Singularity analysis of ore-mineral and toxic trace elements in stream sediments

Qiuming Cheng a,b,*, Frederik P. Agterberg c

a Earth and Space Science and Engineering, York University, Toronto, Canada M3J1P3
b State Key Lab of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan and Beijing, China
c Geological Survey of Canada, Ottawa, Canada K1A0E8

ARTICLE INFO

Article history:
Received 20 December 2006
Received in revised form 28 January 2008
Accepted 1 February 2008

Keywords:
Non-linear theory
Singularity
Mineralization
Mineral potential mapping
GIS

ABSTRACT

Hydrothermal processes in the Earth’s crust can result in ore deposits characterized by high concentrations of metals with fractal or multifractal properties. This paper shows that stream sediments in the neighborhoods of ore deposits also can have singular properties for ore-mineral and associated toxic trace elements. We propose a new local singularity mapping method for assembling element concentration values from stream sediment samples to delineate anomalous areas induced by buried mineral deposits, which are often missed in ordinary geochemical surveys and mapping. Applied to the Gejiu area, Yunnan Province, China, which contains world-class size hydrothermal deposits enriched in tin and other elements, non-linear anomalies for tin and arsenic are identified: (1) many relatively small singularity anomalies in about 10% of the study area; and (2) a large high-concentration anomaly in the eastern part of the area where mining occurs. The ore-mineral and toxic elements within these anomalies describe Pareto-type frequency distributions. Spatial proximity of anomalies of the first kind to the ore deposits (mines and prospective mines) indicates that singularity mapping provides a useful new tool for mineral prospecting. The relation of the second kind of anomaly to mining activities indicates that fractal modeling also can provide useful input for decision-making in environmental protection.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Singular physical or chemical processes may result in anomalous amounts of energy release or mass accumulation that, generally, are confined to narrow intervals in space or time. Singularity is a property of different types of non-linear natural processes including cloud formation (Schertzer and Lovejoy, 1987), rainfall (Veneziano, 2002), hurricanes (Sornette, 2004), flooding (Malamud et al., 1996), landslides (Malamud et al., 2004), and earthquakes (Turcotte, 1997). The end products of these non-linear processes have in common that they can be modeled as fractals or multifractals. Hydrothermal processes in the Earth’s crust can result in ore deposits characterized by high concentrations of metals with fractal or multifractal properties (Cheng and Agterberg, 1996; Agterberg, 1995; Mandelbrot, 1989; Cheng, 2007a, b). Non-linear processes are widespread in nature. Outcomes of such processes are often characterized by positively skewed frequency distributions with Pareto upper-value tails. The largest values may be extreme events. In some applications (e.g. Gutenberg–Richter law for magnitudes of earthquakes), power-law relations describing observed frequency distributions were already in existence long before non-linear explanations (Malamud et al., 2004). Successive non-linear processes generating frequency...
distributions with Pareto tails may be related causally such as rainfall and flooding. Total amount of ore and metals in hydrothermal ore deposits often have Pareto tails (Turcotte, 1997). We show that regional environments of hydrothermal mineral deposits also can exhibit non-linear features for ore-mineral and associated toxic element concentration values of stream sediment samples collected within $26 \times 26$ km$^2$ cells centered on the sampling locations.

