

New Applications of the Model of de Wijs in Regional Geochemistry¹

Frits Agterberg²

This paper is concerned with the lognormal, and its logbinomial approximation, in connection with a three-parameter version of the model of de Wijs. The three parameters are: overall average element concentration value (ξ), dispersion index (d), and apparent number of subdivisions of the environment (N). Multifractal theory produces new methods for estimating the parameters of this model. In practical applications, the frequency distribution of element concentration values for small rock samples is related to self-similar spatial variability patterns of the element in large regions or segments of the Earth's crust. The approach is illustrated by application to spatial variability of gold and arsenic in glacial till samples from southern Saskatchewan. It is shown that for these two elements the model of de Wijs is satisfied on a regional scale but degree of dispersion decreases rapidly toward the local, sample-size scale. Thus the apparent number of subdivisions (N) is considerably less than would be expected if degree of dispersion were to extend from regional to local scale. A random-cut variant of the model of de Wijs produces an empirical frequency distribution of relative element concentration values that can be related to random dispersion index.

KEY WORDS: lognormal distribution, model of de Wijs, multifractal modeling, mass exponents, computer simulation experiments, till samples, southern Saskatchewan.

INTRODUCTION

In large-scale regional environmental or natural resource evaluations it can be important to extrapolate from a relatively small set of element concentration values for samples taken at the Earth's surface to much larger areas or volumes of the Earth's crust. Mineral resource applications of this type were reviewed by Harris (1984). The purpose of this modeling is to consider the frequency distributions of average element concentration values for large blocks in two- or three-dimensional space. Statistical parameters to be estimated include the mean and the block variances. Blocks of different sizes can have the same mean but have different variances. Usually the problem of interest is to learn how many

¹Received: 1 August 2005; accepted: 21 April 2006; Published online: 21 February 2007.

²Geological Survey of Canada, 601 Booth Street, Ottawa, K1A 0E8 Canada; e-mail: agterber@NRCan.gc.ca

large high-concentration blocks there are in a study region without knowing their locations.

In general, the overall average concentration value for a study region is estimated from the original data from within this region or it is taken from a table of global crustal abundance figures. The lognormal frequency distribution often is useful to describe the variability. The lognormal has shape-permanency in that it applies to blocks of different sizes with the same overall mean and spatial covariance structure (Matheron, 1962, 1974; Agterberg, 1984).

During the past 20 years, significant advances have been made in multifractal analysis (see e.g. Lovejoy and Schertzer, 1991; Schertzer and others, 1997; Cheng, 1994). For a general overview of self-similar multifractals, see e.g. Falconer (2003). Multiplicative cascade models help to understand how multifractals can be generated. Such models have relatively simple properties of similarity. Self-similarity means that the process that underlies the generation of a complex multifractal pattern is independent of scale. The generating mechanism proposed by de Wijs (1951) is an example of a multiplicative cascade model that results in approximately lognormal distributions for blocks of different sizes. De Wijs (1951) assumed self-similarity but developed his model before multifractal theory came into existence. It was Mandelbrot (1983) who pointed out that the model of de Wijs is an early example of a multifractal. In this paper it will be shown that methods of multifractal analysis can be used in connection with earlier lognormal statistical approaches to natural resource evaluation. This helps to obtain better estimates of the parameters of the frequency distributions.

The model of de Wijs (1951) applies to self-similar environments where iterative subdivision of larger blocks into smaller blocks results in increasingly different element concentration values. When there is scale-independence, this process would be controlled by a single dispersion index d that remains the same regardless of the size of the block that is being subdivided into parts. For convenience, d is linked to dividing blocks into halves. However, the approach can be generalized for other types of subdivisions into two or more parts (*cf.* Evertsz and Mandelbrot, 1992). It is also possible consider d to be the realization of a random variable D^* (random-cut model to be discussed later in this paper).

There are many geoscience examples of self-similarity. One of these is as follows: in a study based on approximately 75,000 gold determinations for very small blocks, Krige (1966) found an approximate straightline relationship between logarithmic variance and logarithmically transformed size of reef area for gold and uranium values in the Klerksdorp goldfield, South Africa. A linear relationship of this type is to be expected when the model of de Wijs is valid. Size of reef area ranged from 0.1 to 1,000 million square feet. Consequently, in Krige's application, scale-independence applied to square cells with side lengths extending over 5 orders of magnitude.

Krige (1966, Fig. 3) illustrated this concept further by means of gold value contour maps for square areas measuring 500 ft and 10,000 ft on a side. The same moving averaging method was used to construct these gold contour maps, which exhibit similar patterns. Krige's example of self-similarity is based on a very large number of mining assays. Other studies commonly are based on much smaller datasets.

If ξ represents average concentration value of a chemical element in a very large block of rock (e.g. the upper part of the Earth's crust), division of this block into halves results in two smaller blocks with element concentration values equal to $(1 + d)\xi$ and $(1 - d)\xi$, respectively. The ratio of these two different values is $\eta = (1 + d)/(1 - d)$. After n successive subdivisions (for large n), there are 2^n small blocks with concentration values ranging from $(1 - d)^n \xi$ to $(1 + d)^n \xi$. The difference between logarithmically (base e) transformed higher and lower values resulting from subdividing any block into halves is $\ln \eta$. The frequency distribution of the final set of 2^n concentration values is logbinomial. This term was originally used by Matheron (1974). It simply means that the logarithmically transformed values have binomial distribution. When d is relatively large, the logbinomial becomes positively skewed with a thin, long high-value tail. The lower and higher values generated at successive iterations can have random spatial locations with respect to one another.

The model of de Wijs can be used as a starting point for parametric spatial modeling. De Wijs (1951) already derived an equation for the variance of logarithmically transformed element concentration values. It can be written in the form

$$\sigma^2 = \frac{n}{4}(\ln \eta)^2 \quad (1)$$

where σ^2 is the logarithmic variance (base e) of the concentration values. Eq. (1) follows directly from the well-known equation for the variance of a binomial distribution with $p = 1/2$ that is equal to $n/4$, taking account of the fact that the spacing between ordered values along the logarithmic scale is equal to $\ln \eta$. This result will play an important role in this paper because it allows estimation of n from the logarithmic variance after estimation of η by another (multifractal) method. Eq. (1) can be rewritten in the form (cf. Matheron, 1962, p. 309):

$$\sigma^2 = \alpha \cdot \ln \frac{V}{v} \quad (2)$$

where v represents the volume for which the concentration value is determined; V is a larger volume in which v occupies a random position. In one-dimensional applications, the volumes v and V are reduced to line segments, and $V/v = 2^n$.

The constant α satisfies

$$\alpha = \frac{1}{4 \ln 2} [\ln \eta]^2 \quad (3)$$

The organization of this paper is as follows. First, several one-dimensional variants of the model of de Wijs will be illustrated by means of computer simulation experiments. In principle, there are infinitely many ways by means of which a block of rock can be divided into halves. Every subdivision into halves would result in a different value of d (cf. Matheron, 1962). The discrete logbinomial model can be modified into a random-cut model where D^* is a random variable with expected value $E(D^*) = 1$, resulting in different values of d for each subdivision of a block into halves. This generalization of the model of the Wijs results in an empirical method of estimating the frequency distribution of random dispersion index and will be applied later in the paper.

Next, Brinck's geochemical crustal abundance model will be discussed (Brinck, 1974). This model constitutes one of the main earlier applications of the model of de Wijs. It will be shown that Brinck's approach remains applicable but that estimation of the parameters of the model of de Wijs including d can be significantly improved by adopting a multifractal modeling approach.

Finally, the new approach will be illustrated by application to gold and arsenic content of till samples obtained during systematic ultra-low density (1 site per 800 km²) geochemical reconnaissance, Southern Saskatchewan. A summary of results for these examples of application was presented in Agterberg (2005).

COMPUTER SIMULATION EXPERIMENTS

The multiplicative cascade model to be used first is one of the simplest multifractal self-similar models in existence (cf. Cheng and Agterberg, 1996). It has been discussed by many authors including Feder (1988), Evertsz and Mandelbrot (1992), and Falconer (2003). More flexible versions of cascade dynamics are reviewed in Lovejoy and Schertzer (1991), and Schertzer et al. (1997). The multiplicative cascade approach is introduced here because it helps to understand the nature of multifractal models. A one-dimensional computer simulation experiment for it was previously described in Agterberg (1994), and a two-dimensional application in Agterberg (2001a). The number of successive iterations (n) can be odd or even in 1-dimensional applications but is necessarily even in 2-dimensional applications, because otherwise the average concentration value would become biased. The objects for which average element concentration values are considered in this paper are 3-dimensional. However, in 1- or 2-dimensional applications, these objects also are referred to as intervals or cells.

