doi: 10.11720/wtyht.2020.0038
孙尧尧,郝立波,赵新运,等.EM 聚类方法在岩性复杂区水系沉积物地球化学异常圈定中的应用[J].物探与化探,2020,44(6):1306-1312,1321.http://doi.org/10.11720/wtyht.2020.0038
Sun Y Y,Hao L B,Zhao X Y,et al.The application of EM clustering method to the determination of stream sediment geochemical anomalies in areas with variable lithologies[J].Geophysical and Geochemical Exploration,2020,44(6):1306-1312,1321.http://doi.org/10.11720/wtyht.2020.0038

EM 聚类方法在岩性复杂区水系沉积物地球化学异常圈定中的应用

孙尧尧,郝立波,赵新运,陆继龙,马成有,魏俏巧
(吉林大学 地球探测科学与技术学院,吉林 长春 130026)

摘 要: 地球化学背景与异常的划分是化探找矿工作的关键环节,然而,岩性复杂区水系沉积物地球化学异常的圈定常受到岩性背景的影响,若不消除这个影响而直接进行异常圈定,可能会造成一些严重的错误。研究表明,岩性背景问题的本质是多重母体问题,而 Expectation-Maximization(EM)聚类方法能有效地分解多重母体,可在一定程度上消除岩性背景的影响。本文以湖南省某地 1:20 万水系沉积物地球化学数据为例,采用 EM 聚类方法进行多重母体分解,然后进行地球化学异常圈定,并讨论了多重母体分解对异常圈定的影响。结果表明,EM 聚类方法的应用可使异常圈定更加合理,主要体现在虚假异常的削弱和低弱异常的强化方面。

关键词: 地球化学异常;岩性背景;多重母体;EM 算法;水系沉积物
中图分类号: P632 **文献标识码**: A **文章编号**: 1000-8918(2020)06-1306-08

0 引言

在地球化学找矿工作中,异常的圈定是极为重要的环节。目前,在岩性复杂区水系沉积物地球化学异常圈定时仍存在一些亟待解决的问题,其中,岩性背景问题最为突出,尤其是在区域地球化学勘查中[1-4]。若不消除岩性背景差异的影响而直接采用统一的背景上限进行异常圈定,一是可能会将高背景值区的一些正常样品划分为异常样品,从而形成虚假地球化学异常;二是可能会将低背景值区的一些异常样品划分为正常样品,从而掩盖掉部分低弱但重要的地球化学异常[3-5]。

针对岩性背景问题,学者们提出了许多处理方法,如多元线性回归法[6]、滑动平均标准化法[1]和一元线性回归法[3-4]等。然而,这些方法的提出并不是从直接区分岩性背景的角度出发的,或多或少存在一些局限性,因此很难从根本上消除岩性背景的影响。Zhao 等认为岩性背景问题的本质是多重母体问题,因此提出利用 Expectation-Maximization(EM)算法来分解多重母体,进而达到消除岩性背景影响的目的[5]。该方法地质意义明确,对合理圈定岩性复杂区水系沉积物地球化学异常应能起到一定的作用。笔者以湖南省某地 1:20 万水系沉积物地球化学数据为例,尝试将 EM 聚类方法应用到地球化学背景与异常划分中,旨在为岩性复杂区水系沉积物地球化学异常的圈定提供参考。

1 岩性背景问题

研究表明,地球化学背景与异常划分的影响因素主要有岩性背景、地球化学景观条件以及样品粒度等[1-2,6]。但在相同或相近的景观条件下,岩性背景的影响最为突出[3-4]。岩性复杂区往往存在多种岩石类型,不同类型岩石的元素丰度通常存在明显差异。研究表明,岩性单一的地质体的元素含量往

收稿日期: 2020-01-19; **修回日期**: 2020-06-05
基金项目: 中央级公益性科研院所基本科研业务费专项(AS2016P02);国家重点研发计划项目(2016YFC0600606)
作者简介: 孙尧尧(1997-),女,硕士,地球探测与信息技术专业。Email: yaoyaos19@mails.jlu.edu.cn
通讯作者: 魏俏巧(1987-),女,博士,主要从事地球化学教学与研究工作。Email: weiqiaoqiao888@sina.com
万方数据

往服从(近似)正态分布[7-8],其数据可构成一(近似)正态分布的母体。因此,岩性复杂区应存在多个(近似)正态母体。岩石风化形成的水系沉积物对基岩的化学成分具有显著的继承性[9-10],因此岩性复杂区水系沉积物元素含量数据内也应包含多个(近似)正态母体[5]。来源于不同岩性背景的元素含量数据发生混合,即正态母体混合或正态分布叠加,可产生偏态分布,包括左偏、右偏及多峰等分布型式[7-8],这是岩性背景对地球化学异常圈定的直接影响。由此可见,岩性背景问题即多重母体问题,而多重母体问题主要源于岩性背景差异。

