

doi:10.11835/j.issn.1000-582X.2019.251

GPD 建模及其在矿产预测中的应用

秦飞龙^{1a,2}, 刘 剑^{1b}, 颜文勇^{1a}

(1. 成都工业学院 a.大数据与人工智能学院;b.汽车与交通学院,成都 611730;

2. 电子科技大学 数学科学学院,成都 611731)

摘要:探讨矿产资源预测方法。通过研究发现地球化学元素含量分布规律与广义帕累托分布(GPD)理论原理吻合,从而构建矿产资源预测 GPD 模型,在此基础上对模型的阈值和参数选取进行了探讨。将设计的 GPD 模型应用到实际矿区进行矿产预测,结果发现设计出的模型能够有效圈定出地球化学元素异常分布区域,圈定出的区域范围与矿体分布吻合,有利于圈定找矿靶区。

关键词:矿产预测;地球化学异常;GPD;阈值;参数

中图分类号:TM 344.1

文献标志码:A

文章编号:1000-582X(2020)04-064-09

The study of GPD and its application in mineral prediction

QIN Feilong^{1a, 2}, LIU Jian^{1b}, YAN Wenyong^{1a}

(1a. School of Big Data and Artificial Intelligence; 1b. School of Automobile and Communications, Chengdu Technological University, Chengdu 611730, P. R. China; 2. School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, P. R. China)

Abstract: This paper is aimed at exploring the method of mineral resources prediction. Though study we found that the anomalous distribution of geochemical elements was consistent with the generalized Pareto distribution (GPD) of extreme value theory, on the basis of which a new algorithm called the GPD model of geological anomaly was designed. In the new model, the parameters and threshold selection of GPD were discussed and proved. The proposed GPD model of the geological anomaly was applied to identify anomalies in the actual mining area. The results show that the model can effectively identify the geochemical anomaly region of the study area, and the delineated anomaly region coincides with the distribution of ore bodies of actual engineering exploration. The application of the GPD model will be instrumental to the delineation of new prospecting targets.

Keywords: mineral prediction; geochemical anomaly; GPD; threshold; parameters

收稿日期:2019-04-03 网络出版时间:2019-09-18

基金项目:数学地质四川省重点实验室开放基金资助项目(scsxdz2018yb03);四川省科技厅计划项目(2019YJ0375);成都工业学院博士基金项目(2018RC022);成都工业学院实验室开放基金项目(2018ZH10)。

Supported by Opening Fund of Geomathematics Key Laboratory of Sichuan Province(scsxdz2018yb03), the Program of Sichuan Provincial Department of Science and Technology (2019YJ0375), the Doctoral Program of Chengdu Technological University(2018RC022) and Opening Fund of Laboratory of Chengdu Technological University(2018ZH10).

作者简介:秦飞龙(1983—),男,博士,研究方向为数学地质和应用统计,(E-mail)lida_112@163.com。

通讯作者:刘剑(1975—),男,博士,研究方向为石油化工,(E-mail)1051671804@qq.com。

矿产资源在推动人类社会向前发展的过程中起着举足轻重的作用,几乎各行各业都直接或间接依赖于矿产资源^[1]。然而矿产资源是不可再生资源,随着人类对矿产资源需求量日益增大,地球表面及浅部矿产勘查将近枯竭,向深部地球寻求矿产已成为必然趋势^[2],这就要寻求合理的方法进行矿产预测。地球化学元素含量异常分带是进行矿产预测的一种有效的手段^[3],厘清地球化学异常分布规律,就能圈定矿体所在区域。因此选取合理的地球化学异常提取方法至关重要。近年来,学者们在地球化学异常识别研究工作上做了大量研究,取得了诸多成果,为深部盲矿预测提供了宝贵经验。如 Cheng^[4-5]利用多重分形奇异性理论模型来刻画地球化学异常分布;Shen^[6]利用分形模型进行元素异常下限提取;Zuo^[7]采用多重分形谱和接收工作特性曲线相结合的混合方式来识别空间相关的元素组合。事实上每种方法进行地球化学异常提取都需要一定的条件,但地球化学异常形成的地质环境极为复杂,目前难以找到合理的地球化学异常提取方法^[8]。传统的均方差进行地球化学异常提取需要数据服从正态分布,但实际地球化学数据形成的地质环境复杂,其分布并不服从正态分布^[9]。证据权法在矿产预测中有着广泛的应用,但该方法对成矿概率临界值的确定还没有科学方法,常采用专家经验法以及统计分布法等^[10]。实际上,地球化学异常在空间分布上与周围正常元素的分布具有一定的偏离,其观测数据位于整个概率分布的尾部^[11-12],而极值理论中的广义帕累托分布(generalized pareto distribution,GPD)研究的是数据的后尾性分布^[13]。赵鹏大院士也曾在韩国首尔举办的国际统计学会议上将地球化学异常含量分布描述为极值分布^[14]。采用 GPD 进行原生晕地球化学元素异常提取,有利于提取找矿标志,进行矿产预测。目前 GPD 在地球化学异常下限估计方面出现了一些研究成果,如左仁广^[15]提出地质异常属于极值理论研究范畴。然而在利用 GPD 估计地球化学异常值的精度和确定 GPD 的参数方面研究较少,并且极值理论往往只考虑元素含量的频率分布。其实可以将元素含量异常提取后,结合元素异常对应的空间坐标,辅以 MapGis 软件^[16]描绘元素异常含量空间分布规律。因此,笔者结合地球化学异常特性,深入研究 GPD 模型,对其阈值估计和参数估计进行探讨,以此形成地球化学异常提取模型,进行矿产预测。

