

Comparative Studies on the Extraction of Geochemical Anomalies

Guohao Han

Collage of Earth Science, East China Institute of Technology, Nanchang Jiangxi
Email: xiaoheitan100@163.com

Received: Nov. 12th, 2015; accepted: Nov. 26th, 2015; published: Nov. 30th, 2015

Copyright © 2015 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The extraction of geochemical anomalies is based on the reasonable determination of the lower limit of geochemical anomalies. The determination of geochemical anomaly threshold is a fundamental task of exploration and prediction of geochemical mineral exploration. It is advantageous for the delineation of ore mineralization belt and prospecting guide. In this paper, the U, V, Mo elements in 1:200000 soil geochemical data are from Xiangshan area and its surroundings. Its distribution is tested by Q-Q probability map by using SPSS software. It is found that after eliminating extreme values, the elements approximately follow a lognormal distribution. Using the traditional method, the exploratory data analysis (EDA), and the concentration-area (C-A) fractal method, we extract the element of geochemical anomalies, and carry out the contrast research. The results show that: the concentration-area (C-A) fractal method objectively reflects the geochemical anomaly information in the study area, which is a reasonable and effective method to extract geochemical anomalies in the study area.

Keywords

Xiangshan Area, Extraction of Geochemical Anomalies, EDA Technology, C-A Model

地球化学异常提取方法对比研究

韩国豪

东华理工大学地球科学学院, 江西 南昌
Email: xiaoheitan100@163.com

收稿日期：2015年11月12日；录用日期：2015年11月26日；发布日期：2015年11月30日

摘要

地球化学异常的提取是建立在合理的确定地球化学异常下限的基础上的。地球化学异常下限的确定是勘查地球化学矿产勘查和预测的一项基本任务，对于圈定有利成矿带，指导找矿非常关键。文章以相山地区及外围1:20万土壤地球化学数据中的U、Mo、V元素为例，利用SPSS软件，通过Q-Q概率图检验元素含量的分布型式，发现研究区元素剔除极值后，元素近似服从对数正态分布。分别运用传统方法、探索性数据分析(EDA技术)、含量-面积(C-A)分形方法提取元素地球化学异常，并进行对比研究。研究结果表明：基于分形的含量-面积方法，比较客观地反映了研究区的地球化学异常信息，是研究区提取地球化学异常比较合理有效的方法。

关键词

相山地区，地球化学异常提取，EDA技术，C-A模型

1. 引言

近年来，随着地质勘查程度的提高，单纯地依赖野外露头等地质标志直接找矿的难度越来越大。地质找矿工作面临着“难发现、难识别、难利用”的严峻形势。而地球化学方法利用“显微标志”进行矿产勘查，极大的提高了找矿的效率，特别是寻找盲矿和隐伏矿。地球化学异常下限的确定是勘查地球化学的一个基本问题，也是勘查地球化学应用于矿产勘查时决定成败的一个关键性环节[1]。如何合理有效的提取地球化学异常信息，成为影响勘查地球化学在矿产勘查中效果的主要因素。周蒂[2]认为较合理的异常圈定方法应该是：根据研究区的地质、地球化学及景观资料逐点估计地球化学背景值，再从观测数据中减去背景值及随机误差而圈定异常。由于复杂的地质背景和成矿模式的多样性，目前还没有一种可以通用的异常下限计算方法，每种方法都有各自的优势和局限性。

在详实的地质工作基础上，土壤地球化学测量手段具有找矿成本低、见效快、效果好等众多优点，因此一直以来都被人们广泛应用[3]。本文利用相山地区及外围1:20万土壤地球化学勘查数据以及区内地质矿产资料，进行地球化学异常信息的提取。由于研究区分布有我国最大的火山岩性相山铀矿田和金溪熊家山钼矿床，因此选取成矿元素U、Mo元素为主要研究对象，V元素作为辅助元素。分别运用传统方法、探索性数据分析(EDA技术)、含量-面积法(C-A)分别确定地球化学异常下限，做出地球化学异常图，比较地球化学异常信息特征，探讨各方法的合理性与局限性，找出合适研究区的有效方法。