针对上述问题,学者们开展了大量的研究工作,提出了许多处理方法。研究表明,在使用任何数据处理方法划分地球化学背景与异常之前,应先确定元素含量的概率分布型式,且以(近似)正态分布为目标[11]。为满足正态分布的前提,一些数据变换方法被应用到异常圈定中,如对数变换法[12-13]、倒数变换法[14]和 Box-Cox 变换法[15]等,其中对数变换法最为常用。Stanley 研究指出,对数变换及类似的数学变换方法通常难以消除地球化学数据偏态分布的特征[16]。在地球化学数据处理工作中,这一现象是普遍存在的,通常是偏态分布经对数转换后仍不服从(近似)正态分布。然而,这些方法目前仍被广泛使用,可能已经造成了严重的错误。

按地质单元界线区分岩性背景的效果也不是很理想,原因是同一地质单元的岩性组成往往并不单一,即同一地质单元内也可能含有多个(近似)正态母体。因此,一些学者尝试从回归分析的角度来解决岩性背景问题。周蒂提出了多元线性回归分析方法[6],郝立波等提出了一元线性回归分析方法[3]。相比于传统的数据变换处理方法,这些方法具有很大的优势,可以在一定程度上消除岩性背景的影响,较准确地划分出成矿元素的地球化学背景与异常。然而,回归分析方法也有其局限性,因为该类方法仅适用于那些与岩性存在良好线性关系的成矿元素,而不适用于那些受岩性背景影响小或不受影响的元素。另外,线性回归分析要求残差符合正态分布,这也是该方法使用受限的原因之一。

李宝强和孙泽坤提出了滑动平均标准化方法[1],其不足之处是窗口大小的确定具有很大的人为性,若窗口过大,则难以消除元素背景含量的影响,而窗口过小,又会缩小各点的元素含量差异,损失异常信息。至于分形/多重分形方法,对于服从分形分布的元素含量数据,该类方法可能是一种很好的选择[17];然而,该类方法的地质和地球化学意义仍不明确[3-5,19],另外,该类方法也无法区分来自不同岩性母体的样品。同样地,线性判别分析方法[8,20]和概率格纸法[21]等也面临同样的问题,即无法判别样品具体归属于哪一类岩性母体[5]。

解决岩性背景问题的根本途径应是直接区分岩性背景或分解多重母体。Zhao 等研究认为,在岩性复杂区可利用能反映岩性变化的主量和(或)亲石微量元素,通过 EM 聚类算法——一种有效的高斯混合模型聚类方法,对来自不同岩性背景的水系沉积物样品进行分类,以达到区分岩性背景或分解多重母体的目的[5]。基于此,笔者以湖南省某地 1∶20 万水系沉积物地球化学数据为例,尝试采用 EM 聚类方法进行多重母体分解,然后进行地球化学异常圈定,并将其与传统的均值+k 倍标准偏差方法圈定的异常进行比较。

2　方法流程

2.1　样品分类

样品分类主要基于 EM 算法。该算法由 Dempster 等提出[22],是一种重要的统计学分析工具,已被广泛应用于含有隐变量的概率模型参数的极大似然估计中。假设在某一复杂岩性区采集了 n 个水系沉积物样品,其分析结果可设为随机变量 X,$X=(x_1,x_2,x_3,\cdots,x_n)$,样品采自不同的岩性背景区,即 X 内含有多个正态母体,但每个样品的对应母体或类别是未知的。正态母体的分布参数也是未知的,样品的对应类别可设为变量 Z,$Z=(z_1,z_2,z_3,\cdots,z_n)$。EM 算法是先初始化正态母体的分布参数 θ,然后重复以下两个步骤:①根据分布参数 θ 计算样品分类变量 Z 的期望,$Q_i(z_i)=p(z_i\mid x_i;\theta)$;②将概率模型参数的似然函数 $L(\theta)$ 最大化以获得新的分布参数 θ,即

$$\theta:=\arg\max_{\theta}\sum_{i=1}^{n}\sum_{z_i}Q_i(z_i)\ln\frac{p(x_i,z_i;\theta)}{Q_i(z_i)},$$

似然函数 $L(\theta)$ 收敛时,迭代停止。最终,可得到样品的具体分类情况,即在寻找最优参数来极大化样本似然函数的过程中对样本完成最优分类[5]。

在实际问题处理中,应首先根据研究区地质概况选取可以反映岩性变化,且元素间相关性较小的非成矿常量和(或)亲石微量元素作为聚类指标,然后采用 EM 算法对样品进行分类。最优分类数可采用数学衡量指标和地质条件双重限制进行合理判别[5]。数学衡量指标为赤池信息量准则 *AIC*(*akaike's information criterion*),一个衡量模型拟合程

万方数据

度的标准[23]。*AIC* 可以表示为：

$$AIC = 2k - 2\ln(L)。$$

式中：k 是模型的独立参数个数；L 是模型的极大似然函数。按极大似然函数准则，L 越大，模型的拟合度越高，但实际上，在 L 过大的情况下，可能会出现过拟合现象。为了避免过度拟合，*AIC* 准则在追求 L 尽可能大的同时，还要求模型参数要尽可能少，即 k 要尽可能小，换句话说，*AIC* 值越小，模型的拟合情况越好。因此，在有限混合正态分布参数的极大似然估计中，当 *AIC* 值达到最小时可获得最优分类数。采用地质条件限制时，可先根据分类结果绘制分类图，然后将其与地质图进行对比，与地质图吻合度最高的分类图对应的分类数可确定为最优分类数。按统计学的要求，每个分类中样品数应不少于 30 个，若不满足，则需调整分类，或将样品数较少的类合并到与其最近的类中。

