Comparison of mathematical methods of geochemical data processing

A Marfin[1], D V Lychagin [1,2], A Shapovalov[1] and E A Alfyorova[2]

[1] National Research Tomsk State University 634050, Russia, Tomsk, pr. Lenina, 36
[2] National Research Tomsk Polytechnic University 634050, Russia, Tomsk, pr. Lenina, 30

e-mail: dvl-tomsk@mail.ru, marfin1309@gmail.com, shpv@phys.tsu.ru, katerina525@mail.ru

Abstract. We have analyzed the data on concentration of geochemical elements within the 3-DV regional geochemical field. We have determined the fractal dimension index of spider diagrams and found the fractal nature of the distribution of chemical elements along the profile. We have compared these data with the results of the correlation analysis and hierarchical clustering methods. The agreement between the grouping of the elements according to the values of fractal dimension and the results of statistical analysis method was considered satisfactory.

1. Introduction

Geochemical work holds an important position in searching and forecasting activities. Due to the increased accuracy of chemical analysis and its widespread implementation in geological practice, it is necessary to improve the methods of statistical analysis and interpretation of the incoming data. In most cases the initial position is considered the view of the thermodynamic equilibrium of geological object or environment. But this approach is not suitable for systems formed as the result of self-organization. Therefore, the usage of the processing techniques based on the ideas of fractals and self-similarity has increased so far.

In the article by Mr. Li Wan et al., the element distribution dynamics was analyzed by examining its fluctuations [1]. Their findings state the relationship between the scale level and the concentration value, which has a fractal nature. An interesting result was obtained with the evaluation of Ag-Cu-Pb-Zn ores [2]. They determined the element distribution patterns that can be well described by the Weibull distribution with fractal characteristics.

Investigation of fractal dimension allowed Mr. Yan Shi Xie and colleagues determine the self-similar nature of Au-Pb-Zn ore areas (Hunan, China) [3]. Knowledge of elements' behavior allowed the authors to assess the perspective areas and identify anomalies more accurately.

Thus, the increasing attention of researchers to this method of analysis is induced by the fact that it is possible to detect relationships between various elements of geological environment that are generally could not be distinguished with the use of traditional methods [4, 5, 6]. While the application of the ideas of self-similarity in geomorphology, geophysics or sedimentology is a firmly established practice, in geochemical research sphere it is not so widely spread. One of the possible reasons, according to the authors, is the lack of interpretative geological database that could connect the fractal

dimension indexes and the specific environmental conditions. Undoubtedly, the regional works should form the basis of such framework as they provide an opportunity to highlight major patterns of concentration and dispersion of elements. Anyways, the results of different methods should correspond with each other and a real geological environment.

In this context, the aim of the work is fractal and statistical analysis of experimental geochemical data in order to establish the relationship between the methods' parameters and the possibility of a combined application.

2. Results and Discussion

We have analyzed the distribution of elements in the regional geochemical profile 3-DV. Profile band crosses south-eastern part of Central Siberian plate. Minerogenic structure of territory is determined by its geological tectonic position. Geochemical specialization of the crust blocks is related to their evolution and formation history, as well as the set of structural rock complexes (including the ore-bearing ones). There is a clear correspondence between geostructures of certain hierarchical level and the fields of geochemical specialization of the area.

Initial data consists of 762 analyses of geochemical elements. They identify some transit elements (Ti, Mn, Co, Ni, Cu), alkaline-earth metals (Ba, Be), high field strength elements (Pb, Zr), Li and Ga. To calculate the fractal dimension the 2-d images of spider diagrams, which show the dependence of normalized content of chemical elements in the rock on a point on the geochemical profile, i.e. the distance, were used.

By the term "fractal dimension" we define the degree of preservation of self-similarity under the changes of the scale. The diagrams were covered with a grid with the cell size equal to δ. After decreasing δ two times we counted the amount of cells that cover the spider diagram, a similar problem was solved in the example about the coastline [7].

Statistical grouping was performed by comparing the data obtained by hierarchical clustering, calculating the value of D and distribution laws.