2. Singularity and anomaly identification

Outcomes of singular processes are often characterized by multifractal distribution which involves variable singularities and positively skewed frequency distributions with Pareto upper-value tails. In the multifractal context, a measure $\mu$ of the amount of a chemical element in a small area of measuring size $e$ (for example, square of $e$ km on a side) satisfies $\langle \mu(e) \rangle \propto \langle e^\alpha \rangle$, where $\alpha$ stands for “proportional to” when cell size $e$ approaches to zero and $\alpha$ is the singularity index also known as Hölder exponent (Mandelbrot, 1989). This power-law exists usually in statistical sense and represented as expectation $\langle \rangle$. According to the distribution of value of $\alpha$ the entire mapped area can be classified into subsets or fractal each with different singularity and accordingly different fractal dimensions. This is why it has been termed multifractals. The distribution of singularity $\alpha$ in the mapped area can be described by the fractal dimension spectrum function $f(\alpha)$ which implies that majority values of $\alpha$ are close to 2 or the areas with value of $\alpha$ close to 2 showing as normal area with dimension close to 2, whereas the areas with value $\alpha > 2$ or $\alpha < 2$ are more irregular with dimension less than 2. The statistical power-law distribution can be approximated numerically by sampling the areas with difference boxes of variable sizes. For example, in some locations the mean values of pixels at different resolutions might be independent of the size of the pixel (small pixel sizes) within which the values are averaged. In other cases the mean value might proportionally depend on pixel size. The former case indicates non-singular background or areas with singularity index $\alpha$ close to 2 and the latter corresponds to singular component where the values of $\alpha > 2$ or $\alpha < 2$. Singularity property has been commonly observed in geochemical and geophysical quantities (Cheng et al., 1994; Cheng, 1999, 2007a, 2008). Numerically, take a measure $\mu$ as the amount of a chemical element in a square cell measuring $e$ km on a side satisfies, then the power-law relationship between measure and cell size can be expressed as $\mu = ce^\alpha$ where $c$ is a constant, and $\alpha$ is the singularity; then $\alpha$ can be estimated by measuring the slope of the straight-line in a log–log plot of $\mu$ against $e$. In this 2-D application of multifractal

Fig. 1. Geological setting of tin deposits in Gejiu area and map patterns derived from tin concentration values in stream sediment samples. (a) Simplified geology after Yu (2002). Solid lines indicate faults; triangles represent tin mineral deposits; UTM coordinates apply to map corner points. (b) Locations of examples of estimation of local arsenic singularity in Fig. 2. (c) Distribution of tin concentration values in 1000 stream sediment samples using inverse distance weighted moving average method. Eight of 11 large tin deposits fall on large regional anomaly in eastern part of area. (d) Distribution of local tin singularities. Tin deposits tend to occur in places where local singularity is less than 2.
theory, $\mu = \zeta e^2$, where $\zeta = ce^{-2}$ represents average element concentration value in the cell. If element concentration values for samples taken at the surface of a study area are realizations of a stationary random variable with constant population mean, then $\alpha = 2$ representing non-singularity. "Singular" locations where $\alpha < 2$ may indicate anomalous enrichment of the element being studied. Likewise, $\alpha > 2$ may indicate depletion. More discussion about the existence and property of singularity can be found in Cheng (2007b, 2008). To test the power-law relationship and to estimate the values of parameters $c$ and $\alpha$, both measure $\mu$ and density $\zeta$ can be plotted as log-log scales. However, for most of locations where the singularity $\alpha$ is close to 2 so that the relationship between the density $\zeta$ is a constant and independent of $e$, therefore, it is not advantage to estimate the validate the linearity using correlation coefficient index. Whereas in the same situation the measure $\mu$ and $e$ follow power-law relation with exponent 2, which can be validated using correlation coefficient estimated from the linear regression between the values of $\mu$ and $e$. In singular situation where $\alpha$ is not close to 2, then both $\mu$ and $\zeta$ show power-law relations with cell size $e$, therefore, either of these two values can be applied for validation and estimation purpose. However, in most cases due to the values of $\alpha - 2$ is much smaller than $\alpha$ according to the property of singularity, although the estimated values of $\alpha$ and $c$ using either of the two types of values (measure or density) will remain the same, the linearity related to the power-law relationship between measure and cell size is usually better than that related to relationship between density and cell size. Therefore, the authors suggest using the values measure

![Fig. 2. Examples of estimation of local singularity at locations shown in Fig. 1d.](image-url)