2.2 数据标准化

样品分类完成后，对每个分类中的样品成矿元素含量数据进行 3S 检验，将均值±3 倍标准偏差范围外的异常数据剔除。然后，利用剩余数据计算每个分类中成矿元素含量的均值和标准偏差，以此对成矿元素的含量进行 Z 分数标准化，公式如下：

$$y_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j},$$

$$i = 1,2,3,\cdots,n;\ j = 1,2,3,\cdots,m。$$

式中：x_{ij} 为样品 i 中元素 j 的含量；y_{ij} 为样品 i 中元素 j 的含量标准化值；$\bar{x}_j$ 和 S_j 分别为元素 j 的均值和标准偏差。通过标准化，可将来自不同正态母体的元素含量数据校正到统一背景（即均值为 0，标准偏差为 1）。合并标准化后的各个分类的地球化学数据，用于下一步地球化学异常的统一圈定。

2.3 异常圈定

经样品分类和元素含量标准化后，岩性背景差异在一定程度上得到了消除。因此，可在全区采用统一的背景上限进行异常圈定，即直接采用传统的均值加 k 倍标准偏差的方法或其他方法进行异常圈定，k 值的大小可依据实际情况进行适当调整。

3 应用实例

3.1 研究区概况

以湖南省某地 1∶20 万水系沉积物地球化学数据为例，进行了基于多重母体分解的地球化学异常圈定试验。研究区面积约为 4 300 km^2，水系沉积物样品数为 1 050，采样密度约为 1 个样品/4 km^2。水系沉积物化学成分分析项目包括 Ag、As、Au、B、Ba、Be、Bi、Cd、Co、Cr、Cu、F、Hg、La、Li、Mn、Mo、Nb、Ni、P、Pb、Sb、Sn、Sr、Th、Ti、U、V、W、Y、Zn、Zr、Al_2O_3、CaO、Fe_2O_3、K_2O、MgO、Na_2O 及 SiO_2。研究区发育有多个钨锡多金属矿床，成矿地质条件优越。研究区地层复杂，但岩性较为简单，按照出露面积从大到小的排列依次为上古生界灰岩、泥灰岩，下古生界浅变质砂岩，白垩系砂岩，泥盆系砂岩和震旦系砂质板岩。研究区侵入岩主要为印支期和燕山期二长花岗岩，分布于研究区中部偏西处。区内发育有多条走向近 SN 和 NW 的断裂（图 1）。

