

子区中位数衬值滤波法及弱小异常识别

史长义 张金华 黄笑梅

摘 要 提出一种新的数据处理方法——子区中位数衬值滤波法。该法采用子区模拟背景场的空间变异，以EDA技术计算异常临界值，用衬值为基准度量场值的起伏变化，较好地解决了弱小异常的识别问题。初步应用效果良好。

关键词 数据处理；子区中位数衬值滤波法；弱小异常识别；区域化探

SUBREGION MEDIAN CONTRAST FILTERING METHOD AND RECOGNITION OF WEAK ANOMALIES

Shi Changyi, Zhang Jinhua, Huang Xiaomei

(Institute of Geophysical and Geochemical Exploration, Langfang 065000)

Abstract This paper advances a new data-processing technique—subregion median contrast filtering method. It uses subregion to simulate spatial variation of the background, calculates critical values of anomalies with EDA technique, and measures the undulation of field values with the contrast as the datum. This technique fairly satisfactorily solves the problem of recognizing weak anomalies, and the preliminary application proves it to be effective.

Key words data-processing; subregion median contrast filtering method; recognition of weak anomalies; regional geochemical prospecting

自30年代勘查地球化学诞生以来，化探异常的解释推断一直是勘查地球化学界研究活动的前沿。特别是60年代末的国际性“区域化探热”和70年代计算机技术在化探界的普遍应用，给区域化探异常的解释推断带来勃勃生机，相继出现了各种各样的新方法、新手段，特别是计算机系统。但是，异常的圈定、筛选和解释，仍面临着很大的困难。目前化探数据处理有两种倾向：一是利用包括单变量和多变量的各种经典统计学方法，但由于化探数据的复杂性，使得处理结果有时很难解释；二是使用各种简单的数学方法来提取多元素数据中所蕴藏的各种信息，以直观而又易于理解的方式展现，供人们进行主观判断和解释。近十几年，当人们逐渐认识到那些建立在经典统计学基础上的各种传统方法的局限性时，尤其是稳健统计学在化探数据处理中应用后，又有不少确定异常下限的新方法新技术被提出^{〔1〕}，使得异常的识别更具客观性和真实性。

十多年来，我国已完成1·20万区域化探扫面约400万km²，并积累了大量的宝贵资料。这些资料在地质找矿，特别是在找金矿中发挥了巨大作用。但是其潜在的价值尚未得到充分发挥和利用。异常的解释推断和对已取得资料的开发利用是当前区域化探工作面临的两大问题，这直接关系到区域化探能否顺利深入开展与取得经济效益的大小。有关资料证明，不同地区、不同景观条件的地质地球化学背景差异很大，导致不同类型矿

床甚至同一成因类型矿床的异常强度差异较大甚至相差甚远。从目前所谓在有些地区大异常中找不到矿,而在小异常中却找到大矿的原因,除由于埋深和某种特殊景观的影响外,多数是在各子区的地球化学背景差异较大的情况下,采用了全区统一的异常下限来圈定异常,使得高背景区异常变大而较低背景区异常变小或变弱,甚至不能被发现。

目前,在区域化探数据处理中的主要问题是:在处理区化数据,确定异常下限以识别异常时,通常采用的是全区(一幅或几幅图)统一的异常下限,而忽略了不同地质背景、不同地球化学景观区的元素地球化学背景不同甚至差异很大的事实;识别弱小异常的方法技术不成熟;对传统的“ $+2$ (或 3) s 作为异常下限”的局限性认识不足。

针对上述问题,本文提出一种新的数据处理方法——子区中位数衬值滤波法。这种方法的基本思路是把地球化学背景面看成一个连续起伏变化的曲面,不同子区采用不同的异常下限,用窗口代子区,以小窗口为局部异常或噪音,以大窗口为局部背景来拟合地球化学背景变化趋势,用衬值来圈定异常。

1 子区中位数衬值滤波(SAMCF)法

1.1 基本原理

子区中位数衬值滤波(SAMCF)法是在对已知的各种化探数据处理方法进行系统研究的基础上,吸收了EDA技术^[2~5]和滤波技术^[6],开发出的一种处理1:20万区域化探扫描面数据的计算机数据处理方法。这一技术为实现区域化探异常的快速筛选与评价奠定了基础。目前已形成一套从数据处理到绘制单元素、多元素异常图的计算机系统,使数据处理工作在微机上实现了计算机化。

