Journal of Geochemical Exploration 141 (2014) 42–51

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

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

Towards compositional geochemical potential mapping

Raimon Tolosana-Delgado [a,*], K. Gerald van den Boogaart [a,b]

[a] Helmholtz Institute Freiberg for Resource Technology, Halsbrücker Str. 34, 09599 Freiberg, Germany
[b] Institute for Stochastics, TU Bergakademie Freiberg, Prueferstrasse 9, 09599 Freiberg, Germany

ARTICLE INFO

Article history:
Received 15 July 2013
Received in revised form 17 February 2014
Accepted 20 February 2014
Available online 3 March 2014

Keywords:
Geostatistics
clr
Variation-variograms
Centered log-ratio transformation
Balance

ABSTRACT

Mineral potential mapping aims at defining target zones of future local mineral exploration efforts on the basis of regional geochemical and geological surveys. Geochemical potential mapping typically involves using a small set of elements to predict an anomalous presence of a mineral commodity related with them. In this context it is common to work with logarithmically transformed concentrations of elements. This contribution explores a compositionally compliant approach to potential mapping using geochemical evidences, in which the whole set of components is applied a log-ratio transformation before any potential mapping technique is used. In this way, candidate zones can be identified by its high value in a certain ratio of elements, implying that the important information is a contrast between two or more elements, and not an absolute concentration of one of them. Two different potential mapping techniques are used (a method equivalent to the Fisher rule, and a Poisson point process, both accounting for spatial dependence), with different results.

Introduction

Geochemical surveys are common tools in mineral exploration, in several phases and at different spatial resolutions. One of the first steps is potential mapping, where a large district is explored for potential new mining areas. Usually, potential mapping infers some favorability factors from existing deposits in the district (e.g., geochemical or geophysical signature, lithological units and structural elements); these factors are then mapped and they are reconverted to a probability or favorability location by location.

Geochemical information is most commonly used in such favorability mapping after a log-transformation, to account for the common assumption that on large scale, elemental concentrations follow a lognormal law (Ahrens, 1954). Also fractal arguments support that choice (Zuo et al., 2012). Moreover, such mapping is often done by logistic regression techniques (Schaeben, 2012; Singer and Menzie, 2008), as well as with fractal theory (Zuo et al., 2012), singularity mapping (Cheng, 1997, 2007), etc. Most commonly, one element is used to map occurrences of its commodities (e.g., Sn is mapped to search for cassiterite deposits), or sometimes some concomitant elements are used (e.g., As-Ag for gold–silver deposits). To the authors' knowledge, however, no attempt has been made to relate the potential to *low* values of other elements, or to logratios or balances between elements. In general, such logratios will be critically informative when the proxy elements are affected by dilution trends that mask the typical geochemical signature. Fig. 1 shows two hypothetical geological scenarios with a mineral anomaly body emplaced around a faulted area, one of them (right) tilted and buried. If the burial produces a dilution trend on the typical geochemical signature of this body, analysis based on the raw concentrations or logarithmic transformed ones will probably not see the anomaly. On the other hand, an anomaly indicator based on a logratio balance may filter the dilution out, and pinpoint the orebody.

The concept of balance appears in compositional data analysis, and is a way to extract all relevant information from a compositional data set. A compositional data set is a set of variables which inform of the relative importance of a set of parts that constitute a total. In this contribution, the term *part* will be specifically used to denote a compositional variable, while a *variable* will be of an unspecified nature.

Geochemical surveys always deliver compositional information, though this has been seldom taken into account. Theoretical and practical reasons suggest that compositional data should first be applied a log-ratio transformation, prior to any further treatment. Thus, it appears natural to explore the application of conventional potential mapping techniques to a log-ratio transformed geochemical survey.

This paper presents a simple geochemical potential mapping procedure adapted to the compositional nature of this kind of data. The second section presents these characteristics and the basics of the statistical treatment of compositional data. The third section presents briefly a geostatistical procedure to map compositions using these rules. The fourth section develops a geochemical potential mapping procedure loosely based on the Fisher discriminant rule, while the fifth section does the same using a Poisson process model. Both methods account for the spatial dependence between the samples. Finally, the sixth section presents an example where a stream sediment geochemical survey of the Grazer Paleozoikum is used to map potentiality for new Pb–Zn–Ba deposits. Some final remarks and conclusions are given in the last section.

* Corresponding author.
E-mail address: r.tolosana@hzdr.de (R. Tolosana-Delgado).

