



Exploratory Data Analysis Applied in Mapping Multi-element Soil Geochemical Anomalies for Drill Target Definition: A Case Study from the Unpha Layered Non-magmatic Hydrothermal Pb-Zn Deposit, DPR Korea

JANG Gwang-Hyok*, WON Hyon-Chol, HWANG Bo-Hyon and CHOI Chol-Man

Faculty of Geology, Kim Il Sung University, Pyongyang, Democratic People's Republic of Korea

Abstract: A factor analysis was applied to soil geochemical data to define anomalies related to buried Pb-Zn mineralization. A favorable main factor with a strong association of the elements Zn, Cu and Pb, related to mineralization, was selected for interpretation. The median + 2MAD (median absolute deviation) method of exploratory data analysis (EDA) and C-A (concentration-area) fractal modeling were then applied to the Mahalanobis distance, as defined by Zn, Cu and Pb from the factor analysis to set the thresholds for defining multi-element anomalies. As a result, the median + 2MAD method more successfully identified the Pb-Zn mineralization than the C-A fractal model. The soil anomaly identified by the median + 2MAD method on the Mahalanobis distances defined by three principal elements (Zn, Cu and Pb) rather than thirteen elements (Co, Zn, Cu, V, Mo, Ni, Cr, Mn, Pb, Ba, Sr, Zr and Ti) was the more favorable reflection of the ore body. The identified soil geochemical anomalies were compared with the in situ economic Pb-Zn ore bodies for validation. The results showed that the median + 2MAD approach is capable of mapping both strong and weak geochemical anomalies related to buried Pb-Zn mineralization, which is therefore useful at the reconnaissance drilling stage.

Key words: factor analysis, exploratory data analysis, Mahalanobis distance, multi-element, Unpha

Citation: Jang et al., 2021. Exploratory Data Analysis Applied in Mapping Multi-element Soil Geochemical Anomalies for Drill Target Definition: A Case Study from the Unpha Layered Non-magmatic Hydrothermal Pb-Zn Deposit, DPR Korea. *Acta Geologica Sinica (English Edition)*, 95(4): 1357–1365. DOI: 10.1111/1755-6724.14404

1 Introduction

Purely objective separation of geochemical anomalies from background noise has been a major subject of animated controversy in geochemical exploration. Various statistical methods have been widely applied to delineate geochemical anomalies under some initial assumptions related to specific characteristics of the geochemical data (Li et al., 2003; Carranza, 2009, 2010a; Arias et al., 2012). Geochemical anomalies are separated from the background noise by setting a threshold value, which is fixed by the upper and lower limits of standard deviation in a particular population, based on classical statistical methods. Generally, the sign of classical robust statistical methods applied to interpret multi-element geochemical data is that the element concentrations in the crust usually follow a normal distribution. However, since the element concentration in the lithosphere does not usually follow a normal distribution, using classical robust statistical methods for finding a threshold value can lead to false indications of geochemical anomalies (Carranza, 2009). Therefore, a variety of methods have been developed for delineating geochemical anomalies, such as exploratory data analysis (EDA), fractal model etc. (Campbell, 1982; Kürzl, 1988; Chork and Mazzucchelli, 1989; Cheng et al.,

1994; Goncalves et al., 2001; Bounessah and Atkin, 2003; Li et al., 2003, 2004; Reimann and Garrett, 2005; Cheng, 2006; Ali et al., 2007; Carranza, 2009, 2010a, b; Bai et al., 2010; Afzal et al., 2010, 2011, 2012a, b; Arias et al., 2012; Geranian et al., 2013; Heidari et al., 2013). The EDA method (Tukey, 1977) is robust against the non-normality of populations (Kürzl, 1988) and generates results that are less affected by the presence of an outlier (Carranza, 2009). The ‘boxplot’ and ‘median + 2MAD’ (median absolute deviation), two effective approaches in separating geochemical anomalies in EDA method, emphasize the empirical density distribution of the element concentration in geochemical data analysis (Tukey, 1977; Kürzl, 1988; Carranza, 2009). Fractal modelling established by Mandelbrot (1982) was developed geochemical data analysis to separate geochemical anomalies from background by Cheng et al. (1994) and Cheng (1999). EDA and fractal methods have been successfully applied by many researchers to aid in distinguishing anomalies from background noise and interpretation of geochemical data (Inaki et al., 1998; Li et al., 2003, 2004; Cheng, 2006; Ali et al., 2007; Bai et al., 2010; Carranza, 2010a; Carranza and Sadeghi, 2010; Afzal et al., 2010, 2011; Arias et al., 2012).

Unpha is so far the largest layered non-magmatic hydrothermal Pb-Zn deposit in the Democratic People's Republic of Korea (DPRK) with simple forms of minerals

* Corresponding author. E-mail: kh.jang0504@ryongnamsan.edu.kp

such as galena, sphalerite and pyrite, including calcite and dolomite of non-metallic minerals indicated (Choi et al., 2011).

The main objective of this study is to evaluate the effectiveness of the median + 2MAD and fractal modelling in setting thresholds for mapping multivariate soil geochemical anomalies associated with buried Pb-Zn mineralization, that could be utilized for drilling.

2 Geological Setting

2.1 Stratigraphy and dykes

The study area is located in the vicinity of the Unpha deposit in northern Hwanghae Province, the central region of the DPRK, including quasi-knobby and karst fall generated from long-term weathering and denudation of carbonate rocks, with about 95% of the total area being covered by transported and residual soil.

The ore deposit is tectonically located in the southern part of the Sariwon-Sohung Basin in the Haeju subsidence zone. The Jikhyon, Sandangu and Mukchon groups are

distributed throughout the study area with various dykes (Fig. 1). The Jikhyon Group unconformably overlies the Proterozoic stratum and can be subdivided into four members, namely the Jangbong, Obong, Jangsusan and Ansimryong formations. The Jangbong Formation (Pt_{2zn}) consists of conglomerate, conglomeratic quartzite and quartzite, the Obong Formation (Pt_{2ob}) of sericitic siliceous schist, sericitic phyllite and marble, the Jangsusan Formation (Pt_{2zg}) of calcareous schist, sericitic siliceous schist and quartzite, with the Ansimryong Formation (Pt_{2an}) consisting of quartzite, siliceous phyllite and chlorite schist. The Sandangu Group, which unconformably overlies the Jikhyon Group, is subdivided into the Unjoksan, Tokjaesan and Chongsokdu formations. The Unjoksan Formation (Pt_{2un}) is mainly composed of argillaceous limestone, stratiform limestone and dolomitic limestone. The Tokjaesan Formation (Pt_{2dk}) is composed of dolomite, siliceous dolomite and dolomitic limestone, while the Chongsokdu Formation (Pt_{2cn}) consists of limestone and dolomitic limestone. The Mukchon Group is subdivided into the Solhuasan, Okhyon and Rinsan