建立在经典统计学基础上的滤波技术采用低通滤波器(LPF)、高通滤波器(HPF)、框形滤波器(PFF)和柯尔莫哥夫—斯米尔诺夫滤波器(KSF)。最简单的低通滤波器在本质上相当于移动平均方法。高通滤波器的结果等同于移动平均的剩余图,做法是将原始数值减去由某一给定半径的窗口所得出的LPF面。框形滤波器是将中心图元的值与其周围有效半径为 r 处的方形图元环带作对比,如果中心图元的值超过环带的平均值加2倍标准差,则该图元可视为异常。而柯尔莫哥夫—斯米尔诺夫滤波器是将方形环带数据的累积频率分布与内部中心图元块段或图元数据的累积频率分布进行对比,检验准则是2组累积阶梯函数在所研究方向上的可观测差最大,但是其初始假设是在中心和环带内所取数值的总体无显著差异。在对上述4种滤波器的结果进行对比后,Howarth认为用 3×3 中心图元块段窗口的KSF最好,其次是PFF^[6]。

考虑到常规经典统计学方法都有严格的基本假设条件,而它们又常常不能描述化探数据的真实特性,Tukey引入稳健统计学,并提出了一种新的非常规数据处理方法——勘查数据分析(exploratory data analysis)技术^[7],近几年已引起人们重视^[2~5,8]。这种方法不要求任何假设条件,而是根据数据本身所固有的特性来识别异点,确定背景总体和异常总体,且不需要预先对原始数据作任何数学处理。EDA技术中的一套统计参数是数据集的中位数、2个极值点和上、下四分点。它利用下式来识别异点(即异常)

$$F_u = Q_u + 1.5 \times S_h$$

$$F_l = Q_l - 1.5 \times S_h$$

其中 F_U 和 F_L 为异点临界值， F_U 相当于异常（点）下限， F_L 相当于异常（点）上限， Q_U 为上四分点， Q_L 为下四分点， S_n 为内散度（类似于标准离差）。EDA方法中，25%的数据点被认为可能是“野”的，这些“野值”对中位数和节点（上、下四分点）不产生影响，而异点临界值仅由内散度确定。所以，异点临界值就不会受异点的影响。与稳健统计量相反，只要有一个极值点就会给常规经典统计学中的算术平均值和标准离差带来很大的影响，进而可以给建立在常规经典统计学上的由算术平均值和标准离差确定的异常下限带来很大的影响。因此，EDA技术识别异常的一个显著特点就是能够抵抗干扰和“野”数据（当然也可能是异常）的影响。Kürzl已将EDA方法编入计算机资料处理系统，处理Austria的区域水系沉积物资料^{〔3〕}。目前，EDA方法已为越来越多的用户所使用。

子区中位数衬值滤波法是在上述PFF，KSF和EDA技术的基础上开发形成的。它采用子区模拟背景场的空间变异，以EDA技术计算异常临界值，用衬值为基准度量场值的起伏变化。如图1所示，以 3×3 个单位格子（1 20万化探扫面中为 $4\text{km}^2/\text{格子}$ ）作为中心窗口（ W_c ），取以 $r = 4$ 个单位格子为半径的方形环带（ A_n ）中的数据作为中心窗口的背景。以中心窗口（ W_c ）数据子集的中位数（ M_{W_c} ）作为中心格子的趋势值。以方形环带（ A_n ）中的数据子集的中位数作为中心格子的背景值，并计算其负异常上限（ T_{NA} ，即 F_L ）和正异常下限（ TPA ，即 F_U ）。通过计算 M_{W_c} 与 TPA 和 T_{NA} 的比值 CF_P 和 CF_N 来求得中心格子的衬值，凡 $CF_P > 1$ 者为正异常，凡 $CF_N < 1$ 者为负异常。这样，以中心窗口为局部异常或噪声，以远离中心窗口的环带数据作背景，不将中心窗口及其近邻的数据参与背景运算，可以排除局部变化对背景趋势的影响，使背景更加“干净”。

