Journal of Geochemical Exploration 168 (2016) 150–162

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

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

Stepwise regression for recognition of geochemical anomalies: Case study in Takab area, NW Iran

Ahad Nazarpour [a,*], Ghodratolah Rostami Paydar [a], Emmanuel John M. Carranza [b,c]

[a] Department of Geology, College of Sciences, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran
[b] Economic Geology Research Centre, James Cook University, Townsville, Queensland, Australia
[c] Institute of Geosciences, State University of Campinas, Campinas, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 28 October 2015
Revised 19 June 2016
Accepted 3 July 2016
Available online 5 July 2016

Keywords:
Stepwise regression
Geochemical exploration
Geochemical anomaly
Logratio transformation
Takab

ABSTRACT

Stream sediment geochemical data represent compositional materials derived from various sources, including single or multiple lithologic units, soil types, rocks types, etc. In order to delineate geochemical anomalies, stream sediment geochemical data are usually subjected to suitable multivariate analysis, and not simply using univariate threshold values because these are not reliable for delineation of geochemical anomalies in areas with complex geological units. Relationships among multiple major/trace elements and rock types are more important than single major/trace elements for delineation of geochemical anomalies. In this study we present an approach based on robust stepwise multiple regression using values major oxides (SiO_2, Al_2O_3, Fe_2O_3, MnO, and MgO) in stream sediments to predict elemental content related to rock types and to recognize geochemical anomalies. The major/trace element data were subjected to isometric logratio transformation to address the compositional data closure problem. For further examination of the stepwise regression method, its performance was compared to robust principal components analysis (RPCA), median + 2MAD and concentration-area (C-A) fractal methods. The results show that multi-element anomalies obtained by the stepwise regression method, compared to those obtained by the other methods, have stronger spatial association with the known deposits, such as Chichaklo and Ay-Ghale-Si in the Takab 1:25,000 scale geological map (NW) Iran, and the anomalies have stronger spatial correlation with structural features and prospects, and thus can be used as guides to new exploration targets.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Recognition of anomalous and background values in a stream sediment geochemical dataset is one of the basic tasks in mineral exploration. An anomaly can be defined as a concentration of element or metal that is greater than a threshold concentration value (i.e., upper limit of background population). Stream sediments are composite materials derived from the weathering and erosion of one or more sources upstream of a sample site. Therefore, uni-element contents of stream sediments are derived from multiple (usually background and rarely anomalous) sources. In most cases, a major proportion of variation in uni-element contents in stream sediment is due to lithological units underlying the areas upstream of stream sediment sample sites (Carranza, 2010b). Recognition of anomalies from background in a regional-scale stream sediment geochemical dataset is an important stage of mineral exploration to delineate potential areas for detailed investigation at finer scales (Deng et al., 2010; Nazarpour et al., 2015c; Pazand et al., 2011; Rantitsch, 2000; Shamseddin Meigoony et al., 2014; Rezaei et al., 2015).

Various statistical methods have been used to process geochemical data in order to determine threshold values. Statistical quantities, such as the mean, standard deviation (sdev) and percentiles, have been used to define threshold for separating anomalies form background. For example, geochemical anomalies have been defined as values greater than a threshold defined as the 75th or 85th percentile, and mean + 1sdev or mean + 2sdev. Based on such statistical quantities, there are two main groups of methods for determining threshold values: the first group includes frequency-based univariate methods such as mean ± 2sdev (Hawkes and Webb, 1962), probability graphs, box-plot and Q-Q plot (Govett et al., 1975; Miesch, 1981; Sinclair, 1976; Stanley and Sinclair, 1989) and second group includes variance-based multivariate methods to define anomalous multi-element associations (e.g., Aitchison, 1986; Nazarpour et al., 2015a and b).

The application of a single uni-element threshold value, defined by frequency-based methods, to delineate anomalies often results in false-negative anomalies in areas with low background values or false-positive anomalies in areas with high background values, thereby undermining the utility of geochemical exploration to define new targets. A reasonable way to solve this problem is to determine the

---

* Corresponding author.
E-mail addresses: Ahad.nazarpour@gmail.com, A.nazarpour@ahvaziau.ac.ir
(A. Nazarpour).

http://dx.doi.org/10.1016/j.gexplo.2016.07.003
0375-6742/© 2016 Elsevier B.V. All rights reserved.

underlying relationships among geochemical anomalies and plausible causative geological processes. Using rock types to represent geological processes is a common logical solution tool to recognize geochemical anomalies in complex geological settings or in overburden-covered areas (Hao et al., 2014), because there is often a clear relationship between rock types and major oxide content of rocks (Cohen et al., 2012; Reimann and Garrett, 2005). For elimination of lithological effects on uni-element data, one can use a multiple regression model to estimate element values and then subtract these values from measured element values to yield *geochemical residuals* that may or may not be related to anomalous sources (Bonham-Carter and Goodfellow, 1986; Hao et al., 2014).

Another main limitation of the above-mentioned methods is that they do not take into account the variability of spatial-statistical distribution of geochemical data. However, different areas can differ in rock compositions or have experienced different geological processes, which result in different geochemical thresholds. Therefore, the above-mentioned methods are of limited use in situations where there are extensive overlaps between background and anomalous values, or where weak anomalous values are hidden within the strong variance of background (Cheng, 2007).