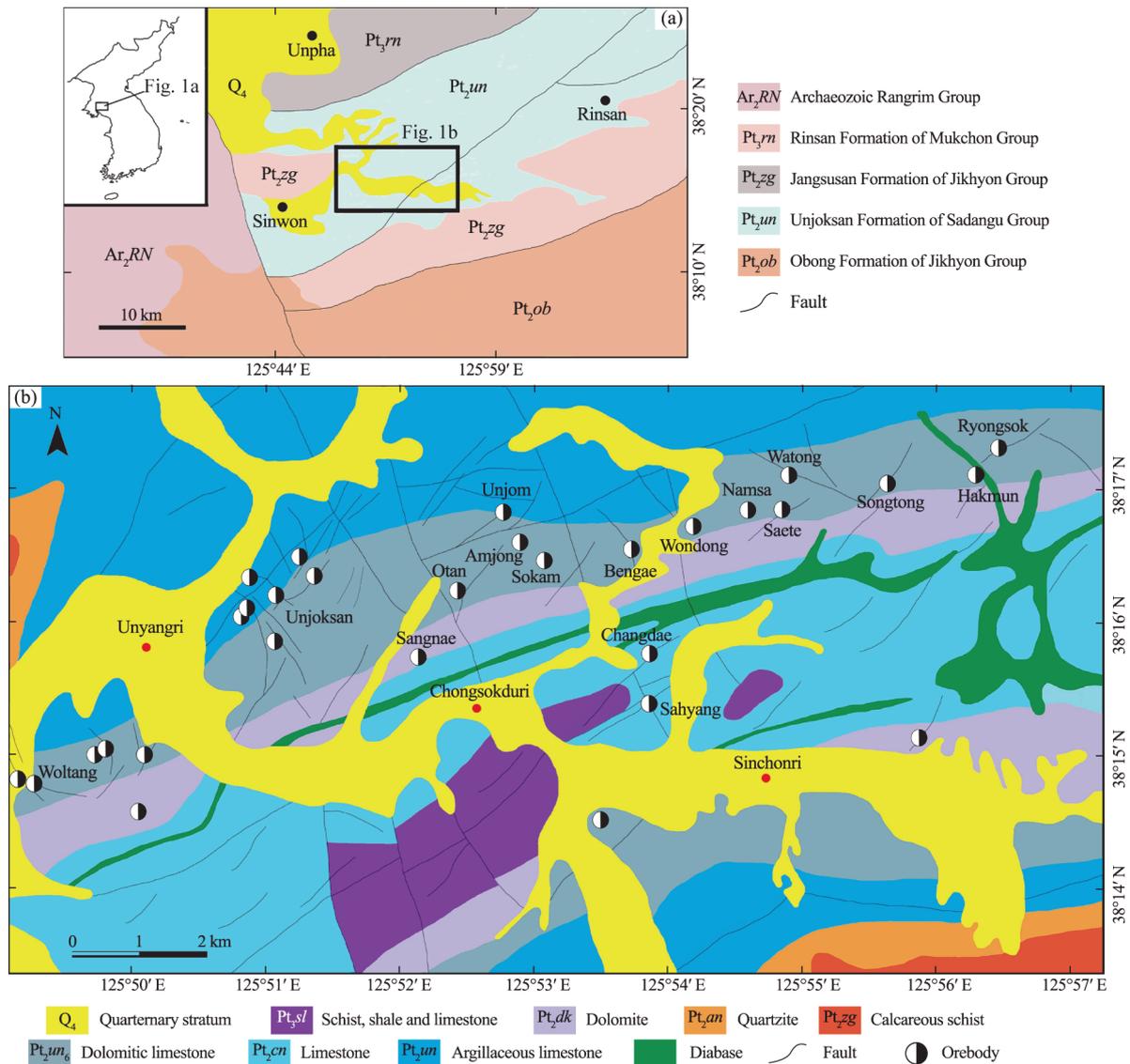


Fig. 1. Geological map of the Unpha exploration area.

formations. The Solhuasan Formation (Pt_3sl) consists of schist, quartzite, shale and limestone, while the Okhyon Formation (Pt_3ok) is limestone and the Rinsan Formation (Pt_3rm) consists of quartzite and phyllite.

There are no big intrusive rocks in the vicinity of the study area, however rocks such as gabbro-dolerite and lamprophyre dykes are widely distributed. These dykes seem most likely related to the formation of the ore-bearing structure that is the Pb-Zn ore body. NE-trending long gabbro-dolerite dykes, formed of sheeted veins and nervation veins, appear in the carbonate rocks, the gabbro-dolerite dykes being chloritized and carbonated around the ore body.

2.2 Orebodies and metallogenesis

The Unjoksan Formation is the primary ore-bearing horizon in the ore deposit region. There are many orebodies present, such as Saete, Namsa, Wondong, Bengae, Sokam, Unjom, Amjong, Otan, Sangnae, Unjoksan, Changdae, Sahyang, Hakmun, Balun, Songtong, Gukhua, Ryongsok, Kwuakgol, Watong, Woltang etc. Some of them, such as Saete, Namsa and Bengae, constitute the major orebodies. Saete orebodies are distributed in the dolomitic limestone stratum and consist of galena, sphalerite and pyrite, the zinc content being higher than the lead. Namsa orebodies are located in the graphite-fill fracture zone associated with the dolomitic limestone stratum and also consist of galena, sphalerite and pyrite. Bengae orebodies are similarly situated in the dolomitic limestone stratum.

Many researchers' opinions on the genesis of the Unpha deposit have been summarised elsewhere (Choi et al., 2011). Firstly, a high content of sulphides such as galena, sphalerite etc. were observed in the country rock, with most orebodies having consistency with a definite horizon or fracture zone. Secondly, specific features of sedimentary deposition, such as a typically impregnated structure, colloidal structure etc. were observed. Thirdly, the orebody is not related to an intrusion, but is linked with a certain stratum (Unjoksan Formation) and the contour of the orebody is vague. Lastly, the isotopic composition of the sulphur and carbon present in the sulfide and carbonate minerals have been determined, along with the physicochemical parameters (T (temperature), f_{O_2} (fugacity of oxygen), pH etc.). As a result, it can be demonstrated that the lead and zinc in the ore deposit were primarily deposited in the stage when the Sadangu Group was deposited, when the Mesozoic era

Hyesan magmatic complex was active, due to the recycling of the hydrothermal solution that occurred due to the thermal effects of the hidden intrusive, pre-depositional lead and zinc being hydrothermally realigned in a state of relatively high f_{O_2} with $(150 \pm 20)^\circ\text{C}$ temperature, pH 6–8, and $\log f_{O_2} > -49$. Therefore, the work of many researchers has clarified that the Unpha deposit is a stratiformic non-magmatogenic hydrothermal Pb-Zn deposit.