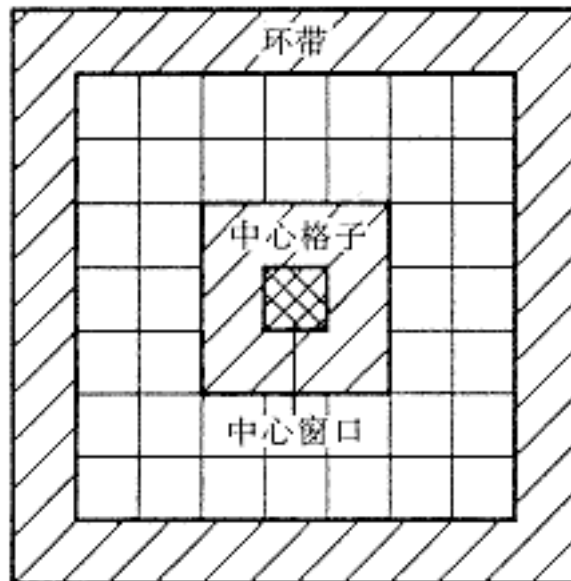


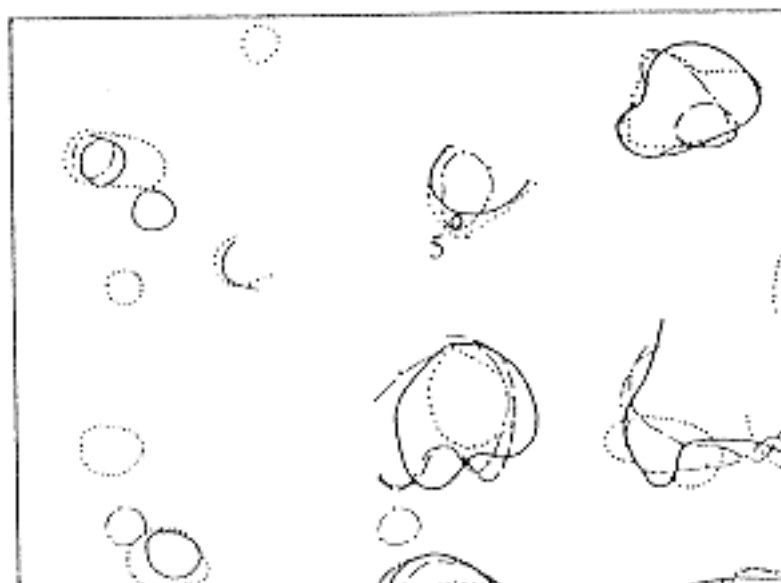
图1 子区中位数衬值滤波示意

SAMCF法用窗口代子区，以小窗口为局部异常或噪音，以大窗口为局部背景，逐点移动，使得每一点或子区的背景不同，把地球化学背景面看成一个连续起伏变化的曲面。即在SAMCF法的处理过程中，自动地把一个大区域划分成很多个小子区（相应于地球化学小子区），每个小子区均具有自己的地球化学背景值和异常下限值。这样，背景高，异常下限就高；背景低，异常下限就低。这就从理论上解决了低背景区和高背景区弱小异常的识别问题，克服了采用全区统一异常下限给异常识别带来的困难，实现了不同的地质地球化学子区采用不同的异常下限圈定异常的要求。此外，经SAMCF法处理，可以同时划分出正异常和负异常。

1.2 滤波窗口大小对比试验

对于滤波技术，选择窗口的合适与否直接关系到数据滤波结果的好坏。当然，滤波窗口的大小应根据数据处理的目的以及研究对象的地质背景来确定。在处理大面积1:20万区化扫面数据时，RESMA系统中采用 2×2 个单位格子(16km^2)为小窗口， 10×10 个单位格子(400km^2)为大窗口取得了满意的效果。史长义等利用RESMA系统处理青海省东昆仑地区1个点/ 16km^2 的区化扫面数据时，经对比研究最终选取 64km^2 为小窗口， 784km^2 为大窗口。在对比了4种不同大小的窗口的滤波处理结果后，Howarth认为以 3×3 个单位格子为小窗口，以 9×9 个单位格子为大窗口的滤波结果最好〔6〕。通过对Bambui地区水系沉积物中Zn数据的处理，Chork认为用 3×3 个单位格子(36km^2)为小窗口和 9×9 个单位格子(324km^2)为大窗口进行滤波，可以有效地识别出已知矿异常，特别是能够发现一些弱异常，可得到令人满意的结果〔9〕。

为了确定利用SAMCF法处理1:20万水系沉积物测量数据的最佳滤波窗口组合，我们利用内蒙某地区1:20万水系沉积物中Au的测量数据，进行了以 36km^2 为小窗口(W_s)， 196km^2 ， 324km^2 ， 484km^2 为大窗口(W_l)和以 100km^2 为小窗口， 484km^2 为大窗口的数据处理对比试验。从图2中可看出以下几点。