The spatial-statistical distribution of geochemical data can be characterized using fractal geometry (Mandelbrot, 1983), which is a branch of non-linear mathematics that has been widely applied in the geosciences (e.g., Afzal et al., 2010; Agterberg et al., 1993; Ali et al., 2007; Carranza, 2008; Cheng et al., 1994; Deng et al., 2010; Sim et al., 1999; Turcotte, 1986; Wei and Yang, 2010). Several fractal and multifractal models, including concentration-area (C-A) (Cheng et al., 1994, Nazarpour et al., 2015a), spectrum-area (S-A) (Cheng, 2004; Cheng et al., 2000; Xu and Cheng, 2001), concentration-distance (C-D) (Li et al., 2003), concentration-volume (C-V) (Afzal et al., 2011) and number-size (N-S) (Agterberg, 1995; Deng et al., 2010; Mandelbrot, 1983; Turcotte, 2002; Wang et al., 2010), have been developed for various applications in the geosciences including analysis of geochemical data.

In addition to the above-mentioned main limitations of frequency- and variance-based methods for anomaly recognition, geochemical data are compositional (i.e., contribution of parts to some whole), which carry exclusively relative information (Aitchison, 1986). For example, if the SiO_2 content of an igneous rock is 68% of the whole weight, then the value of MgO will only be equal to or <32%. This means that geochemical data are not absolute values, but only provide relative information of certain element in a whole sample (Aitchison, 1986). Therefore, compositional data represent a closed number system and should be opened prior to understanding of realistic relationships among compositions (Filzmoser et al., 2009; Carranza, 2011; Nazarpour et al., 2015b). Therefore, it is crucial to apply an appropriate transformation to geochemical data prior to using any method of multivariate analysis. The log-ratio (logarithm of a ratio) transformation methodology proposed by Aitchison (1986) represents a powerful set of techniques to open compositional data. Three log-ratio transformations have been proposed for opening of compositional data: (1) additive log-ratio (alr) transformation (Aitchison, 1986); (2) centered log-ratio (clr) transformation (Aitchison, 1986) and (3) isometric log-ratio (ilr) transformation (Egozcue et al., 2003). These transformations allow for the application of standard statistical methods to transformed data, although with some limitations or modifications. In this study, the stream sediment geochemical data were ilr-transformed prior to statistical analysis.

Finally, exploration geochemical data typically comprise a large set of geochemical variables (e.g., major oxides, trace elements/metals) and the choice of which of these variables can be used as predictor (or independent) and response (or target) variables is a common problem in attempting to describe relationships among such variables through regression analysis. However, considering that lithology is a major source of variation of trace elements/metals and that lithological units are composed of various major oxides (e.g., SiO_2, Al_2O_3, Fe_2O_3, MnO, and MgO), major oxides are typically used as predictor variables in lieu of lithological units. Determining the most significant predictor variables, and therefore reliable estimates of element values, can be achieved through stepwise multiple regression.

This paper focuses on the identification of geochemical anomalies in the Takab 1:25,000 scale geological map sheet by using stream sediment geochemical data to derive geochemical residuals of trace elements, which would indicate areas of enrichment (e.g., due to mineralization) or depletion. Previous researches have derived geochemical residuals by applying stepwise multiple regression using trace elements as dependent variables (or targets for exploration) and areal proportions of lithologic units as independent variables (predictors) or as proxies of the influence of lithology on trace element background concentration (Bonham-Carter and Goodfellow, 1984, 1986; Carranza and Hale, 1997; Moon, 1999; Carranza, 2010a,b). However, we argue that using areal proportion of lithologic units as predictors would depend on the availability of a geological map and the results would vary depending on the scale of the lithologic map used. In this paper, we used major elements as independent variables (predictors) or as proxies of the influence of lithology on trace element background concentration. Results from this proposed methodology are validated using lithogeochemical data and by comparing with results from using the median + 2MAD for exploratory data analysis (EDA), concentration- area (C-A) fractal model and robust principal component analysis (RPCA) as two effective approaches to separate geochemical anomalies from background in stream sediment geochemical compositional data.

2. Methods

2.1. Exploratory data analysis (EDA)

In EDA of geochemical exploration data, the median + 2MAD value was originally used to identify extreme values and act as threshold for further inspection of large data sets (Hawkes and Webb, 1962; Zheng et al., 2014). The EDA was first established by Tukey (1977), was developed further by Kürzl (1988), and then was used by many researchers in modeling of geochemical anomalies (e.g., Ali et al., 2007; Carranza, 2008, 2010a,b; Nazarpour et al., 2014). The *MAD* is the median of absolute deviations of individual dataset values (X_i) from the median of all dataset values (Tukey, 1976):

$$MAD = median|X_i - median(X_i)|. \quad (1)$$

2.2. Concentration-area (C-A) fractal model

The C-A fractal model was first introduced by Cheng et al. (1994) for recognition of geochemical anomalies from background. It has the following general forms (Cheng et al., 1994):

$$A(\rho \le \upsilon) \propto \rho^{-a_1}; A(\rho > \upsilon) \propto \rho^{-a_2} \quad (2)$$

where $A(\rho)$ denotes area with background concentrations (ρ) less than or equal to a threshold concentration (υ) or area with anomalous concentrations (ρ) greater than the threshold concentration (υ), a_1 and a_2 are slopes of straight lines fitted to log-log plots of ρ versus $A(\rho)$.

Cheng et al. (1994) proposed two approaches to calculate $A(\rho)$: (1) the $A(\rho)$ is area enclosed by a contour of concentration value (ρ) on a geochemical map derived by interpolation of the original concentration values using a weighted moving average method; and (2) $A(\rho)$ is obtained by application of the box-counting method to the original concentration values. Distinct patterns, each corresponding to a set of

Download English Version:

https://daneshyari.com/en/article/4456949

Download Persian Version:

https://daneshyari.com/article/4456949

Daneshyari.com