3 Characteristics of the Geochemical Dataset

3.1 Sampling

A total of 15679 soil samples were collected in a regular grid of $100\text{ m} \times 20\text{ m}$ in an area of 120 km^2 .

Soil samples were dried at 60°C and finely ground using an agate swing mill. The chemical analyses were carried out by an X-ray fluorescence (XRF) and atomic absorption (AAS) technique for 35 elements in a geological analytical laboratory. The chemical elements of the study area were as follows: Sr, Co, Zn, Y, Yb, Ag, Cu, Cd, Ce, La, Li, Sn, Mo, Sc, Be, Bi, V, Ba, Ni, Cr, Ga, Mn, Ge, Ti, Pb, Nb, Zr, In, W, Sb, B, P, Tl, Hf and As. Twenty two elements unrelated to Pb-Zn mineralization or showing discontinuous content data were removed from interpretation, therefore, thirteen elements, namely Co, Zn, Cu, V, Mo, Ni, Cr, Mn, Pb, Ba, Sr, Zr and Ti, were used to validate the soil geochemical anomalies.

3.2 Statistical characteristics of the geochemical dataset

The basic statistical parameters and correlation coefficients of the concentrations of the thirteen elements in the soil samples are listed in Table 1 and Table 2.

3.3 Hurst exponent

The Hurst exponent is directly related to the fractal dimension of a process, which gives a measure of the irregular process (Hurst et al., 1965). The Hurst exponent expressing a self-similarity measures the long-range dependence in a time series, which provides a measure of long-term nonlinearity. The expected values of the Hurst exponent, H , vary from 0 to 1. For $H = 0.5$, the cumulative behavior is a random process. $H < 0.5$ expresses anti-persistent behavior and $H > 0.5$ represents fractional Brownian motion with increasing persistence intensity as H approaches 1.

The rescaled range statistic (R/S) analysis can be used

Table 1 Basic statistical parameters of thirteen elements (based on 15,679 samples)

Element	Mean (ppm)	Median (ppm)	Minimum (ppm)	Maximum (ppm)	Variance	Std. Dev.	Skewness	Kurtosis
Co	11.67	6	1	532	227	15.07	10.15	248.63
Zn	85.77	52	1	10367	148832	385.79	20.98	493.72
Cu	15.27	6	1	6120	3568	59.73	57.34	4923.61
V	25.02	6	0.3	1213	1320	36.33	4.17	59.77
Mo	2.93	3	0.2	108	7	2.61	15.63	435.48
Ni	19.12	4	1	526	716	26.77	3.07	29.73
Cr	30.77	6	1	2128	2681	51.78	12.66	378.45
Mn	515.17	413	2	11358	128151	357.98	5.92	95.98
Pb	33.12	15	1	10375	81731	285.89	29.01	921.73
Ba	254.31	203	10	11216	71392	267.19	17.42	564.69
Sr	334.48	312	20	6328	97394	312.08	5.21	52.27
Ti	559.81	152	2	121232	1774027	1331.93	41.92	3509.21
Zr	123.89	60	2	1173	22717	150.72	1.91	2.57

Table 2 Correlation coefficients of thirteen elements

	Co	Zn	Cu	V	Mo	Ni	Cr	Mn	Pb	Ba	Sr	Zr	Ti
Co	1	-0.34	-0.07	0.11	0.61	-0.33	-0.51	-0.67	-0.12	0.71	-0.41	0.11	-0.22
Zn		1	0.78	0.86	-0.43	-0.18	0.09	-0.24	-0.43	-0.17	-0.33	0.67	-0.33
Cu			1	0.77	-0.23	-0.24	0.15	-0.39	-0.05	-0.45	-0.34	0.86	-0.52
V				1	-0.38	-0.36	0.04	-0.23	-0.48	0.02	-0.53	0.67	-0.19
Mo					1	-0.18	-0.29	-0.52	-0.38	0.21	0.12	0.05	-0.14
Ni						1	0.58	-0.07	0.09	-0.18	0.16	-0.17	-0.28
Cr							1	0.21	0.17	-0.39	0.12	0.16	-0.09
Mn								1	-0.31	-0.11	0.32	-0.51	0.73
Pb									1	-0.38	-0.27	0.04	-0.27
Ba										1	-0.27	-0.32	0.28
Sr											1	-0.24	0.26
Zr												1	-0.52
Ti													1

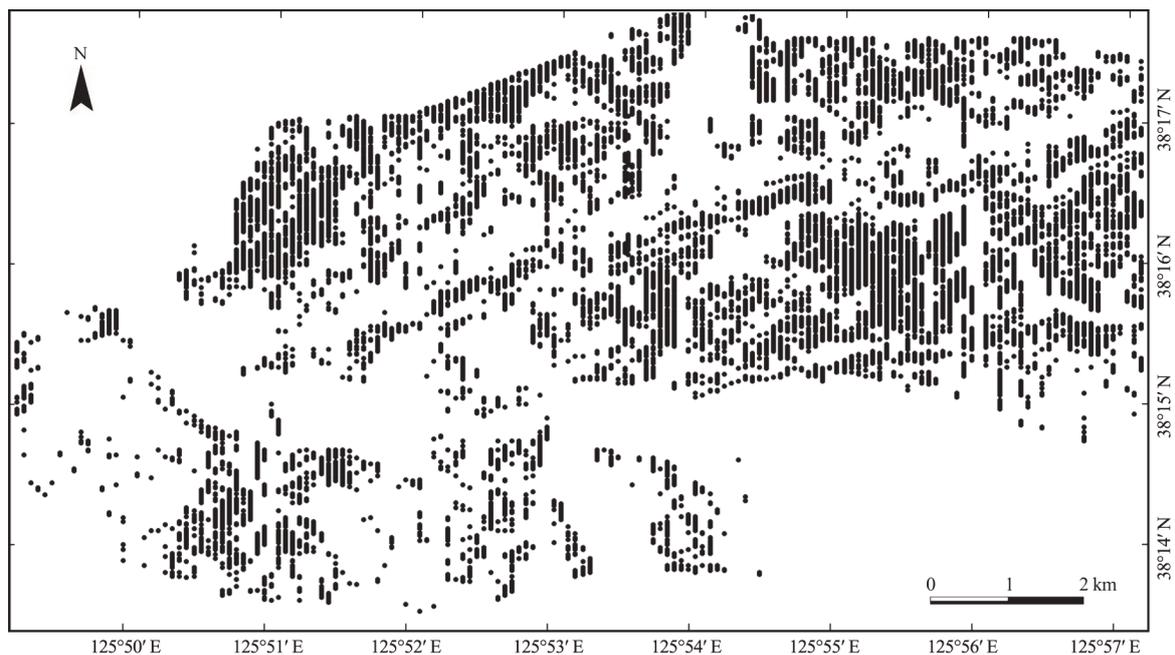


Fig. 2. Soil geochemical sampling map at the Unpha deposit.

to evaluate the Hurst exponent (Mandelbrot and Wallis, 1969). The results obtained by Hurst exponents using R/S analysis for the thirteen elements are shown in Fig. 3.