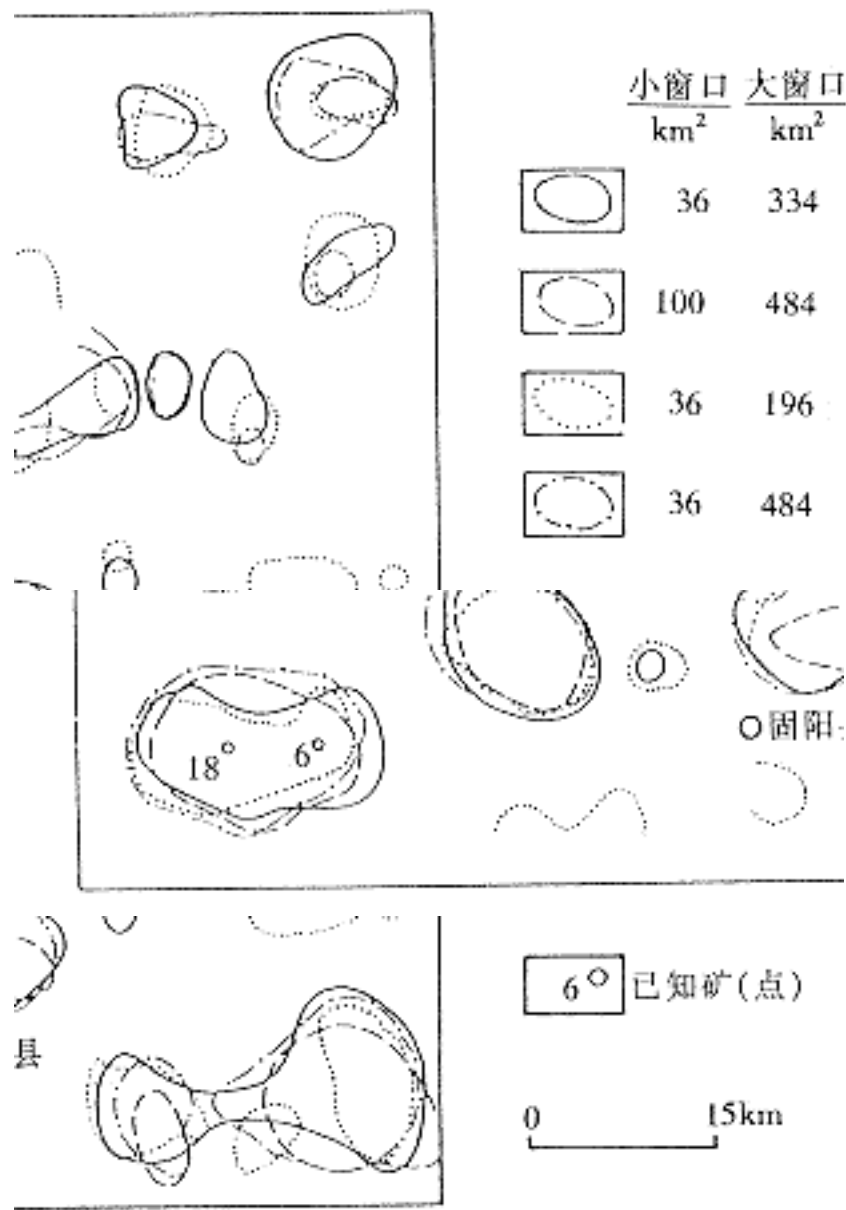


图2 金异常滤波窗口对比

- 1.以 $W_s = 36\text{km}^2$ ， $W_l = 324\text{km}^2$ 为条件的处理结果比较好，已知矿(点)上均有异常出现，其面积有大有小，异常形态也比较规整且与地质背景相关，其分布趋势受地质背景控制。
- 2.当 $W_s = 36\text{km}^2$ ， $W_l = 196\text{km}^2$ 时，虽然已知矿(点)上有异常出现，但从全区看，异常比较零乱。与上一种情况相比，有些异常变小或分解成几个小异常，同时还出现了一些零星小异常。这主要是由于大窗口面积偏小，而遭受局部噪音(异常)的影响所致。
- 3.当 $W_s = 36\text{km}^2$ ， $W_l = 484\text{km}^2$ 时，结果与第一种情况相似，已知矿上均有异常，但5号矿上异常强度和面积较第一种情况小。另外，从全区异常对比看，大异常和强异常的面积二者基本一致，但是小异常的面积则是第一种条件下的大。

