International Journal of Applied Earth Observation and Geoinformation 58 (2017) 157–167

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

An improved data-driven fuzzy mineral prospectivity mapping procedure; cosine amplitude-based similarity approach to delineate exploration targets

Mohammad Parsa [a], Abbas Maghsoudi [a,*], Mahyar Yousefi [b]

[a] Faculty of Mining and Metallurgical Engineering, Amirkabir University of Technology, Tehran, Iran
[b] Faculty of Engineering, Malayer University, Malayer, Iran

ARTICLE INFO

Article history:
Received 17 August 2016
Received in revised form 2 February 2017
Accepted 4 February 2017

Keywords:
Cosine amplitude-based similarity
Fuzzy sets
Mineral prospectivity mapping
Distance distribution analysis

ABSTRACT

Weighting and synthesizing exploration evidence criteria for mineral prospectivity mapping (MPM) are affected by complexity and ambiguity of ore mineralization processes. In this regard, fuzziness could facilitate the modeling of such vague processes for MPM. Furthermore, imprecise selection of the exploration criteria to be used in MPM has negative influence on the efficiency of the generated prospectivity models. In this paper, of various exploration criteria, a coherent set of exploration features were recognized by using the distance distribution analysis. Then, the application of cosine amplitude-based similarity procedure was adapted as a data-driven fuzzy logic approach for predictive mapping of porphyry-Cu prospectivity in Arasbaran metallogenic zone, NW Iran. In addition, a conventional data-driven fuzzy prospectivity model was generated for comparison purpose. Comparison of the two models demonstrated the superiority of the cosine amplitude-based fuzzy procedure for MPM.

1. Introduction

Mineral prospectivity mapping (MPM) involves definition of exploration criteria representing mineral deposits of the type sought, generation of weighted evidence layers corresponding to the defined exploration criteria, synthesis of the weighted evidence layers to generate exploration targets, and evaluation of the generated targets for further exploration surveys (Bonham-Carter et al., 1990).

Due to the diversity and complexity of geological processes, local characteristics of a certain deposit type can be diverse in different areas (Andrada de Palomera et al., 2015; Yousefi and Carranza, 2015a,b). Thus, recognition of efficient exploration criteria to be used in generating weighted evidence layers is a challenging task for MPM. Subsequently, the spatial relationship between known mineral occurrences (KMOs) and diverse exploration criteria could be quantified to discriminate efficient (i.e., those with remarkable positive spatial association with the KMOs) and inefficient criteria (i.e., those with either weak positive or negative spatial association with the KMOs) to MPM (Parsa et al., 2016a; Carranza and Laborte, 2016).

Assigning realistic weights to evidential data, representing their relationships with the deposit-type sought, is another challenging issue for which two major knowledge- and data-driven MPM methods have been used (e.g., Moon, 1990; Porwal et al., 2003; Abedi et al., 2012; Ford et al., 2016; McKay and Harris, 2016; Yousefi and Carranza, 2016). Knowledge-driven methods suit poorly explored areas, while data-driven methods are propitious to well-explored fields (Carranza, 2008). Due to the stochastic and symmetric uncertainties resulting from vague and incomplete understanding of geological processes, the fuzzy set theory (Zadeh, 1965) has been progressively adapted for fuzziness of the mineralization processes to MPM (e.g., An et al., 1991; Porwal et al., 2003; Nykänen et al., 2008; Yousefi and Carranza, 2015a,b; Ford et al., 2016; McKay and Harris, 2016). In this regard, Porwal et al. (2003) proposed a data-driven technique to assign fuzzy membership values (i.e., fuzzy weights) for fuzzy logic MPM. They applied weights of evidence (WofE) method (Bonham-Carter et al., 1990; Cheng and Agterberg, 1999) for quantification of the spatial association between KMOs and classes of evidence layers, in fact, for assigning the fuzzy weights. Despite the successful application of WofE in data-driven fuzzy MPM, there are diverse alternatives for objective assignment of fuzzy membership values (Luo and Dimitrakopoulos, 2003). The cosine amplitude-based similarity procedure is a bivariate approach, which could be applied for appraising the fuzzy interrelationship between two data sets, e.g., a collection of input variables

* Corresponding author.
E-mail address: a.maghsoudi@aut.ac.ir (A. Maghsoudi).

http://dx.doi.org/10.1016/j.jag.2017.02.006
0303-2434/

and a set of training data (Ross, 1995). This method can precisely quantify the resemblances of a dependent variable (here the location of KMOs) and diverse exploration evidence layers to MPM. The quantified values of the resemblance then could be assigned to evidence layers as data-driven fuzzy membership values.

The main objectives of this study are (1) to reduce exploration bias resulting from imprecise selection of exploration criteria, (2) to adapt the cosine amplitude fuzzy procedure (Ross, 1995) for MPM, and (3) to demonstrate its superiority over conventional data-driven fuzzy logic MPM (Porwal et al., 2003). For this, we used an exploration data set of porphyry Cu deposits (hereafter denoted as PCDs) for prospectivity analysis in the northern Urumieh-Dokhtar magmatic belt, NW Iran.

In this paper, based on the generic conceptual model of PCDs (Sillitoe, 2010), a primary set of exploration criteria were collected, and then, by quantification of the spatial association between exploration evidence layers and known PCDs, effective criteria were objectively recognized to MPM. Then, the cosine amplitude-based and the conventional (WofE-based) data-driven fuzzy procedures were applied for weighting and synthesizing exploration criteria. The success-rate and the prediction-rate curves (Agterberg and Bonham-Carter, 2005) were then applied for comparison of the two prospectivity models generated.