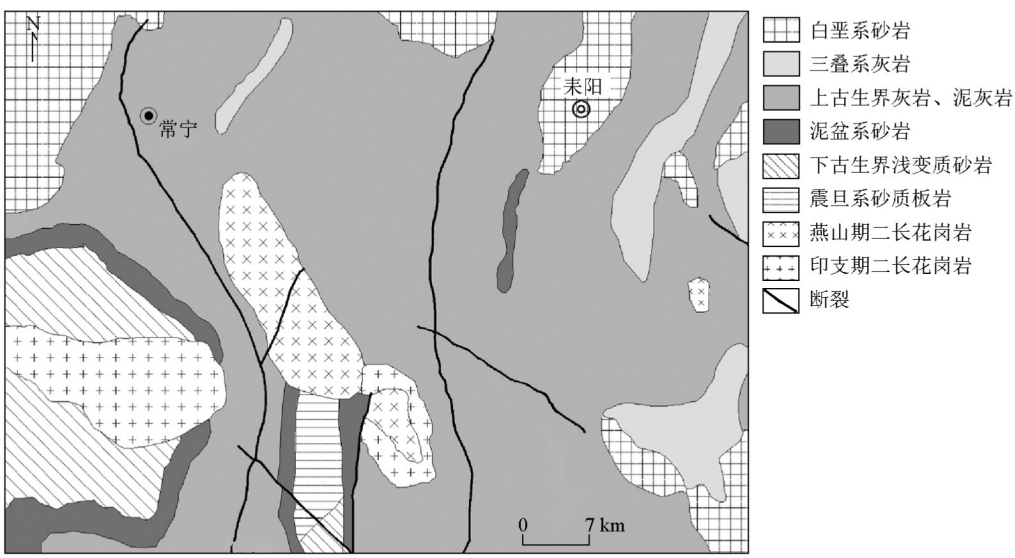

图 1 研究区地质简图

Fig.1 Schematic geological map of the study area

万方数据

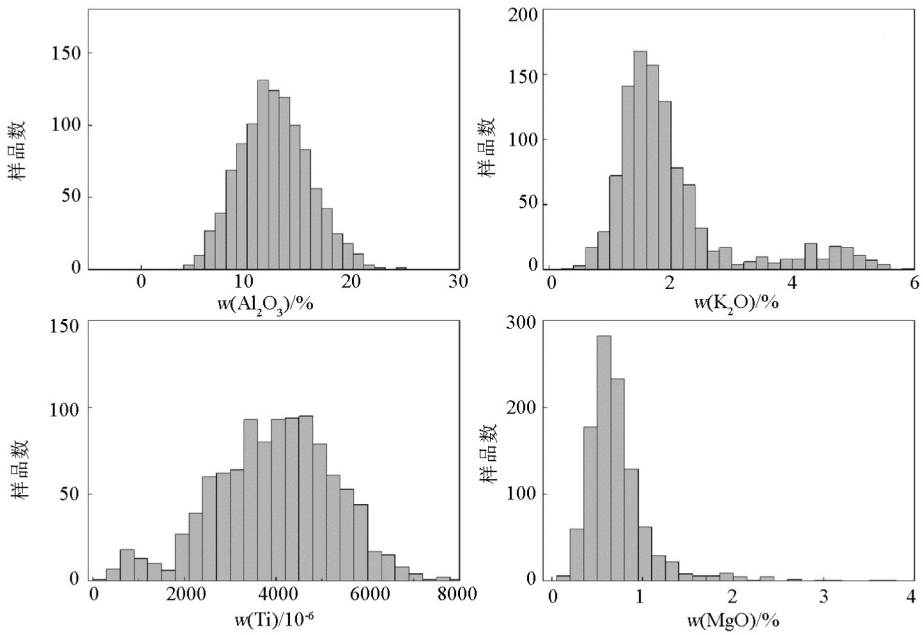

图 2　水系沉积物样品 Al_2O_3、K_2O、Ti、MgO 含量概率分布型式

Fig.2　The frequency distributions of Al_2O_3, K_2O, Ti and MgO of stream sediment samples

研究区岩性组成主要为灰岩、泥灰岩、砂岩和二长花岗岩等。岩性背景差异导致水系沉积物中元素含量的概率分布型式多呈偏态(图 2),如 K_2O 和 MgO 均呈明显的右偏分布,Ti 呈明显的左偏分布,且 K_2O 和 Ti 还具有多峰分布特征;Al_2O_3 虽然呈正态分布,但可能也是多个正态分布叠加的结果。为了合理地划分研究区水系沉积物地球化学背景与异常,需消除岩性背景差异的影响。

3.2　水系沉积物样品分类

首先,选择可以反映岩性变化,且元素间相关性较小的非成矿的常量和亲石微量元素作为分类指标。基于上述原则,本次实验共挑选出 7 种元素,即 Mn、Ti、V、Zr、CaO、K_2O 和 MgO,对其他受岩性影响较小的元素和成矿元素 W、Sn 则不予考虑。利用上述分类指标,通过 EM 算法对样品进行分类,分类过程主要在 MATLAB 中实现。

AIC 值计算结果表明,随着分类数的增大,*AIC* 值先是迅速减小,当分类数增大到一定程度后,*AIC* 值的下降幅度变得越来越小,最后趋于平稳,即 *AIC* 曲线趋于水平(图 3)。根据 *AIC* 判别原则,*AIC* 值达到最小时可获得最优分类数,即分类数越多越好,可在最大程度上区分混合正态母体。然而,在分类过程中还需考虑实际情况,因为分类过细可能会使某些分类中的样品数过少而达不到基本统计要求。

观察发现,分类数 8 前后的 *AIC* 值变化幅度较大,因此可以考虑 8 及以上的分类数;在分类过程中,分类数过多时部分类别中的样品数过少,甚至

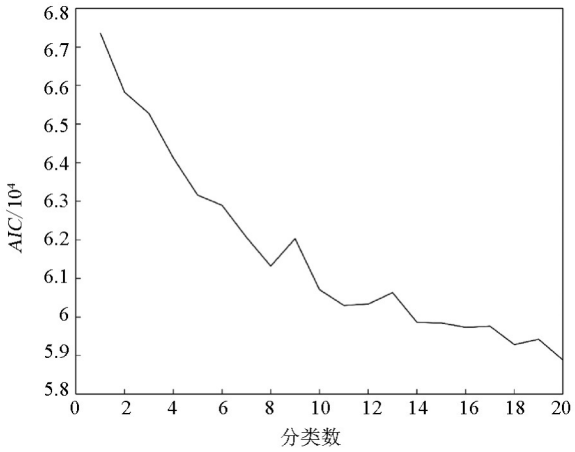

图 3　分类数与 *AIC* 值关系曲线

Fig.3　Relationship between number of clusters and *AIC* values

只有个位数样品;另外,研究区岩性组成相对简单,因此分类数不宜过大。基于上述理由,在分类数 8 的基础上将一些样品数过少的类合并到了附近的类中,最终将样品分为 5 类(图 4)。由于 EM 算法在计算过程中容易出现局部最优解的情况,因此在每次分类过程中进行了多次重复计算。将分类结果成图,比较分类图与地质图(图 1)的吻合程度,发现分类结果较为合理。整体上,分类界线与研究区主要地质单元界线吻合较好,但也存在同一地质单元内出现多种分类的现象,如图幅中部的上古生界灰岩、泥灰岩分布区出现了第 1 类和第 3 类样品,这是同一地质单元内具有不同岩性单元的真实反映。另外,还存在同一分类区域对应着不同地质单元的现象,如图幅中第 4 类样品分布区对应着白垩系砂岩、

