motive force for element migration and aggregation, but also supply the matter source(Craw et al., 2006). As Mao et al.(2002) and chen et al.(2008)mentioned , most gold deposits are related to Mesozoic granites in Xiaoqinling–Xiong'ershan region. In this paper, the granite vein and the Yanshanian granite porphyry were extracted as the evidence layers and the distance from the intrusive contact was used as the indicator criterion. However, the further from the intrusive, the less the possibility of mineralization , thus, the inverse square of the distance from intrusive was taken as the evidence value in each cell (Fig. 2a). It is generally accepted that faults are important channels for the movement of geological fluids (Pirajno, 2010). Without faults, there could be little migration and enrichment of elements, as a result, impossible to generate gold deposits as well. Consequently, we take proximity to fault as the evidence value, whose acquisition way is similar to that of heat source(Fig. 2b).

3.2.2 Geochemical evidences

142 Stage factor analysis (SFA), proposed based on principal component analysis by Yousefi et al. (2012,
143 2014) as an optimized method, was performed on 4036 samples of 10 elements to obtain multi-element
144 geochemical anomalies factors based on statistical product and service solutions (SPSS) platform.

145 **Table 1**

146 Rotated factor matrix of staged factor analysisor data of samples from the study area. Loadings in bold
147 represent the selected elements based on threshold of 0.6 (the absolute threshold value) for each stage.

First main phase						Second main phase			
First stage			Second stage			Third stage		Fourth stage	
Element	F1	F2	Element	F1	F2	Element	F1	Element	F1
Au	-0.1	**0.726**	Au	-0.07	**0.719**	Cu	**0.698**	Au	**0.681**
Ag	0.48	0.592	As	0.153	**0.803**	Zn	**0.797**	As	**0.819**
As	0.118	**0.785**	Cu	**0.719**	-0.029	Mo	**0.758**	Pb	**0.783**
Sb	0.184	0.304	Pb	0.291	**0.713**	Cd	**0.838**		
Cu	**0.723**	-0.002	Zn	0.72	0.384				
Pb	0.257	**0.685**	Mo	**0.771**	0.072				
Zn	**0.692**	0.354	Cd	**0.812**	0.16				
Mo	**0.761**	0.1							
W	0.266	0.235							
Cd	**0.803**	0.195							

148 From the results of staged factor analysis (Table 1), two indicator factors (Fig. 2c and Fig.2d) were
149 obtained to reflect the presence of orogenic gold deposits, F1 (Cu-Zn-Mo-Cd) and F2 (Au -As-Pb),
150 respectively. Elements with high factor loading values (greater than 0.6 with bold) in third and fourth

stage of the two factors could be used as indicator factors based on geochemical criteria to define exploration targets.

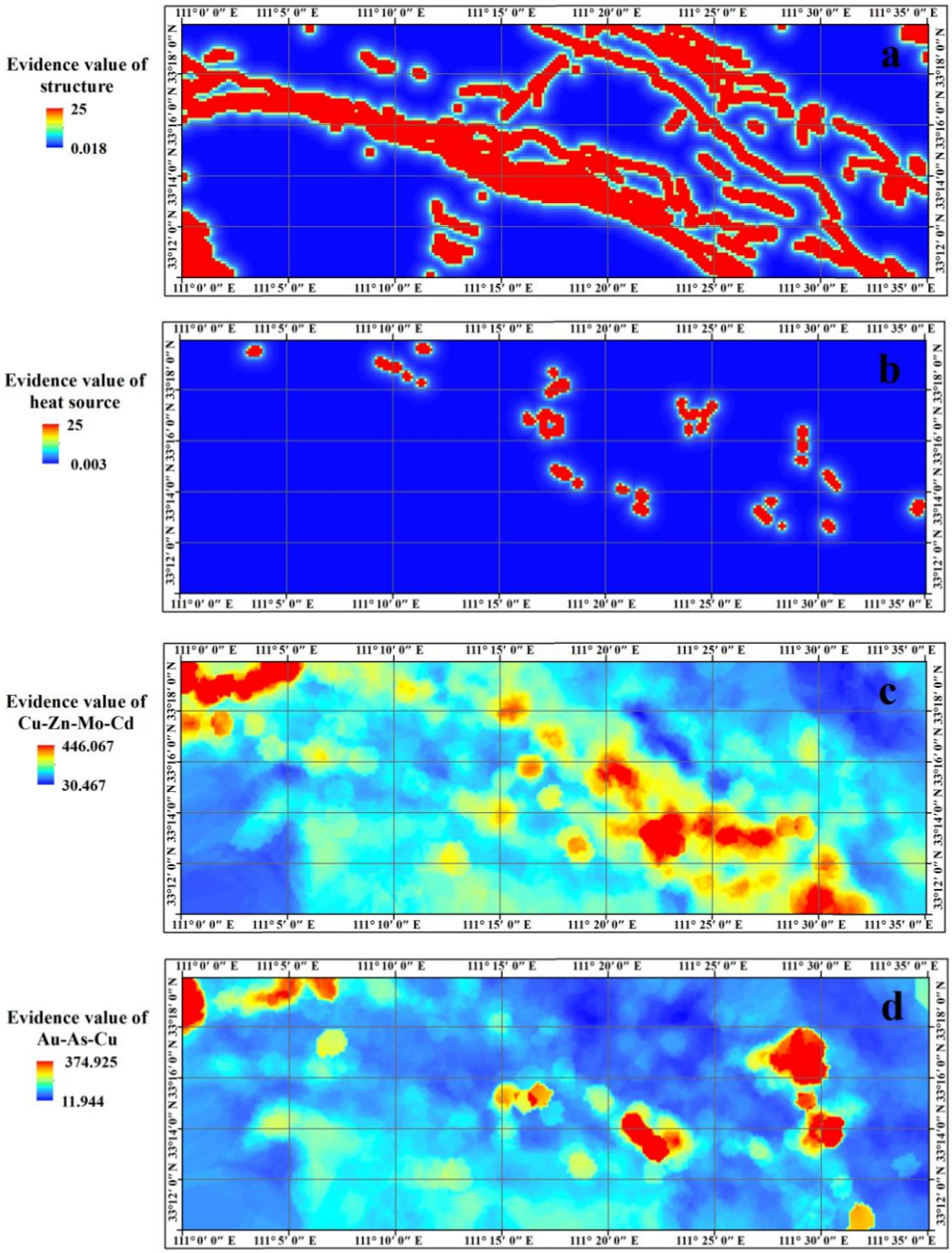


Fig.2. Evidence values of (a) structure; (b) heat source; (c) Cu-Zn-Mo-Cr factor and (d) Au-As-Pb factor.