2. 元素含量的分布型式

地球化学数据处理过程中，首先要判断元素含量的分布型式。在勘查地球化学中，检验正态分布型式的常用方法有偏度、峰度检验法、柯尔莫戈洛夫-斯米尔诺夫准则检验法和概率格纸检验法[4]。本文利用SPSS统计软件，采用Q-Q概率图对元素含量进行分布检验。其原理是：用变量数据分布的分位数与所指定分布的分位数之间的关系曲线来进行检验，当符合指定分布时，图中各点(近似)成一条直线。

元素的分布型式是受空间分布、成矿作用、地质背景及污染状况等诸多因素影响。一般情况下，常量元素服从正态分布，而微量元素服从对数正态分布。由于研究U、Mo、V三种元素都为微量元素，因此在计算背景值与异常下限之前，先对各个元素的含量值取对数，并检验其否服从对数正态分布。利用SPSS软件做出各元素的对数正态分布Q-Q图(见图1)。

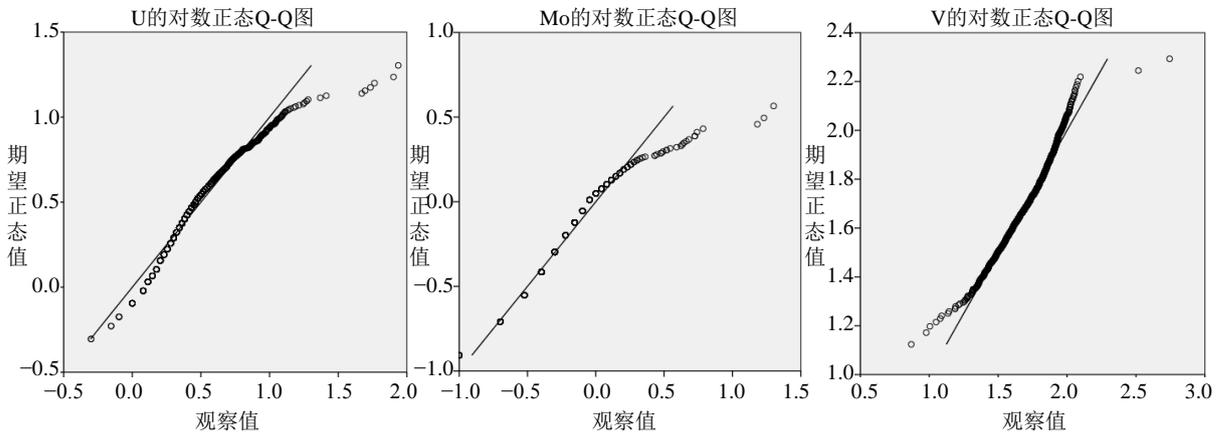


Figure 1. The Lognormal Q-Q figure of 3 elements
图 1. 3 种元素的对数正态 Q-Q 图

从图 1 可以看出，U、Mo、V 元素含量分布的分位数与对数正态分布的分位数之间的关系偏离直线，不服从对数正态分布。用迭代法逐次剔除一些特高值和特低值，再检验元素的分布型式。

由图 2 知，U、Mo、V 对数正态 Q-Q 图中大多数点分布在一条直线上，只有少数点偏离直线，所以 U、Mo、V 元素近似服从对数正态分布。

3. 元素土壤地球化学特征

对研究区土壤地球化学 U、Mo、V 元素进行了参数统计(表 1)，数据为原始数据经均值加减三倍标准差迭代剔除后的剩余数据的基本统计量。

从表 1 知，研究区元素背景质量分数的分布特征如下：U 元素的平均质量分数高于地壳丰度值，其富集系数远大于 1，表明 U 元素在区内相对富集；Mo、V 元素的平均质量分数低于地壳丰度值，其富集系数均小于 1，表明 Mo、V 元素在区内相对亏损。从变异系数可以发现，U 元素的变异系数较 Mo、V 的变异系数大，表明区内 U 元素有不同程度的分异和局部富集，可能是由成矿作用的叠加造成。特别是相山地区，铀矿床众多，U 元素具强分异特点，矿化信息强，有显著的地球化学富集，是研究区内主要的成矿与找矿指示元素。