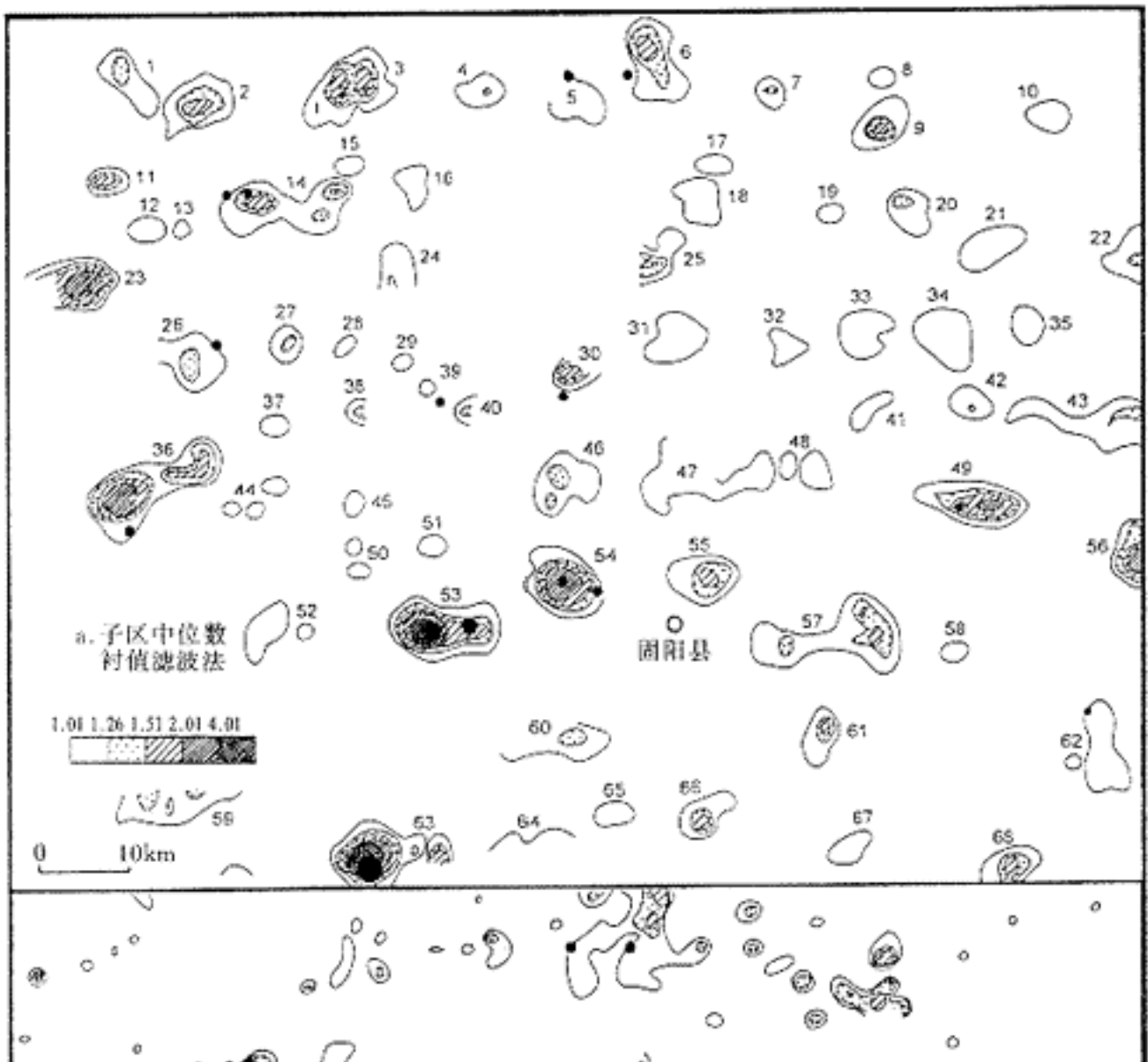
4. 当 $W_s = 100\text{km}^2$, $W_l = 484\text{km}^2$ 时, 6号和18号矿上出现了异常, 而5号矿上无异常, 与第一种情况相比, 异常明显减少, 异常面积也变小。这主要是由于小窗口面积过大, 而使得局部异常水平降低。

因此, 以 $W_s = 36\text{km}^2$, $W_l = 324\text{km}^2$, 可作为利用SAMCF法处理1—20万水系沉积物测量数据的最佳滤波窗口组合。

需要说明的是, 与其它数据处理方法一样, 空白区和图幅边界效应会对边缘数据的处理结果产生一些影响。

2 初步应用效果——SAMCF法之特点与弱小异常识别

图3为在内蒙古某地区25900 km^2 范围内水系沉积物测量的Au异常。其中图3a为用SAMCF法处理所得衬值异常, 图3b为常规法(以 $T = +1.65s$ 为异常下限)所圈出异常。