4. Method and result

4.1 Logistic function

Since different evidence layers have values of diverse orders of magnitude, they are incomparable and cannot be superimposed to generate a prospecting map. Therefore, we used S-shape logistic function to transform the evidence values to obtain fuzzy numbers with a range of 0-1, which could be proper proxies of the evidence values.

$$F_{EV}=\frac{1}{1+e^{-s(EV-i)}} \quad (1)$$

S-shaped logistic function, as showed in equation (1)(Yousefi and Nykänen, 2016), where EV and F_{EV} are evidence value and fuzzy score while s and i are unknown parameters, has been used by a number of geologists as an efficient tool for mineral prospectivity mapping and most of which have obtained anticipative results. However, as Yousefi and Nykänen (2016) mentioned, the selection of the parameters s and i in logistic function is often subjective due to the fact that different experts often have different preferences. Moreover, some subjectively selected parameters cannot get continuous fuzzy numbers. That is to say, such experience-based parameters sometimes give different evidence values with similar or concentrated fuzzy numbers, thus, resulting in the fuzzy score a poor representative of the source data.

Therefore, Yousefi and Nykänen (2016) suggested to define the maximum and minimum values of F_{EV} as close to 1(0.99) and close to 0(0.01), which represents the most and least important evidence value separately. With EV_{max} and EV_{min}(the maximum and minimum values of EV)as well as F_{EVmax} and F_{EVmin} (the maximum and minimum values of F_{EV}) known, Equations set (2) and (3)(Yousefi and Nykänen, 2016) are easily to be solved to get s and i. In this way, a data-driven fuzzy score is obtained, which avoids the distortion of conversion caused by subjective bias of experts.

$$\begin{cases} F_{EVmax}=\frac{1}{1+e^{-s(EV_{max}-i)}} \\ F_{EVmin}=\frac{1}{1+e^{-s(EV_{min}-i)}} \end{cases} \quad (2)$$

178 $$\begin{cases} s=\frac{9.2}{EV_{max}-EV_{min}} \\ i=\frac{EV_{max}+EV_{min}}{2} \end{cases} \quad (3)$$

179

180 *4.2 Data transformation*

181 For Orogenic gold deposits, faults are usual play a role of activity space or even storage space for
182 ore-containing hydrotherm (Groves et al., 2000; Herbert et al., 2014; Carranza et al., 2015). It is generally
183 believed that the further from the fault, the lower the degree of profitability of the mineralization, so the
184 inverse square of proximity to fault is applied as the evidence value. When the evidence value is 0(the
185 cell contains the fault itself), its inverse square does not exist, so we manually assign it the maximum
186 value. Based on the GIS platform, proximity to fault of each cell is computed, and the maximum and
187 minimum of whose inverse square are calculated to be 25 and 0.018, respectively. Using equations (2)
188 and (3), the corresponding s and i is acquired as 0.3683 and 12.5090. Proximity to intrusive is used as
189 the data set of heat source. Since the further from the intrusive contact, the lower the degree of
190 profitability of the mineralization, we applied the inverse square as the evidence value as well. For the
191 nonuniformity of sampling points, the ordinary kriging interpolation method was conducted on the two
192 multi-element (Cu-Zn-Mo-Cd factor and Au-As-Pb factor) to obtain the evidence value of each cell on
193 the basis of verifying that they conform to the normal distribution.

194 After obtaining the evidence values of the above layers, formula (1) was performed to obtain the
195 fuzzy score in each cell (the maximum value is 0.99, and the minimum value is 0.01), which represents
196 the favorable degree of mineralization. The parameters used in the conversion process are shown in Table
197 2 and the obtained fuzzy score layers are revealed in Fig.3.

198 **Table 2**

199 Parameter values calculated for each evidence layer.

Evidential layer	s	i
Structure	0.3683	12.5090
Heat source	0.3680	12.5015
Cu-Zn-Mo-Cd-	0.0221	238.2672
Au -As- Pb	0.0253	193.4347