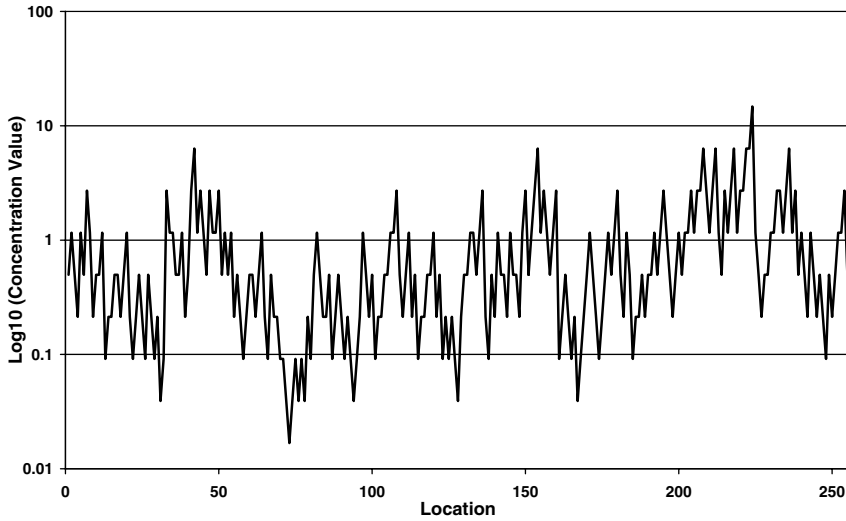


Figure 1. Artificial sequence of 256 values generated by version of model of de Wijs described in the text. Dispersion index $d = 0.4$; overall mean $\xi = 1$.

A line segment of length L can be partitioned into $N(\varepsilon)$ cells (intervals) of equal size ε . Let $\mu_i(\varepsilon)$ denote the mass of a quantity measured for the i th cell of size ε in $(0, L]$ with $i = 1, 2, \dots, 1/4, N(\varepsilon)$. A simple stochastic version of the multiplicative cascade model in $\mathfrak{R}^1$ then is as follows. At the first stage ($k = 1$) in a process of n stages, the interval $(0, L]$ with mass ξL is subdivided into two equal intervals: (a) $(0, L/2]$ with mass $(1 + D)\xi L$, and (b) $(L/2, L]$ with $(1 - D)\xi L$, where D is a random variable with $\text{pr}(D = d) = \text{pr}(D = -d) = 1/2 (d > 0)$. At stage 2 these two intervals are halved again with new masses for the halves defined in the same way as at stage 1. The process is repeated at stages $k = 3, 4, \dots, 1/4$. At stage k the i th subinterval with concentration value $X_i(\varepsilon)$ has size $\varepsilon = L/2^k$, and $E\{X_i(\varepsilon)\} = \xi$. The frequency distribution of $X_i(\varepsilon)$ is logbinomial, tending to become lognormal in the center as $\varepsilon \rightarrow 0$ and, depending on the direction of ordering, slightly weaker than lognormal in both tails.

Figure 1 shows a realization of this process for $n = 8$, $\xi = 1$, and $d = 0.4$. Figure 2 is a lognormal Q-Q plot of the $2^n = 256$ element concentration values that were generated. The straightline fits well in the centre but small departures in both tails illustrate that the logbinomial has slightly thinner tails than the lognormal. The stochastic simulation algorithm used to obtain this result, and other simulation experiments to be described in this paper, are unconditional. Conditional versions of these algorithms that would honor existing measurements do not currently exist.

Suppose that $d = 0.4$ is replaced by a normal random variable D with expected value $E(D) = 0.4$ and variance $\sigma^2(D) = 0.01$. Thus every block divided

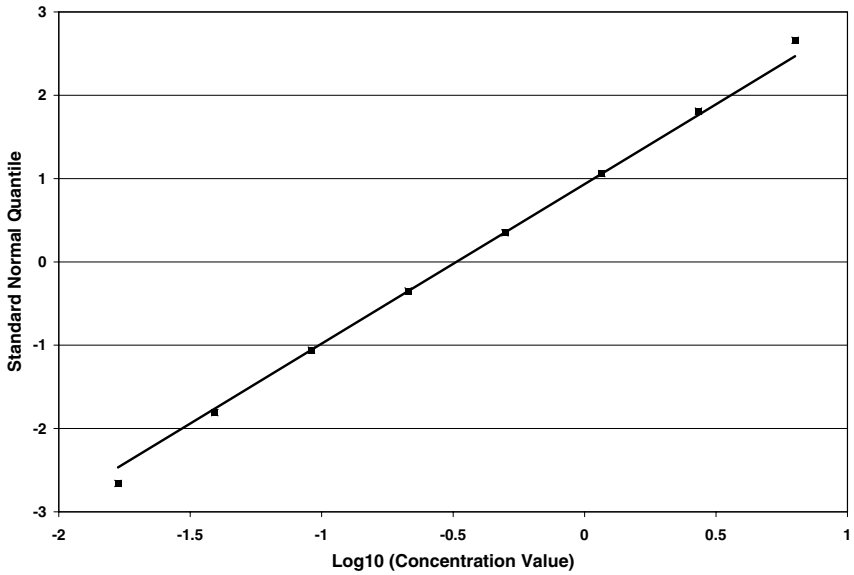


Figure 2. Cumulative frequency distribution (lognormal Q-Q plot) of 256 values of Figure 1. Solid squares represent discrete logbinomial distribution. Straightline is lognormal distribution fitted to central part of logbinomial. Note that tails of logbinomial are slightly thinner than lognormal tails.

into halves has a different value of d that is a realization of D . If, for example, $d = 0.3924$ for division of a block with concentration value X , then the two halves would have concentration values equal to $1.3924X$ and $0.6078X$, respectively. Without loss of generality, $E(X) = \xi$ can be set equal to 1 as before.

Figures 3 and 4 show a realization and a Q-Q plot for this process with $n = 8$, $\xi = 1$, $E(D) = 0.4$ and $\sigma(D) = 0.1$. This new computer simulation experiment produces results that are close to those for the discrete, logbinomial model of Figures 1 and 2. It could be argued that the example of Figures 3 and 4 is more realistic because fewer restrictions are imposed on the concentration values. However, this increased flexibility goes at the expense of the introduction of the new parameter $\sigma^2(D)$. Introduction of a new parameter can be avoided by using an approach in which $E(D)$ and $\sigma(D)$ are interconnected as shown in the next section.

RANDOM-CUT VARIANT OF THE MODEL OF DE WIJS

In the original model of de Wijs illustrated in Figures 1 and 2, the dispersion index d is a constant. The lower and higher values generated at successive iterations were assigned random spatial locations with respect to one another by introducing

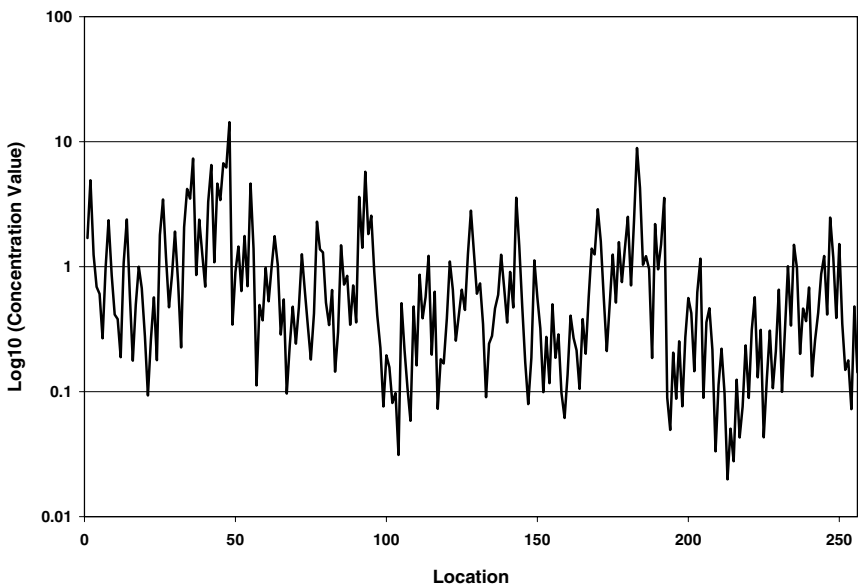


Figure 3. Computer simulation experiment of Figure 1 repeated after replacing constant dispersion index $d = 0.4$ by normal random variable D with $E(D) = 0.4$ and $\sigma(D) = 0.1$. Although 256 values differ from one another, overall pattern of spatial variability resembles pattern of Figure 1.