2. Methods

2.1. Distance-distribution analysis

Diversity of geological processes in different areas results in the variety of the key geological and geochemical indicators, for prospecting a certain mineral deposit-type sought (Yousefi and Carranza, 2015a). Therefore, the primary criteria, which selected based on the generic characteristics of the deposit-type sought, should be evaluated to recognize efficient exploration criteria in an area under prospecting.

For assessing the degree of spatial association between exploration criteria and KMOs, the distance distribution analysis (DDA: Berman, 1977) has been used (e.g., Carranza, 2009a). This method involves with simultaneous construction of two curves, namely: (1) the cumulative relative frequency distribution of distances from every location to a set of geological particulars and (2) the cumulative relative frequency distribution of distances from the locations of KMOs to the same set of geological particulars. Appearing the latter curve above the former one denotes a positive spatial association between the KMOs and the geological particulars. On the other hand, if the former curve appears above the latter curve, it denotes that there is a negative spatial association between the KMOs and the geological particulars. This is because the former curve is representative of a random distribution of cells around a region, while the latter curve represents underlying geological processes and determines the distribution pattern of mineral deposits. The difference between the latter and the former curves (D), at every distance to the geological particulars, denotes how each geological particular is associated with the mineralization of the type-sought. In this regard, positive values of D indicate positive spatial association and negative values of D indicate negative spatial association of geological particulars with the KMOs. The maximum value of D denotes the strongest spatial coincidence between the KMOs and the geological particulars (Carranza, 2009a).

2.2. Data-driven fuzzy MPM

The fuzzy set theory (Zadeh, 1965) has been successfully adapted to MPM (e.g., Nykänen et al., 2008; Lusty et al., 2012; Elliott et al., 2016). As Porwal et al. (2003) mentioned, fuzzy MPM involves with three general steps of (1) fuzzification, (2) fuzzy processing, and (3) defuzzification. Fuzzification is the process of converting input variables, here spatial evidence values, into relative degrees of membership in a [0,1] range (Zadeh, 1965). Fuzzy processing refers to synthesizing the fuzzified evidence values (or layers) to a continuous fuzzy prospectivity model, and defuzzification is the process of discretization of the continuous synthesized model to a crisp and interpretable set through selecting proper threshold values. In the following subsections, the mathematical basis of conventional and cosine amplitude-based fuzzification procedures are described.

2.2.1. Conventional data-driven fuzzy procedure

The WofE procedure measures the spatial association of different classes of evidence layers and KMOs, using Student's t-values, according to the below equation (Bonham-Carter et al., 1990):

$$t = \frac{C}{\sigma} = \frac{W^{+} - W^{-}}{\sqrt{\sigma^{2}(W^{+}) + \sigma^{2}(W^{-})}} \tag{1}$$

where, t is the Student's t-value, C is the total contrast and σ is the standard deviation of C. The "C" is the difference between the positive weight (W^{+}) and the negative weight (W^{-}) for map patterns where indicator exploration feature is present or absent, respectively. The values of $\sigma^2(W^+)$ and $\sigma^2(W^-)$ are the variances of W^+ and W^-, respectively. The larger t-values indicate stronger spatial correlations between the indicator exploration features and KMOs (Bonham-Carter et al., 1990). Details about the calculation of W^+, W^-, $\sigma^2(W^+)$ and $\sigma^2(W^-)$ can be found in related literature (e.g., Bonham-Carter et al., 1990).

Based on the t-value, a linear fuzzy membership function has been defined for fuzzification of discretized evidence layers according to Eq. (2) (Porwal et al., 2003):

$$\mu_{ij} = \begin{cases} 0.01 \quad if \;\; t_{ij} = t_{\min} \; AND \; t_{\min} < 0 \\ 0.5 - \dfrac{t_{ij}}{2 \times t_{\min}} \quad if \;\; t_{\min} < t_{ij} \le 0 \\ 0.5 + \dfrac{t_{ij}}{2 \times t_{\max}} \quad if \;\; 0 < t_{ij} \le t_{\max} \end{cases} \tag{2}$$

where, μ_{ij} is the fuzzified value of the i^{th} class of the j^{th} evidence layer, t_{ij} is its corresponding t-value, and $t_{\min}$ and $t_{\max}$ are, respectively the minimum and maximum t-values of the total data set.

2.2.2. Cosine amplitude-based fuzzy procedure

The cosine amplitude-based procedure, quantifies the degree of similarity between a dependent variable (here the presence and the absence of mineralization) and an independent one (exploration evidential data) in a [0,1] range (Ross, 1995). The quantified degrees of similarity then could be considered as fuzzy membership values of the exploration evidence layers (Ercanoglu and Gokceoglu, 2004).

In MPM, an evidential layer, X, which has been classified in m different categories, can be described by an array of m vectors, as:

$$X = \{x_1, x_2, \ldots, x_m\} \tag{3}$$

Each class of evidence layer, x_i in the above expression is a vector of size n, and can be expressed by:where, n is the number of cells within the i^{th} class.

The fuzzy membership score, μ_{ij}, is determined via a pairwise comparison between a class of evidence layers (the independent variable), x_i, and the dependent variable, x_j, (i.e., the distribution of cells containing KMOs and cells without KMOs). Based on the cosine amplitude-based procedure, the data-driven fuzzy mem-