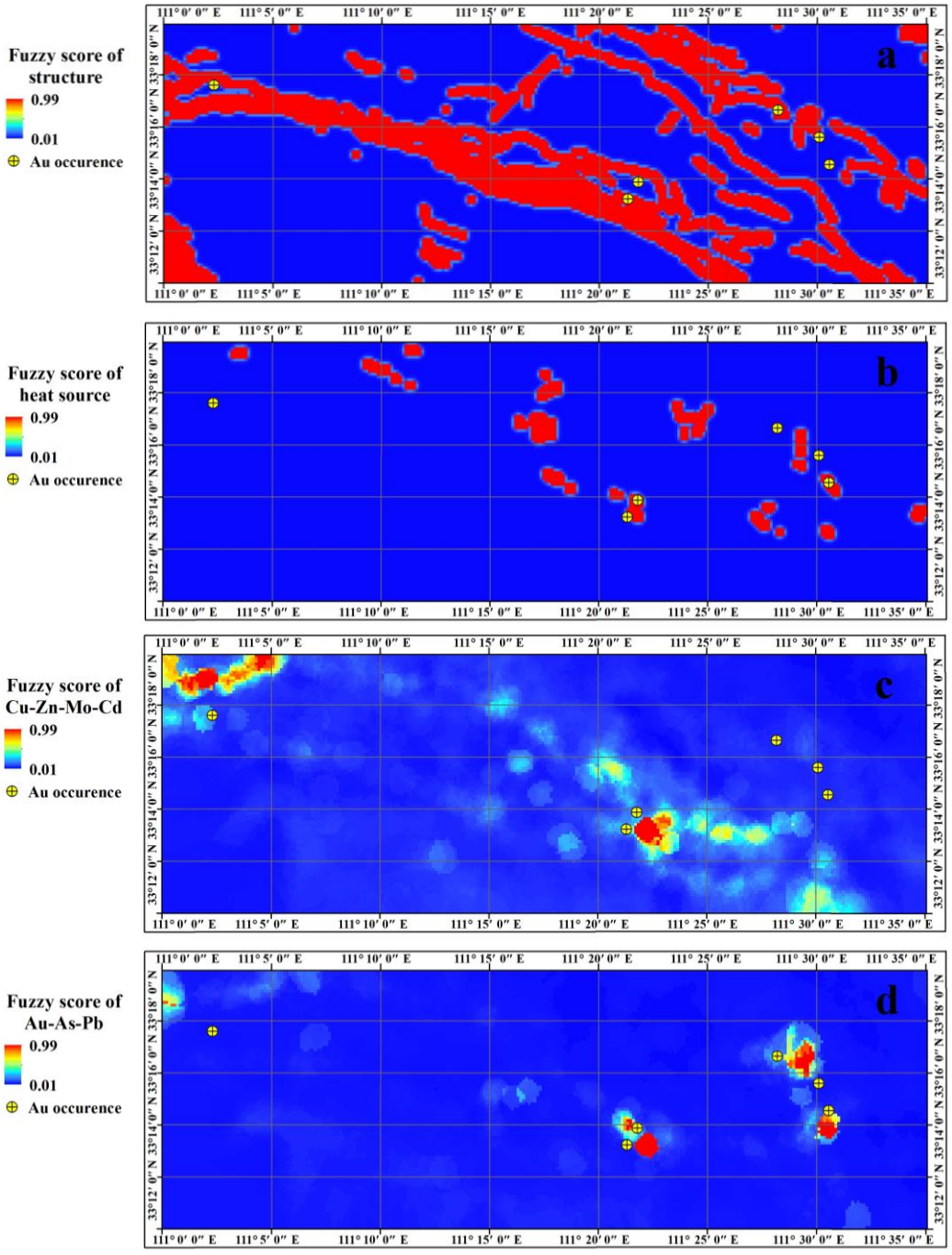


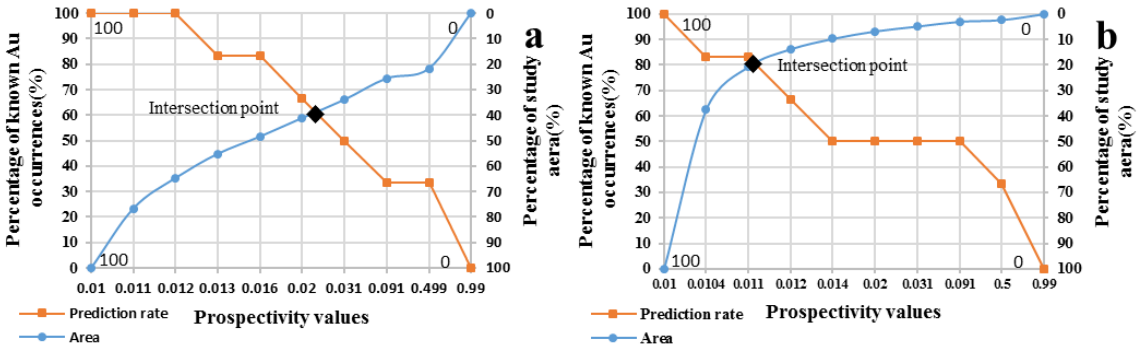
Fig.3. Fuzzy score of: (a)structuret; (b)heat source; (c) Cu-Zn-Mo-Cd factor and (d) Au-As-Pb factor.

In order to check whether the fuzzy score could be a good representation of the mineralization favorability, the obtained Fuzzy score layers (Fig.3a, 3b, 3c and 3d) were compared with the evidence values layers (Fig2a, 2b, 2c and 2d) for a comparable purpose, respectively. The comparison revealed both of them are highly similar and the former have a slight convergence than the latter only in areas where the evidence values have unobvious anomaly, therefore better to identify the real anomalies.

This is consistent with research of Bishop (2006) and Yousefi et al.(2014) that non-linear transformation(like logistic function)gains an optimal decision boundary between different classes of a variable for classification, thus boosts stronger discrimination between anomaly and background values.So we believe that the data-driven logistic function used in this paper is appropriate for its ability of transforming evidence values of different magnitudes to be fuzzy scores with 0 to 1 range and retaining the relative importance.

4.3 Evaluation of fuzzy evidence layers

Before conducting fuzzy overlay, we performed P-A plot, which is able to evaluate the mineralization advantage of each layer objectively, to quantify their prediction ability. This method employs the ratio of predicted gold deposits to total versus the ratio of the accumulated area to total as indicators based on the known gold deposits (Yousefi and Carranza, 2016).



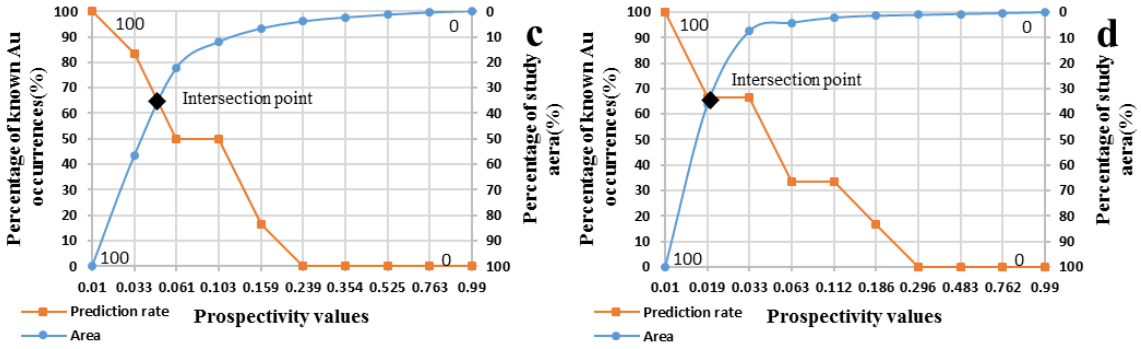


Fig.4. Prediction rate-Area(P-A) plot of each layer: (a)structure; (b)heat source; (c) Cu-Zn-Mo-Cr factor and (d) Au-As-Pb factor.

The P-A plot of evidence layer were shown in Fig.4.The intersection of the two curves represents the prediction ability of the evidence layer. The higher the intersection, the stronger the prediction ability, and the closer it is to mineralization.

4.4 Integration

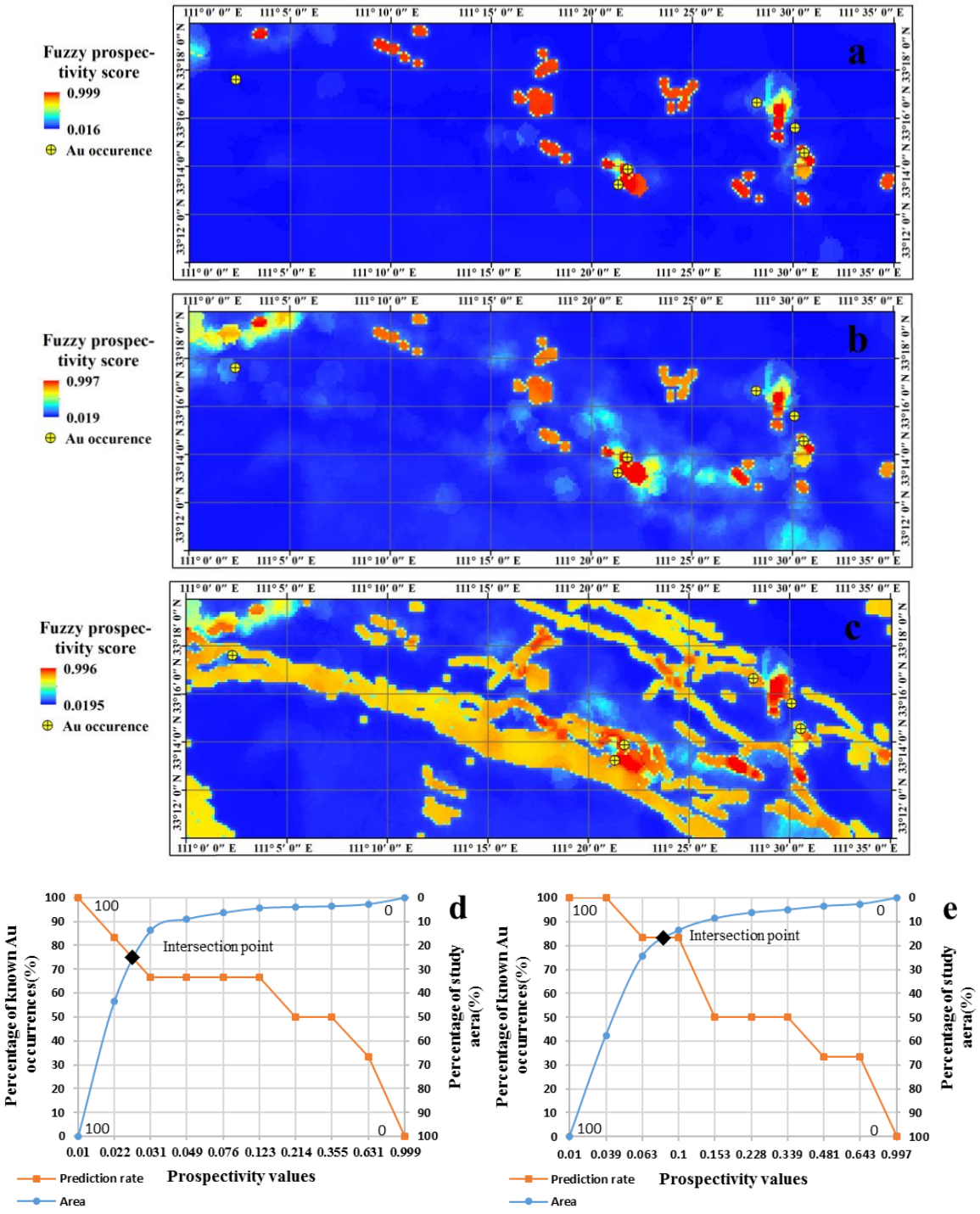
As can be seen from Table 3, the prediction capability of each fuzzy evidence layers was quite different. The best one is the heat source, with a value of 81, significantly exceeds other layers. In order to better find out the relationship between each fuzzy evidence layer and mineralization, we conducted fuzzy overlay with a γ value of 0.95 (Bonham-Carter, 1995) to integrate these layers.