4 Factor Analysis

Factor analysis is a technique for describing relationships between variables in a low-dimensional space. Two main methods, principal factor analysis (PFA) and the maximum likelihood method (ML), exist for extracting the common factors in factor analysis. PFA works with a reduced correlation or covariance matrix, whereas ML uses a complicated statistical optimization procedure to extract the factors. In common with many other statistical methods, factor analysis is very sensitive to non-normality of populations. Therefore, it is important to test whether or not all variables follow a normal distribution prior to factor analysis.

Geological and mineralization processes generate diverse geochemical features that are expressed in more than one element in anomalies. Regional geochemical data practically never show a normal distribution, often showing compositional and multivariate characteristics,

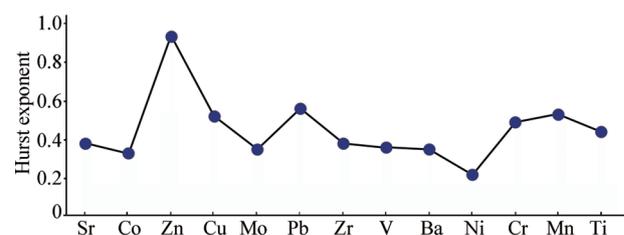


Fig. 3. Hurst exponent curve line of element distribution.

therefore a data transformation is crucial to reveal their interelement relationships hidden in multivariate datasets prior to analysis (Aitchison, 1986; Reimann et al., 2002; Egozcue et al., 2003; Filzmoser et al., 2009; Carranza, 2009, 2011; Zuo et al., 2012, 2013a, b; Buccianti and Grunsky, 2014; Zuo, 2014). In order that all entered variables come as close as possible to a normal distribution, the much more widespread log-transformation, among many different transformations (square root, logit, etc.), is used to result in a nearly normal distribution.

The factor loadings are carried out by an orthogonal-rotation of the coordinate system. There are many different

methods for factor rotation such as Varimax, Promax, Oblimin or Quartimin. Varimax and Promax are orthogonal rotations, but Oblimin and Quartimin are oblique rotation methods. The rotated factors are not correlated in the former case and the rotated factors can be correlated in the latter case. The results of the Varimax orthogonal rotation are more stable than the others (Reimann et al., 2002). The factor analysis achieved by using the PFA method and Varimax rotation are given in Table 3.

Six factor loadings, representing various element associations and mineralization processes, were extracted. The total explained variance is 87.5% from six factors. In summary, we interpret factor F1 to be a Co-Mo factor, factor F2 as a V-Ni-Cr-Mn-Ba-Zr-Ti factor, factor F3 as a Sr factor, factor F4 as a Zn factor, factor F5 as a Cu factor and factor F6 as a Pb factor. The Mahalanobis distance, a measure of distance between two points in space defined by correlated variables such as Zn, Cu and Pb from factor analysis, was used to interpret the soil geochemical anomalies at the Unpha exploration area.

5 Concentration-area (C-A) Fractal Modeling

Concentration-area (C-A) fractal modeling proposed by Cheng et al. (1994), could be used to separate the geochemical anomalies from the background noise. The general form of the C-A model is as follow:

$$A(\rho \leq v) \propto \rho^{-d}$$

where $A(\rho)$ represents the area with concentration values greater than the contour value ρ ; v indicates the threshold; and d is the characteristic exponent or fractal dimension. $A(\rho)$ with element concentrations greater than the value ρ usually forms a power-law relation. The breaks between straight-line segments on this plot and the corresponding values of ρ have been used as cut-offs to separate geochemical values into different causal factors, such as geological differences and geochemical processes (Cheng et al., 1994; Carranza, 2009; e.g. Afzal et al., 2010; Arias et al., 2012; Zuo et al., 2012, 2013a, b).

The C-A fractal model was applied to the Mahalanobis distances from the soil geochemical data and two line segments, fitted to the C-A log-log plot, were generated (Figs. 4a, b). The break between the straight-line segments and the corresponding values of the Mahalanobis distances

have been used as threshold values to delineate the geochemical anomaly map. The threshold on this anomaly map is indicative of two populations, which is interpreted as the geochemical background and anomalies of lead and zinc, or associated mineralization elements.

6 Exploratory Data Analysis (EDA)

The boxplot and median + 2MAD techniques of the

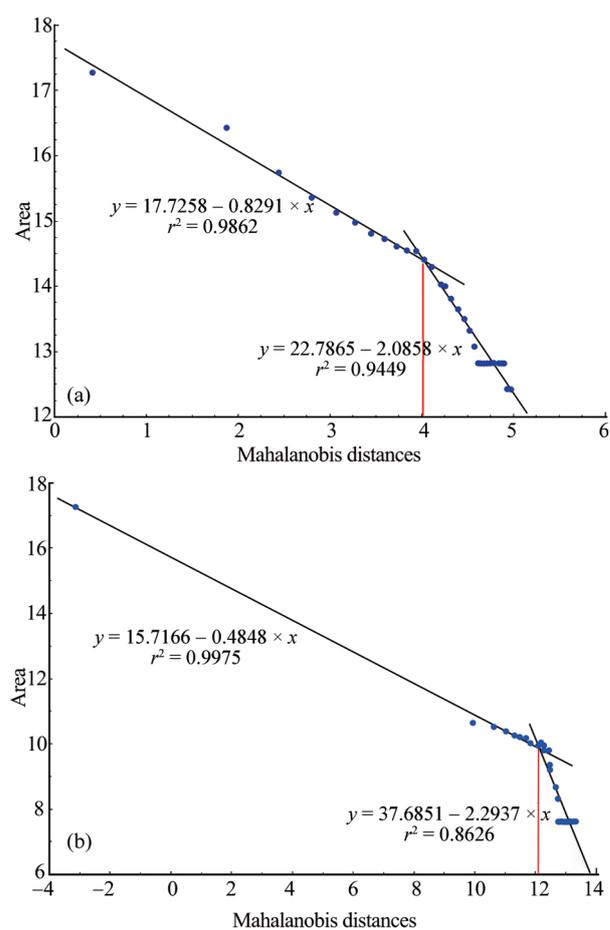


Fig. 4. C-A log-log plot of the MD_1 (a) and MD_2 (b), showing the thresholds with red lines.