Fig. 2. Examples of estimation of local singularity at locations shown in Fig. 1d. Cell size is length of side of square cells with half-widths set equal to 1, 3, 5, ..., 13 km. Plots on left and right show relations between values $\mu$ and $\zeta$ and cell size $e$, respectively. Value of $\mu$ is sum of concentration values for all $2 \times 2$ km cells contained within larger cells centered on a location and value of $\zeta$ is $\mu/e^2$; base of logarithms is 10. Straight-line fitted by ordinary least squares method (log value regressed on log cell size); $R^2$ is multiple correlation coefficient squared. Typically all seven points fall on the best-fitting straight line although small deviations occur (e.g. second point at Location 2). Slopes of straight-lines fitted to values in plots on left provides estimate of local singularity $\alpha$ and those calculated from plots on right provide estimate of $\alpha - 2$. 
and ceiling size to do the validation and estimation. Both cases will be shown in the case study in this paper.

Where the landscape permits this, stream sediments are the preferred sampling medium for reconnaissance geochemical surveys concerned with mineral exploration (Plant and Hale, 1994). During the 1980s and 1990s, government-sponsored reconnaissance surveys covering large parts of Austria, the Canadian Cordillera, China, Germany, South Africa, UK, and USA were based on stream sediments (Darnley et al., 1995). These large-scale national projects, which were part of an international geochemical mapping project (Darnley, 1995), generated vast amounts of data and continue to be a rich source of information. We have applied singularity analysis to data from about 7800 stream sediment samples collected as part of the Chinese regional geochemistry reconnaissance project (Xie et al., 1997). For illustration, we use about 1000 stream sediment tin concentration values from the Gejiu area in Yunnan Province. This area of about 4000 km² contains 11 large tin deposits. Several of these, including the Laochang and Kafang deposits, are tin-producing mines with copper extracted as a by-product. These hydrothermal mineral deposits also are enriched in other chemical elements including silver, arsenic, gold, cadmium, cobalt, iron, nickel, lead, and zinc. Applications to be described here are restricted to tin, arsenic, and copper. Tin and copper are the ore elements of most interest for mineral prospecting whereas arsenic is a toxic element. Water pollution due to high arsenic, lead, and cadmium concentration values is considered to present one of the most serious health problems especially in underdeveloped areas where mining is the primary industry such as in the Gejiu area. Knowledge of the characteristics of spatial distribution of ore elements and associated toxic elements in surface media is helpful for the planning of mineral exploration as well as environmental protection strategies.

The Gejiu mineral district (Fig. 1a) is located along the suture zone of the Indian Plate and Euro-Asian plates on the southwestern edge of the China sub-plate,
approximately 200 km south of Kunming, capital of Yunnan Province, China. The Gejiu Batholith with outcrop area of about 450 km² is believed to have played an important role in the genesis of the tin deposits (Yu, 2002). The ore deposits are concentrated along intersections of NNE–SSW and E–W trending faults.

Stream sediment sample locations in the Gejiu area are equally spaced at approximately 2 km in the north–south and east–west directions. Every sample represents a composite of materials from the drainage basin upstream of the collection site (Plant and Hale, 1994). Regional trends are captured in a moving average map of tin concentration values from within square cells measuring 26 km on a side (Fig. 1c). Several parameters have to be set for use of the inverse distance weighted moving average method (Cheng, 2003). In our application, each square represents moving average for a square window measuring 26 km on a side with influence of samples decreasing with distance according to power-law function with exponent set equal to 2. Original sample locations were 2 km apart both in the north–south and east–west directions. It shows a large anomaly in the eastern part of the Gejiu area surrounding most large tin deposits including the mines. The three large tin deposits in the central part of the Gejiu area are recent discoveries. To illustrate in detail how singularities were estimated 12 different locations were arbitrarily selected on the map as shown on Fig. 1b. In total, 1056 local tin singularities were estimated (Fig. 1d) by assembling the tin concentration values from within the same 26 × 26 km² cells used for the moving average map (Fig. 1c). Fig. 2 illustrates in detail how singularities were estimated at the 12 different locations. For comparison purpose we show the results obtained using two types of values: measure and density. The plots on the left show the results of measure $\mu$ and cell size $\varepsilon$ and the plots on the right provide the results of density $\zeta$ and cell size $\varepsilon$. It can be seen that both types of plots showing linear relationship between values $\mu$ or $\zeta$.
and cell size $e$. As explained previously, the values of $\alpha$ estimated using these types of plots should be the same while the correlation coefficients related to the plots of $\mu$ and $\epsilon$ are higher than those related to values of $\zeta$ and $\epsilon$. The main difference between the patterns of Figs. 1c and d is that lower and higher local tin singularity values are more evenly distributed across the Gejiu area than the tin concentration values themselves. The preceding analysis was repeated for arsenic, which is a toxic element. Its moving average and local singularity patterns (Fig. 3) are similar to those obtained for tin (Fig. 1).