Table 3

Prediction ability of each evidence layers.

Evidential layer	% of known Au occurrences	% of study area
Fig.3a (structure)	61	39
Fig.3b (heat source)	81	19
Fig.3c (Cu-Zn-Mo-Cr)	65	35
Fig.3d (Au-As-Pb)	67	33

The fuzzy evidence layers with prediction ability greater than or equal to 81, 67, 65 and 61are integrated separately to obtain 3 overlay maps (Fig.5), which were then estimated by P-A plot and their evaluation results are shown in Table 4.



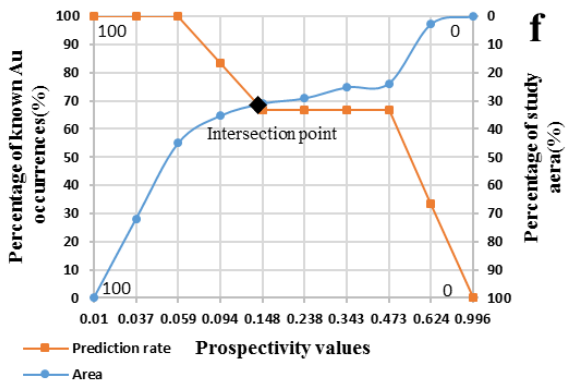


Fig.5. Perspectivity score of (a) integrated by Fig.3b and 3d, (b) integrated by Fig.3b, 3d and 3c, and (c) integrated by Fig.3b, 3d, 3c and 3a as well as P-A plot of Fig.5a, 5b and 5c.

Though there are many similarities between the Fig.5a and 5b, the latter has slightly stronger predictive value Fig.5d, 5e and Table 4). However, we also noticed that the corresponding forecasting capacity from Fig. 5d to 5f possessed a trend of increasing first and then declining. Among them, Fig.5e reached a maximum with a prediction rate of 84%, at this point, the evidence layers for integration were Fig.3b, 3c and 3d with prediction rate of 81%, 65% and 67%, respectively. This phenomenon is weird, because several layers with lower prediction ability have a better result when integrated. In spite of this, it is in line with previous study (Yousefi, 2015). What is interesting is that the prediction rate of Fig.5c dropped significantly when fault is added for integrating. This may be caused by the multi-phase superposition of tectonic movements in this area, leading to the development of a large number of faults, while there is no magmatic hydrothermal activity in some faults.

Table 4

Prediction ability of each integrated layers.

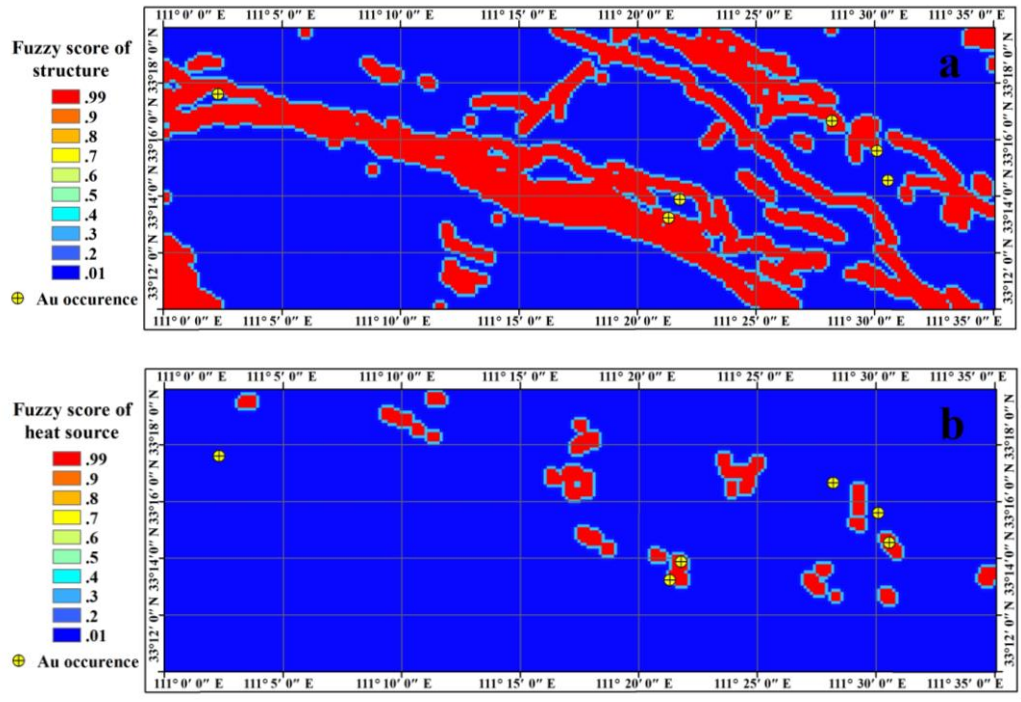
Fuzzy prospectivity map	% of known Au occurrences	% of study area
Fig.5a	75	25
Fig.5b	83	17