4. 地球化学异常信息提取

结合元素分布型式检验的结果，分别运用传统方法、探索性数据分析(EDA 技术)、含量-面积法(C-A)确定 U、Mo、V 元素的异常下限值，并做出各元素出地球化学异常图，以探讨研究区异常提取方法的科学性与合理性。

4.1. 传统的统计学方法

传统的统计学方法基于地球化学数据符合正态分布或是对数正态分布。如果化探数据符合正态分布，则采用平均值加 2 倍标准差的方法计算异常下限。如果化探数据不符合正态分布，则采用平均值 \pm 3 倍标准差迭代剔除直至无离群点数据，随后检验数据的分布型式。

由图 2 知，经多次迭代剔除数据后，U、Mo、V 元素近似服从对数正态分布。以平均值加 2 倍标准差计算异常下限，得到研究区各元素特征值数据见表 2。

由于 U、Mo、V 元素异常含量相对变化幅度不大，分别以异常下限值的 1, 1.5, 2 倍作为异常的外带，中带，内带勾绘异常等值线。

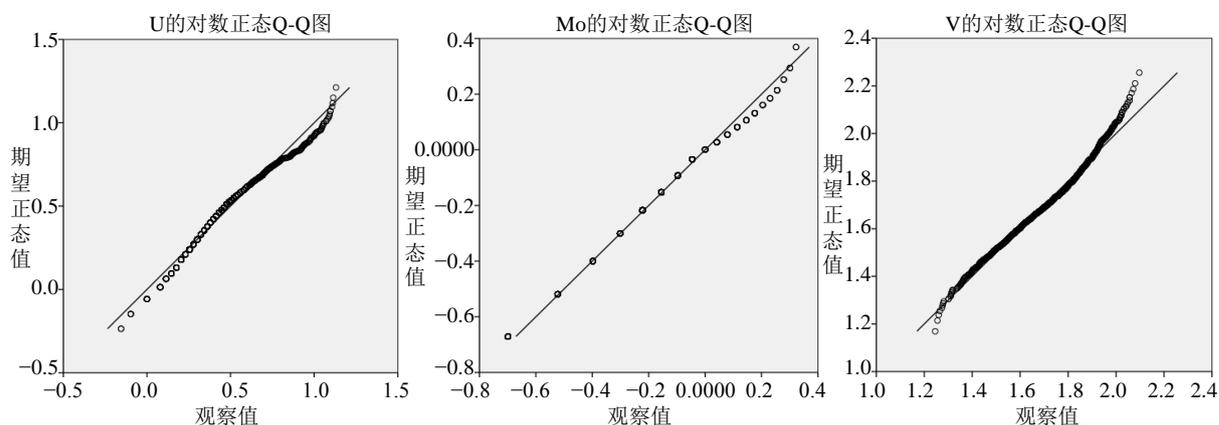


Figure 2. The Lognormal Q-Q figure after the data processing of 3 elements

图 2. 数据处理后 3 种元素的对数正态 Q-Q 图

Table 1. Soil geochemical characteristic parameters of the study area

表 1. 研究区土壤地球化学特征参数

特征值	平均值	最小值	最大值	标准差	变异系数	地壳丰度	富集系数
U	3.4	0.5	13.5	2.1	0.6	1.85	1.8
Mo	0.6	0.2	2.1	0.3	0.5	1.1	0.5
V	55.1	17.6	125.0	19.7	0.4	85	0.6

注: 1. 元素含量为 10^{-6} ; 2. 变异系数为标准差与平均值之比; 3. 地壳丰度据黎彤[5]; 4. 富集系数为平均值与地壳丰度之比。

Table 2. Characteristics value of U, Mo, V elements after iterative elimination

表 2. U、Mo、V 元素经过迭代剔除后的特征值统计

特征值	U	Mo	V
元素数据个数	1507	1507	1507
剔除数据个数	19	47	14
平均值	3.4	0.6	55.1
标准差	2.1	0.3	19.7
异常下限	7.6	1.2	94.5

注: 元素含量为 10^{-6} 。

以异常下限值的 1, 1.5, 2 倍作为异常值的外带、中带、内带下限值(表 3)。通过 MapGIS 软件按照外带、中带、内带下限值做出单元素异常图(图 3)。

