

# 以先验地质资料为基础的放射性参数统计方法

黄旭钊, 范正国, 梁月明, 徐昆

(中国国土资源航空物探遥感中心, 北京 100083)

**摘要:** 根据地层单元统计放射性含量的均值、方差、变异系数, 可以了解放射性元素的分布特征及迁移变化规律, 为放射性场的地质解释, 进行地质填图提供依据。在 MAPGIS 平台上, 采用 VC++ 开发语言, 研制了以先验地质资料为基础的航空伽马能谱参数的计算机自动统计软件, 它不仅能够统计每个地层单元的放射性元素的均值、方差、变异系数, 同时可根据变异系数的大小, 确定现有地质图中地层单元岩性定名的准确性或边界的准确性。

**关键词:** 均值; 方差; 变异系数; 地质填图; 相关系数; 谱系图; GIS

**中图分类号:** P631.4      **文献标识码:** A      **文章编号:** 1000-8918(2004)05-0439-04

以先验地质资料为基础, 对航空伽马能谱测量参数进行统计分析, 研究区内各种岩性和地层的放射性元素分布特点, 总结其分布规律, 作为区分岩性、岩性填图、修改现有地质图的依据。众所周知, 航空伽马能谱数据量较大, 传统的统计方法是将与测量比例尺相同的测区地质图的各主要地质单元的界线标绘在同比例尺的总量及钾、铀、钍含量的平面剖面图或平面等值线图上, 然后按照地层单元和一定的网格在同一位置上分别读取相应的总量和各元素含量值, 并计算其均值、方差、变异系数。可见工作量相当大, 难以保证统计数据的正确性。因此, 由计算机来完成上述工作, 是要亟待解决的问题。近年来, 作者研制了通过控制测线号和测点号来提取多边形内的数据, 然后由计算机自动统计放射性参数的方法, 但是查找多边形的测线号和测点号相对来说还是比较麻烦的, 并且以此来确定复杂的地层边界也不够准确。地理信息系统的发展与应用, 为解决上述问题提出了新的思路, 即通过判断一个点与任意不规则多边形的关系来查找落入任意地层边界内的放射性数据, 并以此来统计其特征参数。

## 1 放射性参数统计方法

### 1.1 基本统计方法

由于地理信息系统的空间数据具有空间拓扑关系, 提供了空间点、线、面之间的包含、覆盖、相离和相接等空间关系的描述。它集中存储了以下内容:

空间分布位置信息; 属性信息; 拓扑空间关系信息。地理要素之间的空间区位关系可抽象为点、线(或弧)、多边形(区域)之间的几何关系。文中所涉及的是点与多边形之间的空间位置关系(图 1), 即点在区域外部; 点在区域的边界上; 点在区域内, 可以记数和统计。当然, 图 1 所示的多边形区域是一种极为简单的情况, 实际上多边形区域还存在凸多边形和凹多边形的情况, 在判断点与多边形的空间关系时有所不同的。矢量结构的 GIS 能够利用其存储的信息, 通过计算每个点相对于多边形线段的位置, 进行点是否在一个多边形中的空间关系判断。

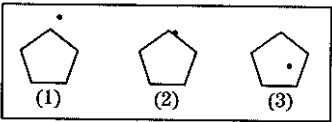


图 1 点与面拓扑关系示意

矢量化之后的地质图中的图元是具有上述多边形性质的空间数据, 而测区内每个测点都可以用一对坐标和观测值来表示, 所以, 观测数据与地质图元之间存在着上述拓扑关系。因此, 所要解决的问题归结为判断每个测量数据点属于哪个地层单元, 然后对同一地层单元的所有测点统计计算的问题。这里需要注意的是, 地质图坐标投影与放射性测量数据点坐标投影必须一致。统计单元的确定以现有地质图所提供的地层和岩体作为统计单元。均值、方差、变异系数的计算公式如下:

均值：
$$\bar{X} = \sum_{i=1}^n x_i / n, \tag{1}$$

方差：
$$S = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / n}, \tag{2}$$

变异系数：
$$C_v = S / \bar{X}. \tag{3}$$

式中  $n$  为某地层单元内放射性数据的总点数,  $x_i$  为第  $i$  号测点的放射性数据值(总道、钾、铀、钍含量中的一个), 方差和变异系数是统计单元内各测点的能谱含量距其平均值的相对偏离程度的参数。

