Journal of Geochemical Exploration 139 (2014) 207–216

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

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

Estimating the average concentration of minor and trace elements in surficial sediments using fractal methods

Tuhua Ma [a], Changjiang Li [a,*], Zhiming Lu [b]

[a] Zhejiang Information Center of Land and Resources, 310007 Hangzhou, China
[b] Computational Earth Science Group (EES-16), Los Alamos National Laboratory, Los Alamos, NM 87545, USA

ARTICLE INFO

Article history:
Received 29 January 2013
Accepted 5 August 2013
Available online 14 August 2013

Keywords:
Fractal distribution
Fractal averaging
Minor and trace elements
Censored data
Environmental geochemistry

ABSTRACT

The methods chosen to calculate the average value of the concentration for any geochemical element should depend on the probability distribution of the element abundance data. In this study, a fractal-based method was introduced to estimate the mean concentrations of geochemical elements that follow fractal frequency distributions. The fractal-based method has been tested on two abundance datasets for Ag, As, Au, Cu, Pb, Zn, Ce, Cr, and U from 529 floodplain sediment samples in China and from 10,927 stream sediment samples in Zhejiang Province, China. We compared the fractal method with other methods, including the arithmetic averaging, geometric averaging, and median, and found that there exist large discrepancies among these averages. The results show that the average calculated using the fractal-based method is always smaller than the arithmetic average and also generally smaller than the geometric mean and the median. The discrepancies may be attributed to the fact that the datasets follow a fractal distribution rather than a normal or a lognormal distribution. This study indicates that calculated arithmetic mean, geometric mean, or median may overestimate the average concentrations for elements that follow a fractal distribution.

1. Introduction

Estimating abundances of geochemical elements in the Earth's crust is always a challenge for geochemists, and it has attracted the attention of geochemists for at least 100 years. Element abundances are pivotal background values for exploration of mineral resources and determination of environmental pollution levels. In all existing models for estimating chemical composition of the crust (e.g., Clarke, 1889; Clarke and Washington, 1924; Gao et al., 1998; Goldschmidt, 1933; Taylor, 1964; Taylor and McLennan, 1985; Wedepohl, 1995), element abundances were derived from the averages of the compositions of surface exposures. The key difficulties in deriving element abundances include: (1) the tremendous geochemical heterogeneity of the crust, which calls for a need to devise a method for the generalization of particular data (Yaroshevskii, 2007); and (2) the reliability of the estimated mean concentration. Sediments from floodplains (Darnley et al., 1995; Xie and Cheng, 1997) and from continental river discharges (Yaroshevskii, 2007) have been considered as an “average sample” for materials of the crust exposed on erosion surfaces of continents, to overcome the first difficulty. However, the second problem still exists so far. Based on the assumption that concentrations in Earth's rocks and sediments follow normal distributions, average concentrations of geochemical elements were calculated by arithmetic averages (Rock, 1988). It has been found (Ahrens, 1954a,b) that many elements, especially trace elements, do not follow a normal distribution, but instead show a skewed or a tailed distribution. Calculating the arithmetic mean for skewed data will result in a biased (over) estimate of the central value (Filzmoser et al., 2009a). In this case, data are usually transformed by taking their logarithms, and then the normal model is used to calculate the geometric mean for the dataset. However, in many cases, a geometric mean is not appropriate because logarithms of data may still exhibit a heavy skewed or a tailed distribution (Chapman, 1976; Iqbal and John, 2010). Filzmoser et al. (2009a), based on the concept of compositional data analysis, suggested that when using an ilr-transformation for original data, it is possible to transform the computed arithmetic mean of ilr-transformed data back to the original data scale. However, as shown in Fig. 3 of Filzmoser et al. (2009a), the distribution of Na_2O is left-skewed; and it remains heavily left-skewed after ilr-transformation.

The heavy skewed or tailed distribution mentioned above, together with similar empirical discovery in many other application fields, has resulted in the formulation of the fractal theory (Mandelbrot, 1983). Original measured data from geochemical surveys can be regarded as a height field defined over a certain domain. The pattern of a height field is termed as geochemical landscape (geochemical surface). For many minor or trace elements, the geochemical landscapes, like the length of the famous Mandelbrot's coastline that commonly vary with the scale of observation (sampling density), are undetermined (Li et al., 2002, 2004), that is, in a given area, the denser one takes samples, the more details one can obtain. This implies that minor or

* Corresponding author.
E-mail addresses: zjigmr@mail.hz.zj.cn (C. Li), Zhiming@lanl.gov (Z. Lu).

0375-6742/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gexplo.2013.08.008

trace element abundance data probably follow fractal (power-law) distributions. In fact, many studies have provided strong support for the hypothesis that the distribution of minor or trace element abundance data is fractal (e.g., Allègre and Lewin, 1995; Bölviken et al., 1992; Cheng et al., 1994; Li et al., 2002, 2003, 2004; Lima et al., 2003).

If the abundance of a geochemical element follows a fractal distribution, the average value should be estimated with a fractal-based method rather than conventional methods such as arithmetic averaging or geometric averaging.

In this study, fractal averaging, a method based on fractal frequency distributions, is introduced to calculate the average concentration of geochemical elements. Hereinafter, the mean derived from the fractal averaging is called "fractal mean". This fractal-based method has been tested on two abundance datasets for Ag, As, Au, Cu, Pb, Zn, Ce, Cr, and U from 529 floodplain sediment samples that cover nearly most of the land surface of China and from 10,927 stream sediment samples collected in Zhejiang Province, China. We compared the fractal method with other methods, including the arithmetic averaging, geometric averaging, and median, and found that there exist large discrepancies among these averages.