All elements excepting lithium were combined in pairs by hierarchical clustering using complete-linkage method. Following pairs were formed: Cu-Ba, Ni-Be, Ga-Pb and Mn-Co, Zr-Ti. Element grouping plan is shown in Figure 1.

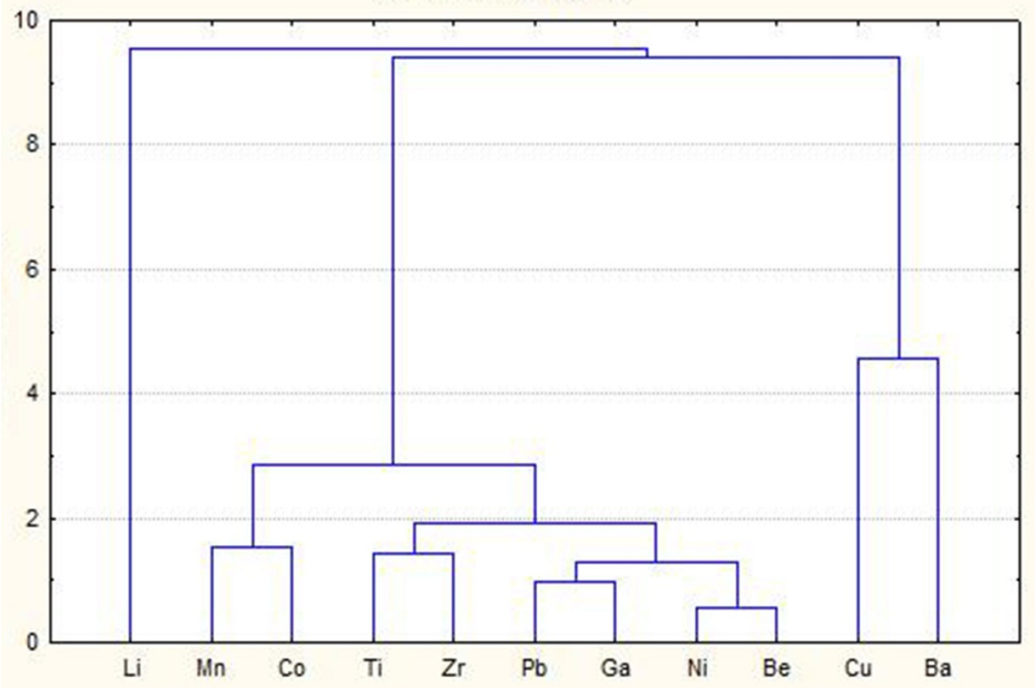

Figure 1. Dendrogram of element grouping

Correlation between the elements was studied; the results are shown in Table 1 and Figure 2.

Table 1. Correlation matrix.

	Ba	Be	Co	Cu	Ga	Li	Mn	Ni	Pb	Zr	Ti
Ba	1	0.13	-0.21	0.76	0.78	-0.14	0.13	-0.07	0.56	0	-0.03
Be	0.13	1	-0.5	0.17	0.28	-0.17	-0.3	-0.7	-0.3	0.26	0.39
Co	-0.21	-0.5	1	-0.03	-0.23	0.3	0.56	0.86	0.29	0.25	0.16
Cu	0.76	0.17	-0.03	1	0.77	0.15	0.25	0.05	0.51	0.09	0.12
Ga	0.78	0.28	-0.23	0.77	1	0.02	0.06	-0.18	0.55	0	0.24
Li	-0.14	-0.17	0.3	0.15	0.02	1	0.13	0.33	0.01	0.01	0.13
Mn	0.13	-0.3	0.56	0.25	0.06	0.13	1	0.61	0.59	0.36	-0.25
Ni	-0.07	-0.7	0.86	0.05	-0.18	0.33	0.61	1	0.43	0.21	-0.1
Pb	0.56	-0.3	0.29	0.51	0.55	0.01	0.59	0.43	1	0.14	-0.1
Zr	0	0.26	0.25	0.09	0	0.01	0.36	0.21	0.14	1	0.27
Ti	-0.03	0.39	0.16	0.12	0.24	0.13	-0.25	-0.1	-0.1	0.27	1