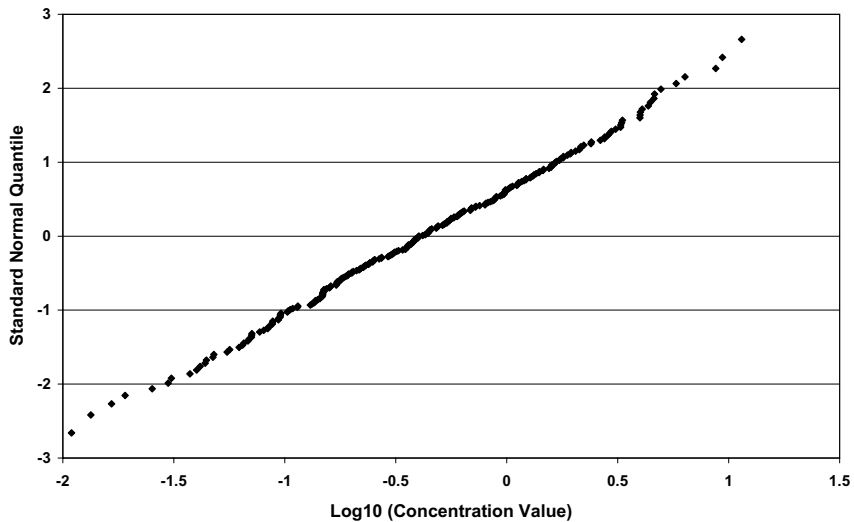


Figure 4. Lognormal Q-Q plot of 256 values of Figure 3. Pattern resembles straightline for lognormal in Figure 2.

the random variable D with $\text{pr}(D = d) = \text{pr}(D = -d) = 1/2 (d > 0)$. A different way of formulating this type of randomness is to introduce a new random variable $D^* = 1 + D$ that assumes either the value d or the value $-d$ with equal probabilities. Consequently, $E(D^*) = 1$. In the variant of the model of de Wijs resulting in Figures 3 and 4, D^* is bimodal. Its two separate peaks, which satisfy normal distributions with $\sigma(D) = 0.1$, are centered about 0.4 and 1.4, respectively. It would be more realistic to replace this bimodal model by a unimodal frequency density model centered about $E(D^*) = 1$.

We can also ask whether it would be possible to estimate the frequency distribution of D^* . The following idea was originally proposed by Matheron (1962, p. 310–311). It is based on the concept of relative concentration value $Y_{12} = X_1/X_2$ where X_1 represents concentration value of a small block with volume V_1 that is randomly located within a larger block with concentration value X_2 and volume V_2 . In our computer simulation experiments, V_1 is exactly half of V_2 but V_1 can be much smaller than half of V_2 .

If there is self-similarity, Y_{12} , on average, only depends on the ratio V_1/V_2 of the block volumes. It is independent of the volumes themselves. Suppose that smaller blocks are contained within larger blocks so that

$$V_1/V_2 = V_2/V_3 = \dots = V_n/V_{n-1} = \lambda.$$

Consequently, $V_n/V_1 = \lambda^{n-1}$. For the examples in $\mathfrak{N}^1$, $V_n/V_1 = 2^n$. All relative concentration values $Y_{i,i+1} (i = 1, \dots, n-1)$ are realizations of the same random variable, and $Y_{1p} = X_1/X_p = Y_{12}Y_{23} \dots Y_{n-2,n-1}$. Suppose that the logarithmic variance of $Y_{i,i+1}$ is written as h^2 , then σ^2 representing the logarithmic variance of Y_{1n} satisfies

$$\sigma^2 = (n-1)h^2 = \alpha^* \ln\{V_n/V_1\}, \text{ where } \alpha^* = h^2/\ln \lambda.$$

This expression is equivalent to Eq. (2) for the logbinomial model of the original model of de Wijs. However, the underlying rationale is different. We now assume $D^* = 1 + D$ to be a unimodal random variable with expected value equal to 1 and variance $\sigma^2(D^*)$. If D^* has normal frequency distribution, its mean deviation is equal to $\sigma\sqrt{2/\pi}$. This statistic would be equivalent to the constant dispersion index d in the original model of de Wijs.

Suppose that the mean deviation is set equal to the dispersion index $d = 0.4$ used for constructing Figures 1 and 2. Then $\sigma = 0.5013$. However, D^* can not be negative nor greater than 2. In the following computer simulation experiment, occurrence of very small or large values of D^* is prevented by setting random normal numbers less than 0.1 equal to 0.1, and values greater than 1.9 equal to 1.9.

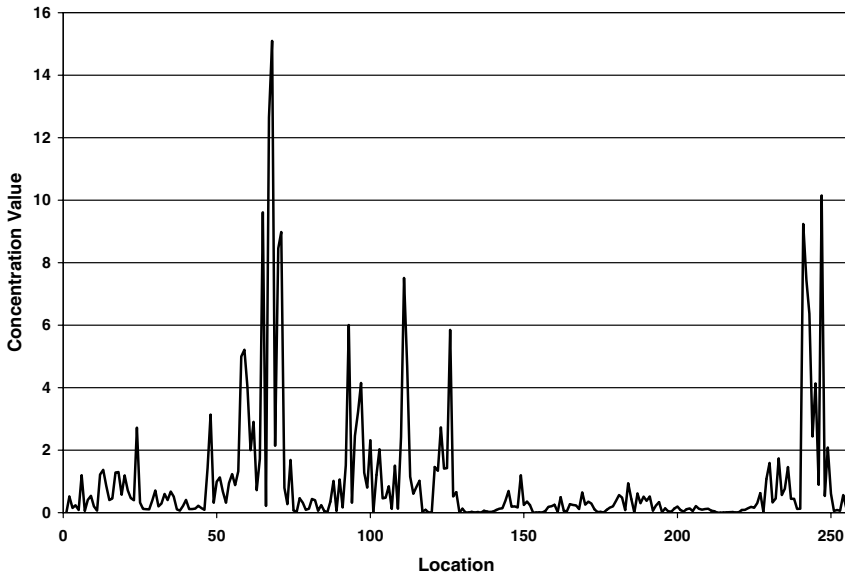


Figure 5. Computer simulation experiment of Figure 1 repeated after replacing constant dispersion index $d = 0.4$ by approximately normal random variable D^* with $E(D^*) = 1$ and $E\{|D^* - E(D^*)|\} = 0.4$, which is equivalent to $\sigma(D^*) = 0.51$ (see text).

The resulting approximately normal distribution has standard deviation slightly less than 0.5013. Replacement of this value by 0.51 results in a slightly different, approximately normal distribution with mean deviation equal to 0.40. Figures 5 and 6 show results for this random-cut model with $n = 8$, $\xi = 1$, $E(D) = 0.40$ and $\sigma(D) = 0.51$. Arithmetic instead of logarithmic scale for concentration values is used in Figure 6.

The preceding computer simulation experiments illustrate that the original model of de Wijs (Figures 1 and 2) with a single value for the dispersion index d produces results that are similar to more realistic models in which D is a random variable resulting in different values of d for every division of blocks into halves. In most of the remainder of this paper we will continue to use the original model of de Wijs because it provides valid approximations for large values of n . Moreover, multifractal theory is well established for the original model of de Wijs and can be used without difficulty.

However, the alternative approach of allowing $D^* = 1 + D$ to be a random variable allows empirical estimation of the frequency distribution of relative concentration values $Y_{12} = X_1/X_2$ where X_1 is the concentration value of a small block with volume V_1 that is randomly located within a larger block with concentration value X_2 and volume V_2 . If $V_1/V_2 = 0.5$, measurements of Y_{12} would

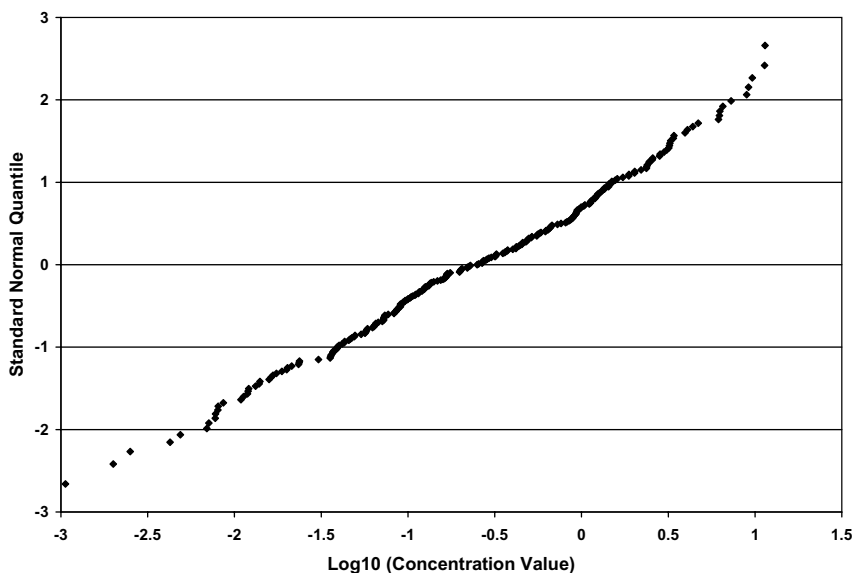


Figure 6. Lognormal Q-Q plot of 256 values of Figure 5. Pattern resembles straightline for lognormal in Figure 2.

provide direct estimates of values that are realizations of the random variable D^* . This approach, with V_1 much smaller than V_2 , will be applied to Au and As concentration values of till samples at the end of this paper.