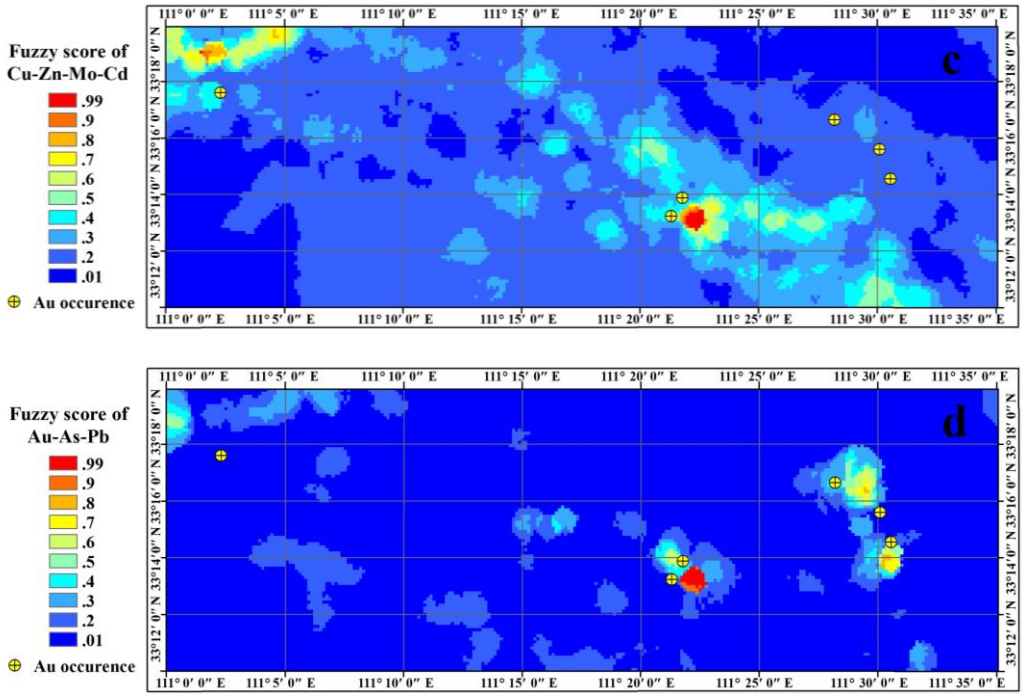
Fig.5c	68	32

In summary, it is manifest that the heat sources, Cu-Zn-Mo-Cd factor and Au-As-Pb factor are the most important factors in the mineralization process. Consequently, the corresponding Fig.5b could be used as the ideal perspectivity map in the study area.

4.5 Fuzzy prospectivity score obtained in knowledge-driven way

Carranza (2008) suggested that two or more methods should be conducted in MPM to get a proper metallogenic target area. Thus, in order to further demonstrate the superiority of logistic functions, we employed the knowledge-driven method to divide the evidence layers on the four evidence layers of Fig.2 into 10 classes at specific intervals, and assigned a mineralization favorability weight to each class in a linear way based on expert judgment for comparison purpose (Fig. 6).





256 **Fig.6.** Fuzzy score obtained by knowedge-driven way: (a) structure; (b)heat source; (c) Cu-Zn-Mo-Cr

257 factor and (d) Au-As-Pb factor.

258 Subsequently, with fuzzy gamma($\gamma = 0.95$) operation was performed, the fuzzy prospectivity score

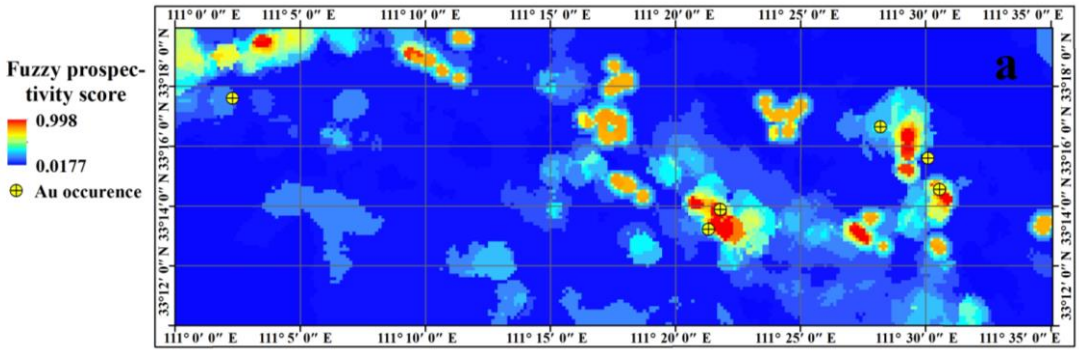
259 and the matching P-A plot were obtained (Fig.7a, 7b). According to Fig. 7b, the intersection value is 77,

260 obviously lower than that of Fig.5b (83). The above results indicate that that the weighting of the evidence

261 layer by using the logistic function not only avoids subjective judgment but also has a higher prediction

262 rate compared with the traditional discrete linear method. This is consistent with the findings of Yousefi

263 and Carranza (2015a).



264

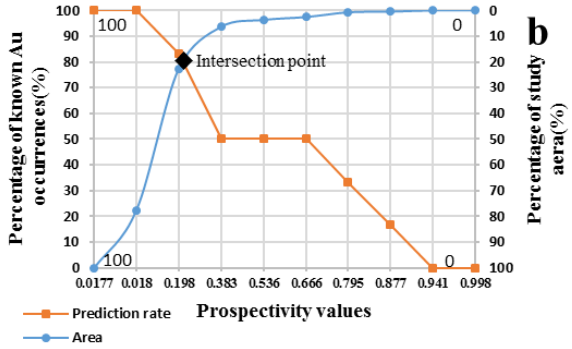


Fig.7. (a) Fuzzy prospectivity score integrated by Fig 6b, 6c and 6d; (b) P-A plot of Fig.7a.

4.6 To determine the thresholds

As Mandelbrot, B.B.(1962,1983,1985) and Carlson(1991) mentioned, in many cases, ore deposits are characterized by aggregation and fractal distribution. Therefore, in order to figure out the high, moderate, low and weak areas of the mineralization more accurately, we conducted the C-A method to defuzzify Fig. 5b so as to obtain a prospectivity map. This method was proposed by Cheng (1994) and has been used and approved by a number of geologists (Cheng et al., 1994, 1996; Cheng and Agterberg, 1995; Xiang et al., Wei and Pengda, 2002; Weiberg et al., 2004; Hodkiewicz et al., 2005; Ford and Blenkinsop, 2008; Raines, 2008; Satyabrata Behera,2019; Chen, 2019). It applies the logarithm of the concentration and the logarithm of the corresponding area greater or equal to concentration as the X-axis and Y-axis separately, which can reflect the fractal characteristics inherent in the prediction map. According to this method, in the log-log graph, the consistent slope represents a fractal dimension, and the concentration value corre sponding to the fractal point could be used as the threshold for differentiating the favorable, unfavorable and intermediate areas of mineralization (Yousefi and Nykänen, 2016).

In this paper, the logarithm of the fuzzy score and the logarithm of the cumulative area were taken as the X-axis and Y-axis, respectively (Fig.8a). Three inflection points were obtained (-1.41, -1.11, -0.19) and the corresponding fuzzy prediction values (0.039, 0.078, 0.643) were acquired, which then were used

to divide the study area into four parts (Fig.8b). The result, high potential area accounts for 2.5% of the study area with 2 known Au occurences contained, moderate potential area accounts for 16.1% of the study area with 3, low potential area accounts for 38.4% of the study area with 1 and low potential area accounts for 43% of the study area with none, would be an ideal metallogenic prediction map.