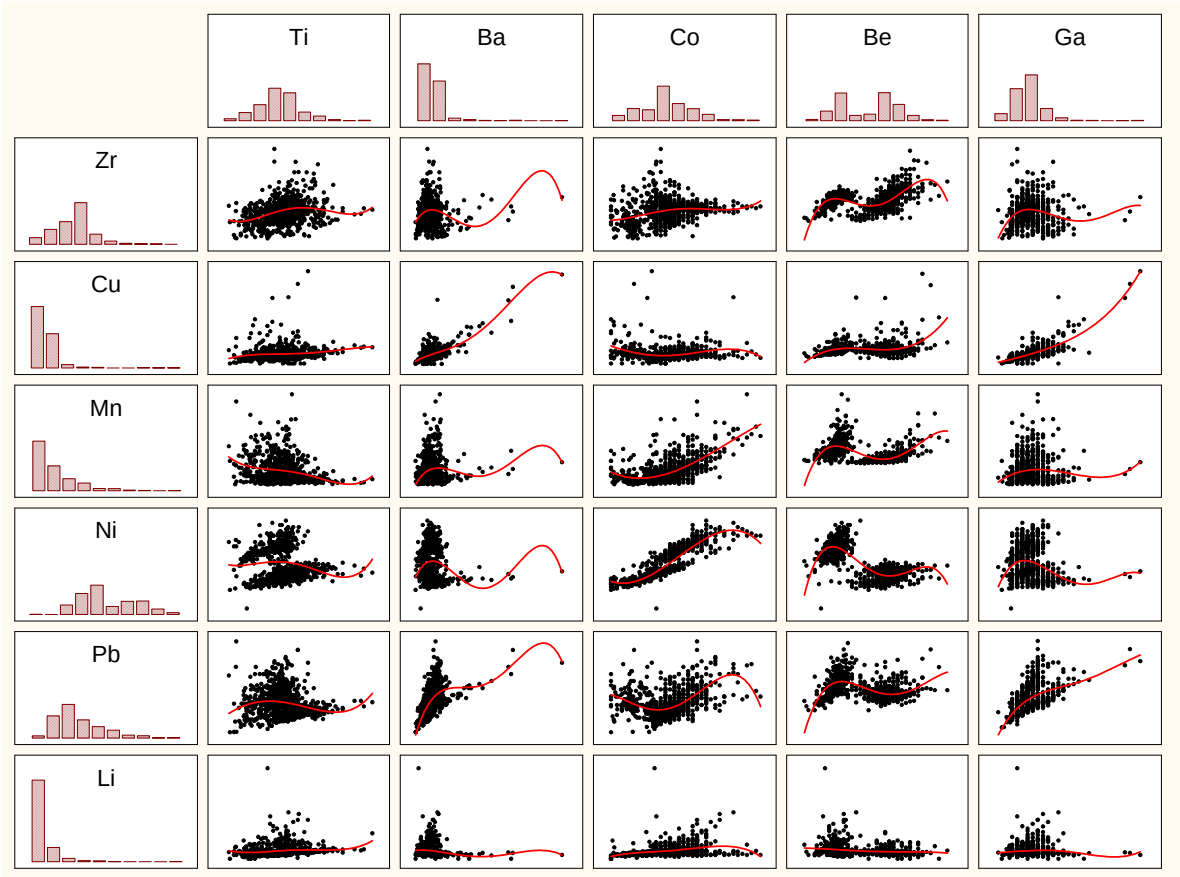

Figure 2. Graphic display of some correlation pairs.

It is possible to observe the presence of high values of the correlation coefficient in some pairs, e.g. Ba-Ga, Ba-Cu, Co-Ni, Mn-Ni and others. It is worth nothing that not all of the values are important.

For instance, the Ga-Pb pair provides seemingly moderate correlation coefficient, but it is clearly visible on Figure 2 that this dependence is poorly modeled by polynomial approximation. There is no connection between the geochemical classification of elements and their distribution laws. Also, we must keep in mind that the correlation itself does not state that one variable affects the other or that they are casually linked. We should also be aware of the possible effect of the third factor.

Therefore, the analysis of the correlation coefficient does not allow us to distinguish groups of elements and reconcile the data with the results of cluster analysis. On this basis, we proposed a method for calculating the cell fractal dimension of the distribution diagrams (spider diagrams), one of which is shown in Figure 3. Grouping in pairs is made considering the values of the fractal dimension of spider diagrams (Table 2).