1.2 再统计方法

将具有相似特征的同一套地层或岩体的数据归并成一组, 而特征相差甚远的该套地层或岩体归并成另一组, 将变异系数较大的数据剔除掉, 然后再按照新的分组做统计, 称为再统计。再统计的计算公式与基本统计方法一样, 所不同的是数据集合不同。与基本统计方法相比, 再统计要复杂得多, 因为它不是基本统计结果的简单合并, 而是根据需要重新分组之后用原始数据再做统计计算, 那么, 分组就成为相当重要的因素。我们所采用的分组方法是: 首先计算各地层岩性之间的相关性, 即根据基本统计结果的均值和方差, 计算其相关系数, 以表示研究对象之间的相关性, 然后进行动态聚类分析, 并将聚类结果用谱系图表示出来。谱系图的形成是通过各地层之间相互对比, 根据相似性指标的大小逐步并类, 直至最后并为一大类为止, 形成一个分类级别由小到大的分类系统。根据谱系图再结合测区实际情况, 就可以对地层单元进行分组了。相关系数的计算公式为:

$$r_{ij} = \frac{\sum_{i=1}^n (x_{ti} - \bar{x}_t)(x_{ji} - \bar{x}_j)}{\sqrt{\sum_{i=1}^n (x_{ti} - \bar{x}_t)^2 \sum_{i=1}^n (x_{ji} - \bar{x}_j)^2}}. \tag{4}$$

式中  $t, j = 1, 2, \dots, p$  为地质单元编号 ( $t \neq j$ );  $i = 1, 2, \dots, n$  为变量号, 如航空伽马能谱数据为例,  $i = 1, 2, 3, 4$  分别代表 Tc、K、U、Th。  $r$  的取值范围  $[-1, 1]$ ,  $|r|$  值越大, 表示 2 个地质单元之间关系越密切,  $|r|$  值越小, 表示 2 个地质单元之间关系越不密切。

再统计结果在实际应用中有着重要的作用, 使我们能够更加清晰地了解该地区地层或岩体放射性核素含量分布的普遍规律, 对于地质填图有着积极的指导作用。

2 放射性参数统计方法的实现

根据上述原理和思路, 我们在 MAPGIS 平台上, 采用 VC++ 开发语言, 研制了以先验地质资料

为基础的航空伽马能谱参数的计算机自动统计软件, 其流程如图 2 所示。

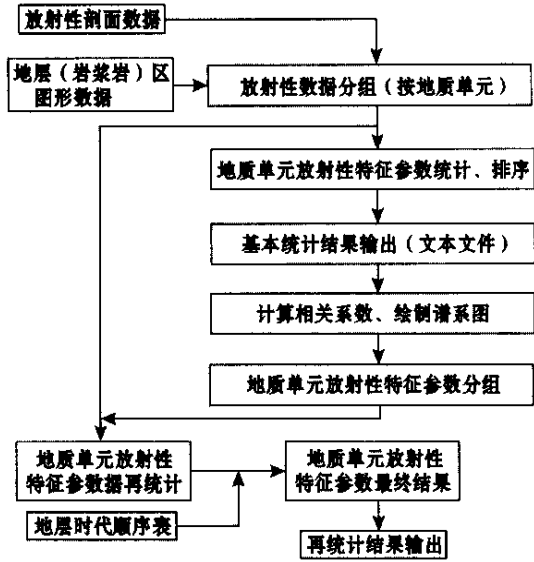


图 2 放射性参数统计流程

2.1 放射性数据分组

存储在数据库中的放射性数据或文本形式的剖面数据, 不能被直接使用, 要按照所输入的区图元重新组织数据, 该过程接受一个输入, 产生一个输出, 目的是按地质单元进行数据分组。分组方法就是前面所论述的通过判断每个测量数据点属于哪个地层单元, 然后将同一地层单元的所有测点分组存放。