2. Methodology

In general, a fractal (power-law) distribution is of the form

$$p(x) = Cx^{-\alpha}, \tag{1}$$

where $p(x)$ is the number of objects with size x, and C and α are constants, which can be determined from a dataset. The scaling exponent α could be a fraction and is usually called the fractal dimension. The density function diverges as x approaches zero, so there must be some low bound (denoted as $x_{\min}$) for this distribution. The normalization constant C can be found from the constraint $\int_{x_{\min}}^{\infty} p(x)\mathrm{d}x = 1$, i.e., $C = (\alpha - 1)x_{\min}^{\alpha-1}$, and the density function becomes

$$p(x) = \frac{\alpha-1}{x_{\min}}(x/x_{\min})^{-\alpha} \tag{2}$$

In many practical applications, one of the methods to study data is to calculate the cumulative distribution function (Newman, 2005). The probability $P(x)$ that the size X has a value greater than x is

$$P(x) = P(X>x) = \int_x^{\infty} p(x')\mathrm{d}x'. \tag{3}$$

If the distribution is fractal, substituting (2) into (3) and integrating it yields

$$P(x) = (x/x_{\min})^{-(\alpha-1)}. \tag{4}$$

Thus the cumulative distribution function $P(x)$ also follows a fractal distribution, but with a flatter slope (scaling exponent), which is 1 less than the exponent obtained from Eq. (1). The cumulative distribution has an advantage in that it can reduce statistical fluctuations without losing any information (Newman, 2005).

By definition, the ensemble mean for the fractal distribution can be derived as

$$\langle x \rangle = \int_{x_{\min}}^{\infty} xp(x)\mathrm{d}x = \frac{\alpha-1}{2-\alpha}x_{\min}\left(\frac{x}{x_{\min}}\right)^{-\alpha+2}\Bigg|_{x_{\min}}^{+\infty}. \tag{5}$$

It can be seen from Eq. (5) that the mean of the fractal distribution depends on the scaling exponent α, which leads to several scenarios. If

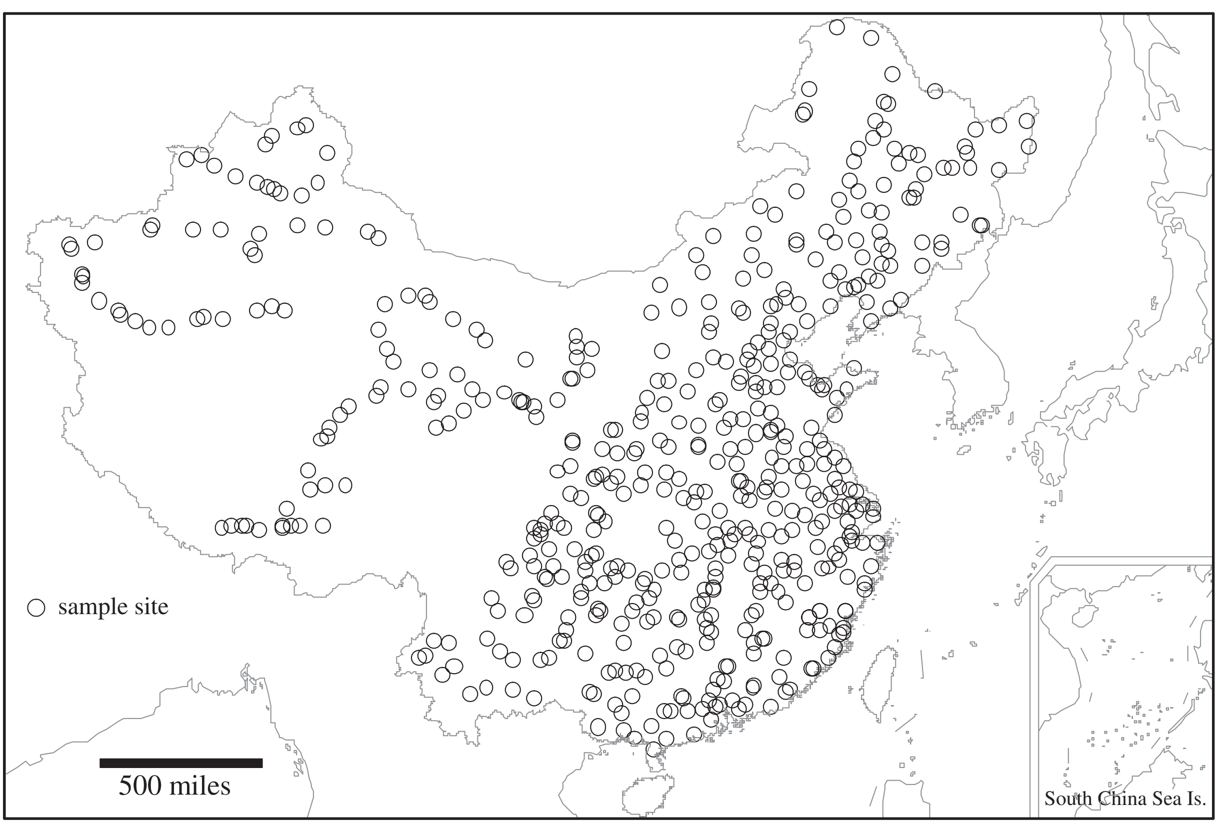

Fig. 1. Sampling locations of floodplain sediments for a wide-spaced geochemical survey in China.