Table 3 Varimax rotated factor loadings applied to 13 variables

Element	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Co	0.884241	0.366785	-0.058389	0.062708	0.076848	0.14508
Zn	0.056642	0.10646	0.018536	0.980818	0.025501	0.074278
Cu	0.119772	0.236421	-0.009941	0.025624	0.91853	0.093387
V	0.46402	0.690806	-0.180194	0.136302	0.264334	0.175197
Mo	0.958068	0.02109	0.00968	0.012163	0.056049	0.118218
Ni	0.575682	0.705899	-0.142112	0.112829	0.213511	0.160051
Cr	0.450625	0.794326	-0.111517	0.078508	0.152637	0.100715
Mn	-0.199951	0.599969	0.251045	-0.111335	-0.261817	0.330429
Pb	0.255951	0.11614	0.068332	0.096658	0.102004	0.91387
Ba	0.039987	0.715884	0.539522	0.076814	-0.079647	0.090564
Sr	-0.063155	-0.181036	0.926654	0.010088	-0.003184	0.057656
Zr	0.116604	0.840479	-0.023268	0.021614	0.144958	0.004732
Ti	0.119456	0.845755	-0.238104	0.127419	0.236978	0.045813
Eigenvalue	5.786593	1.744157	1.373938	0.970478	0.814676	0.686043
Total (%)	44.51225	13.41659	10.56875	7.46521	6.26674	5.27725
Cumulative	44.51225	57.92884	68.4976	75.96281	82.22954	87.5068

The values in bold show the important elements in the factor loadings.

EDA approach have been widely applied to delineate geochemical anomalies (Zhang and Selinus, 1998; Reimann et al., 2002; Carranza, 2009; Chiprés et al., 2009). A boxplot is built around the median (Tukey, 1977; Carranza, 2009) and the MAD is estimated as follows (Tukey, 1977).

$$\text{MAD} = \text{median}[|X_i - \text{median}(X_i)|]$$

where the $|X_i - \text{median}(X_i)|$ are absolute differences between X_i and median of these values.

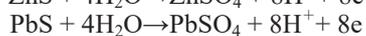
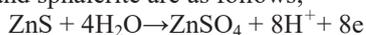
Since the threshold defined by the median + 2MAD method is lower than that defined by the boxplot method (Reimann and Garrett, 2005), the median + 2MAD method was applied to the Unpha soil geochemical data to define the thresholds between the multi-element anomalies and the background noise (Table 4).

In comparison with the C-A fractal method (Fig 4), the threshold by median + 2MAD method is lower (Table 4).

7 Discussion

Fig. 3 shows the Hurst exponents of various elements with high fit quality at the study area. The values of the Hurst exponents of Zn, Y and Ag range from 0.62 to 0.93, all greater than 0.5, indicating both persistent phenomena and better continuity of mineralization. In contrast, the Hurst exponents of other elements such as Sr, Co, Li, Mo, Be, V, Ba, Ni, Ga and Zr, range from 0.22 to 0.38, all less than 0.5, indicating the presence of anti-persistent phenomena. The Hurst exponents of Cu, Mn, Pb, Cr, Ti, La and Sn are mostly around 0.5, indicating random distribution. Zinc and lead are the main elements related to mineralization at the study area. Comparing the Hurst exponents of zinc and lead with each other, the Hurst exponent of zinc is 0.93, further beyond 0.5, representing both a persistent distribution and good continuous mineralization, but lead is 0.56, which is close to 0.5, indicating a mostly random distribution. The Hurst exponents of zinc and lead show that zinc has a most homogeneous distribution and lead distribution has the greatest randomness in the soil dispersion field in the study region, which is in accordance with the geochemical characteristics of zinc and lead.

Essentially, sulfide oxidation is an electrochemical process. For example, electrochemical reactions generated from galena and sphalerite are as follows;



where the solubilities of ZnSO_4 and PbSO_4 are 35.0 and 0.0042 at 20°C, respectively (David, 2003). Because of the greater solubility of ZnSO_4 relative to PbSO_4 , most

zinc ions are beyond the oxidation zone. Moreover, the electrode potential of galena is 0.30 ± 0.10 V, and that of sphalerite is -0.05 ± 0.01 V, that is, the former is greater than the latter, sphalerite therefore oxidizing earlier than galena under the association of these two sulfides (Ryss, 1983). Zinc having the large-scale homogeneous anomalous field and lead having the small scale non-homogeneous anomalous field, this geochemical process explains why the Hurst exponent of zinc is greater than lead.

The second factor (Table 2), with 13.4% of the total data variance, indicated that elements such as V, Ni, Cr, Mn, Ba, Zr and Ti, did not reflect the Pb-Zn mineralization. The elements indicated by the second factor are components of gabbro-dolerite, which is in accordance with the Pb-Zn ore body distribution spatially, and we therefore interpreted the second factor as gabbro-dolerite.

The first factor, with 44.5% of the total data variance, indicated Co and Mo, then the third factor with 10.6% of the total data variance indicated Sr. We found that Co, Mo and Sr are negatively correlated with the main elements such as Pb and Zn from Table 2. Therefore, the first and third factor implies the elements that are negatively related to the main elements.

Two classes— MD_1 and MD_2 —of multivariate soil geochemical anomalies were mapped by the median + 2MAD and C-A fractal method at the Unpha exploration area. There is a similarity between the multivariate anomalies of MD_1 and MD_2 mapped by the median + 2MAD method (Figs. 5, 6), however, there were differences between the soil anomalies of MD_1 and MD_2 mapped by the C-A fractal method (Figs. 7, 8). Due to the lower threshold values determined by the median + 2MAD method (Table 3), the area of the mapped soil anomalies is larger than the anomalies mapped by the C-A fractal method.

The four multivariate anomaly maps (Figs. 5–8) revealed almost uniformly NE–SW trending multivariate soil geochemical anomalies at the study area. Unlike the C-A fractal method, the median + 2MAD method shows not only the strong multivariate anomalies, but also delineated relatively smaller anomalies in soil geochemical dispersion of the study area (Figs. 5, 6). The strong multivariate anomalies are mainly associated with outcropping Pb-Zn mineralized zones, whereas the moderate multivariate anomalies could be related to the buried Pb-Zn mineralization. The overlapped results between the soil anomalies and the economic Pb-Zn ore bodies are given in Table 5.

The results showed that there is 78.9% overlap between the soil anomalies by the median + 2MAD method on the Mahalanobis distances defined by thirteen elements and the economic Pb-Zn ore bodies, the soil anomalies by the median + 2MAD method on the Mahalanobis distances defined by three elements showed 84.2% overlap with the economic Pb-Zn ore bodies (Table 5), the latter being better than the former. Additionally, there is a 68.4% overlap between the soil anomalies shown by the C-A fractal method on the Mahalanobis distances defined by thirteen elements and the economic Pb-Zn ore bodies, the

Table 4 The thresholds to separate anomalies from background in soil samples

Method	MD_1		MD_2	
	Background	Anomalies	Background	Anomalies
C-A fractal	1.51–56.1	56.1–168.4	0.03–187962.0	187962–623680.3
Median + 2MAD	1.51–10.4	10.4–168.4	0.03–29.4	29.4–623680.3

MD_1 : Mahalanobis distance in the space defined by thirteen variables (Co, Zn, Cu, V, Mo, Ni, Cr, Mn, Pb, Ba, Sr, Zr and Ti); MD_2 : Mahalanobis distance in the space defined by three variables (Zn, Cu and Pb).