万方数据

泥盆系砂岩和下古生界浅变质砂岩。总的来说,将水系沉积物样品分为 5 类较符合实际情况,其中,第 1 类样品对应灰岩,第 2 类样品对应燕山期二长花岗岩,第 3 类样品对应泥灰岩,第 4 类样品对应砂岩,第 5 类样品对应印支期二长花岗岩。

分类后绝大多数母体中的样品元素含量都服从(近似)正态分布(图 5),说明分解多重母体取得了较好的效果,在确定地球化学背景与异常的过程中可直接采用传统的均值加 k 倍标准偏差的方法。值得注意的是,某些区域内可能存在水系沉积物混合物,是由来自不同岩性背景区的水系沉积物混合形成的。研究表明,多重母体分解后,这些水系沉积物混合物中的多数元素含量仍不服从(近似)正态分布[5]。针对这种情况,目前还没有较好的处理方法。在这种情况下可作取对数、平方根、倒数等变换或直接当作正态分布处理,这是一个策略选择问题。

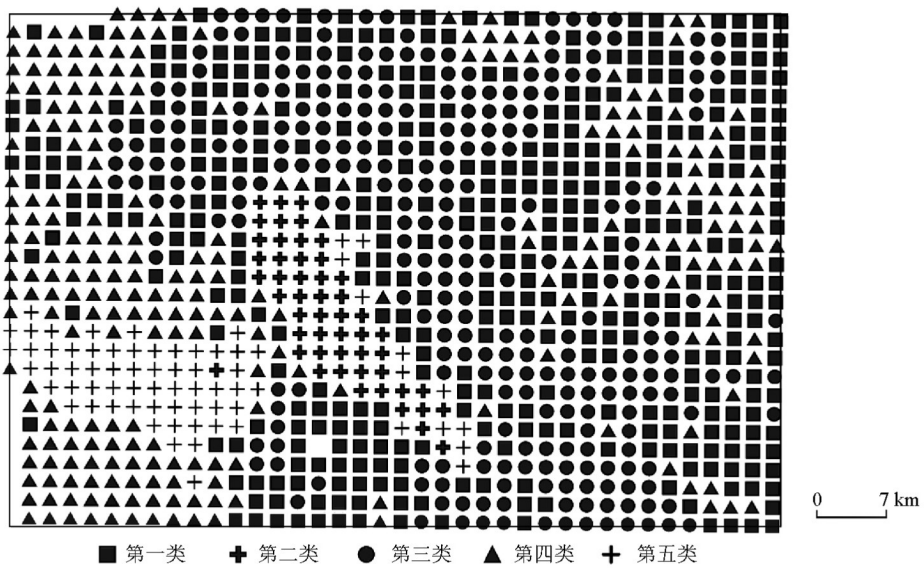

图 4 分类结果

Fig.4 Optimal classification map

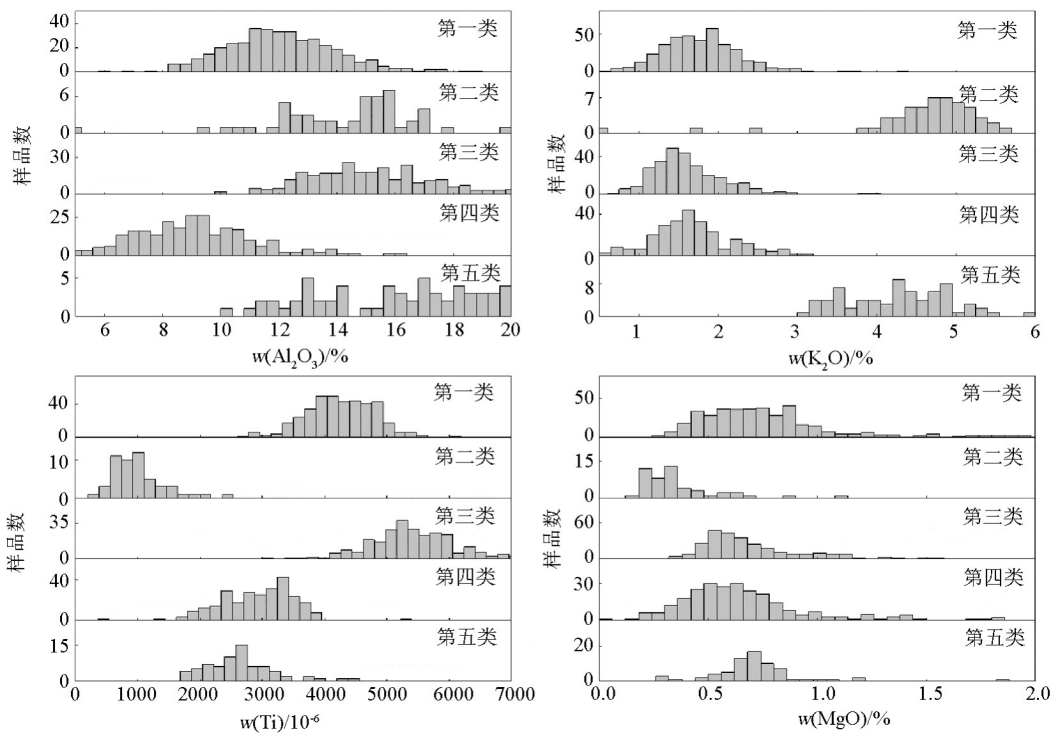

图 5 分类后 Al_2O_3、K_2O、Ti、MgO 含量概率分布型式

Fig.5 The frequency distributions of Al_2O_3, K_2O, Ti and MgO of separated samples

万方数据

3.3　异常圈定及效果评价

为了便于比较,笔者采用 EM 方法和传统的均值加 k 倍标准偏差的方法分别圈定了 W 和 Sn 的地球化学异常(图 6)。由于 W 和 Sn 的概率分布型式为非正态,因此采用传统方法处理之前先对原始数据进行了对数变换,然后采用 3S 检验法对所有样品数据进行检验,剔除异常样品后再计算均值和标准偏差,以此进行 Z 分数标准化。