BRINCK'S GEOCHEMICAL CRUSTAL ABUNDANCE APPROACH

The logbinomial was used in mineral resource evaluation studies by Brinck (1974). A comprehensive review of Brinck's approach can be found in Harris (1984). The original discrete model of de Wijs is assumed to apply to a chemical element in a large block of the Earth's crust with known average concentration value ξ commonly set equal to the element's crustal abundance.

Chemical analysis is applied to blocks of rock that are very small in comparison to the large block targeted for study. Let $n = N$ represent the maximum number of subdivisions of the large block. Suppose that the element concentration values available for study: (1) constitute a random sample from the population of 2^N very small blocks, and (2) show an approximate straightline pattern on their lognormal Q-Q plot. The slope of this line then provides an estimate of σ from which η (and d) can be derived by means of Eq. (1). Brinck (1974) set 2^N equal

to the average weight of the very small block used for chemical analysis divided by total weight of the environment targeted for study.

Later applications of Brinck's geochemical crustal abundance approach include Ruzicka (1976) and Garrett (1986). These authors used other methods for estimating the maximum number of subdivisions of the environment (N). In his study of sedimentary uranium deposits, Ruzicka (1976) based the estimate of N on sedimentological considerations assuming that the environment had been sorted naturally into small homogeneous domains. Garrett (1986) estimated N from maximum "barrier" concentration values chosen for chemical elements in particular deposit types or regions. Both authors acknowledged that estimating the number of subdivisions (N) constitutes a major problem in applications of Brinck's approach (also see Harris, 1984; Chapter 8, p. 184–222).

Lognormality of Trace Elements in Rocks

One of the basic assumptions in the preceding geochemical abundance model is that trace elements are lognormally distributed. Ahrens (1953) postulated lognormality as the first law of geochemistry. In general, it can not be assumed that the element concentration values for very small blocks of rock collected from a very large environment satisfy a single lognormal frequency distribution model. However, the lognormal model often provides a valid first approximation especially for trace elements. Reasons why the lognormal model may not be applicable include the following: Concentration values for all constituents form a closed number system and this prevents major constituents to be lognormally distributed. Also, discrete boundaries (contacts) between different rock types commonly exist in regional applications and mixtures of two or more lognormals would occur if rock types have lognormals with different parameters.

Three types of generating mechanisms or explanations have been suggested for lognormality. The first one is reviewed by Aitchison and Brown (1957): processes during which random increases in value are proportional to value can result in lognormal distributions, in the same way that processes subject to conditions underlying the central limit theorem of mathematical statistics lead to normal distributions. The second type of explanation was first advocated by Vistelius (1960): mixtures of populations with mean values that are proportional to standard deviations can result in positively skewed distributions that resemble lognormal distributions even if the original populations are normal. Finally, multiplicative cascade models such as the model of de Wijs used in this paper can help to explain lognormality.

Allègre and Lewin (1995) provide an overview of geochemical distributions that is wider than the approach taken in this paper, which is restricted to the lognormal and its logbinomial approximation. Geochemical distributions can have

Pareto instead of lognormal tails (also see Turcotte, 2002), or they may be mixtures of different types of distributions. It is also possible to work with a trinomial multiplicative cascade model (Gonçalves, 2001).

A relatively simple generalization of the lognormal model is to assume that the concentration values for a chemical element in a large region or environment originate from two different lognormal populations representing background and anomalies, respectively. The largest concentration values then primarily represent the anomalies. This type of modeling either uses lognormal Q-Q plots (see Sinclair, 1991), or use is made of concentration-area (CA) log-log plots (Cheng et al., 1994) to distinguish between the two separate populations. The approach advocated in this paper assumes lognormality of concentration values that are larger than average. It remains possible that (a) the relatively small concentration values (background) represent a mixture of different populations, and (b) the largest values satisfy a Pareto distribution with a tail that is thicker than lognormal.

Self-similarity and the Multifractal Paradox

Multifractal modeling offers an independent method to verify validity of the model of de Wijs and to estimate d . In principle, the value of n could be made infinitely large. According to Eq. (1), the logarithmic variance then also becomes infinitely large and the frequency distribution of the element concentration values would cease to exist. Application of the method of moments in multifractal analysis (see later) results in a multifractal spectrum that is a limiting form for infinitely large n . The frequency distribution corresponding to this limiting form can not exist in reality because it has infinitely large variance. In practice, any set of element concentration values for very small blocks of rock collected from a very large environment has a frequency distribution with finite logarithmic variance.

Suppose that the generating process of subdividing blocks under the same dispersion index (d) ceases to be operative for blocks that are larger than the very small blocks used for chemical analysis. In examples of application to be discussed in this paper, d at the regional scale does not apply at the local scale (for small blocks used for chemical analysis). Locally, d is either zero or much smaller than d at the regional scale. Under these conditions, an apparent maximum number of subdivisions N can be estimated. Self-similarity at scales exceeding a critical lower limit results in a model of de Wijs with three parameters: ξ , d , and N .

The multifractal method that can be used for estimating d was previously introduced by Agterberg (2001a, b). It uses the method of moments and is based on results originally obtained by Evertsz and Mandelbrot (1992) and clearly explained in several textbooks including Feder (1988) and Falconer (2003). A brief summary is as follows. When the mass of a quantity is measured in cells, power moment sums can be formed by raising all mass measurements to powers q and summing

over all cells. A property of self-similar multifractals is that power moment sums are related to cell side according to power laws with mass exponents $\tau(q)$ (cf. Cheng and Agterberg, 1996). In two-dimensional applications, each cell mass measure is the product of cell concentration value and cell area. If the condition of self-similarity is satisfied, any mass exponent $\tau(q)$ can be estimated as the slope of the best-fitting straightline for a specific value of q . The first derivative of $\tau(q)$ with respect to q provides an estimate of the so-called singularity $\alpha(q)$. The singularity has its minimum value ($\alpha_{\min}$) for the cell with the largest possible concentration value. Artificial upper limit truncation effects should be avoided. For the applications described in this paper, maximum element concentration values are significantly less than 100%.

Validity of the model of de Wijs for the larger concentration values results in

$$\lim_{q \rightarrow \infty} \frac{d\tau(q)}{dq} = \alpha_{\min} = \log_2 \left(\frac{\eta + 1}{\eta} \right)^2 \quad (4)$$

In 2-dimensional applications, chemical element measures are formed by multiplying average cell concentration values by cell areas. Raising these measures to relatively high powers q filters out the influence of smaller concentration values. Thus our estimate of d is primarily based on parameters estimated from the relatively large concentration values of an element. The lower-value tails of the observed frequency distributions could be mixtures of separate populations, and the exact nature of these mixtures can be left undetermined.

If the model of de Wijs is satisfied, the mass exponents $\tau(q)$ become linearly related to q when q is large. The parameter d then can be estimated from the slope of the resulting straightline. Together with the estimate of logarithmic variance this yields an estimate of apparent maximum number of subdivisions N when Eq. (1) is used. It is noted that d also can be estimated from the multifractal spectrum, which normally is the end product of applications of the method of moments. However, in the case history described in the next section, use of Eq. (4) produced better results more quickly.

In an earlier application of this approach to gold in bedrock of the Mitchell-Sulphurets area of northwestern British Columbia (cf. Cheng, 1994), Agterberg (2001b) estimated $d = 0.6$ and $N = 20$ with $\xi = 3.1$ ppm Au. The approximate size of this study area was 10 km by 10 km. Ten subdivisions in the east-west and north-south directions (yielding $N = 20$) would result in a square cell measuring 10 m on a side. Thus, N is much smaller than expected if self-similarity with constant d would continue to exist down to hand-specimen scale.