由图 3 知通过传统的统计学方法获得的元素异常形态各异, 以圆状、片状及尖棱状为主, 面积大小不一。U 和 Mo 元素异常带呈北东向展布, V 元素异常区域主要分布在研究区的西北部位。该方法获得的 U、Mo、V 元素异常下限值较低, 圈出的异常面积过大, 给下一步异常查证工作带来了很大的困难。

4.2. 探索性数据分析

美国统计学家 Tukey (1977)提出探索性数据分析技术(简称 EDA 技术), 该技术利用稳健统计学和非参数统计学, 是一种非常规统计学方法; 并引入箱型图, 从箱型图中可迅速看出数据的结构和特点。这种方法不需要检验数据的分布型式而直接获取背景值和异常下限。其原理为: 取中位数为背景值, 上异点为异常下限值。上异点 = $2.5 \times$ 上四分点 - $1.5 \times$ 下四分点。

Table 3. Threshold values of U, Mo, V
表 3. U、Mo、V 元素的下限值

特征值	U	Mo	V
异常下限值	7.6	1.2	94.5
外带下限值	7.6	1.2	94.5
中带下限值	11.4	1.8	-
内带下限值	-	-	-

注：元素含量为 10^{-6} 。

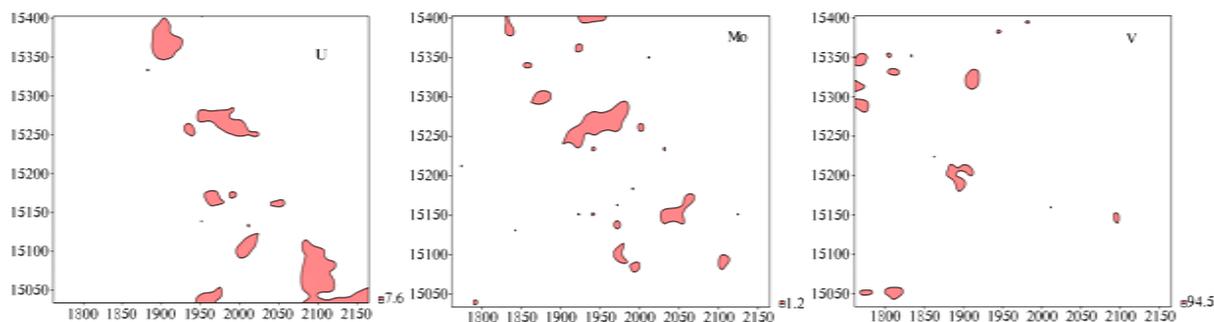


Figure 3. Anomaly maps of 3 elements delineated by traditional statistical method
图 3. 传统方法圈出的 3 种元素异常图

利用 SPSS 软件分别做出 U、Mo、V 元素的箱型分析图(图 4)和相关数据统计(表 4)。

由 EDA 技术求得的 U、Mo、V 三种元素的下限值比传统的统计学方法求取的下限值稍小。李佑国(2007)研究认为当元素不服从正态或对数正态分布时，用 EDA 方法获得的上异点值作为异常下限，具有明显的优势。

以 EDA 技术获取的异常下限值的 1, 1.5, 2 倍作为异常值的外带、中带、内带，做出各单元异常图(见图 5)。

由于 EDA 技术是对数剧排序，通过特定位置的数据进行分析的方法，能够排除极值的影响。由图 5 知 EDA 技术所圈定的各元素地球化学异常面积大，异常形态各异；异常浓度分带明显，浓集中心强。除此之外，EDA 技术还圈出了许多传统方法漏掉的异常信息，如在研究区的东北和东南部位，各圈出 V 元素异常。