1 GPD 理论基础

广义帕累托分布(GPD)亦称阈值模型,它是对样本中充分大的观测数据进行分布拟合。假设 x_1, x_2, \dots, x_n 为一组随机变量并且属于同一分布,如果存在一个较大的数据 μ ,并且有固定常数 $\beta > 0$,GPD 可表达为^[17]

$$G(x; \xi, \beta) = \begin{cases} 1 - (1 + \xi)^{-\frac{x-\mu}{\beta}} & \xi \neq 0, \\ 1 - e^{-\frac{x-\mu}{\beta}} & \xi = 0. \end{cases} \quad (1)$$

式中: ξ 为形状参数, μ 为样本阈值, β 为尺度参数,当 $\xi \geq 0$ 时, $x \geq \mu$;当 $\xi < 0$ 时, $\mu \leq x \leq -\frac{\beta}{\xi} + \mu$ 。尺度参数 $\beta=1$ 时,称为 GPD 的标准形式。取不同的形状参数 ξ ,可得出 GPD 的标准分布函数图(图 1),由图 1 可知,GPD 为后尾分布。

在式(1)中,需要确定其尺度参数 β ,形状参数 ξ 和阈值 μ 。然而阈值 μ 的估计至今未有明确的方法,其中比较有效的阈值估计方法是 MEF 法^[12,18](mean excess function),假定 μ_0 为初始阈值,对任意阈值 $\mu(\mu > \mu_0)$,MEF 定义为^[18]

$$E(\mu) = \frac{\beta\mu + \xi\mu}{1 - \xi}, \quad (2)$$

式中: $E(\mu)$ 是关于阈值 μ 的一个线性函数。如果已知样本数据 $x_1, x_2, \dots, x_n, x_i$ 为大于阈值 μ 的数据,则 $E(\mu)$ 统计方法可表示为^[19]

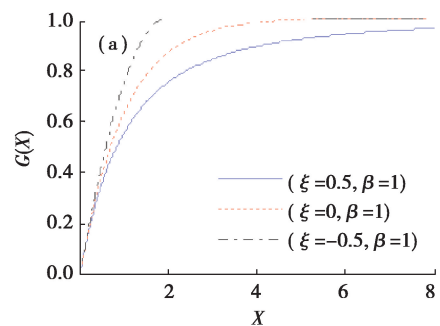


图 1 GPD 的标准分布形式
Fig. 1 The standard distribution function of the GPD

$$E(\mu) = \frac{\sum_{i=1}^n (x_i - \mu)}{N_n}, \quad (3)$$

式中: n 为样本总数; N_n 为超过阈值 μ 的数据个数。在 $(\mu, E(\mu))$ 的分布中选取充分大的 μ_0 , 使得当 $x \geq \mu_0$ 时, $E(\mu)$ 为近似线性函数, 该 μ_0 定义为阈值^[19]。而 GPD 的形状参数和尺度参数可由极大似然函数法进行估计^[20]。然而, 在实际数据处理中, 通过矩法估计得出的 GPD 的参数较似然函数估计的结果更合理^[19]。矩法估计是通过样本矩来反应总体矩, 以此对参数进行估计, 即

$$\begin{cases} \xi = [1 - (\bar{t}/\delta)^2]/2, \\ \beta = \bar{t} \times (1 - \xi). \end{cases} \quad (4)$$

式中: \bar{t}, δ 分别为 $t_i (t_i = x_i - \mu \geq 0)$ 的平均值和标准差。

2 地球化学异常提取 GPD 模型构建

地球化学数据的异常区域分布在观察数据的尾部, 这与 GPD 对极大值或极小值数据分布拟合吻合, 因此可以通过 GPD 构建地球化学异常模型, 确定地球化学找矿标志, 进行矿产预测。对一组地球化学数据 $X = \{x_1, x_2, \dots, x_n\}$, GPD 建模如下:

第 1 步 数据准备。

第 2 步 采用 QQ 图检验数据是否正态分布, 是否满足后尾性分布^[19]。

第 3 步 采用式(4)估计 GPD 的形状参数 ξ 和尺度参数 β 。

第 4 步 在数据中选取不同阈值 μ , 采用式(3)计算出 $E(\mu)$ 并绘制 $(\mu, E(\mu))$ 散点图。当 $x \geq \mu$ 时, $E(\mu)$ 近似为线性分布, 则该处的 μ 即为阈值, 亦为地球化学元素异常下限。

第 5 步 把第 3 步估计出的参数 ξ, β 和第 4 步估计出的阈值 μ 代入式(1), 得出 GPD 超阈值分布函数。

第 6 步 利用 PP 图^[19] 进行诊断检验, 检验 GPD 超阈值理论分布与数据实际分布是否一致, 如果不一致, 返回第 4 步, 估计新的阈值, 直到理论分布与实际分布拟合程度吻合为止。

在上述构建的模型中, MEF 法估计阈值比较直观, 但仍然具有一定的局限性, 因为该方法只能估计出阈值位于某个区间内, 不能估计出具体的精确值。在本文中构建了一个新的稳定性估计法。