In order to illustrate the preceding methods of approach, a new case study will be presented in the next section. Brink's approach is mainly graphical. Likewise, model verification in fractal-multifractal modeling generally consists of the fitting of straightlines on log-log graphs. Our approach in the following case-history

Table 1. Gold and Arsenic in Till Samples, Southern Saskatchewan: Results of Graphical Estimation of Parameters for the Model of de Wijs. Gold Concentration Values in ppb; As Concentration Values in ppm

	Gold	Arsenic
Straightline on Q-Q plot: intercept	0.8006	−5.3035
Ditto, slope	0.4222	2.7611
Mean	2.48	7.29
Logarithmic variance	5.61	0.131
Cell data Q-Q plot: intercept	−0.2121	12.399
Ditto, slope	0.8784	6.0591
Logarithmic variance of cell averages	1.296	0.0272
Straightline on Tau vs <i>q</i> plot: intercept	−5.3035	−1.4232
Ditto, Slope	2.7611	1.8075
Dispersion index (Eq. 3)	0.433	0.069
Apparent number of subdivisions—till samples	26.12	27.5
Number of subdivisions, cell data	6.03	5.71

study also will be mainly graphical with ordinary least squares straightline fitting applied to sequences of consecutive data points.

**APPLICATION TO AU AND AS IN TILL SAMPLES FROM
SOUTHERN SASKATCHEWAN**

The approach is illustrated by application to gold and arsenic content of till samples obtained during systematic ultra-low density (1 site per 800 km²) geochemical reconnaissance, Southern Saskatchewan. There are 389 observed gold concentration values ranging from below detection limit to 77 ppb. As many as 228 Au values are below detection limit (=2 ppb). If these are set equal to 1 ppb, the arithmetic mean gold concentration value becomes 2.91 ppb. This estimate slightly overestimates true mean gold concentration value because most gold values below 2 ppb are less than 1 ppb. A better estimate of mean gold concentration value to be derived later is 2.48 ppb (cf. Table 1). In total, only about one third of the 389 observed values exceed this average. The corresponding median gold concentration value is as low as 0.15 ppb Au.

Another minor source of uncertainty is as follows. Small gold concentration values just above the detection limit are subject to limited resolution. This topic was recently reviewed by Grunsky (2005) for Au in relation to As. Fortunately; cumulative frequencies for larger (>5 ppb) are not affected by these types of uncertainty associated with the smallest gold concentration values only.

For arsenic, all observed values are above the detection limit with overall mean of 7.3 ppm and ranging from about 2 to 44 ppm. The main difference between frequency distributions of Au and As is that gold has positively skewed

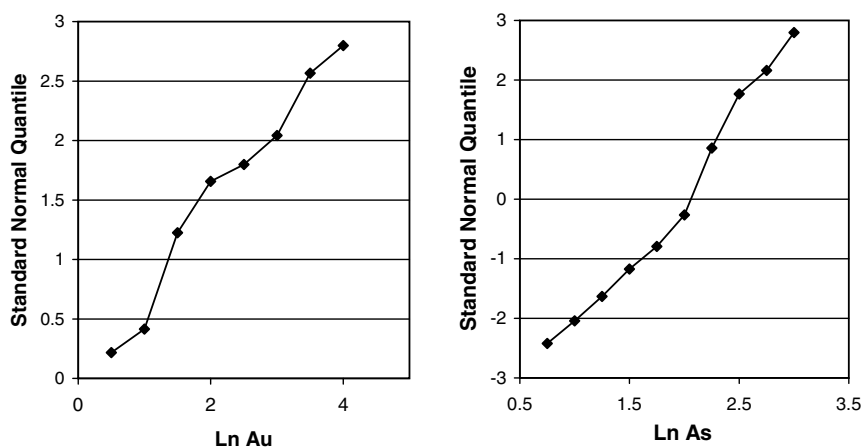


Figure 7. Cumulative frequency distributions (lognormal Q-Q plots) of gold and arsenic in 379 till samples, Southern Saskatchewan. Units of Au and As concentration values are ppb and ppm, respectively.

distribution whereas arsenic shows approximately symmetric, normal distribution. Most diagrams in this section are for Au only. Similar results (not shown as diagrams) were obtained for As. Estimates for both Au and As are listed in Table 1.

Lognormal Q-Q plots for Au and As are shown in Figure 7. The first step in Brinck's approach consists of fitting a straightline on the Q-Q plot. Figure 8 shows our straightline fit for Au as obtained by ordinary least squares. The range of gold values to which the straightline was fitted is narrower than the range in Figure 7 for several reasons: (1) Q-Q values for the largest gold values were not used because their uncertainty is greater than that of values closer to the median. Using the relatively largest values could create artificial distortions when all observations are assigned equal weights as in ordinary least squares; (2) as pointed out before, cumulative frequencies for small Au concentration values near detection limit are subject to limited resolution bias. This problem is circumvented when larger concentration values are used only; and (3) although the data set is rather small, the pattern for Au in Figure 7 suggests that there could be two or more populations (cf. Sinclair, 1991). In a two-population model, the range of gold values in Figure 8 would represent the population with larger mean that is of greater interest than lower background gold values. Admittedly this graphical approach is subjective. However, the statistics obtained by it agree with results to be obtained later (see discussion of Figs. 11, 12, and 14). Also, logarithmic (base e) mean ($=2.58$ ppb) and variance ($=0.125$) of the gold values plotted in Figure 8 and modeling the logarithmically transformed values between 2.0 and 3.2 as a doubly truncated normal distribution yields approximately the same results.

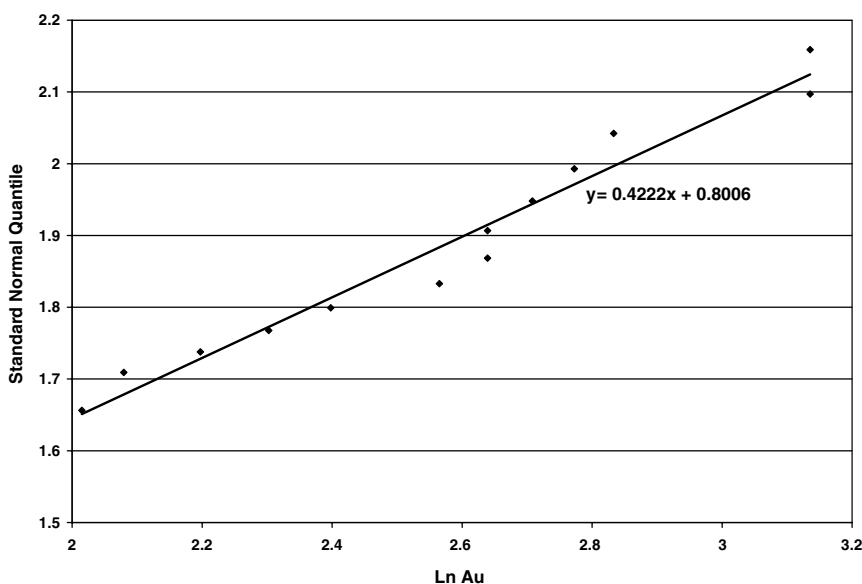


Figure 8. Straightline fitted to part of lognormal Q-Q plot for 379 gold concentration values. Inverse of slope provides estimate of logarithmic (base e) standard deviation.

Because natural logarithms are used, the standard deviation of logarithmically transformed concentration values can be set equal to the inverse of the slope of the best-fitting straightline. It amounts to 2.37 ppb for Au. Difference in shape of frequency distribution is reflected in the fact that the logarithmic variance of Au is more than 40 times as large as the logarithmic variance of As (see Table 1).

Original sampling design for the geochemical reconnaissance study in Southern Saskatchewan is described by Garrett and Thorleifson (1995). An 80×80 km grid was selected as the basic sampling cell. It was successively subdivided into 40×40 km, 20×20 km and finally 10×10 km grid cells for sampling with randomly selected 1×1 km target cells. Compositional data for many other elements and minerals were determined as well. Gold and arsenic were selected for this study because their frequency distributions seem to be unimodal, and these elements had been used in other multifractal studies (Cheng, 1994; Agterberg, 2001a and b).