4.3. 含量-面积分形方法

成秋明等(1994)提出了元素含量-面积法(C-A 分形) [6]。

$$A(\rho > c) \propto c^{-\beta} \tag{1}$$

其中， $A(\rho > c)$ 表示元素含量(浓度 ρ) 大于某一含量阈值 c 的区域面积， β 为奇异性指数， \propto 代表成正比。对(3-1)式进行变形有：

$$A(\rho > c) = Kc^{-\beta} \tag{2}$$

(2)式两边取自然对数得：

$$\ln(A(\rho > c)) = -\beta \ln(c) + \ln(k) \tag{3}$$

Table 4. Characteristics value of U, Mo, V
表 4. U、Mo、V 元素的特征值统计

元素	样品个数	最小值	最大值	下四分位数	上四分位数	下限值
U	1507	0.5	86.6	2.1	4.1	7.1
Mo	1507	0.1	20.0	0.4	0.7	1.2
V	1507	7.4	125.0	39.5	60.5	92.0

注：元素含量为 10^{-6} 。

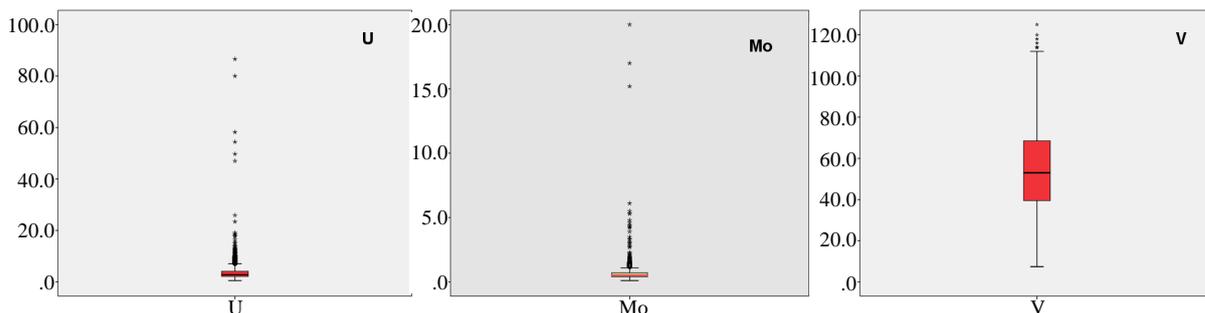


Figure 4. Box-type analysis chart of U, Mo, V elements

图 4. U、Mo、V 元素含量箱型分析图

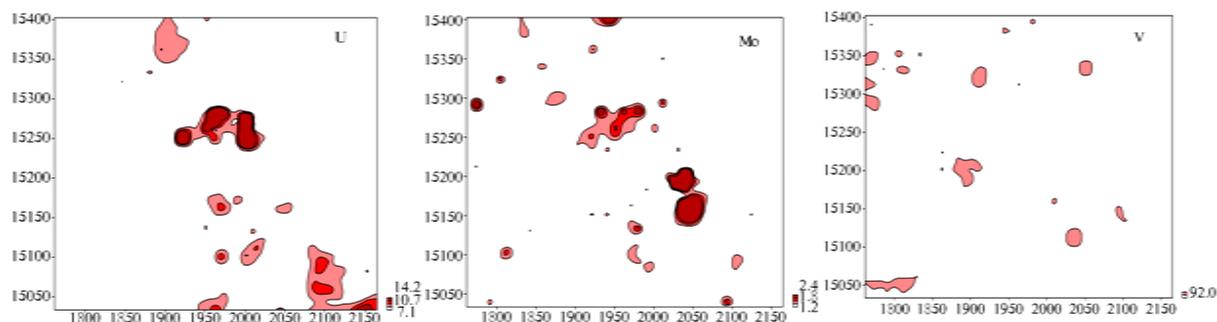


Figure 5. Anomaly maps of 3 elements delineated by EDA

图 5. EDA 方法圈出的 3 种元素异常图

由(3)式知：含量大于某一阈值的区域面积与这个阈值在双对数坐标系 ($\text{Log}(c) - \text{Log}(A)$) 中成线性关系，直线的斜率即为分形的维数。由于地球化学元素分布满足多重分形分布，所以同一元素在双对数坐标系下，会呈现多重线性关系。即有多个分形维数存在。采用最小二乘法分段拟合双对数散点图，直线间的交点为地球化学场发生突变的临界点[7]-[11]，找出合适两条直线交点处对应的元素含量值 c_0 作为异常下限点。