A histogram of all local tin singularities is unimodal (Fig. 4a). Strength of spatial correlation between a point pattern and a contour map can be estimated by means of the weights of evidence method (Agterberg, 1989; Bonham-Carter, 1994). A Student’s $t$-value diagram (Fig. 3b) can be used to express statistical significance of strength of spatial correlation between (a) point pattern of 11 tin deposits (Fig. 1a), and (b) local tin singularity map (Fig. 1d). Although there are relatively few tin deposits, $t$-values near the peak (where $t = 4.84$ at $\alpha = 1.925$) exceed $t_{0.05} = 2.0$ representing 95% confidence level for statistical significance indicating positive correlation between the two patterns. In total, 93 local tin singularities have $\alpha < 1.925$. Their combined area measures only 8.8% of total study area, but 61.3% of the tin deposits occur within this relatively small low-singularity sub-area. The logarithm (base $e$) of number of local tin singularities exceeding $\alpha$ is linearly related to $\alpha$ (Fig. 3c) with slope of approximately 2.3.

Likewise, there is an approximate straight-line relationship (slope $= -3.2$) between logarithmically transformed cumulative area and largest tin concentration values (Fig. 3d). This is an example of a concentration-area ($C$–$A$) plot (Cheng et al., 1994). The pattern on a $C$–$A$ plot is automatically broken into a number of successive straight-line segments. For tin (Fig. 4d), $\log_e \alpha$ of the third segment extends from 7.414 to 8.637 with best-fitting straight-line: $y = 1.64 x + 3.00$. It represents a Pareto frequency distribution for the highest concentration values.

Other chemical elements enriched in the hydrothermal tin deposits show patterns similar to those obtained for tin (Figs. 4a–d) as illustrated for arsenic (Figs. 4e–h) and...
concentration value ($\xi$) of a chemical element in the block then can be written as $(1+d)\xi$ for one half and $(1-d)\xi$ for the other half so that total mass is preserved. The coefficient of dispersion $d$ is independent of block size. The multifractal patterns generated by this cascade have many local maxima and minima. Strictly speaking, a cascade does not result in point-wise convergence to any value of $z$ because it performs a random walk at any scale. Instead of values at points, our local singularities are average values of $z$ within small cells. Fig. 6a illustrates a simplest form of multiplicative cascade process, any cell containing a chemical element in 1-, 2-, or 3-D space with average concentration value $\xi$ set equal to unity is divided into two halves with element concentration values $(1+d)\xi$ and $(1-d)\xi$, respectively. For the first cell at the beginning of the process, $\xi$ can be set equal to unity. The index of dispersion ($d$) is independent of cell size. In 2-D space, two successive subdivisions into quarters result in 4 and 16 cells with concentration values. The maximum element concentration value after $k$ subdivisions is $(1+d)^k$, and the minimum value is $(1-d)^k$; $k$ is even in 2-D applications in order to preserve mass but the frequency distribution of all concentration cannot be distinguished from that arising in 1- or 3-D applications of this multiplicative cascade model. In a random cascade, larger and smaller values are assigned to cells using a discrete random variable. Multifractal patterns generated by a random cascade have more than a single maximum. The frequency distribution of the element concentrations at any stage of this process is called "logbinomial" because logarithmically transformed concentration values satisfy a binomial distribution. The logbinomial converges to a lognormal distribution although its upper and lower value tails remain weaker than those of the lognormal (Agterberg, 2007a). Notation can be simplified by using indices that are powers of $(1+d)$ and $(1-d)$, respectively; for example, $(1+d)^k (1-d)^k$ is written as $31$ in the 16-cell matrix on the left in the next row. If at each stage of subdivision, the location of higher and lower concentration cells is determined by a Bernoulli-type random variable, the arrangement of cells may become as shown in the 16-cell matrix on the right of Fig. 6a. Because of its property of self-similarity, the model of de Wijs was recognized to be a multifractal by Mandelbrot (1989) who adopted this approach for applications to the Earth’s crust.