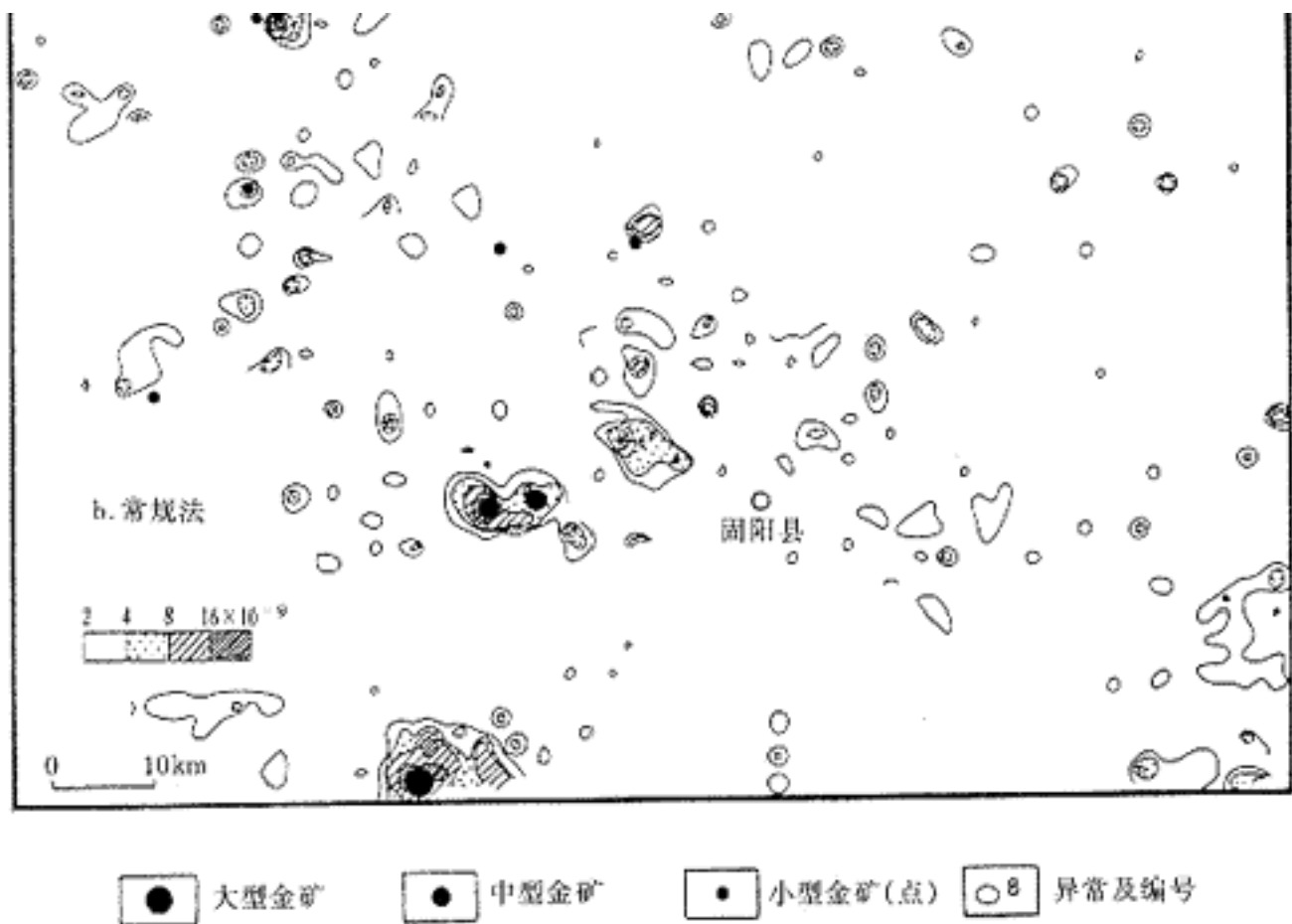


图3 内蒙某地区子区中位数衬值滤波法与常规法金异常

显然,2种方法所圈异常对已知矿(大、中、小型)均有较好显示。但是,从总体分布规律、异常规模和强度分析,二者存在明显差异。衬值异常图比较规整,异常规模也较大,在已知矿和非已知矿区均有强异常,总体分布规律性较好,反映异常受NW, NE向2组构造控制而呈带状分布,显示出地质背景控矿特征。常规异常图上,除已知矿异常为强异常以及未知异常的孤点异常外,基本为中—弱异常且面积较小,总体分布规律与衬度值异常相比较差。

结合本区背景分布(图4)分析,衬值法不仅在高背景区圈出了显著异常,而且在低背景区也发现了很多规模大、强度高的异常;相反,常规法对低背景区异常,特别是弱小异常反映较差。SAMCF法衬值异常与常规法异常相比,具有如下优势。