首先，把研究区 U、Mo、V 元素数据在 MAPGIS 中生成点位图，再把点位图转化为矢量图。然后再进行多重分形 IDW 插值处理，将得到的栅格数据进行含量 - 面积分形异常分解。得到各元素含量 - 面积双对数坐标图(见图 6)。双对数关系图可用三条直线或四条直线来拟合，反映了元素含量的空间分布存在多个无标度区即元素服从多重分形分布。直线间的交点为地球化学场发生突变的临界点。应尽量保证各分段直线的拟合度，以提高分界点的客观性，真实的反映研究区的地质背景。

一般来讲，第一段斜率较小属于背景区。第二段以上为主要含量区间，通常为异常区。第二段直线反映由区域地质作用形成的区域异常，第三段和第四段直线反映局部地球化学异常[12]。根据研究区的元

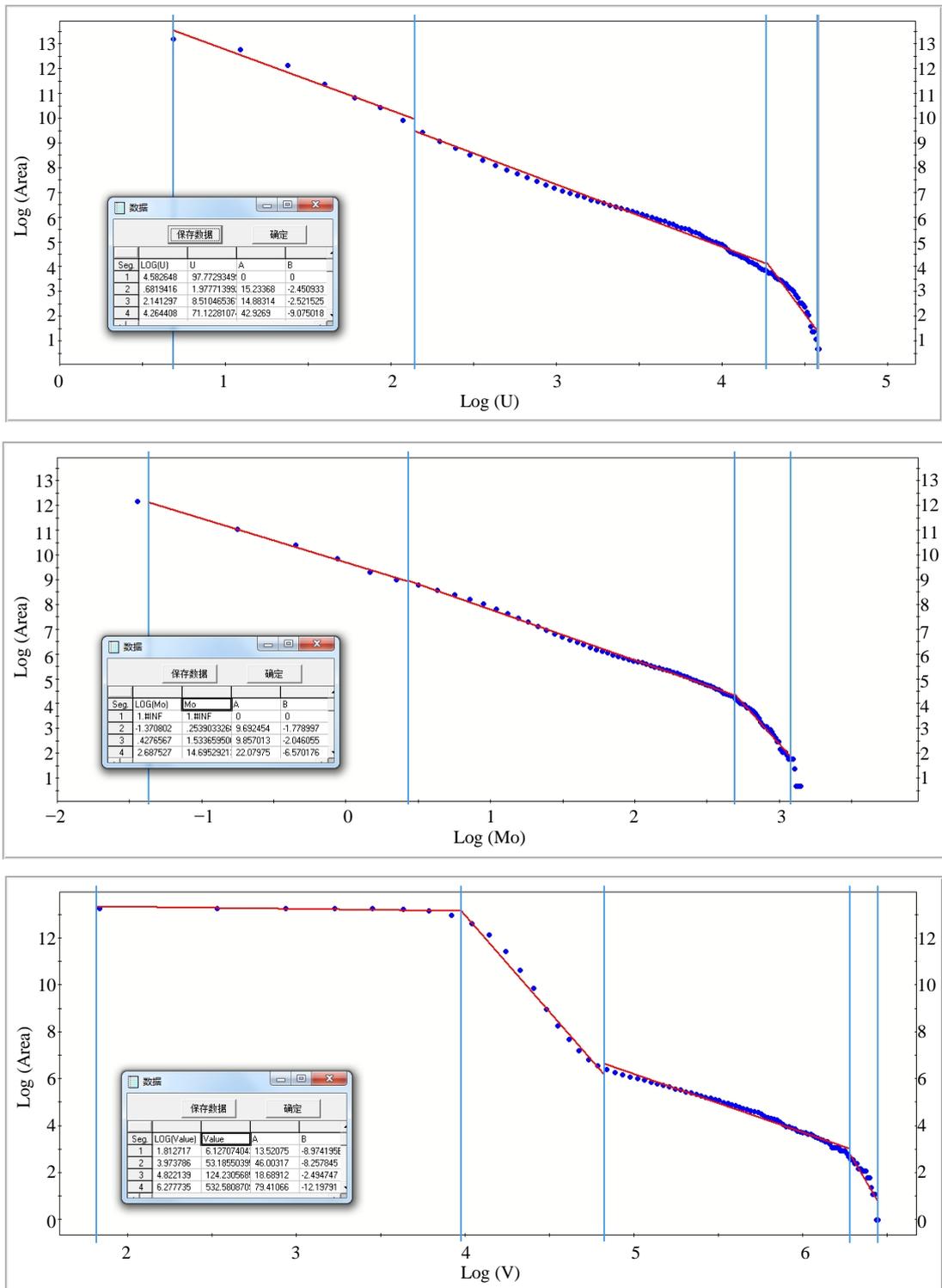


Figure 6. Log-Log plot for 3 elements representing relationships between areas and concentration
 图 6. 3 种元素含量值与对应面积值的对数关系图

素地质背景和成矿特征，选取 U、Mo 元素的异常下限值为第一段直线和第二段直线的交点；V 元素的异常下限值为第二条直线和第三条直线的交点。从分段拟合表中找出不同直线交点处所对应元素的含量值

(见图 6)。得到 U、Mo、V 元素的下限值分别为 8.5 ppm、1.5 ppm、124.2 ppm。

以 C-A 分形方法取得的异常下限为阈值，得到各元素异常图(见图 7)。