Although there was a known Au occurrence located in the low potential area, we noticed it was so close to the moderate potential area. This may be attribute to the substitution of point position for area projection. This Au occurrence was, acturally, an alteration zone about 0.76km in length with a trend near east-west, while projected only its center on the horizontal. Therefore, it can be seen from the prospectivity map that the gold occurrence is away from the favorable metallogenic area less than 2 cells (400). In reality, the gold occurrence was partially contained by the moderate areas.

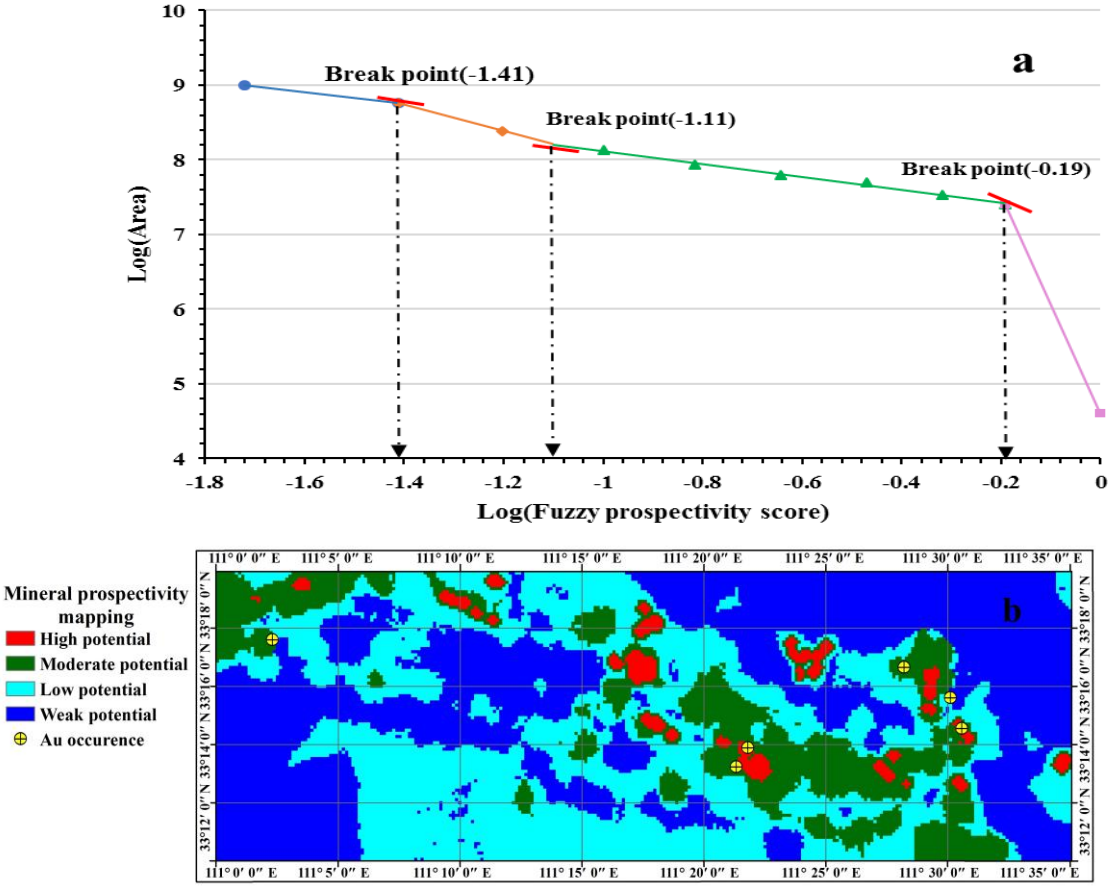


Fig.8. (a) Concentration-Area (C-A) model and (b) prospectivity map generated by Fig.5b.

5. Discussion

Different evidence layers cannot be directly compared and integrated when used for mineral prospectivity mapping because of their different dimensions. Data-driven based logistic function, which is able to transform the evidence values of different magnitude into fuzzy values with the range of 0-1 so as to represent their relative importance of mineralization, could be a suitable means. There is an assumption in this process that the largest and smallest evidence values are set a fuzzy score close to 1 (0.99) and close to 0 (0.01), respectively. This assumption is based on practical experience, in which experts assign weight to evidence values in fuzzy prospective predictions (Yousefi and Nykänen, 2016). Obtaining s and i by solving equations, as a data-driven method, is able to factually reflect the relative importance of evidence values.

In this paper, a total of 4 fuzzy layers of geology and geochemistry were evaluated by P-A plot. The results demonstrate that the heat source possess the highest prediction rate, which is consistent with the strong control of orogenic gold deposits by geological hydrothermal. The prediction rate of faults also reaches 61%, which has obvious positive correlation of mineralization. At the same time, both the two multi-element geochemistry l evidence layers have a prediction rates greater than 65%, which is in line with the SFA analysis.

When overlying, a fuzzy gamma operation is used with a recommended value of 0.95 (Bonham-Carter, 1995). The results revealed that the heat source, Au-As-Pb factor and Cu-Zn-Mo-Cd factor had a relatively high prediction rate, which was then integrated to get a perspectivivity map with a maximum prediction rate of 83%. This is higher than each of the evidence layers or integrating all of them. However, the prediction ability from Fig.5a to 5c has a tendency of going up first and declining latter. We wonder whether there would be a specific prediction rate, elevates the final prediction ability with evidence layer

integrated over it and vice versa.

6. Conclusions

(1)The data-driven based logistic function has an excellent ability of converting evidence values of different scales into fuzzy scores with a range of 0-1, and the relative importance of the obtained fuzzy scores are approximately in line with the original evidence values. Meanwhile, Due to the fact of reducing the influence of subjective preferences, data-driven based logistic function has abetter prediction effect than that of traditional knowledge-driven method.

(2) In this paper, the data-driven logistic function and the P-A evaluation were jointly applied to mineralization prediction .The result, heat source P-A plot has the highest predictive ability (81%), indicating the strong correlation between mineralization and the intermediate-acid intrusive rock (vein), which is in line with the general characteristics of orogenic gold deposits.

(3) The mineralization prediction map was generated for the study area. The metallogenic target spread roughly along the northwest – southeast (consistent with the regional tectonic), which shows the high correlation between gold mineralization and tectonic activities, and provides guidance for further large-scale exploration .

Acknowledgements

This study was supported by the National Natural Science Foundation of China(Nos. U1903207, 41972176 and 41872248) as well as A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions. The authors would like to thank Professor Mahyar Yousefi for his patient guidance. We also express special thanks to Dr. Du Xishihui for her warm suggestions. We appreciate the anonymous reviewers for their constructive comments.

336 **References**

337 Abedi M, Norouzi, G.H., Fathianpour, N., 2013. Fuzzy outranking approach: A knowledge-driven
338 method for mineral prospectivity mapping. *International Journal of Applied Earth Observation and*
339 *Geoinformation* 21,556-567.

340 Abedi M., Norouzi, G.H., Bahroudi, A., 2012. Support vector machine for multi-classification of mineral
341 prospectivity areas. *Computers & Geosciences* 46, 272-283.

342 Behera S., Panigrahi M. K., Pradhan A., 2019. Identification of geochemical anomaly and gold potential
343 mapping in the Sonakhan Greenstone belt, Central India: An integrated concentration-area fractal
344 and fuzzy AHP approach. *Applied Geochemistry* 107, 45-57.