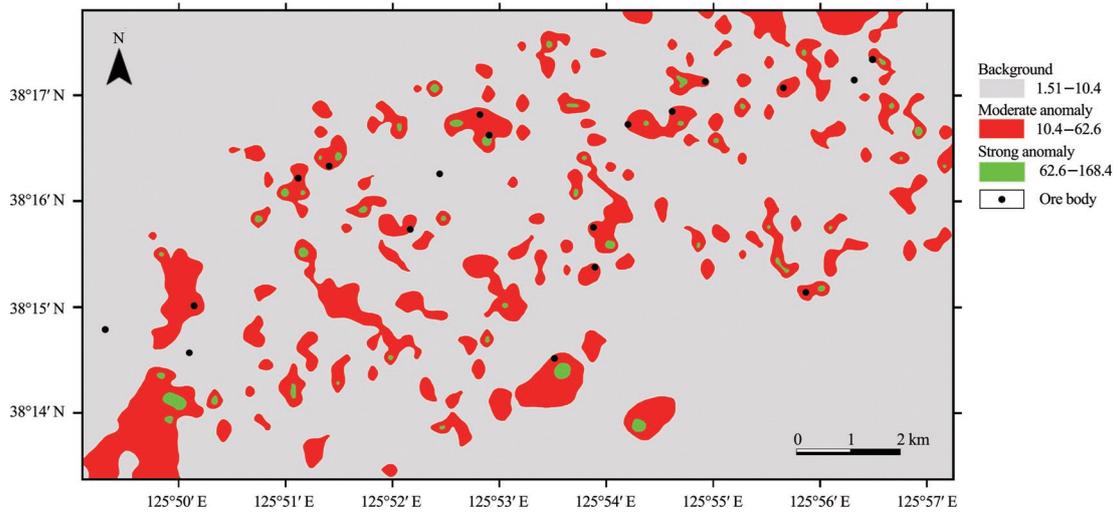


Fig. 5. Multi-element soil geochemical anomaly map of MD₁, based on the median + 2MAD method.

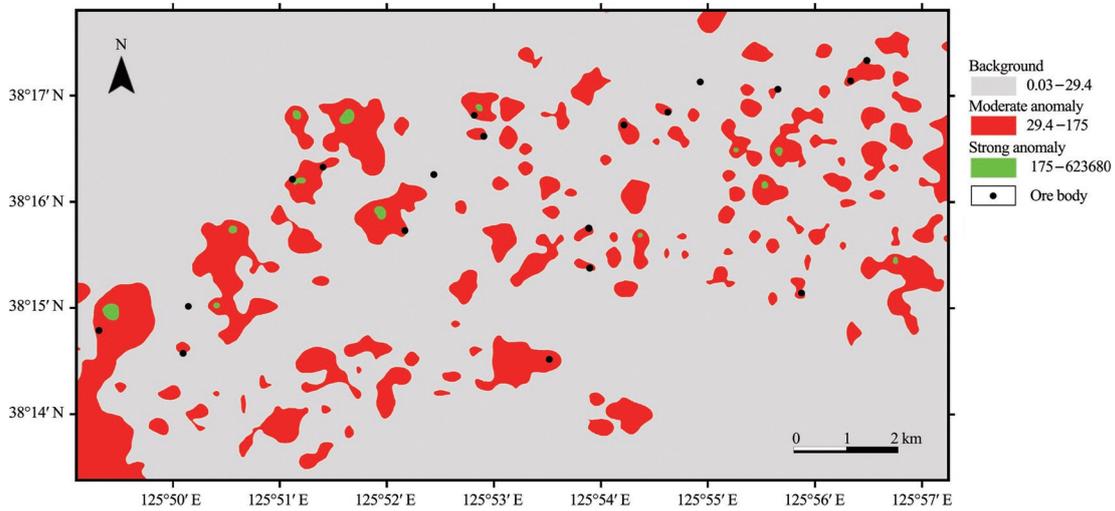


Fig. 6. Multi-element soil geochemical anomaly map of MD₂, based on the median + 2MAD method.

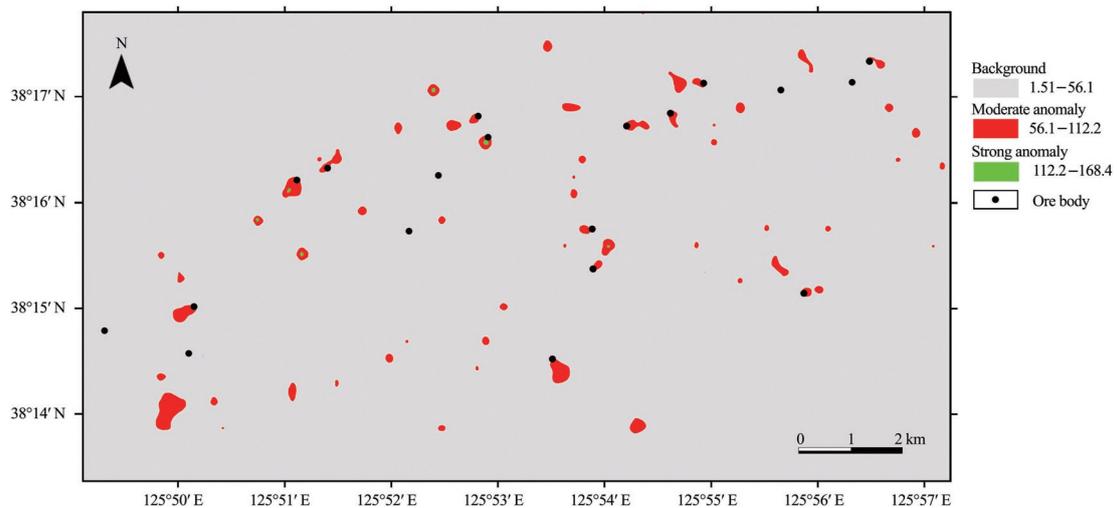


Fig. 7. Multi-element soil geochemical anomaly map of MD₁, based on the C-A fractal method.

soil anomalies by the C-A fractal method on the Mahalanobis distances defined by three elements showing a 52.6% overlap with the economic Pb-Zn ore bodies, the

former being better than the latter.