从图 7 可以看出，基于 C-A 分形方法圈出的 U 和 Mo 元素异常空间位置的分布与前两种方法圈出的空间位置基本一致；而且异常面积显著减小，异常形态各异。V 元素异常区域仅分布在研究区的东南和东北部位，并且异常面积小，异常形态呈圆状、椭圆状。

4.4. 三种方法的对比认识

分别运用传统方法、EDA 技术、C-A 方法确定 U、Mo、V 元素的异常下限值(见表 5)，并且比较研究各方法提取的异常特征。

传统的统计学方法处理数据时具有很大的局限性，特别是在地质条件复杂地区，元素往往不服从正态分布，此时用传统的统计学方法获得的异常不能合理地突出异常特征。从表 5 可以看出，该方法确定的异常下限值低于 C-A 分形方法确定的下限值。其原因可能是研究区地质条件比较复杂，导致元素在不同的区域具有不同的分散富集特征。而且通过传统的迭代剔除方法处理时，部分样品的低值点和高值点被剔除，致使异常信息容易被遗漏。

EDA 技术充分利用了研究区的数据信息，不需要剔除极值点元素的数据，保证了数据的完整性。但由于研究区植被覆盖严重，元素被第四系覆盖物所稀释，大部分的样品含量值较低，致使通过 EDA 技术确定的异常下限值比其他两种方法低。EDA 技术圈出的异常面积较大，给下一步的异常查证工作带来比较大的困难。

C-A 分形方法不仅考虑了地球化学场的空间相关性，而且还考虑了几何形态和尺度不变性[13]，将背景与异常的形成认为是两个相互独立的过程。因此由 C-A 分形方法提取的地球化学异常，更能客观的反映研究区的矿床地球化学异常特征。从区内矿产分布看，由 C-A 分形方法圈出的异常区域与研究区矿床(矿点)吻合度较好。