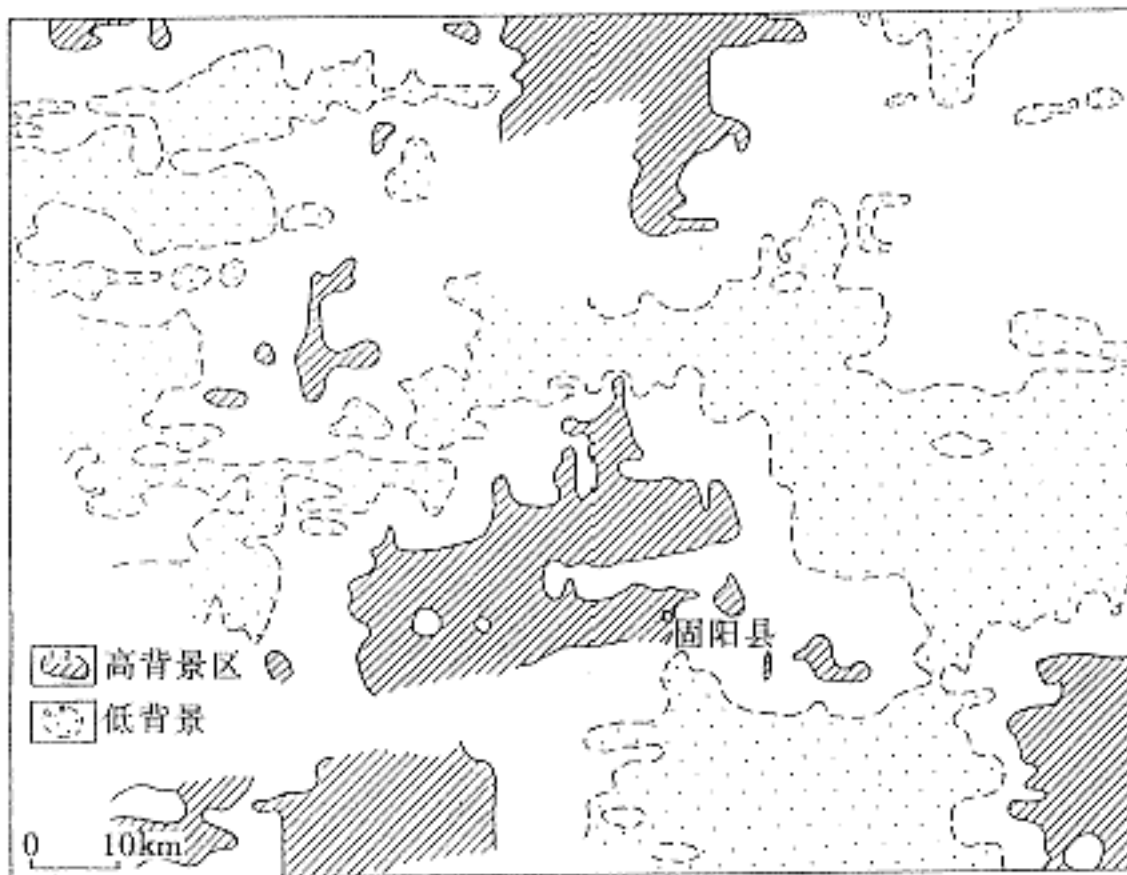


图4 内蒙某地区金背景分布

1.保持了强异常。图5a中衬值异常与常规异常基本一致，但衬值异常较常规异常更规整。图3中常规异常图上的强异常和矿异常，在衬值异常图上仍表现为强异常，虽多数在规模和强度上两者基本一致，如30，53，54，63，62，14，6，9号等异常，但有些异常面积也有所增加，如26，56号。

2.强化了弱、小异常。图5b中常规异常较弱，SAMCF使其得到强化。同样，在图3中常规异常图上的弱、小异常，通过子区中位数衬值滤波后不仅强度有所增大，而且异常面积也有增加，如23，3，53，36，66，43号等异常。



图5 子区中位数衬值滤波法与常规法金异常对比

1—第四系；2—第三系；3—砂岩、砂砾岩夹泥岩；4—混合岩、板岩、石英岩；5—斜长片麻岩、角闪片岩、大理岩夹石英岩；6—斜长角闪片岩、大理岩；7—花岗岩；8—石英闪长岩；9—花岗闪长岩；10—闪长岩；11—断层；12—地质界线；13—金矿

3.发现了新异常。图5c中常规法基本没圈出异常，而SAMCF却在此成矿有利部位上发现了很强的异常。另外，在图3的常规法异常图上基本无异常或只有单点弱异常，子区中位数衬值滤波后却发现了很强且面积较大的异常(图3a)，如1，2，25，49，43，67号等衬值异常，在常规法异常图上表现得很小或没显示。

4.衬值异常有利于不同异常的对比。采用局部背景来圈定异常可克服由地质背景不同而造成的地球化学背景变化给异常识别所带来的影响。

通过对图3a中3，36，49号异常的野外查证，在3个异常区均发现了金矿化，尤其是在后2个异常区中，不仅发现了一些高含量的金矿化点，而且还发现了若干条含金矿构造破碎带和蚀变破碎带。尤其是在36号金异常区，通过槽探揭露，目前已发现3条Au?Ag矿化带和4条Au?Ag矿体，已求出Au的D+E级储量约410kg和Ag的D+E级储量约27t，并有望实现大的找矿突破。这些异常的查证结果除说明利用SAMCF法处理数据在识别异常方面的有效性外，还证明该法在处理和识别弱小异常方面的优越性。