由于样本数据的形状参数 ξ 与尺度参数 β 是固定的, 在 MEF 法确定的阈值区间内, 任选一个初始阈值 μ_0 。对任意的阈值 $\mu > \mu_0$, 有 $\beta(\mu)_{\text{GPD}} = \sigma_{\text{GEV}} + \xi(\mu - \mu_0)$ ^[18]。因此有

$$\beta(\mu) = \sigma + \xi(\mu_0) \text{ 或 } \beta(\mu) = \beta(\mu_0) + \xi(\mu - \mu_0). \quad (5)$$

若令

$$\beta^*(\mu) = \beta(\mu) - \xi(\mu), \quad (6)$$

$\beta^*(\mu)$ 显然与阈值 μ 无关, 称 $\beta^*(\mu)$ 为修正的阈值函数。因此, 如果 μ_0 选取合理, 数据超出量 $x_i - \mu_0 (x_i > \mu_0)$ 服从 GPD 分布, 那么对任意的阈值 $\mu > \mu_0$, $\beta^*(\mu)$ 不会随着 μ 的变换发生改变。

3 实际数据处理与矿产预测

研究区鸡冠咀铜金矿床为中国重要的矽卡岩热液型矿床之一, 其东经为 $114^\circ 54' 42'' \sim 114^\circ 55' 45''$, 北纬为 $30^\circ 04' 45'' \sim 30^\circ 05' 50''$ 。区内地质构造复杂, 主要包括: 褶皱构造、断裂构造及断裂侵入接触构造。岩浆岩主要为喷出岩和侵入岩两大类型, 侵入活动主要在燕山早期、燕山晚期及喜山早期进行。经接触交代变质作用发育为石榴子石、透辉石、金云母及绿泥石等矽卡岩矿物, 通过接触热变质作用发育为白云石大理岩。区内围岩蚀变强烈, 主要包括高岭石化、绿泥石化及矽卡岩化等。矿区有 4 个主矿体群和 14 个主矿体, 主矿体位于大理岩残留体的层间破碎带和大理岩与侵入岩的断裂接触部位。2 条主断裂 F_1 和 F_3 把区域划分为 A, B, C, D 4 个部分(图 2)。区内矿产资源丰富, 尤其位于 22 线和 34 线之间的 VII 号矿体, 有着巨大的找矿潜力^[12, 19]。笔者选取 26 号勘探线 VII 矿体 Mo, Ba, Zn 成矿元素的含量进行矿产预测研究。根据区域矿体前缘→矿体中部→矿体尾部走势(图 8~10), 结合 3 种元素面金属含量随着深度变换情况可知(图 3), Ba 元素面金属含量随着深度变换呈现降低趋势, 属于前缘晕元素; Zn 元素面金属含量随着深度变换呈现升高趋势, 属于尾晕元素; Mo 元素面金属含量随着深度变换近似不变, 属于近矿晕元素。但实际各元素含量随着深度