数据标准化后,将背景上限确定为均值加 2 倍标准偏差,然后进行异常圈定。EM 方法计算步骤为:对每个子类的样品数据先采用 3S 检验方法剔除异常样品,计算均值和标准偏差,然后对各子类中的数据进行标准化,以消除岩性背景差异的影响,合并标准化之后的各个子类数据按统一背景上限进行异常圈定。

结果表明,两种方法圈定的地球化学异常存在明显差异,消除岩性背景后圈定的异常无论从强度上还是形态上均发生了显著变化。总的来说,分类后圈定的 W 和 Sn 的异常与已知矿床(点)的位置对应关系更好,而常规方法圈定的异常有偏移现象。另外,本文方法圈定的异常强度得到了明显增强,不仅发现了一些传统方法没有发现的低弱异常,如 B1、B2、B3、B4 和 B5,还消除了岩性背景高值区形成的虚假异常,如二长花岗岩分布区的 A1 和 A2。这些虚假异常是由岩性背景造成的。统计发现,不同岩性分布区的元素背景值差异较大(表 1),砂岩分布区的元素背景值最低,其次是灰岩和泥灰岩分布区,而燕山期二长花岗岩分布区的元素背景值最高,印支期二长花岗岩次之,若全区统一进行异常圈定,显然会压低砂岩分布区的异常,放大二长花岗岩分布区的异常。

受岩性背景影响较大的 W 和 Sn,两种方法圈定的异常差异较大。由于二长花岗岩分布区的元素背景值明显高于其他岩性分布区,因此传统方法圈定的 W、Sn 异常多位于二长花岗岩分布区,而在灰岩、泥灰岩及砂岩分布区几乎没有显示。相比较而言,本文方法消除了岩性背景差异的影响,圈定出的 W、Sn 异常主要分布于二长花岗岩岩体的周边,且在砂岩、灰岩及泥灰岩分布区也有显示。另外,本文方法圈定的异常规模也较大,强度也较高,与已知矿床(点)的对应关系更好,说明利用该方法圈定的异常更为合理。