345 Billa, M., Cassard, D., Lips, A.L.W., Bouchot, V., Touliere, B., Stein, G., Guillou-Frottier, L., 2004.
346 Predicting gold-rich epithermal and porphyry systems in the central Andes with a continental-scale
347 metallogenic GIS. *Ore Geology Reviews* 25, 39-67.

348 Bonham-carter G. F. 1995. Geographic Information Systems for Geoscientists: Modelling with GIS.
349 Computer Methods in the Geosciences 13. Pergamon Press, New York.

350 Bonham-Carter G.F., 1995. Geographic Information Systems for Geoscientists: Modelling with GIS.
351 *Computer Methods in the Geosciences 13*. Pergamon Press, New York.

352 Bureau of geology and mineral resources of Henan province, 1989. *Regional Geology records of Henan*
353 *province*.geological press, Beijing(in Chinese with English abstract).

354 C., Groves, D.I., Ojala,V.J., Eilu,P., Gardoll,S.J., 2008. Reconnaissancescale conceptual fuzzy-logic
355 prospectivity modeling for iron oxide copper–gold de-posits in the northern Fennoscandian
356 Shield,Finland. *Australian Journal of Earth Sciences* 55, 25-38.

357 Carlson, C.A., 1991. Spatial distribution of ore deposits. *Geology* 19, 111-114.

Carranza, E. J. M., 2008. *Geochemical anomaly and mineral prospectivity mapping in GIS*. In Handbook of exploration and environmental geochemistry (Vol. 11, p. 351). Amsterdam: Elsevier.

Carranza, E.J.M., 2010. Improved wildcat modeling of mineral prospectivity. *Resource Geology* 60, 129-149.

Carranza, E.J.M., Hale, M., 2001. Geologically-constrained fuzzy mapping of gold mineralization potential, Baguio district, Philippines. *Natural Resources Research* 10,125-136.

Carranza, E.J.M., Hale, M., 2002. Wildcat mapping of gold potential, Baguio district, Philippines. *Transactions of the Institutions of Mining and Metallurgy,* Section B: Applied Earth Science 111, 100-105.

Carranza, E.J.M., Hale, M., Faassen, C., 2008a. Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. *Ore Geology Reviews* 33, 536-558.

Carranza, E.J.M., Laborte, A.G., 2015. Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Computers & Geosciences* 74, 60-70.

Carranza, E.J.M., van Ruitenbeek, F.J.A., Hecker, C., van der Meijde, M., van der Meer, F. D., 2008b. Knowledge-guided data-driven evidential belief modeling of mineral prospectivity in Cabo de Gata, SE Spain. *International Journal of Applied Earth Observation and Geoinformation* 10, 374-387.

Cassard, D., Billa, M., Lambert, A., Picot, J.C., Husson, Y., Lasserre, J.L., Delor, C., 2008. Gold predictivity mapping in French Guiana using an expert-guided data-driven approach based on a regional-scale GIS. *Ore Geology Reviews* 34, 471-500.

Chen D., Wei J., Wang W., Shi W., Li H., Zhan X., 2019. Comparison of Methods for Determining the Thresholds of Geochemical Anomalies and the Prospecting Direction-A Case of Gold Deposits in

the Gouli Exploration Area, Qinghai Province. *Mineral* 9(6), 368; doi:10.3390/min9060368.

Chen, Y., Pirajno F., Qi J., 2008. The Shanggong gold deposit, Eastern Qinling Orogen, China: Isotope geochemistry and implications for ore genesis. *Journal of Asian Earth Sciences* 33:252-266.

Cheng Q., Agterberg, F.P., Ballantyneb S.B., 1994. The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration* C 51, 109-130.

Cheng, Q., Agterberg, F.P., 1999. Fuzzy weights of evidence and its application in mineral potential mapping. *Natural Resources Research* 8, 27-35.

Cheng, Q., Agterberg, F.P., Bonham-Carter, G.F., 1996. Fractal pattern integration for mineral potential estimation. *Nonrenewable Resources* 5, 117-130.

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

Craw, D., Begbie, M., MacKenzie, D., 2006. Structural controls on Tertiary orogenic gold mineralization during initiation of a mountain belt, New Zealand. *Mineralium Deposita* 41(7), 645-659.

Du X., Zhou K., Cui Y., Wang J., Zhang N., Sun W., 2016. Analytical hierarchy process (AHP) and prediction-area (P-A) plot for mineral prospectivity mapping: a case study from the Dananhu metallogenic belt, Xinjiang, NW China. *Arabian Journal of Geosciences* 9, 298, doi:10.1007/s12517-016-2316-y.

Ford, A., Blenkinsop, T.G., 2008. Combining fractal analysis of mineral deposit clustering with weights of evidence to evaluate patterns of mineralization: application to copper deposits of the Mount Isa Inlier, NW Queensland, Australia. *Ore Geology Reviews* 33,435-450.

Ford, A., Miller, J.M., Mol, A.G., 2016. A comparative analysis of weights of evidence, evidential belief functions, and fuzzy logic for mineral potential mapping using incomplete data at the scale of

investigation. *Natural Resources Research* 25, 19-33.

Goldfarb, R.J., Groves, D.I., Garwin S., 2001. Orogenic gold and geologic time: a global synthesis. *Ore Geology Reviews* 18, 1-75.

Groves, D.I, Goldfarb, R.J., Gebre-Mariam, M., Hagemann, S.G., Robert, F., 1998. Orogenic gold deposits: A proposed classification in the context of their crustal distribution and relationship to other gold deposit types. *Ore Geology Reviews* 13, 7-27.

Groves, D.I., Goldfarb, R.J., Knox-Robinson, C.M., Ojala, J., Gardoll, S., Yun, G.Y., Holyland, P., 2000. Late-kinematic timing of orogenic gold deposits and significance for computer-based exploration techniques with emphasis on the Yilgarn Block, Western Australia. *Ore Geology Reviews* 17, 1-38.

Hengl T., 2006. Finding the right pixel size. *Computers & Geosciences* 32, 1283-1298.

Herbert, S., Woldai, T., Carranza, E.C.M., van Ruitenbeek, F.J.A., 2014. Predictive mapping of prospectivity for orogenic gold in Uganda. *Journal of African Earth Sciences* 99, 666-693.

Hodkiewicz, P.F., Weinberg, R.F., Gardoll, S.J., Groves, D.I., 2005. Complexity gradients in the Yilgarn Craton: fundamental controls on crustal-scale fluid flow and formation of world-class orogenic-gold deposits. *Australian Journal of Earth Sciences* 52,831-841.

Kerrich, R., Goldfarb, R., Groves, Garwin S., Jia Y., 2000, The characteristics, origins, and geodynamic settings of supergiant gold metallogenic provinces. *Science in China Series D: Earth Sciences* 43(1), 1-68.

Kerrich, R., Wyman D., 1990. Geodynamic setting of mesothermal gold deposits: An association with accretionary tectonic regimes. *Geology* 18 (9), 882-885.