变换出现不同的折转(图 3),说明矿床在不同中段表现多期次热液的叠加现象,指示深部存在盲矿体可能,下面将构建的模型进行盲矿体预测。

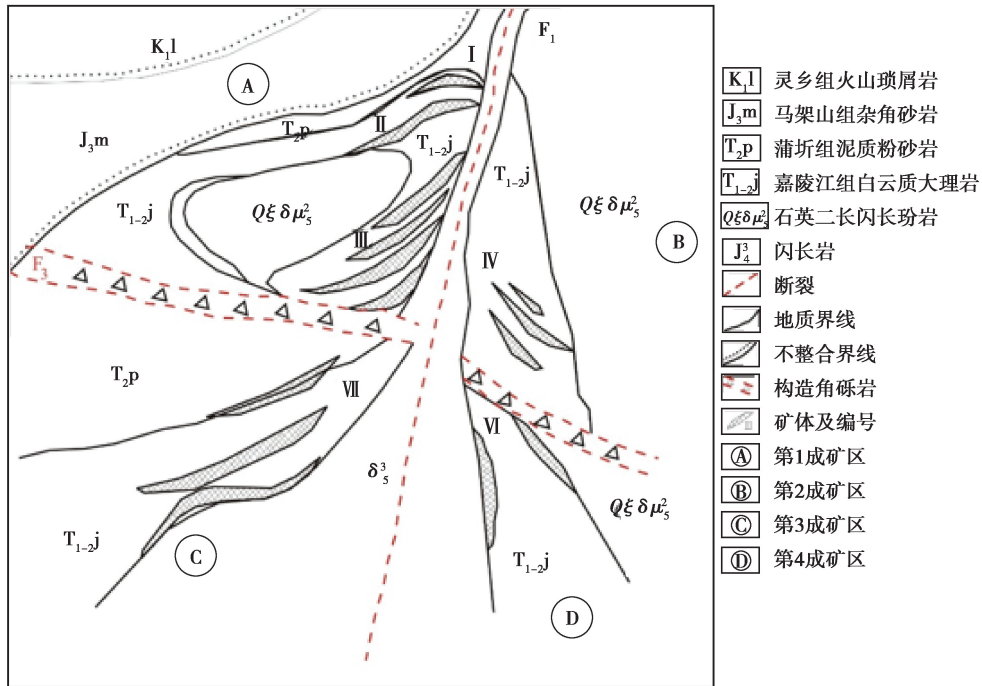


图 2 鸡冠咀铜金矿床模式示意图

Fig. 2 The pattern of the Jiguanzui Cu-Au mining deposit

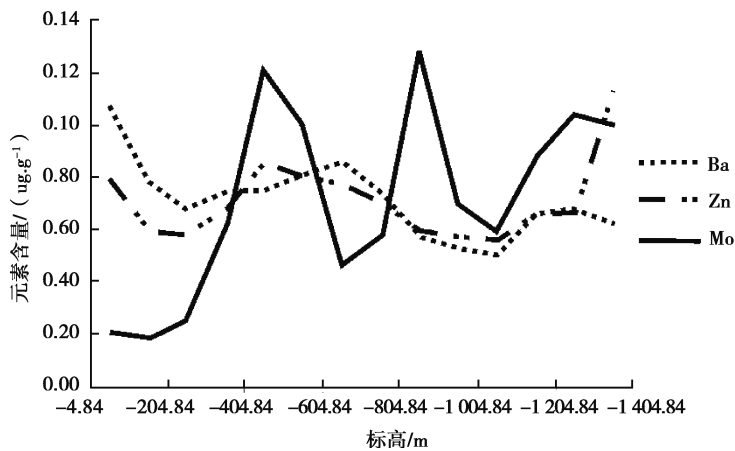


图 3 元素含量随深度变化曲线

Fig. 3 The variation curves of element content with depth

由 3 种元素含量统计分析(表 1)可知,3 种元素含量偏度都是正数,不服从正态分布,通过 QQ 图(图 4)进一步发现 3 种元素含量均为右偏,表现后尾分布特性,符合提出的算法模型适用条件。

表 1 元素统计量结果

Table 1 Element statistics

元素	均值	最小值	最大值	标准差	协方差	变异系数	偏度	峰度
Mo	25.79	5.31	467.75	13.24	183.36	0.49	5.81	82.89
Ba	307.45	42.19	1 584.92	98.42	9 461.27	0.29	0.69	4.15
Zn	98.89	26.21	380.59	24.44	587.45	0.25	0.53	6.32

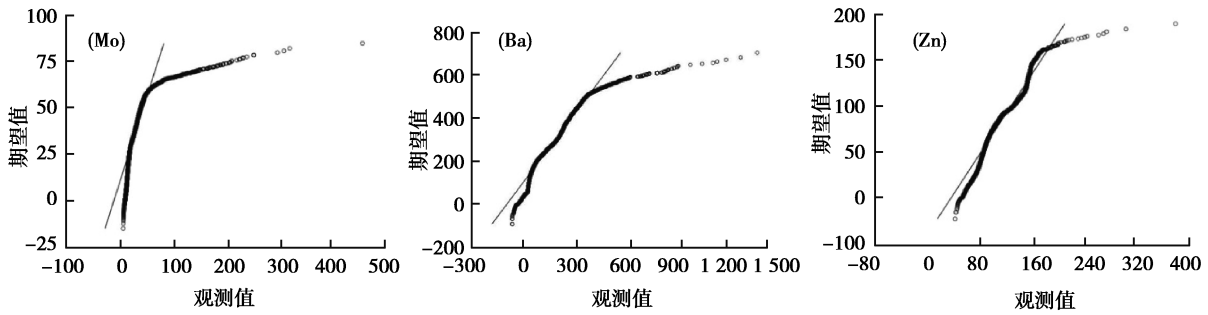


图 4 Mo, Ba, Zn 元素的偏态性检验

Fig. 4 The skewness test of Mo, Ba and Zn

通过设计的模型分别得出 Mo, Ba, Zn 3 种元素含量的超额均值函数图(图 5),由图 5 可知,3 种元素含量高于区间值 $[35.87, 38.21]$, $[378.42, 388.72]$, $[135.22, 140.17]$ 以后,其 MEF 均表现线性递增,即为阈值区间。为了选取精确的阈值,在 3 个区间分别均匀插入 50 个点,利用修正函数得出 $\beta^*(\mu)$ 与阈值的变换情况(图 6),由图 6 知 Mo, Ba, Zn 元素含量分别在 36.29, 382.56, 138.47 后估计量 $\beta^*(\mu)$ 不会随着 μ 变换而改变,即为阈值,利用式(5)得出 Mo, Ba, Zn 的尺度参数分别为 6.94, 48.66, 3.86, 形状参数分别为 0.37, 0.17, 0.30, 从而 3 种元素含量 GPD 分布为

$$G_{Mo}(x) = 1 - \left(1 + \frac{0.37}{6.94}(x - 36.29)\right)^{-\frac{1}{0.37}}, G_{Ba}(x) = 1 - \left(1 + \frac{0.17}{48.66}(x - 382.56)\right)^{-\frac{1}{0.17}},$$

$$G_{Zn}(x) = 1 - \left(1 + \frac{0.30}{3.86}(x - 138.47)\right)^{-\frac{1}{0.30}}.$$

确定 GPD 后,利用 PP 图进行诊断性检验(图 7)发现 3 种元素超额值含量均依附在直线左右,进而元素含量的 GPD 分布与实际地层中分布具有较高吻合性,说明地球化学异常下限选取合理。