According to the results of calculation of the fractal dimension (Table. 2) Li diagram has the lowest index. In comparison with the fractal dimension of the other charts, the value does not belong to any group; low correlation coefficient with other elements and these dendrograms can serve as an indirect confirmation of the isolated behavior of the element (Figure 1). At the same time, the difference of fractal dimension values in Cu-Ba, Ni-Be, Ga-Pb and Mn-Co plots is of the order of 10^{-3}.

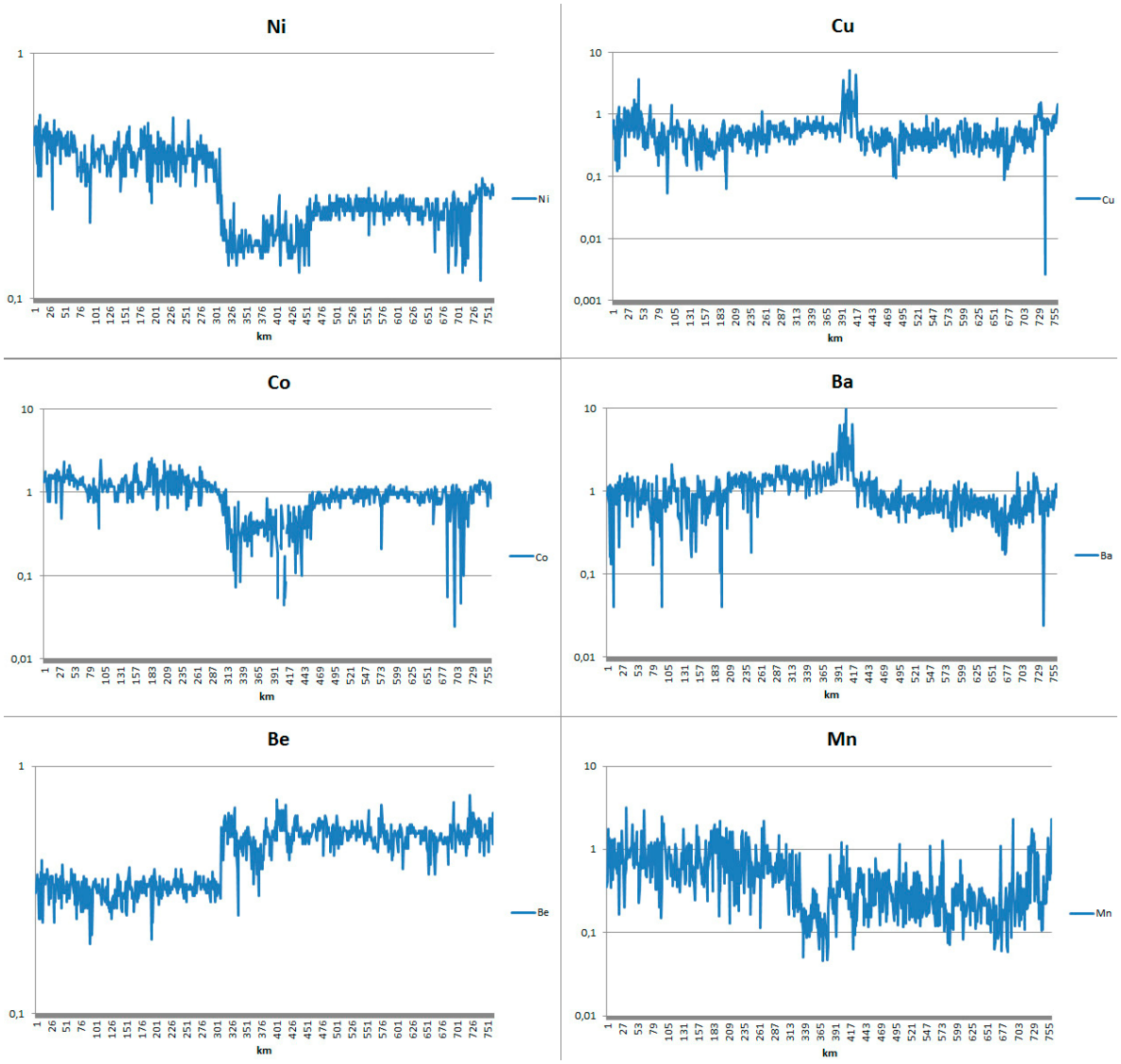

Figure 3. Spider diagrams of Ni, Cu, Co, Ba, Be, Mn

Zr-Ti pair provides the greatest value of ΔD and an absence of correlation which is followed by a low determination coefficient. Thus, the communication within the group is implicit, if existent at all, or may be a subject to more complex patterns.

The Cu-Ba group is characterized by a low ΔD rate, log-normal distribution, high determination value and is well modeled by the polynomial approximation. It is interesting that these elements have the lowest value of the fractal dimension of spider diagrams after lithium.

Ni-Be, Pb-Ga and Ti-Zr groups can be described as moderately related, with a normal distribution law, minor correlation, and, as a result, average determination rate. The curious fact is that in this group, Ti-Zr pair stands in the dendrogram in a separate subgroup, which is also evident through minimal determination and maximal ΔD values.

The last group, Mn-Co, is characterized with average determination coefficient, rather small ΔD, and different distribution laws for Mn and Co.

Table 2. The results of fractal and statistics analysis.

Chemical element	Fractal dimension, *D*	*ΔD*	Distribution law	Determination coefficient, R^2
Ba	1.49897	0.00718	Log-normal	0.58
Cu	1.49179		Log-normal	
Ni	1.50768	0.00068	Normal	0.49
Be	1.50700		Normal	
Mn	1.50878	0.00576	Log-normal	0.31
Co	1.51454		Normal	
Ti	1.50230	0.01884	Normal	0.07
Zr	1.52114		Normal	
Pb	1.52683	0.00202	Normal	0.30
Ga	1.52885		Normal	
Li	1.47607		Log-normal	-

3. Conclusion