http://dx.doi.org/10.1016/j.gexplo.2014.02.026
0375-6742/

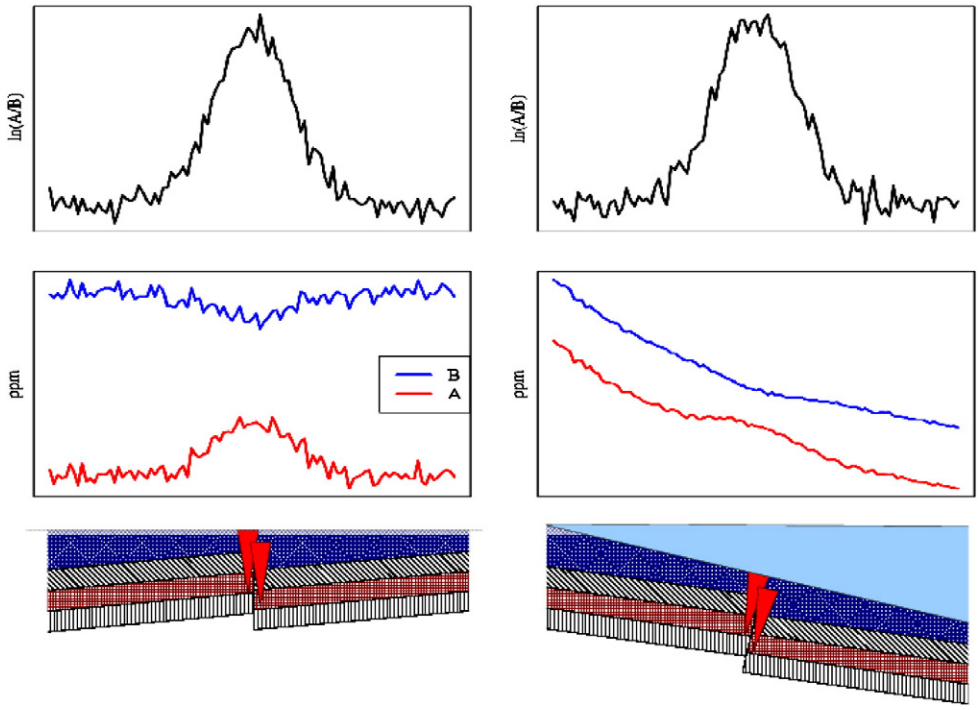

Fig. 1. Schematic illustration of a situation in which the use of logratios for potential mapping may become critical. Geological sketches (bottom), raw geochemical signature of two elements (center) and logratio signature (top).

Basics of compositional data analysis

A data set is considered of compositional nature if its variables describe the relative importance of some components forming a whole. Typically, this relative importance is described in % or other proportion unit (parts in one, ppm, ppt, etc.). Geochemical data sets are archetypical examples of compositional data, where the whole set of parts describes how an ideal 100 g of mass (of soil, of a certain soil layer, of tissues of moss, or whatever support was analyzed) should be split among the elements of the periodic system. Such data sets are formed by non-negative parts (as a negative proportion of one element is impossible), and at each datum, all parts should sum up to 100% or less (as a larger sum would imply that 100 g of total mass contains more than 100 g of chemical elements).

If a composition $\mathbf{z} = [z_1, z_2, \ldots, z_D]$ has D components, these conditions can be mathematically stated as $z_i \geq 0$ and $\sum_{i=1}^{D} z_i \leq 100$ %. If we identify this vector of D numbers with a point in the D-dimensional real space R^D, the set of points satisfying these conditions, and thus being possible compositions, is called the D-part simplex, denoted S^D.

Chayes (1960) was the first geochemist to call the attention to the fact that correlations within a compositional data set are fundamentally spurious, related to the constant sum constraint, and not to any natural process. After his warning, a series of papers appeared discussing the several influences of this spurious correlation on many of the most common statistical methods (Butler, 1976; Darroch, 1969; Pawlowsky-Glahn, 1984).

Aitchison (1982) was the first to identify the source of these problems, namely the fact that compositions convey only relative information. This can be mathematically stated by the principle of *scale invariance*: if a vector $\mathbf{z}$ conveys only relative information, then the results of a correct analysis should be the same, whether one analyzes $\mathbf{z}$ or $p\mathbf{z}$, for any constant positive scaling factor p.

A second important condition is *subcompositional coherence*: as an analysis never involves all possible components, it is required that results do not depend on which parts have not been taken into account; or, in other words, if some components are random, results with respect to the other non-random components should be stable, independently on whether the random components were considered in the analysis or not. Classical Pearson correlation coefficients are not subcompositionally coherent, as Chayes (1960) showed.

To comply with these two conditions, ratios of components should be analyzed, instead of the raw components. And, due to mathematical reasons, it is better to analyze logratios than ratios (Aitchison, 1997). Several families of logratios have been proposed as standard tools, from which we use here three: the pairwise logratio transformation, the centered logratio transformation and the isometric logratio transformation. A pairwise logratio transformation expresses a composition with the $D(D-1)$ possible pairwise logratios $\zeta_{ij} = \ln(z_i/z_j)$. The centered logratio transformation (Aitchison, 1986) computes D scores out of the same number of components as $\zeta_i = \ln(z_i/g(\mathbf{z}))$, by using the geometric average of all of them $g(\mathbf{z}) = \sqrt[D]{\prod_j z_j}$. This can be written in matrix notation as

$$\mathrm{clr}(\mathbf{z}) =: \zeta = \mathbf{H} \cdot \ln \mathbf{z}, \quad \mathbf{H} = \mathbf{I}_D - \frac{1}{D} 1_D \cdot 1_D^t,$$

where the logarithm should be applied component-wise, the matrix $\mathbf{I}_D$ is the $D \times D$ identity matrix and 1_D is a (column-)vector of D ones. The inverse of the clr transformation is simply

$$\mathrm{clr}^{-1}(\zeta) =: C[\exp(\zeta)] = \mathbf{z}, \tag{1}$$

where the closure operation

$$C(\mathbf{z}) = \frac{100}{1_D^t \cdot \mathbf{z}} \mathbf{z}$$

forces the argument to a constant sum of 100% without changing the ratios between the components.