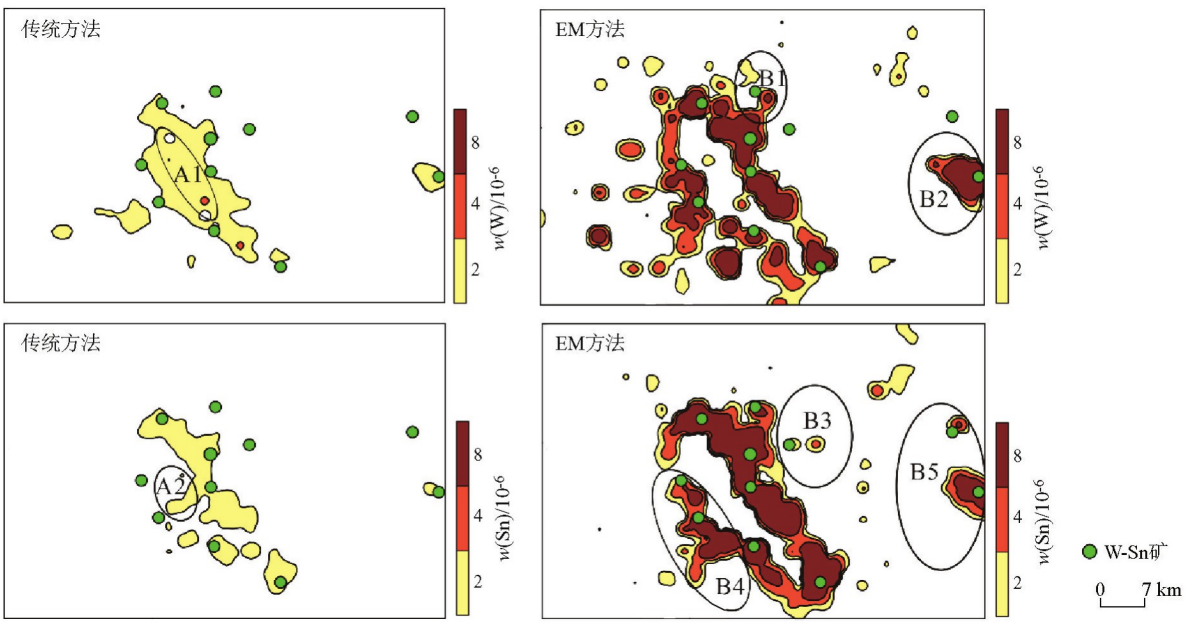

图 6　传统方法(左侧)与 EM 方法(右侧)圈定的 W、Sn 地球化学异常

Fig.6　Geochemical anomalies of W and Sn delineated by traditional method(left) and the EM method(right)

表 1　分类样品中 W、Sn 背景值

Table 1　Background values of W and Sn of separated samples

背景值	第一类样品(391)	第二类样品(52)	第三类样品(278)	第四类样品(259)	第五类样品(70)
$w(W)/10^{-6}$	3.4	55	4.9	2.7	18
$w(Sn)/10^{-6}$	5.2	31	6.2	2.7	17

注:括号内为样品数。

万方数据

4 结论

1）采用 EM 分类方法能够快速、有效地区分来自不同岩性背景的水系沉积物样品，可在一定程度上消除岩性背景的影响。该方法以水系沉积物对基岩化学成分显著的继承性为基础，具有明确的地质意义。

2）消除岩性背景影响后，不仅可以圈定出低弱地球化学异常，还可以削弱虚假地球化学异常。

3）分类完成后，对各类样品元素含量进行正态分布检验，若服从（近似）正态分布，则可直接采用传统的均值加 k 倍标准偏差的方法或其他方法圈定地球化学异常，若不服从（近似）正态分布，则可作取对数、平方根、倒数等变换或直接当作正态分布处理，这是一个策略选择问题。

参考文献（References）：

[1] 李宝强，孙泽坤. 区域地球化学异常信息提取方法研讨[J]. 西北地质，2004，37(1)：102－108.
Li B Q，Sun Z K. Study on the method of geochemical anomalies analysis[J]. Northwestern Geology，2004，37(1)：102－108.

[2] 程志中，谢学锦. 岩石中元素背景值变化对地球化学成矿预测的影响[J]. 中国地质，2006，33(2)：411－417.
Cheng Z Z，Xie X J. Influence of variation in element background values in rocks on metallogenic prognosis in geochemical maps[J]. Geology in China，2006，33(2)：411－417.

[3] 郝立波，李巍，陆继龙. 确定岩性复杂区的地球化学背景与异常的方法[J]. 地质通报，2007，26(12)：1531－1535.
Hao L B，Li W，Lu J L. Method for determining the geochemical background and anomalies in areas with complex lithology[J]. Geological Bulletin of China，2007，26(12)：1531－1535.

[4] Hao L B，Zhao X Y，Zhao Y Y，et al. Determination of the geochemical background and anomalies in areas with variable lithologies [J]. Journal of Geochemical Exploration，2014，139：177－182.

[5] Zhao X Y，Hao L B，Lu J L，et al. Origin of skewed frequency distribution of regional geochemical data from stream sediments and a data processing method [J]. Journal of Geochemical Exploration，2018，194：1－8.

[6] 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.
Zhou D. Unit-wise adjustment of geochemiacl background data and its significance in geochemical anomaly delineation[J]. Geophysical and Geochemical Exploration，1986，10(4)：263－273.

[7] Vistelius A B. The skew frequency distributions and the fundamental law of the geochemical processes [J]. The Journal of Geology，1960，68：1－22.

[8] Govett G J S，Goodfellow W D，Chapman R P，et al. Exploration geochemistry：distribution of elements and recognition of anomalies [J]. Mathematical Geology，1975，7：415－446.

[9] 郝立波，马力，赵海滨. 岩石风化成土过程中元素均一化作用及机理：以大兴安岭北部火山岩区为例[J]. 地球化学，2004，33(2)：131－138.
Hao L B，Ma L，Zhao H B. Elemental homogenization during weathering and pedogenesis of volcanic rocks from North Da Hinggan Ling[J]. Geochimica，2004，33(2)：131－138.

[10] 郝立波，陆继龙，马力. 浅覆盖区土壤化学成分与基岩化学成分的关系及其意义——以大兴安岭北部地区为例[J]. 中国地质，2005，32(3)：477－482.
Hao L B，Lu J L，Ma L. Relation between the chemical compositions of residual soils and bedrocks in shallow overburden areas and its significance：A case study of the northern Da Hinggan Mountains[J]. Geology in China，2005，32(3)：477－482.