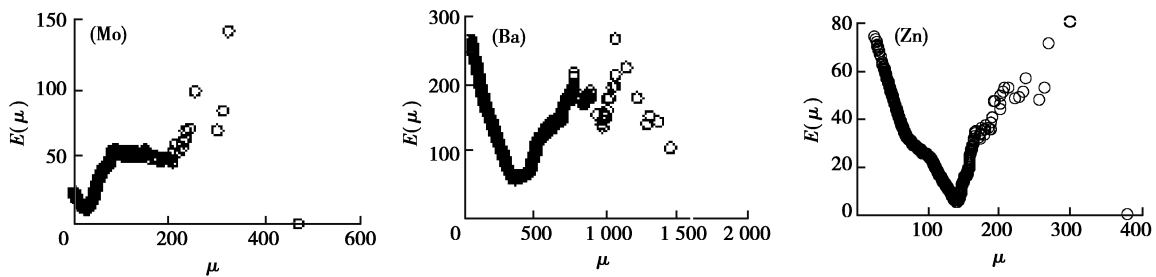


图 5 Mo, Ba, Zn 元素超额均值函数图

Fig. 5 The MEF diagram of Mo, Ba and Zn

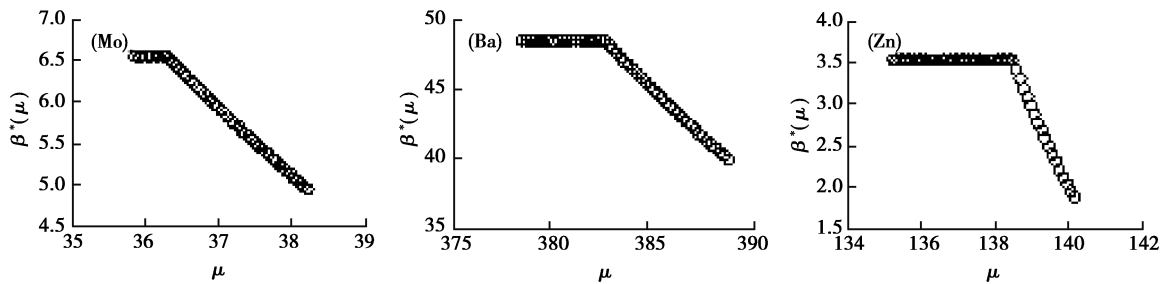


图 6 Mo, Ba, Zn 元素的修正尺度参数 $\beta^*(\mu)$ 随 μ 的变换关系

Fig. 6 The relations between $\beta^*(\mu)$ and thresholds μ of Mo, Ba and Zn

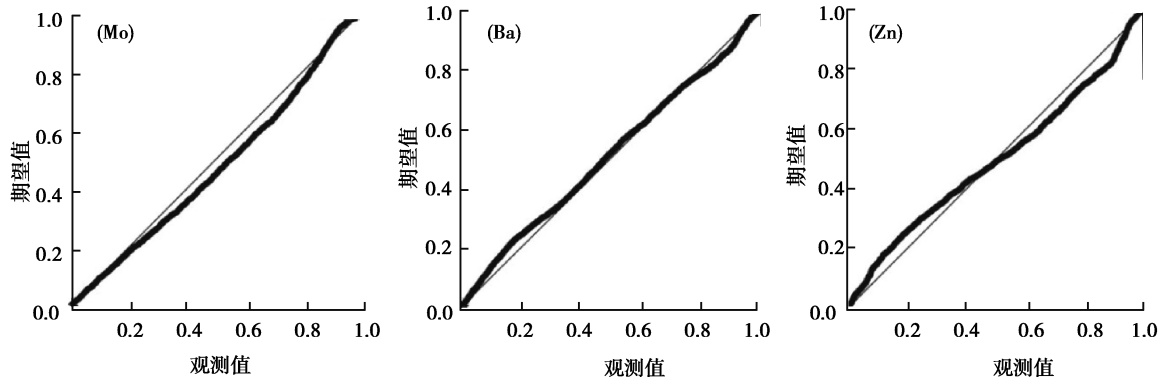


图 7 Mo, Ba, Zn 元素 GPD 分布拟合检验
 Fig. 7 The GPD fitting of Mo, Ba and Zn

根据元素异常下限得出 Mo, Ba, Zn 元素含量空间分规律(图 8~10),由图 8~10 知元素含量分带能够反映矿体走势。另外,元素 Ba 不仅在矿体顶部分布,在矿体下面也存在富集,该现象为前缘晕的反常分带(图 8);元素 Zn 在矿体上方与前缘晕元素出现叠加现象,反映出元素多期次成矿特性,并且在矿体下方与前缘同样具有共存现象,说明深部存在找矿潜力;元素 Mo 分带较好地刻画了矿体分布特征,在深部具明显的伸现象,特别是在钻孔 KZK10 和 KZK11 以下 1 100 m 左右,具有强异常分带现象,该异常很可能是钻孔 ZK2620 下方 1 300 m 钼矿体的延伸,说明此处可能存在盲矿体。