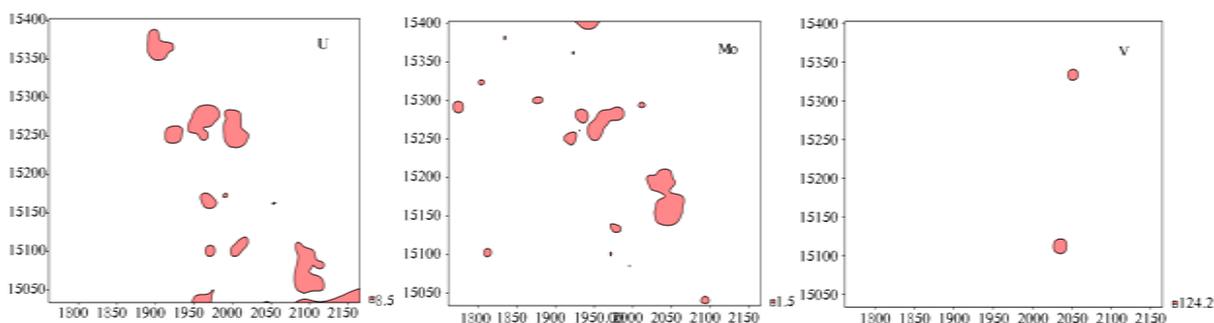


Figure 7. Anomaly maps of 3 elements delineated by C-A

图 7. C-A 方法圈出的 3 种元素异常图

Table 5. Threshold values of 3 elements with different methods

表 5. 不同方法获得的 3 种元素的下限值

异常提取方法	U	Mo	V
传统方法	7.6	1.2	94.5
EDA 技术	7.1	1.2	92.0
C-A 方法	8.5	1.5	124.2

注：元素含量为 10^{-6} 。

5. 结论

1) 通过 Q-Q 概率图对研究区 U、Mo、V 元素含量进行分布型式检验, 发现各元素均不服从正态或对数正态分布。由元素含量 - 面积双对数坐标图, 可以发现研究区元素含量的空间分布存在多个无标度区, 表明元素具有多重分形的特征。

2) 传统方法剔除了极值点数据, 破坏了数据的完整性, 致使传统方法圈定的异常往往遗漏矿化信息。由于研究区植被覆盖严重, 元素被第四系覆盖物所稀释, 大部分的样品含量值较低, 因此 EDA 技术确定的异常下限值低, 圈出的异常面积较大, 给今后的异常查证工作带来比较大的困难。

3) C-A 分形方法理论依据是地球化学场具有分形结构, 有效克服了传统方法的局限, 更能体现地球化学数据的原始特征, 是研究地球化学异常比较合理的方法。该方法确定的异常下限值高, 圈出的异常面积小, 减少了异常查证面积, 利于迅速定位找矿靶区。

参考文献 (References)

- [1] 龚庆杰, 张德会, 韩东昱. 一种确定地球化学异常下限的简便方法[J]. 地球与环境, 2001, 29(3): 215-220.
- [2] 周蒂. 分区背景校正法及其对化探异常圈定的意义[J]. 物探与化探, 1986, 10(4): 263-273.
- [3] 熊亮, 朱杰勇, 朱林生, 等. 土壤地球化学测量扬子地台西缘会泽铅锌矿带筇竹寺组中的应用[J]. 物探与化探, 2010, 34(4): 467-471.
- [4] 蒋敬业, 程建萍, 祁士华, 等. 应用地球化学[M]. 武汉: 中国地质大学出版社, 2006.
- [5] 黎彤, 袁怀雨, 吴胜昔, 等. 中国大陆壳体的区域元素丰度[J]. 大地构造与成矿学, 1999, 23(2): 101-107.
- [6] Cheng, Q.M., 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.
[http://dx.doi.org/10.1016/0375-6742\(94\)90013-2](http://dx.doi.org/10.1016/0375-6742(94)90013-2)
- [7] 韩东昱, 龚庆杰, 向运川. 区域化探数据处理的几种分形方法[J]. 地质通报, 2004, 23(7): 714-719.
- [8] 张磊, 申纬. 分形方法在澳大利亚新南威尔土地区地球化学异常下限确定中的应用[J]. 地质通报, 2009, 28(23): 245-249.
- [9] 王志刚. 分形方法在海南屯昌地区地球化学找矿中的应用[J]. 地质调查研究, 2012, 35(4): 315-320.
- [10] 徐明钻, 朱立新, 马生明, 等. 多重分形模型在区域地球化学异常中的应用探讨[J]. 地球学报, 2010, 31(4): 611-618.
- [11] 杨茂森, 黎清华, 杨海巍. 分形方法在地球化学异常分析中的运用研究:以胶东矿集区为例[J]. 地球科学进展, 2005, 20(7): 809-814.
- [12] 张焱, 周永章. 多重地球化学背景下地球化学弱异常增强识别与信息提取[J]. 地球化学, 2012, 41(3): 278-291.
- [13] 成秋明, 张生元, 左仁广, 等. 多重分形滤波方法和地球化学信息提取技术与进展[J]. 地学前缘, 2009, 16(2): 185-198.