[11] Reimann C，Filzmoser P，Garrett R. Background and threshold：critical comparison of methods of determination [J]. Science of the Total Environment，2005，346：1－16.

[12] Ahrens L H. A fundamental law of geochemistry [J]. Nature，1953，172：1148.

[13] Miesch A T，Chapman R P. Log transformation in geochemistry [J]. Mathematical Geology，1977，9：191－198.

[14] Hoyle M H. Transformations：An introduction and a bibliography [J]. International Statistical Reviews，1973，41：203－223.

[15] Box G E P，Cox D R. An analysis of transformations [J]. Journal of the Royal Statistical Society：Series B，1964，26：211－252.

[16] Stanley C. Numerical transformation of geochemical data：1. Maximizing geochemical contrast to facilitate information extraction and improve data presentation [J]. Geochemistry：Exploration，Environment，Analysis，2006，6：69－78.

[17] Cheng Q M，Xu Y G，Grunsky E. Integrated spatial and spectrum method for geochemical anomaly separation [J]. Natural Resources Research，2000，9：43－51.

[18] Cheng Q M. A new model for quantifying anisotropic scale invariance and for decomposition of mixing patterns [J]. Mathematical Geology，2004，36：345－360.

[19] 谢淑云，鲍征宇. 地球化学场的连续多重分形模式[J]. 地球化学，2002，31(2)：191－200.
Xie S Y，Bao Z Y. Continuous multifractal model of geochemical fields[J]. Geochimica，2002，31(2)：191－200.

[20] Whitehead R E S，Govett G J S. Exploration rock geochemistry—detection of trace element halos at heath steele mines (N.B.，Canada) by discriminant analysis [J]. Journal of Geochemical Exploration，1974，3：371－386.

[21] Sinclair A J. Selection of threshold values in geochemical data using probability graphs [J]. Journal of Geochemical Exploration，1974，3：129－149.

[22] Dempster A P，Laird N M，Rubin D B. Maximum likelihood from incomplete data via the EM Algorithm [J]. Journal of the Royal Statistical Society：Series B，1977，39：1－38.

[23] Akaike H. Statistical predictor identification [J]. Annals of the Institute of Statistical Mathematics，1970，22：203－217.

万方数据

下转 1321 页

The application of multiple geophysical methods to the study of deep metallogenic regularity in the Sanhexu mining area, the Xianghualing orefield, Hunan Province

QU Li-Jun, WANG Qing, LI Bo, YAO Wei

(*Geophysical and Geochemical Institute of Hunan Province, Changsha* 410116, *China*)

Abstract: In this study, the surface anomaly distribution characteristics were quickly grasped by performing high-precision magnetic methods on the ground in the Sanhexu mining area of the Xianghualing orefield. Preliminary prediction anomalies are mainly caused by northeastward fault zones. The magnetic anomalies were analyzed by using CSAMT, and a continuous high-resistance uplift was found in the middle and deep parts. The anomaly is inferred to be granite. On such a basis, the authors began inversion of the magnetic anomaly, and it is found that the magnetic object is layered at the edge of the granite and is located in a low-resistance anomaly zone. According to the results of geophysical inference, the favorable metallogenic sites were selected for drilling and exposure. The boreholes ZK502, ZK1601, and ZK3001 successfully verified the inferred granite and revealed high level tin-lead-zinc polymetallic orebody in the sandstone of the Tiaomajian stratum. Based on the results of drilling, geophysical prospecting, and geological background, the authors carried out a comprehensive analysis and study of the metallogenic regularity of the mining area, established a metallogenic model, and delineated the next target for ore prospecting. The above two geophysical methods have good effects on the prediction of deep mineralization.

Key words: high precision magnetic method; CSAMT; deep prospecting; metallogenic model

(本文编辑:王萌)

上接 1312 页

The application of EM clustering method to the determination of stream sediment geochemical anomalies in areas with variable lithologies

SUN Yao-Yao, HAO Li-Bo, ZHAO Xin-Yun, LU Ji-Long, MA Cheng-You, WEI Qiao-Qiao

(*College of Geo-Exploration Science and Technology, Jilin University, Changchun* 130026, *China*)

Abstract: The separation of anomalies from backgrounds is a critical step in geochemical prospecting. However, the determination of stream sediment geochemical anomalies is always affected by lithologic backgrounds in areas with variable lithologies. If this influence is not eliminated prior to anomaly determination, some serious errors may occur. In fact, a problem of lithologic background is essentially a problem of multiple population, which can be effectively solved by the clustering method based on the Expectation-Maximization (EM) algorithm. In this study, the authors applied the EM clustering method to a 1 ∶ 200,000 scale stream sediment survey, and then discussed the influence of separating multiple populations on anomaly determination. A practical example demonstrates that geochemical anomalies of stream sediments in lithologically complex regions can be determined in a more reasonable way by using the EM clustering method. This is mainly reflected in two aspects: on the one hand, strong but false anomalies can be eliminated, and on the other hand, weak but important anomalies can be uncovered.

Key words: geochemical anomaly; lithology background; multiple population; EM algorithm; stream sediments

(本文编辑:蒋实)

万方数据