3. Non-linear processes and power-law relations

3.1. Multiplicative cascade models

Multifractal cascade models play a fundamental role in quantifying turbulent intermittency and other non-linear processes (Scherzter et al., 1997). Two relatively simple 2-D multiplicative cascade models are the model of de Wijs (1951) (Agterberg, 2007a) and Turcotte’s “fractal” cascade (Turcotte, 1997). These models are graphically illustrated in Figs. 6–8. In the original model of de Wijs, any block of rock is divided into two equal parts. The copper (Fig. 5). For arsenic, the second straight-line segment in the C–A plot (Fig. 4h) extends from 7.147 to 7.955 with best-fitting straight-line: $y = 25.6069-3.0178x$. In addition to the C–A method, several other methods for non-linear spatial information extraction have been developed during the past 10 years to aid in the analysis of regional geochemical and geophysical map data (Cheng, 1999). These include integrated spatial and spectral analysis for geochemical anomaly separation (Cheng et al., 2001), eigenvalue–eigenvector analysis of multifractal fields (Li and Cheng, 2004; Cheng, 2005), and use of local singularities for spatial interpolation (Cheng, 2006). Most of these methods have been incorporated in the software package GeoDAS (Cheng, 2003) that has also been used to obtain the results in this report.
associated with a fractal instead of a multifractal. More detailed analysis about the Turcotte’s model can be found in Lovejoy and Schertzer (2007). The slope $b$ of the straight-line representing this Pareto distribution on log–log paper satisfies $b = \frac{1}{C_0 \log_2(1 + d)}$. It is noted that this application differs from the so-called $a$ model (Schertzer et al., 1997) that also results in concentration frequency distributions with Pareto tails. Applicability of the $a$-model to stream sediment geochemistry remains to be investigated.

Already in the 1980s, Schertzer and Lovejoy (1985) had pointed out that the $p$-model can be regarded as “micro-canonical” version of their $\alpha$-model in which the strict condition of local preservation of mass is replaced by a more general condition of preservation of mass within wider neighborhoods (preservation of ensemble averages). Applicability of the $\alpha$-model to stream sediment geochemistry remains to be investigated.

Fig. 6b shows Turcotte’s model: after each subdivision, only the half with larger concentration is further subdivided.
into halves with concentration values equal to \((1+d)\xi\) and \((1-d)\xi\). This simplifies the process as illustrated for 16 cells in 2-D space. At each stage of this process the concentration values have a Pareto-type frequency distribution. In analogy with Turcotte’s derivation for blocks in 3-D space (Turcotte, 1997), it can be shown that a fractal dimension \(D\) equal to \(2 \log_2(1+d)\) can be defined for this process. The previously introduced parameter \(b\) can be converted into either the fractal dimension \(D\) \(\left(=\frac{1}{C_0^{D/2}}b\right)\) or the index of dispersion \(d\) \(\left(=2^{-D/2}-1\right)\) characterizing the non-linear process.

3.3. Dispersion index \(d\) estimation

In general, a higher index of dispersion means greater spatial variability. Linear relationship between variables in singularity (e.g. Fig. 3c) or C–A plots (e.g. Fig. 3d) means that the element concentration values have Pareto frequency distributions. Thus, the parameter \(b\) can be estimated in two ways: either as the slope of the best-fitting straight-line on a C–A plot, or from that on the local similarity plot. The results of these two methods of

**Fig. 6.** Multiplicative cascade models resulting in lognormal and Pareto frequency distributions. (a) Logbinomial multifractal model of de Wijs (Mandelbrot, 1989; de Wijs, 1951; Agterberg, 2007a, b). (b) Turcotte’ model (1997).