For the case study described in this paper, a modified grid was used for southern Saskatchewan resulting in a square-shaped study area consisting of 64 square cells measuring slightly less than 80 km on a side (see Fig. 9). The purpose of this re-definition of boundaries of study area was to minimize edge effects in multifractal analysis. Data from outside the square study area were not used. In total, 290 of the original 389 (on average, about 4.5/cell) were retained. None of

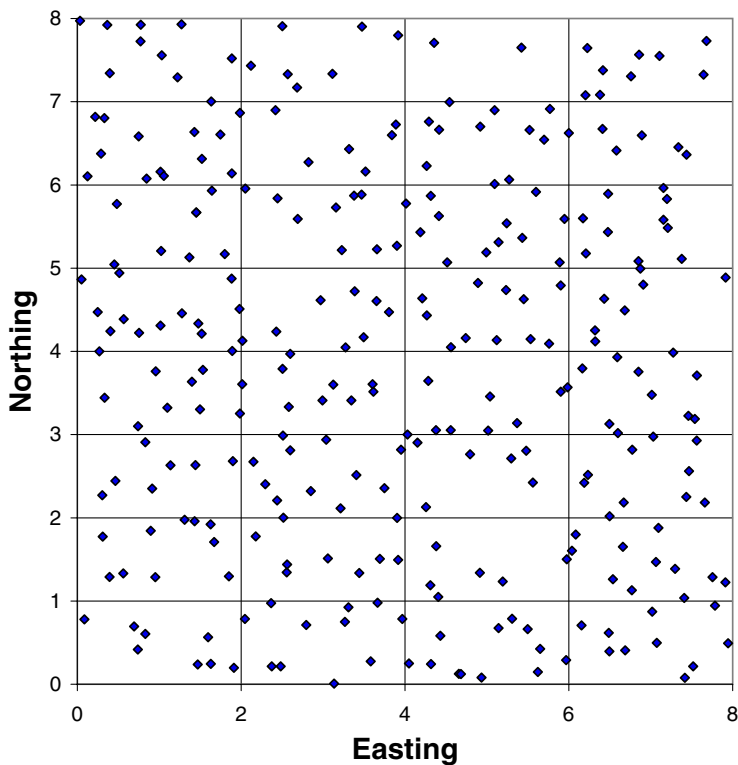


Figure 9. Study area of 64 cells with locations of till samples. Unit of distance for east-west and north-south directions is 78.125 km.

the cells inside our study area are empty. Figure 10 is a Q-Q plot of the 64 gold cell average concentration values with best-fitting straightline.

The square arrays of Au and As cell average concentration values were subjected to multifractal analysis using the method of moments. Figure 11 shows power moment sum results for gold. For relatively large values of q there is approximate linear relationship between the mass exponent $\tau(q)$ and q for Au (Fig. 12). The slopes of the best-fitting straightlines are 0.962 for Au and 1.808 for As, respectively (Table 1).

By means of Eq. (4), these slopes were converted into dispersion index estimates of $d = 0.433$ for Au, and $d = 0.069$ for As. Straightline fitting to lognormal Q-Q plots for the 64 cell average concentration values gave logarithmic variances of 1.292 for Au, and 0.029 for As, respectively. Application of Eq. (1) then yields estimates of n equal to 6.0 for Au, and 5.7 for As, respectively. Both estimates are nearly equal to 6 representing their expected value for the model of de Wijs

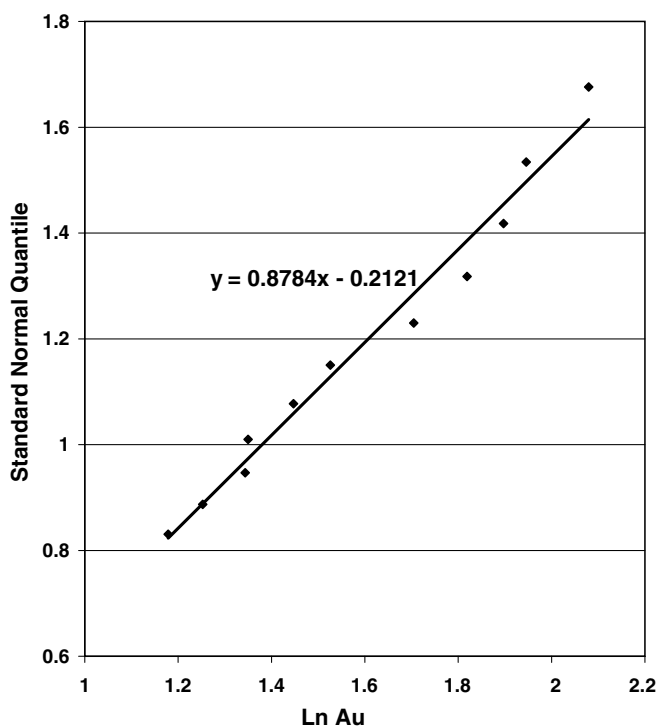


Figure 10. Straightline fitted to part of lognormal Q-Q plot for 64 average gold concentration values for cells.

because $2^6 = 64$. This agreement between results indicates that the model of de Wijs is approximately satisfied on a regional scale for both gold and arsenic.

Logarithmic variances estimated from the 379 original till samples were 2.369 for Au and 0.362 for As, respectively. Using Eq. (1) this yields estimates of N equal to 26.1 for Au, and 27.5 for As, respectively. These apparent maximum numbers of subdivisions of the environment would correspond to square cells measuring approximately 75 m, and 40 m on a side, respectively. Clearly, these cells are much larger than the very small areas where the till was actually sampled for the purpose of chemical analysis. It shows that, for both Au and As, the regional model of de Wijs does not apply at the local sampling scale. This result is similar to the one previously obtained for Au in the Mitchell-Sulphurets area (see before).

Logbinomial distributions for gold and arsenic are shown in the lognormal Q-Q plots of Figures 13 and 14. Because the logbinomial is discrete, the preceding estimates of N were rounded off to the nearest integer. The parameters of the distributions shown are $\xi = 1.82$ ppb, $d = 0.433$, and $N = 26$ for gold (Fig. 13A);

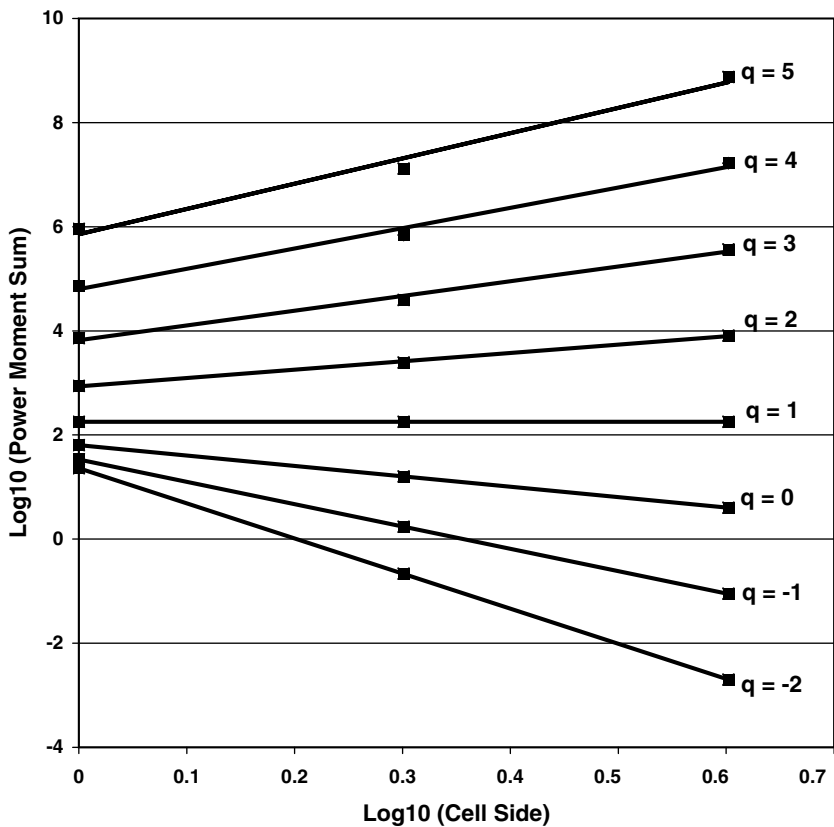


Figure 11. Method of moments applied to 64 average gold concentration values for cells. A property of self-similar multifractals is that power moment sums of cell masses raised to powers of q are related to cell side according to power laws with mass exponents $\tau(q)$ (cf. Cheng and Agterberg, 1996). Each mass exponent $\tau(q)$ is estimated as the slope of the best-fitting straightline for a specific value of q . Because of uncertainties associated with gold detection limit, straightlines for positive values of q can be used only. Cell side unit is same as in Figure 9.

and $\xi = 7.05$ ppm, $d = 0.069$, and $N = 28$ for arsenic (Fig. 13B). Mean values ξ were determined by forcing the logbinomials and their lognormal approximations through the centers of clusters of points on lognormal Q-Q plots for original data. These results remain approximate but are interesting because they permit shape comparison of logbinomial and lognormal tails.