These imply that the multi-element soil anomaly shown by the median + 2MAD method is better than the C-A

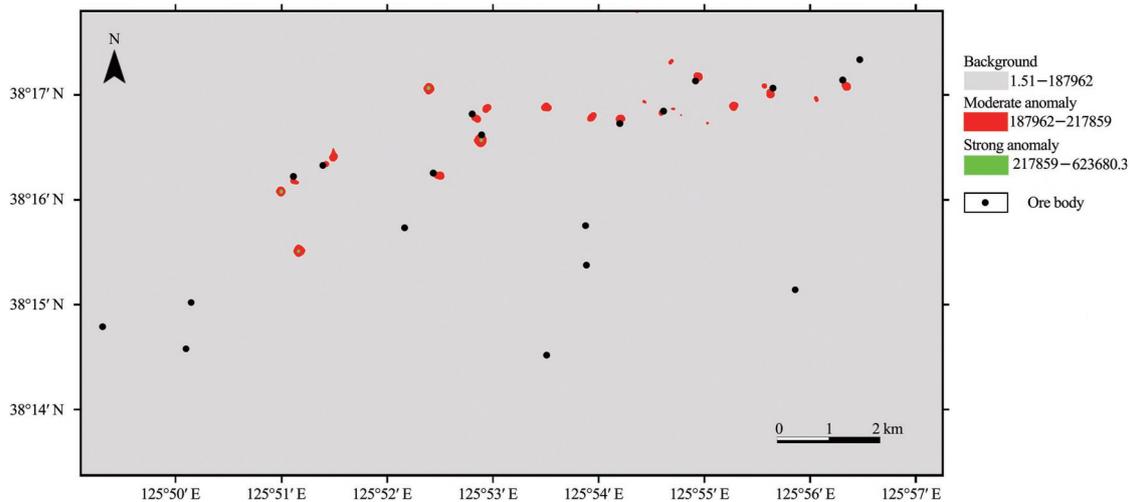


Fig. 8. Multi-element soil geochemical anomaly map of MD₂, based on the C-A fractal method.

fractal method as a favorable reflection of ore body at the study area. The explanation for this is due to the low threshold values defined by the median + 2MAD method compared to the C-A fractal method (Table 4). Therefore, the weak anomalies related to ore bodies may not be detected by the C-A fractal method.

The higher overlap between the soil anomalies shown by the median + 2MAD method on the Mahalanobis distances defined by three elements rather than thirteen elements and the economic ore bodies is due to the three indicator elements (Zn, Cu and Pb) related to the ore body being selected from robust factor analysis (Table 3). Therefore, classifying the elements by a proper statistical method in consideration of geological features is important to detect an anomaly related to an ore body.

8 Conclusions

(1) A robust factor analysis, applied to the soil geochemical data of the Unpha Pb-Zn deposit, showed that the 4th, 5th and 6th factors indicated Zn, Cu and Pb respectively, the main elements related to mineralization, and a second factor is associated with gabbro-dolerite.

(2) By comparing the identified multivariate soil geochemical anomalies, it was concluded that the median + 2MAD, rather than the C-A fractal method, is the effective method in distinguishing the soil geochemical anomalies from the background. Anomalies indicated by both methods, median + 2MAD and C-A fractal, delineated a coherent NE-SW trending multivariate soil geochemical anomaly at the study area.

(3) The soil anomaly shown by the median + 2MAD method on the Mahalanobis distance defined by three principal elements rather than thirteen elements is the most accurate reflection of the ore body. Therefore, this method is useful for reconnaissance drilling or a semi-detailed drilling survey.

(4) It is important to classify the elements appropriately to delineate a multivariate soil geochemical anomaly. Except for the factor analysis suggested in this paper, there are other robust statistical methods that could be used for

classifying the elements. This method is useful for the soil geochemical dispersion field and possibly warrants further research in the lithological dispersion field.

Acknowledgments

This work was carried out at the Geochemical Exploration Research Institute of Department of Geology, Kim Il Sung University. We thank Dr. Ryu Tong-Gwon for their collaboration in the work.

Manuscript received Feb. 14, 2019

accepted Nov. 15, 2019

associate EIC: XU Jifeng

edited by Jeff LISTON and GUO Xianqing

References

- Afzal, P., Khakzad, A., Moarefvand, P., Omran, N.R., Esfandiari, B., and Alghalandis, F.Y., 2010. Geochemical anomaly separation by multifractal modeling in Kahang (Gor Gor) porphyry system, Central Iran. *Journal of Geochemical Exploration*, 104: 34–46.
- Afzal, P., Alghalandis, Y.F., Khakzad, A., Moarefvand, P., and Omran, N.R., 2011. Geochemical anomaly separation by multifractal modeling in Kahang (Gor Gor) porphyry system, Central Iran. *Journal of Geochemical Exploration*, 108: 220–232.
- Afzal, P., Alghalandis, F.Y., Moarefvand, P., Omran, N.R., and Haroni H.A., 2012a. Application of power-spectrum-volume fractal method for detecting hypogene, supergene enrichment, leached and barren zones in Kahang Cu porphyry deposit, Central Iran. *Journal of Geochemical Exploration*, 112: 131–138.
- Afzal, P., Zarifi, A.Z., and Yasrebi, A.B., 2012b. Identification of uranium targets based on airborne radiometric data analysis by using multifractal modeling, Tark and Avanligh 1:50 000 sheets, NW Iran. *Nonlinear Process in Geophysics*. 19: 283–289.
- Aitchison, J., 1986. *The Statistical Analysis of Compositional Data*. Chapman and Hall, London, 1–416.
- Ali, K.H., Cheng, Q., and Zhijun, C., 2007. Multifractal power spectrum and singularity analysis for modelling stream sediment geochemical distribution patterns to identify anomalies related to gold mineralization in Yunnan Province, South China. *Geochemistry: Exploration, Environment, Analysis*, 7 (4): 293–301.
- Arias, M., Gumiel, P., and Martín-Izard, C., 2012. Multifractal