**Fig. 7.** Logbinomial pattern for \(d = 0.4\) and \(N = 14\). Increasing number of subdivisions for model of de Wijs (as in Fig. 6) to 14 resulted in 128 \(\times\) 128 pattern shown in which values greater than 4 were truncated. Frequency distribution of all 214 values is logbinomial and approximately lognormal except in the highest-value and lowest-value tails that are thinner than lognormal. When number of subdivisions becomes large, end product cannot be distinguished from that of multiplicative cascade models in which dispersion index \(D\) is modeled as a continuous random variable with mathematical expectation equal to 1 instead of Bernoulli variable allowing values \(+d\) and \(-d\) only. Lognormal model may provide good approximations for the regional background distribution of trace elements. However, it does not result in a highest-value tail with Pareto distribution as observed for tin in Gejiu area.
estimation will be different if we are dealing with different kinds of anomalies. Suppose that parameters describing the two separate fractal applications are identified by the subscripts 1 (for local singularity anomalies) and 2 (for the high-concentration anomaly). From $\beta_2 = -3.1576$ (estimated slope of best-fitting line in Fig. 3d), it follows immediately that $d_1(Sn) = 0.095$. C–A diagrams for arsenic (Fig. 4h) and copper (Fig. 5b) yield estimates of $\beta_2$ equal to $-3.0178$ (As) and $-3.0856$ (Cu). Conversion into $d_2$ gives estimates of $d_2/As = 0.258$ and $d_2/Cu = 0.252$, respectively. The close resemblance of these results suggests that all three elements have the same dispersion index ($d_2 \approx 0.25$). The estimates $d_1(As) = 0.082$ and $d_1(Cu) = 0.087$ are slightly less than $d_1(Sn) = 0.095$. Spatial variability within the broad regional anomaly with $d_2 \approx 0.25$ (for Sn, As, and Cu) exceeds that for the low-singularity anomalies with $d_1(Sn) = 0.095$, $d_1(As) = 0.082$, and $d_1(Cu) = 0.087$. Combined area of anomalies with local singularities less than 2 occupies about 10% of total study area for tin and other elements enriched inside and outside the tin deposits. Although the anomalies satisfy a Pareto-type frequency distribution, it can be concluded from the shapes of singularity histograms (e.g. Fig. 3a) for the entire study area that relative frequencies of other local singularities ($>2$) decrease with increasing $\alpha$. Observed frequencies close to zero are reached for singularities of about 3. The frequency distribution of concentration values in the remaining 90% of the area has a shape that is positively skewed with zero frequency at zero concentration value. Thus relative frequencies first increase toward a peak and then decrease for larger tin concentration values. This larger (90%) part of the Gejiu area can be regarded as background on which the low-singularity anomalies are superimposed. However, this background includes the broad regional high-value anomaly in the eastern part of the Gejiu area represented by the Pareto distributions for largest observed values in the C–A diagrams. This secondary fractal process reflects pollution due to past mining activities.

4. Conclusions

Singularity mapping for tin, arsenic, and other elements enriched in Gejiu mineral district show many local anomalies where $\alpha < 2$. Because known tin deposits are spatially correlated with these anomalies, local singularity mapping can be a useful tool in prospecting for buried mineral deposits. Fractal modeling of the observed highest concentration values in the eastern part of the Gejiu mineral district can help to assess and reduce pollution associated with future mining activities.

Acknowledgments

We thank S. Lovejoy (McGill University, Montreal) and Zhao Pengda (China University of Geosciences) for helpful suggestions. The research was financially supported by a Distinguished Young Researcher Grant (40525009), a Strategic Research Grant (40638041) awarded by the Natural Science Foundation of China and a High-Tech Research and Development Grant (2006AA06Z115) by the Ministry of Science and Technology of China.

References