The dispersion index of gold ($=0.433$) is much larger than that of arsenic ($=0.069$), but the patterns on Q-Q plots in Figure 13 for these two elements are similar. Both are approximately lognormal over the ranges of observed values. Outside these ranges the logbinomial upper (and lower) tails become increasingly

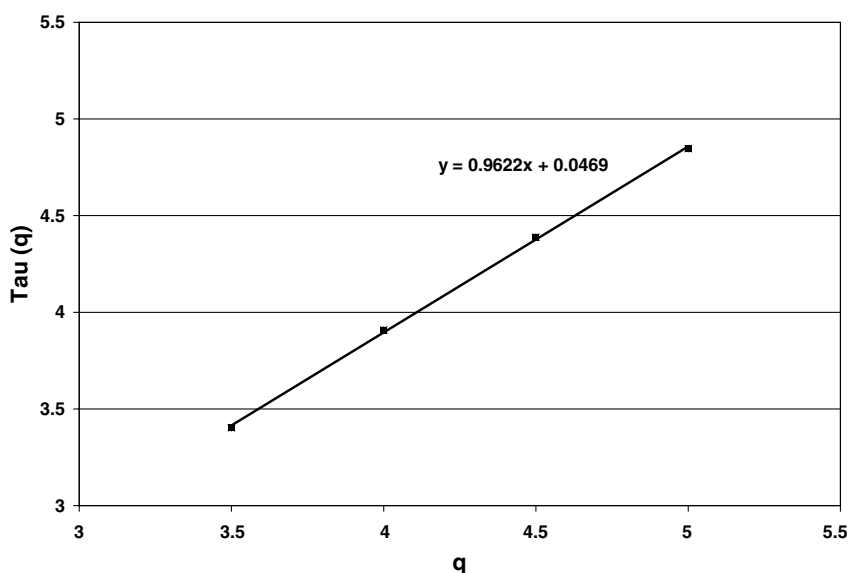


Figure 12. Mass exponent Tau (q) versus relatively large values of q . Slope of straightline is used to estimate dispersion index $d = 0.433$ for gold using Equation (4).

thinner than their lognormal approximations. However, the resulting differences have relatively little practical significance. For example, the logbinomial would predict that 1.45% of all possible till samples have gold concentrations greater than 32 ppb, whereas the corresponding lognormal prediction (=1.56%) is hardly different from this.

Application of Random-Cut Model

Finally, the random-cut model is applied to the gold concentration values. The underlying theory was explained before in discussion of the third computer simulation experiment (Figs. 5 and 6).

Relative element concentration values $Y_{1n} = X_1/X_{n2}$ were calculated for gold, where X_1 is the concentration value of a till sample with volume V_1 that is randomly located within a cell belonging to the grid of cells shown in Figure 9 with concentration value X_n cell size V_n . Figure 14 is a lognormal Q-Q plot of these 290 relative Au concentration values. The straightline provides a good fit except on the left side of Figure 14 where there is some bias related to the Au detection limit and at the ends where frequencies of relative gold concentration values are small.

The logarithmic variance derived from the straightline of Figure 14 is 0.5717. This illustrates that positive skewness of the relative gold concentration values is

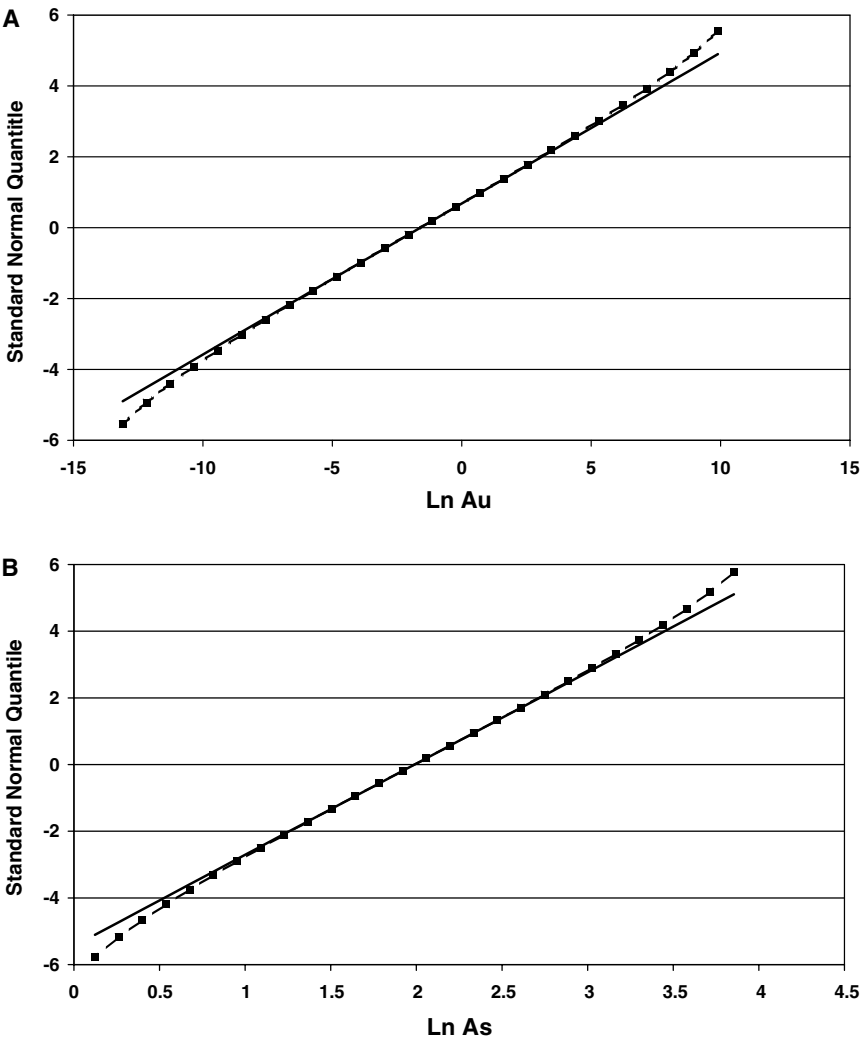


Figure 13. Lognormal Q-Q plots for discrete logbinomial frequency distributions of (A) gold and (B) arsenic concentration values resulting from model of de Wijs. (A) For Au: overall average element concentration value $\xi = 1.82$ ppb, dispersion index $d = 0.433$, and apparent number of subdivisions of the environment $N = 26$; (B) for As: $\xi = 7.05$ ppm, $d = 0.069$, and $N = 28$.

significant. Setting V_n/V_1 equal to 2^{20} in Eq. (2) yields an estimated value of $\alpha^* = 0.04124$. The logarithmic variance of Y_{12} would amount to $h^2 = 0.5717/19 = 0.03009$. Because it can be assumed that Y_{12} , like Y_{1n} , has lognormal distribution, its variance is estimated to be $\sigma^2(Y_{12}) = 0.2326$. Because this value is rather small,

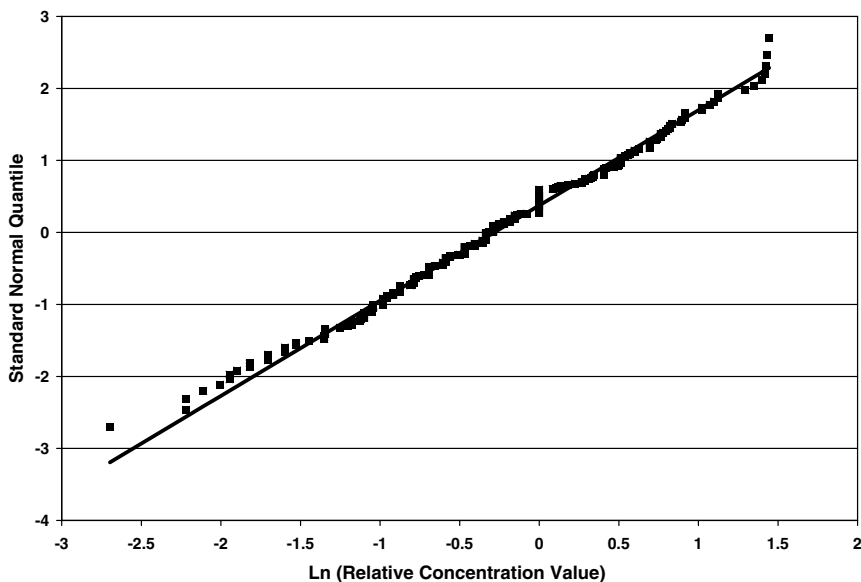


Figure 14. Lognormal Q-Q plot for relative gold concentration values. Best-fitting straightline indicates approximate lognormality and is used to estimate mean deviation of dispersion index variable D^* in random-cut model. See text for further explanation.

Y_{12} is approximately normally distributed with standard deviation $\sigma(Y_{12}) = 0.482$. It follows that mean deviation from $E(Y_{12}) = 1$ amounts to m.d. = 0.39. This would be a crude estimate of average dispersion index in the random-cut model. It is close to $d = 0.43$ estimated for the original logbinomial model of de Wijs (Table 1).

Although m.d. and d are different parameters we would expect them to be approximately equal to one another because of convergence of logbinomial and lognormal distributions for increasing number of subdivisions (n). It was attempted to repeat the preceding analysis for Ar but the 290 values $Y_{1n} = X_1/X_{n2}$ for As do not show a simple straightline pattern, and m.d. could not be estimated with sufficient precision.