利用SAMCF法的处理结果，还建立了福建紫金山陆相火山岩型Cu?Au矿田的区域地质地球化学异常结构模式〔10〕。

子区中位数衬值滤波(SMCF)法的特点如下。

1.由于利用了EDA技术来计算异常下限，使得SAMCF法与一般常规数据处理方法不同，它不受原始数据分布律的制约，也不需要原始数据作任何预处理，而是直接利用原始数据来识别异常。

2.以衬值而非真实含量来圈定异常，便于不同元素地球化学场的对比和综合，以及在全国范围内的对比。

3.把地球化学背景面看成一个连续起伏变化的曲面，以小窗口为局部异常或噪音，以大窗口为局部背景。以窗口代子区，用子区来模拟背景场的空间变异，可更细致地逼近地球化学场的自然起伏。

4.以局部背景而非全区背景为对比标准计算衬值异常下限和划分异常，即不同子区(或背景区)采用不同的异常下限，这样既可突出强异常及压抑背景噪音的影响，又可强化高背景区弱小异常，特别是更有利于发现低背景区的弱小异常。

5.据SAMCF法可同时圈定出正异常和负异常。

6.这套方法简单、快速，在一般微机上即可运行。

3 认识与结论

1.SAMCF是一种建立在非常规统计学基础上的数据处理方法，不受原始数据分布律的制约，可直接利用原始数据来识别异常。把地球化学背景面看成一个连续起伏变化的曲面，以小窗口为局部异常或噪音，以大窗口为局部背景来模拟背景场的空间变异。采用动态背景即不同子区(或背景区)采用不同的异常下限来圈定衬值异常，不仅可以突出强异常及压抑背景噪音的影响，而且还可以强化高背景区弱小异常，特别是更有利于发现低背景区的弱小异常。

2.以中心窗口为局部异常或噪声,以远离中心窗口的环带数据作背景,不将中心窗口及其近邻的数据参与背景运算,可以排除局部变化对背景趋势的影响,使得背景更加“干净”。

3.由我国十几个不同类型不同景观区大中型Cu多金属矿1 20万水系沉积物测量数据以及上述内蒙某地区化探扫面数据的子区中位数衬值滤波处理结果证明,子区中位数衬值滤波法基本上没有漏掉已知矿异常的现象,相反,还强化了高背景区弱小异常,尤其是发现了新的高背景区和低背景区的弱小异常。比较好地解决了弱小异常的识别问题。事实表明,子区窗口的选择是比较合适的,这套数据处理方法基本上是成功的。它适用于全国不同地区不同景观区的1 20万水系沉积物测量数据的处理和异常筛选。

作者简介:史长义,男,高级工程师,1962年9月生,河北昌黎人。1983年毕业于 河北地质学院地质系,1990年毕业于中国地质大学(武汉)应用地球化学系,获理学硕士学位。现任国土资源部物化探研究所化探方法室副主任,中国地质学会勘查地球化学专业委员会秘书长。主要从事矿产勘查地球化学方法技术研究和勘查工作,在国内 外发表论文30多篇,专著(合)一部。曾获省部级科技成果二、三等奖各一次。

作者单位:国土资源部地球物理地球化学勘查研究所,廊坊 065000

参考文献

- 1 史长义.异常下限与异常识别之现状.国外地质勘探技术,1995(3):19~25
- 2 K ü rzl H.Exploratory data analysis:recent advances for the interpretation of geochemical data. J Geochem Explor,1988, 30: 309~322
- 3 K ü rzl H.Data analysis and geochemical mapping for the regional stream sediment survey of Austria. J Geochem Explor,1989, 32:349~352
- 4 史长义.化探数据解释推断的新方法——EDA技术.国外地质勘探技术,1991(1):38~41
- 5 史长义.勘查数据分析(EDA)技术的应用.地质与勘探,1993(11):52~58
- 6 Howarth K J.化探中的统计学和数据分析.北京:冶金工业出版社,1990
- 7 Tukey J W.Exploratory data analysis.Addison W.Reading MA, 1977.506p
- 8 Howarth R J.Statistical applications in geochemical prospecting: a survey of recent developments.J Geochem Explor,1984, 21: 41~61
- 9 Chork C Y.用勘查数据分析和稳健统计学对化探数据作空间滤波.见:第三届勘查地球化学学术讨论会论文选编, 1988.263~267
- 10 史长义,张金华,黄笑梅.福建紫金山陆相火山岩型铜金矿田区域地质地球化学异常结构模式.物探与化探,1996(3)

收稿日期:1998-06-29