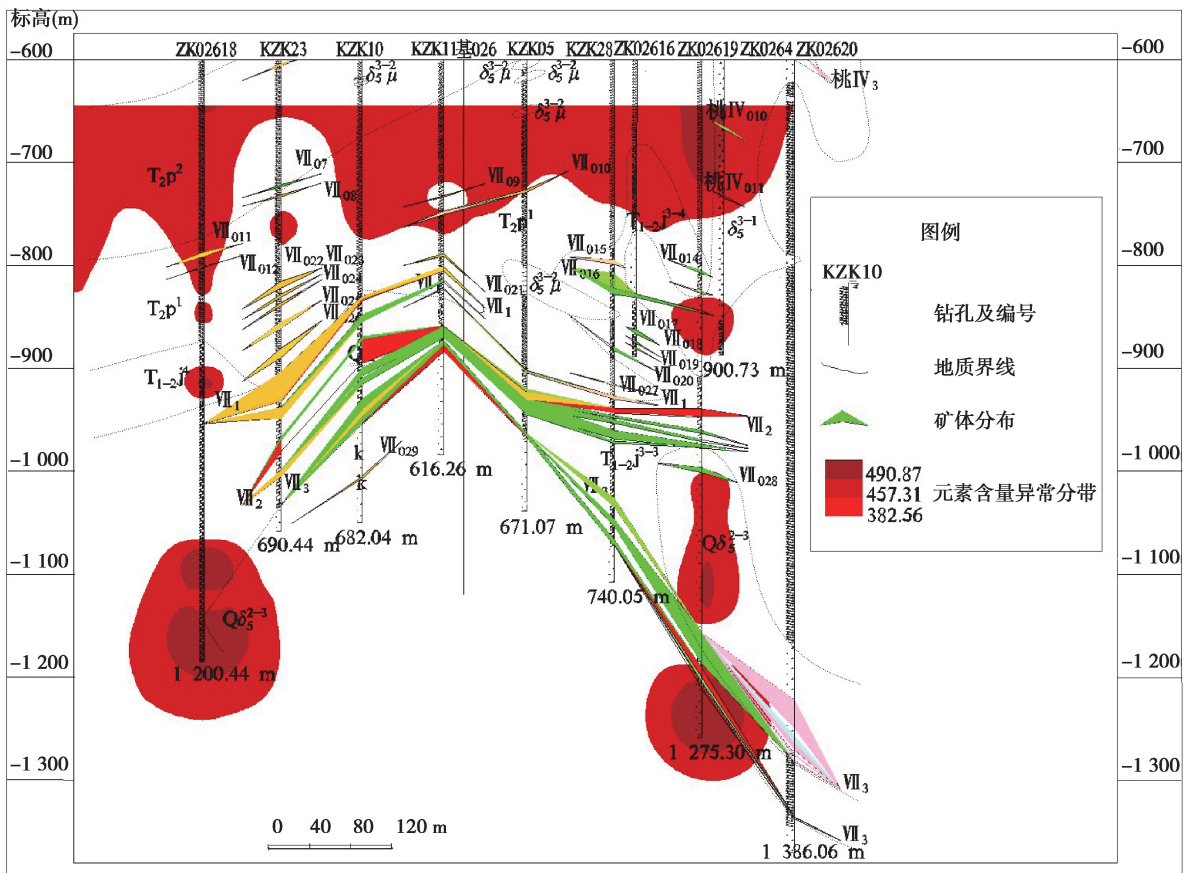


图 8 Ba 元素异常空间分布
 Fig. 8 The identified anomaly region of Ba

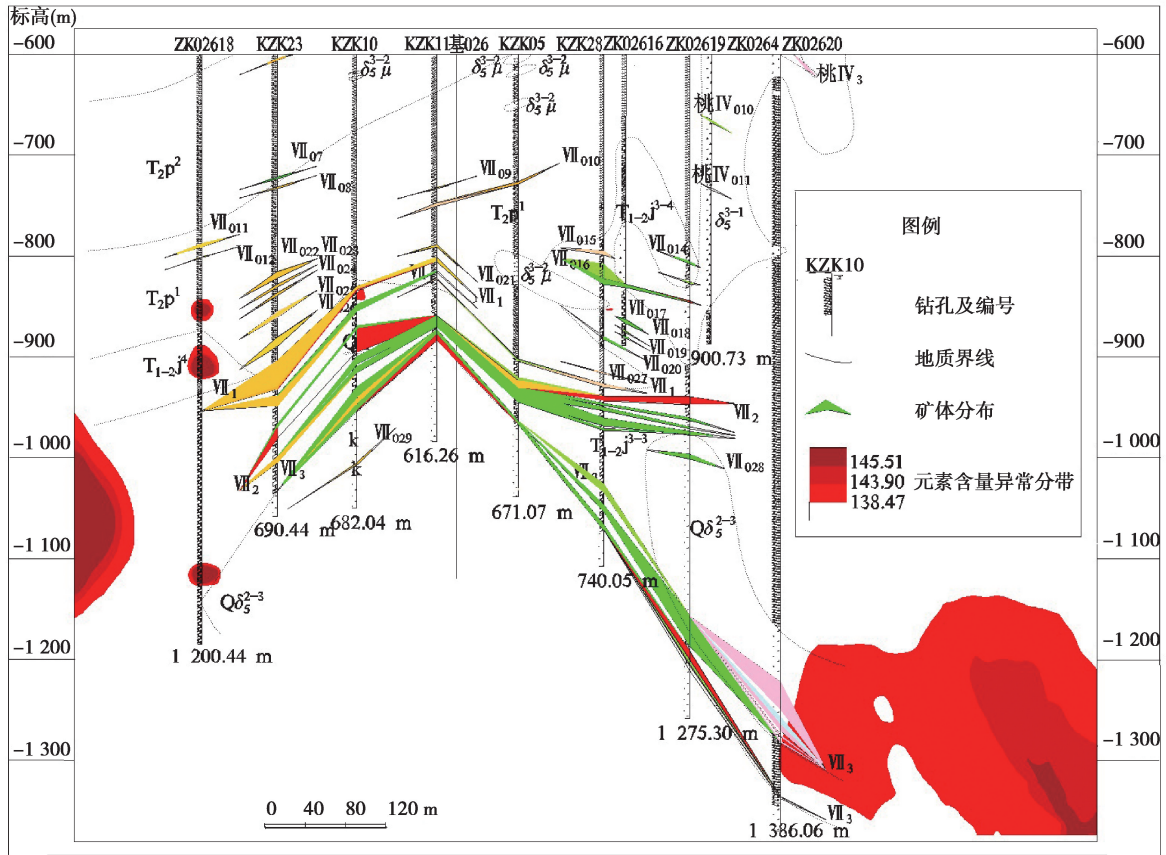


图 9 Zn 元素异常空间分布

Fig. 9 The identified anomaly region of Zn

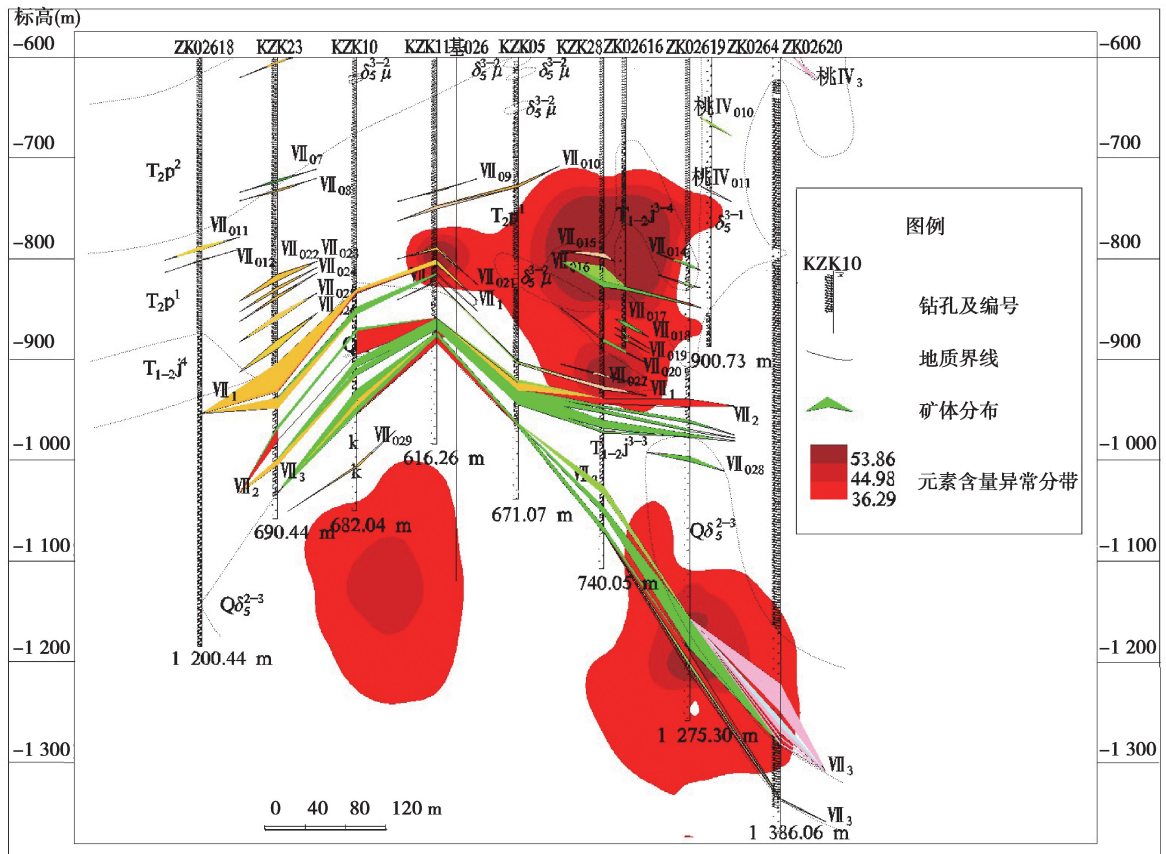


图 10 Mo 元素异常空间分布

Fig. 10 The identified anomaly region of Mo

4 结 论

以实际矿区为研究对象,通过构建地球化学异常 GPD 模型研究原生晕地球化学元素异常分布规律,有利于深部矿产预测。具有如下结论:

1)通过原生晕地区化学元素含量异常分布具有后尾性特征,结合 GPD 原理构建了一套地球化学元素异常提取模型,并对模型的参数和阈值进行了探讨和验证,同时还设计了修正尺度函数来确定元素具体异常下限值。

2)将所设计的 GPD 模型进行实际地球化学元素含量异常提取,得出 Mo, Ba, Zn 元素含量异常下限分别为 36.29, 382.56, 138.47, 通过诊断检验表明元素的理论分布与实际分布具有一致性,从而设计的模型具有合理性。

3)通过设计的模型得出 Mo, Ba, Zn 元素异常分带与矿体走势吻合,根据元素异常分布规律可以确定矿体深部存在较大找矿潜力,尤其位于钻孔 KZK10 和 KZK11 深度以下 1 100 m 左右具有强异常,可能存在钨盲矿体。

参考文献:

- [1] 陶建格, 沈镭. 矿产资源价值与定价调控机制研究[J]. 资源科学, 2013, 35(10): 1959-1967.
TAO Jiange, SHEN Lei. The value of ore resources and pricing regulation mechanisms[J]. Resources Science, 2013, 35(10): 1959-1967. (in Chinese)
- [2] 姚磊, 吕志成, 陈辉, 等. 再谈矿山深部及外围找矿新发现及意义[J]. 南京大学学报(自然科学), 2018, 54(2): 296-307.
YAO Lei, LU Zhicheng, CHEN Hui, et al. A reappraisal on the new discovery of deep ore exploration of mines and adjacent areas and its significances[J]. Journal of Nanjing University (Natural Science), 2018, 54(2): 296-307. (in Chinese)
- [3] Xiong Y H, Zuo R G, Wang K X, et al. Identification of geochemical anomalies via local RX anomaly detector[J]. Journal of Geochemical Exploration, 2018, 189: 64-71.
- [4] Cheng Q M. Non-linear theory and power-law models for information integration and mineral resources quantitative assessments[M]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008: 195-225.
- [5] Cheng Q M. Singularity analysis of magmatic flare-ups caused by India - Asia collisions[J]. Journal of Geochemical Exploration, 2018, 189: 25-31.
- [6] Shen W. Fractal invariable distribution and its application in large-sized and super large-sized mineral deposits[J]. Geoscience Frontiers, 2011, 2(1): 91-95.
- [7] Zuo R G. Selection of an elemental association related to mineralization using spatial analysis[J]. Journal of Geochemical Exploration, 2018, 184(A): 150-157.
- [8] Wang J, Zuo R G. Identification of geochemical anomalies through combined sequential Gaussian simulation and grid-based local singularity analysis[J]. Computers & Geosciences, 2018, 118: 52-64.
- [9] 陈健, 李正栋, 钟皓, 等. 多种地球化学异常下限确定方法的对比研究[J]. 地质调查与研究, 2014, 37(3): 187-192.
CHEN Jian, LI Zhengdong, ZHONG Hao, et al. Comparison of multiple methods to determine the geochemical anomaly threshold[J]. Geological Survey and Research, 2014, 37(3): 187-192. (in Chinese)
- [10] 严冰, 阳正熙, 周莉, 等. 证据权法成矿预测模型结合分形模型在成矿预测中的应用研究[J]. 矿业研究与开发, 2012, 32(1): 29-33, 122.
YAN Bing, YANG Zhengxi, ZHOU Li, et al. Application of the combination of weights-of-evidence model and fractal model in metallogenic prediction[J]. Mining Research and Development, 2012, 32(1): 29-33, 122. (in Chinese)
- [11] Tian M, Wang X Q, Nie L S, et al. Recognition of geochemical anomalies based on geographically weighted regression: a

- case study across the boundary areas of China and Mongolia[J]. *Journal of Geochemical Exploration*, 2018, 190: 381-389.
- [12] 秦飞龙. 原生晕地球化学和岩心高光谱粗糙集耦合建模与深部矿产预测研究[D]. 成都: 成都理工大学, 2017.
QIN Feilong. Coupling modeling of the primary halo geochemistry and the core hyperspectral with the rough set model and its studying in the prediction of deep mineral resources[D]. Chengdu: Chengdu University of Technology, 2017. (in Chinese)
- [13] Mohie El-Din M M, Sadek A, Mohie El-Din M M, et al. Characterization of the generalized Pareto distribution by general progressively Type-II right censored order statistics[J]. *Journal of the Egyptian Mathematical Society*, 2017, 25(4): 369-374.
- [14] Chen Y Q, Zhao P D, Chen J G, et al. Application of the geo-anomaly unit concept in quantitative delineation and assessment of gold ore targets in Western Shandong Uplift Terrain, Eastern China[J]. *Natural Resources Research*, 2001, 10(1): 35-49.
- [15] 左仁广. 基于地质异常的矿产资源量化预测与不确定性评价[D]. 武汉: 中国地质大学, 2009.
ZUO Renguang. Geoanomaly-based mineral resources quantitative prediction and uncertainty evaluation[D]. Wuhan: China University of Geosciences, 2009. (in Chinese)
- [16] 王晓华. 基于 MapGIS 与 Grapher 联合绘制地球化学剖面图[J]. *吉林地质*, 2017, 36(4): 75-77.
WANG Xiaohua. Geochemical profile drawing based on MapGIS and Grapher[J]. *Jilin Geology*, 2017, 36(4): 75-77. (in Chinese)
- [17] Rootzén H, Segers J, Wadsworth J L. Multivariate generalized Pareto distributions: parametrizations, representations, and properties[J]. *Journal of Multivariate Analysis*, 2018, 165: 117-131.
- [18] 花拥军. 极值理论[M]. 北京: 科学出版社, 2011.
HUA Yongjun. Study on extreme value theorem[M]. Beijing: Science Press, 2011. (in Chinese)
- [19] Qin F L, Liu B L, Guo K, et al. Using EVT for geological anomaly design and its application in identifying anomalies in mining areas[J]. *Mathematical Problems in Engineering*, 2016(3): 1-11.
- [20] Castillo J D, Serrà I. Likelihood inference for generalized Pareto distribution[J]. *Computational Statistics & Data Analysis*, 2015, 83: 116-128.

(编辑 张 莘)