CONCLUDING REMARKS

When element concentration values randomly collected from a large region are approximately lognormally distributed, this may be an indication of scale independence or self-similarity of their spatial variability. Because of its discrete nature, the original model of de Wijs does not apply locally. It can be replaced by

a locally more realistic random-cut model. However, in large regions, the random-cut model converges to approximately the same distribution as the model of de Wijs.

The logbinomial distribution resulting from the original model of de Wijs converges to lognormal form except in the tails of the frequency distributions where the logbinomial is slightly weaker. This model was used extensively by Brinck (1974) for estimating the frequency distribution of elements in segments of the Earth's crust. Now a new tool is available for this type of crustal abundance studies because multifractal modelling can be used to estimate regional dispersion indices of approximately lognormally distributed elements.

The examples for glacial till samples from southern Saskatchewan used in this paper differ in that the frequency distribution for Au is positively skew with a thin long upper tail whereas As is nearly normally distributed. The model of de Wijs was shown to apply to both elements on the regional scale (square cells measuring 80 km on a side and larger) with dispersion index $d = 0.433$ for Au, and $d = 0.069$ for As. At the local sampling scale d rapidly decreases to zero or a small value. Consequently, there exist an apparent maximum number of subdivisions (N) in the model of de Wijs that can be estimated in addition to average abundance (ξ) and d . A random-cut version of the model of de Wijs was successfully applied to gold in Southern Saskatchewan till.

The methods developed in this paper apply to elements with unimodal frequency distribution. Multifractal modeling can produce useful information on separating anomalies from geochemical background (fractal C-A method, Cheng et al., 1994; and S-A method, Cheng, 2001). Further work on the methods developed in this paper can consist of the modeling of mixtures of spatial frequency distributions of element concentration values. The aim of another modification of the approach described in this paper can be to model approximately lognormal distributions with Pareto tails.

ACKNOWLEDGMENTS

This paper partly resulted from lectures and workshops on lognormal and Pareto distributions presented at James Cook University (Townsville), University of Western Australia (Perth), Cape Town University, University of Science and Technology (Trondheim), University of Girona, University of Nancy—CRPG, and Delft University of Technology, when I was 2004 IAMG Distinguished Lecturer, and during spring 2005 at China University of Geosciences (Wuhan). Scientists who have offered helpful suggestions to ideas presented in this paper include Qiuming Cheng, Shaun Lovejoy, Thomas Blenkinsop, Christien Thiart, Vera Pawlowsky, Gloria Mateu, Jean-Jacques Royer, Richard Sinding-Larsen, and Gert-Jan Weltje. Thanks are also due to R.G. Garrett, Geological Survey of Canada, who kindly provided the till sample data used for the examples. Critical

comments by Nils Gustavsson, Geological Survey of Finland, and an anonymous reviewer have resulted in significant clarifications of presentation of the material. Of course, any shortcomings remain my responsibility. Financial support from NSERC and the State Key Lab for Geological Processes and Mineral Resources (Wuhan) is gratefully acknowledged.

REFERENCES

- Agterberg, F. P., 1984, Use of spatial analysis in mineral resource evaluation: *Math. Geol.*, v. 16, no. 6, p. 565–589.
- Agterberg, F. P., 1994, Fractals, multifractals, and change of support, *in* Dimitrakopoulos, R., ed., *Geostatistics for the next century*: Kluwer, Dordrecht, p. 223–234.
- Agterberg, F. P., 2001a, Multifractal simulation of geochemical map patterns, *in* Merriam, D. F. and Davis, J. C., *Geologic modeling and simulation: Sedimentary Systems*: Kluwer, New York, p. 327–346.
- Agterberg, F. P., 2001b, Aspects of multifractal modelling: *Proceedings of IAMG Annual Meeting, Cancun, September, 2001*, CD Rom.
- Agterberg, 2005, Application of a three-parameter version of the model of de Wijs in regional geochemistry: *Proceedings of IAMG' 05 held in Toronto, Canada, August 21–26, 2005*, China University of Geosciences Printing House, Wuhan, China, v. 1, p. 291–296.
- Ahrens, L. H., 1953, A fundamental law of geochemistry: *Nature*, v. 172, p. 1148.
- Aitchison, J., and Brown, J. A. C., 1957, *The lognormal distribution*: Cambridge University Press, Cambridge, 176 p.
- Allègre, C. J. U., and Lewin, E., 1995, Scaling laws and geochemical distributions: *Earth Planet. Sci. Lett.*, v. 132, p. 1–13.
- Brinck, J. W., 1974, The geochemical distribution of uranium as a primary criterion for the formation of ore deposits, *in*: *Chemical and physical mechanisms in the formation of uranium mineralization, geochronology, isotope geology and mineralization*: International Atomic Energy Agency Proceedings Series, no. STI/PUB/374, p. 21–32.
- Cheng, Q., 1994, Multifractal modelling and spatial analysis with GIS: gold potential estimation in the Mitchell-Sulphurets Area, northwestern British Columbia: Unpublished doctoral thesis, University of Ottawa, Canada, 268 p.
- Cheng, Q., 2001, Selection of multifractal scaling breaks and separation of geochemical and geophysical anomaly: *Jour. China Univ. of Geosc.*, v. 12, p. 54–59.
- Cheng, Q., and Agterberg, F. P., 1996, Multifractal modeling and spatial statistics: *Math. Geol.*, v. 28, no. 1, p. 1–16.
- Cheng, Q., Agterberg, F. P., and Ballantyne, S. B., 1994, The separation of geochemical anomalies from background by fractal methods: *J. Geochem. Explor.*, v. 51, no. 2, p. 109–130.
- De Wijs, H. J., 1951, Statistics of ore distribution, part I: *Geologie en Mijnbouw*, v. 13, p. 365–375.
- Evertsz, C. J. G., and Mandelbrot, B. B., 1992, Multifractal measures (Appendix B), *in* Peitgen, H.-O., Jurgens, H. and Saupe D., eds., *Chaos and fractals*: Springer Verlag, New York, p. 922–953.
- Falconer, K., 2003, *Fractal geometry—mathematical foundations and applications* (2nd ed.): Wiley, New York, 337 p.
- Feder, J., 1988, *Fractals*: Plenum, New York, 283 p.
- Garrett, R. G., 1986, Geochemical abundance models: an update, 1975 to 1987: *U.S. Geological Survey Circular 980*, p. 207–220.
- Garrett, R. G., and Thorleifson, L. H., 1995, Kimberlite indicator mineral and till geochemical reconnaissance, southern Saskatchewan: *Geological Survey of Canada Open File 3119*, p. 227–253.

- Gonçalves, M. A., 2001, Characterization of geochemical distributions using multifractal models: *Math. Geol.*, v. 33, no. 1, p. 41–61.
- Grunsky, E. C., 2005, Assessment of geochemical data—Strategies for evaluation and interpretation: Proceedings of IAMG' 05 held in Toronto, Canada, August 21–26, 2005, China University of Geosciences Printing House, Wuhan, China, v. 1, p. 1–14.
- Harris, DeVerle P., 1984, *Mineral Resources Appraisal*: Clarendon Press, Oxford, 445 p.
- Krige, D. G., 1966, A study of gold and uranium distribution patterns in the Klerksdorp gold field: *Geoexploration*, v. 4, p. 43–53.
- Lovejoy, S., and Schertzer, D., 1991, Multifractal analysis techniques and the rain and cloud fields from 10^{-3} to 10^6 m, in Schertzer, D. and Lovejoy, S., eds., *Non-linear variability in geophysics*: Kluwer, Dordrecht, p. 111–144.
- Mandelbrot, B. B., 1983, *The fractal geometry of nature* (updated and augmented edition): W. H. Freeman and Company, New York, 468 p.
- Matheron, G., 1962, *Traité de géostatistique appliquée*: Mém. Bur. Rech. Géol. Minières, v. 14, p. 1–333.
- Matheron, G., 1974, Les fonctions de transfert des petites panneaux: Note Géostatistique 127, Centre de Morphologie Mathématique, Fontainebleau, France.
- Ruzicka, V., 1976, Uranium resources evaluation model as an exploration tool—exploration for uranium deposits: International Agency for Atomic Energy Proceedings Series, no. 434, p. 673–682.
- Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y., and Marsan, D., 1997, Multifractal cascade dynamics and turbulent intermittency: *Fractals*, v. 5, no. 3, p. 427–471.
- Sinclair, A. J., 1991, A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited: *J. Geochem. Explor.*, v. 41, no. 1, p. 1–22.
- Turcotte, D. L., 2002, Fractals in petrology: *Lithos*, v. 65, nos. 3–4, p. 261–271.
- Vistelius, A. B., 1960, The skew frequency distributions and the fundamental law of geochemical processes: *J. Geology*, v. 68, no. 1, p. 1–22.