This work exposes that within the regional geochemical profile analyzed elements demonstrate correlative behavior. Rate of correlation between the specific elements is different. To determine its characteristics a set of different statistic tools was used, such as the correlational and hierarchical analysis and geometrical approach to the study of the spider diagrams. It was possible to trace the relationship between the index of the fractal dimension and the results of the dendrograms analysis and the comparison with the standard statistical indexes. Basing on the properties of elements groups discussed above, it is possible to presume the direct relationship between the value of D and the coefficient of determination.

Therefore, calculation of the fractal dimension of spider diagrams can be an effective tool in conjunction with standard statistics methods used to identify the nature of the relationship between the chemical elements. It is also worth nothing that very often, along the fractal, come the multifractal characteristics of the geological environment [8]. These facts demand the applying the synergetic view on geology and geochemistry. That means studying processes in non-equilibrium systems from the position of self-organization [9]. Some of these processes can be revealed and examined by comparing the data obtained by different statistical methods.

Acknowledgements

The author thanks Siberian Research Institute of Geology, Geophysics and Mineral Resources for providing the information on the 3-DV profile.

This study (research grant No 8.1.76.2015) was supported by The Tomsk State University Academic D.I. Mendeleev Fund Program in 2015.

References

[1] Wan L, Chen P, Gong Z 2013 *Applied Mechanics and Materials* **249-250** 26
[2] Yanshi X, Jianwen Y, Kaixuan T, Liang C, Yang H and Yonghui L 2015 *Advanced Materials Research* **1092-1093** 1398
[3] Wan L, Chen P and Hu X 2012 *Applied Mechanics and Materials* **204-208** 4851
[4] Panahi A, Cheng Q 2004 *Mathematical Geology* **36 7** 827
[5] Cheng Q 1999 *Computers & Geosciences* **25** 949
[6] Xie S, Cheng Q, Xing X, Bao Z and Chen Z 2010 *Geoderma* **160** 36
[7] Feder J 1989 *Fractals, Plenum Press* 283
[8] Xie S, Bao Z 2004 *Mathematical Geology* **36 7** 847
[9] Haken H 1977 *Synergetics* (Berlin: Springer-Verlag)