Knox-Robinso C.M., 2000.Vectorial fuzzy logic: a novel technique for enhanced mineral prospectivity mapping, with reference to the orogenic gold mineralisation potential of the Kalgoorlie Terrane,

Western Australian. *Australian Journal of Earth Sciences* 47, 929-941.

Liu Y, Cheng Q, Xia Q, Wang X, 2014. Mineral potential mapping for tungsten polymetallic deposits in the Nanling metallogenic belt, South China. *Journal of Earth Science* 25, 689-700.

Liu Y, Cheng Q, Xia Q, Wang X, 2015. The use of evidential belief functions for mineral potential mapping in the Nanling belt, South China. *Frontiers of Earth Science* 9, 342-354.

Luo, J., 1990. Statistical mineral prediction without defining a training area. *Mathematical Geology* 22, 253-260.

Luo, X., Dimitrakopoulos, R., 2003. Data-driven fuzzy analysis in quantitative mineral resource assessment. *Computers & Geosciences* 29, 3-13.

Lusty, P.A.J.,S cheib, C., Gunn, A.G., Walker, A.S.D., 2012. Reconnaissance-scale prospectivity analysis for gold mineralization in the Southern Uplands-Down-Longford Terrane, NorthernIreland. *Natural Resources Research* 21, 359-382.

Mandelbrot, B.B., 1962. *Statistics of natural resources and the law of Pareto*: Watson Research Center, New York. I.B.M. Research Note NC-146, 131 p.

Mandelbrot, B.B., 1983, *The fractal geometry of nature*: New York, W.H. Freeman and Company, 468 p.

Mandelbrot, B.B., 1985, Self-affine fractals and fractal dimension. *Physica Scripta*, v. 32, p. 257-260.

Mao, J., Goldfarb, R.J., Zhang, Z., Xu W., Qiu Y., DENG J., 2002. Gold deposits in the Xiaoqinling–Xiong'ershan region, Qinling Mountains, central China. *Mineralium Deposita* 37, 306-325.

Moon, W.M., Ping, A.N., 1991. Integration of geological, geophysical, and remote sensing data using fuzzy set theory. *Geoinformatics FCE CTU* 2 (2), 171-176.

Pirajno, F., 2010. Intracontinental strike-slip faults, associated magmatism, mineral systems and mantle dynamics: examples from NW China and Altay-Sayan (Siberia). *Journal of Geodynamics* 50, 325-

346.

Porwal, A., Carranza, E.J.M., Hale, M., 2003a. Artificial neural networks for mineral potential mapping: A case study from Aravalli Province, Western India. *Natural Resources Research* 12, 155-171.

Porwal, A., Carranza, E.J.M., Hale, M., 2003b. Extended weights-of-evidence modelling for predictive mapping of base metal deposit potential in Aravalli province, western India. *Exploration and Mining Geology* 10(4), 273-287.

Porwal, A., Carranza, E.J.M., Hale, M., 2004. A hybrid neuro-fuzzy model for mineral potential mapping. *Mathematical Geology* 36, 803-826.

Porwal, A., Carranza, E.J.M., Hale, M., 2006. A hybrid fuzzy weights-of-evidence model for mineral potential mapping. *Natural Resources Research* 15, 1-14.

Raines, G.L., 2008. Are fractal dimensions of the spatial distribution of mineral deposits meaningful? *Natural Resources Research* 17, 87-97.

Roy, R., Cassard, D., Cobbold, P.R., Rossello, E.A., Billa,M., Bailly, M., Lips, A.L.W., 2006. Predictive mapping for copper–gold magmatic-hydrothermal systems in NW Argentina: use of a regional-scale GIS, application of an expert-guided data-driven approach, and comparison with results from a continental-scale GIS. *Ore Geology Reviews* 29, 260-286.

The second geology prospecting institute of Henan bureau of geology and mineral exploration and development, 2015. *1: 50,000 mineral prospect survey report of Bailang, Jingziguan, Xixia, Qiyu, Siwan, Xichuan and Yuandian, Henan province*(in Chinese).

Wang, G., Carranza, E.J.M., Zuo, R., Hao, Y., Du, Y., Pang, Z., Sun, Y., Qu, J., 2012. Mapping of district-scale potential targets using fractal models. *Journal of Geochemical Exploration* 122, 34-46.

Xiang, Z., Gu, X., Wang, E., Wang, X., Zhang, Y., Wang, Y., 2019. Delineation of deep prospecting targets

468 by combining factor and fractal analysis in the Kekeshala skarn Cu deposit, NW China. *Journal of*
469 *Geochemical Exploration* 198, 71-81.
470 Yousefi, M. 2017.Recognition of an enhanced multi-element geochemical signature of porphyry copper
471 deposits for vectoring into mineralized zones and delimiting exploration targets in Jiroft area, SE
472 Iran. *Ore Geology Reviews* 83, 200-214.
473 Yousefi, M., Carranza E.J.M., 2015a. Fuzzification of continuous-value spatial evidence for mineral
474 prospectivity mapping. *Computers & Geosciences* 74, 97-109.
475 Yousefi, M., Carranza E.J.M., 2015b. Prediction–area (P–A) plot and C–A fractal analysis to classify and
476 evaluate evidential maps for mineral prospectivity modeling. *Computers & Geosciences* 79, 69-81.
477 Yousefi, M., Carranza E.J.M., 2016. Data-driven index overlay and boolean logic mineral prospectivity
478 modeling in greenfields exploration. *Natural Resources Research* 25, 3-18.
479 Yousefi, M., Carranza, E.J.M., Kamkar-Rouhani, A., 2013. Weighted drainage catchment basin mapping
480 of stream sediment geochemical anomalies for mineral potential mapping. *Journal of Geochemical*
481 *Exploration* 128, 88-96.
482 Yousefi, M., Kamkar-Rouhani, A., Carranza, E.J.M., 2012. Geochemical mineralization probability
483 index (GMPI): a new approach to generate enhanced stream sediment geochemical evidential map
484 for increasing probability of success in mineral potential mapping. *Journal of Geochemical*
485 *Exploration* 115, 24–35.
486 Yousefi, M., Kamkar-Rouhani, A., Carranza, E.J.M., 2014. Application of staged factor analysis and
487 logistic function to create a fuzzy stream sediment geochemical evidence layer for mineral
488 prospectivity mapping. *Geochemistry: Exploration, Environment, Analysis* 14, 45–58.
489 Yousefi, M., Kreuzer, O.P., Nykänen, V., Hronsky, J. M.A. 2019. Exploration information systems – A

490 proposal for the future use of GIS in mineral exploration targeting. *Ore Geology Reviews* 111,
491 103005, https://doi.org/10.1016/j.oregeorev.2019.103005.
492 Yousefi, M., Nykänen V., 2016. Data-driven logistic-based weighting of geochemical and geological
493 evidence layers in mineral prospectivity mapping. *Journal of Geochemical Exploration* 164, 94-106.
494 Zadeh, L. A., 1965, *Fuzzy sets: IEEE Information and Control* v. 8, no. 3, p. 338-353.
495 Zuo, R., 2011. Regional exploration targeting model for Gangdese porphyry copper deposits. *Resource*
496 *Geology* 61, 296-303.