- analysis of geochemical anomalies: A tool for assessing prospectivity at the SE border of the Ossa Morena Zone, Variscan Massif (Spain). *Journal of Geochemical Exploration*, 122: 101–112.
- Bai, J., Porwal, A., Hart, C., Ford, A., and Yu, L., 2010. Mapping geochemical singularity using multifractal analysis: Application to anomaly definition on stream sediments data from Funin Sheet, Yunnan, China. *Journal of Geochemical Exploration*, 104: 1–11.
- Bounessah, M., and Atkin, B.P., 2003. An application of exploratory data analysis (EDA) as a robust non-parametric technique for geochemical mapping in a semi-arid climate. *Applied Geochemistry*, 18(8): 1185–1195.
- Buccianti, A., and Grunsky, E., 2014. Compositional data analysis in geochemistry: Are we sure to see what really occurs during natural processes? *Journal of Geochemical Exploration*, 141: 1–5.
- Campbell, N.A., 1982. Statistical treatment of geochemical data. In: Smith, R.E. (ed.), *Geochemical Exploration in Deeply Weathered Terrain*. CSIRO Institute of Energy and Earth Resources, Floreat Park, Western Australia, 141–144.
- Carranza, E.J.M., 2009. Geochemical Anomaly and Mineral Prospectivity Mapping in GIS. *Handbook of Exploration and Environmental Geochemistry*, 11. Elsevier, 1–351.
- Carranza, E.J.M., 2010a. Catchment basin modelling of stream sediment anomalies revisited: Incorporation of EDA and fractal analysis. *Geochemistry: Exploration, Environment, Analysis*, 10(4): 365–381.
- Carranza, E.J.M., 2010b. Mapping of anomalies in continuous and discrete fields of stream sediment geochemical landscapes. *Geochemistry: Exploration, Environment, Analysis*, 10: 171–187.
- Carranza, E.J.M., 2011. Analysis and mapping of geochemical anomalies using logratiotransformed stream sediment data with censored values. *Journal of Geochemical Exploration*, 110: 167–185.
- Carranza, E.J.M., and Sadeghi, M., 2010. Predictive mapping of prospectivity and quantitative estimation of undiscovered VMS deposits in Skellefte district (Sweden). *Ore Geology Reviews*, 38: 219–241.
- Cheng, Q., 1999. Spatial and scaling modeling for geochemical anomaly separation. *Journal of Geochemical Exploration*, 65 (3): 175–194.
- Cheng, Q., 2006. Multifractal modelling and spectrum analysis: Methods and applications to gamma ray spectrometer data from southwestern Nova Scotia, Canada. *Science in China Series D Earth Sciences*, 49(3): 283–294.
- Cheng, Q., Agterberg, F.P., and Ballantyne, S.B., 1994. The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration*, 51: 109–130.
- Chiprés, J.A., De la Calleja, A., Tellez, J.I., Jimenez, F., Cruz, C., and Guerrero, E.G., 2009. Geochemistry of soils along a transect from Central Mexico to the Pacific Coast: A pilot study for continental-scale geochemical mapping. *Applied Geochemistry*, 24: 1416–1428.
- Choi, B.S., He, D.H., Yun, J.S., Sin, Y.H., and Jo, S.J., 2011. *Josenjijilchongse (6) (A series of Korean Geology, part 6)*, Industry Press, Pyongyang, 190–211 (in Korean).
- Chork, C.Y., and Mazzucchelli, R.H., 1989. Spatial filtering of exploration geochemical data using EDA and robust statistics. *Journal of Geochemical Exploration*, 34(3): 221–243.
- David R.L., 2003. *Handbook of Chemistry and Physics*, CRC Press, 1131–1132.
- Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35: 279–300.
- Filzmoser, P., Hron, K., Reimann, C., and Garrett, R., 2009. Robust factor analysis for compositional data. *Computers & Geosciences*, 35: 1854–1861.
- Geranian, H., Mokhtari, A.R., and Cohen, D.R., 2013. A comparison of fractal methods and probability plots in identifying and mapping soil metal contamination near an active mining area, Iran. *Science of the Total Environment*, 463: 845–854.
- Goncalves, M.A., Mateus, A., and Oliveira, V., 2001. Geochemical anomaly separation by multifractal modeling. *Journal of Geochemical Exploration*, 72: 91–114.
- Heidari, M., Ghaderi, M., and Afzal, P., 2013. Delineating mineralized phases based on lithochemical data using multifractal model in Touzlar epithermal Au-Ag (Cu) deposit, NW Iran. *Applied Geochemistry*, 31: 119–132.
- Hurst, H.E., Black, R.P., and Simaika, Y.M., 1965. *Long-term Storage: An Experimental Study*. Constable, London, 1–145.
- Inaki Y., Francisco V., and Jose-Miguel H., 1998. Anomaly threshold estimation and data normalization using EDA statistics: Application to lithochemical exploration in Lower Cretaceous Zn-Pb carbonate-hosted deposits, Northern Spain. *Applied Geochemistry*, 13: 421–439.
- Kürzl, H., 1988. Exploratory data analysis: Recent advances for the interpretation of geochemical data. *Journal of Geochemical Exploration*, 30(1–3): 309–322.
- Li, C., Ma, T., and Shi, J., 2003. Application of a fractal method relating concentrations and distances for separation of geochemical anomalies from background. *Journal of Geochemical Exploration*, 77: 167–175.
- Li, C., Ma, T., and Cheng, J., 2004. A fractal interpolatory approach to geochemical exploration data processing. *Mathematical Geology*, 36(5): 593–606.
- Mandelbrot, B.B., 1982. *The Fractal Geometry of Nature*. Freeman, San Francisco, 1–460.
- Mandelbrot, B.B., and Wallis, J.R., 1969. Some long-run properties of geophysical records. *Water Resources Research*, 5: 321–340.
- Reimann, C., and Garrett, R.G., 2005. Geochemical background-concept and reality. *Science of the Total Environment*, 350(1–3): 12–27.
- Reimann, C., Filzmoser, P., and Garrett, R.G., 2002. Factor analysis applied to regional geochemical data: Problems and possibilities. *Applied Geochemistry*, 17(3): 185–206.
- Ryss, Y.S., 1983. *Geoelektricheskoye Metodi Razvedki (Electrogeochemical Method in Mineral Exploration)*, Moscow, 1–249 (in Russian).
- Tukey, J.W., 1977. *Exploratory Data Analysis*. Addison-Wesley, Reading.
- Zhang, C., and Selinus, O., 1998. Statistics and GIS in environmental geochemistry—Some problems and solutions. *Journal of Geochemical Exploration*, 64: 339–354.
- Zuo, R., 2014. Identification of geochemical anomalies associated with mineralization in the Fanshan district, Fujian, China. *Journal of Geochemical Exploration*, 139: 170–176.
- Zuo, R., Carranza, E.J.M., and Cheng, Q., 2012. Fractal/multifractal modelling of geochemical exploration data. *Journal of Geochemical Exploration*, 122: 1–3.
- Zuo, R., Xia, Q., and Wang, H., 2013a. Compositional data analysis in the study of integrated geochemical anomalies associated with mineralization. *Applied Geochemistry*, 28: 202–211.
- Zuo, R., Xia, Q., and Zhang, D., 2013b. A comparison study of the C-A and S-A models with singularity analysis to identify geochemical anomalies in covered areas. *Applied Geochemistry*, 33: 165–172.

About the first and corresponding author



JANG Gwang-Hyok male, born in 1977 in Pyongyang City, DPR Korea; masters, graduated from Kim Il Sung University; senior lecturer of Department of Geochemistry, Faculty of Geology, Kim Il Sung University; he is now engaged in the study of exploration geochemistry and environmental geochemistry using 3S (GPS, GIS and RS) techniques. E-mail: kh.jang0504@ryongnamsan.edu.kp.