2.2 放射性参数统计、排序及输出

按照公式(1)、(2)、(3)和数据分组结果统计 Tc、K、U、Th 的均值、方差和变异系数, 并输出文本文件。该文件是按照地质单元(区图元)编号自动排序的, 而地质图在成图时不一定将相同地层或岩体单元顺序编号, 所以相同名称的地层或岩体往往不能排在一起, 给进一步的比较、分析带来一定困难。该软件采用二叉树链表法按图元名称升序进行排序, 同时输出结果文件。二叉树是用来排序的一种有效方法, 它是树状结构的一种, 其定义如下: 由有限个节点所构成之集合, 此集合可以为空; 二叉树的根节点可分为 2 个子树, 称为左子树和右子树, 左子树和右子树亦叫二叉树(图 3)。二叉树中的每个

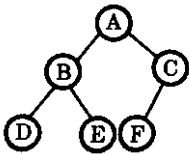


图 3 二叉树示意

节点代表一个地质单元, 构造二叉树时, 以地质单元

名称为链进行构造,再以中序遍历方式输出该二叉树,就得到按地质单元名称进行排序的放射性数据基本统计结果。由于不同测区地质单元数量是不同的,在具体实现该方法时,考虑到软件对不同测区的适应性,采用链表法动态构造二叉树。

### 2.3 相关系数计算与谱系图绘制

根据计算的每个地层单元的均值、方差,按照公式(4)计算其相关系数,然后进行动态聚类并绘制谱系图。表 1 为新疆某地部分地层单元的放射性数

表 1 新疆某地部分地层单元的放射性数据

| 统计点数 | K 均值 | U 均值 | Th 均值 | 图元编号 | 图元名称  |
|------|------|------|-------|------|-------|
| 215  | 1.89 | 1.67 | 5.34  | 508  | 乌苏群   |
| 145  | 1.99 | 2.45 | 4.37  | 480  | 梧桐窝子组 |
| 324  | 2.01 | 2.06 | 5.56  | 459  | 乌苏群   |
| 139  | 2.40 | 2.91 | 6.84  | 321  | 乌苏群   |
| 32   | 1.79 | 2.28 | 5.02  | 531  | 梧桐窝子组 |
| 976  | 2.09 | 2.15 | 4.27  | 712  | 梧桐窝子组 |
| 97   | 2.16 | 2.82 | 4.86  | 385  | 梧桐窝子组 |
| 28   | 2.18 | 3.18 | 5.13  | 445  | 梧桐窝子组 |
| 744  | 2.25 | 2.10 | 5.93  | 186  | 西域组   |

据,图 4 为根据该数据绘制的谱系图。从图中可以看出:480 和 712 号地质体相关系数最高,首先联结成组,其次联结 459 和 186 号地质体、385 和 445 号地质体,再次,则将 508 与 459 和 186 号地质体联结,依此类推……,最后在相关系数为 0 处,联结成一个大组。该谱系图为文本文件,可以直观地反映各地层单元之间的相关程度。

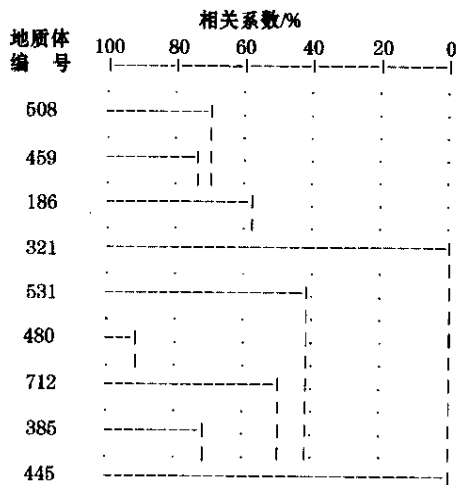


图 4 谱系

### 2.4 放射性特征参数分组

将统计结果和谱系图进行分析、对比,一般能够发现以下信息:分布范围较大的地层,有时变异系数较大,与地层中岩性分布不均匀有关;对于岩体来

说,有些岩体变异系数较大,说明这些单元可能包括 2 种或 2 种以上的岩性;也可能岩体界线划分不准确或岩性定名有误。在进行地质填图时,可以通过图元编号找到对应的图元,并认真分析,综合其它资料给出合理的解释。对于这样的数据在进行再统计时,原则上是应该予以剔除的,因此,在进行再统计之前,需对初次统计的结果按照一定的原则进行归并、分组、剔除。不同的地区分组的原则可能不同,可以说,初次统计结果、谱系图及研究区的具体情况分析是归并、分组的基本依据。

### 2.5 放射性数据再统计

初次统计结果反映了每一个个体的特征,为了了解研究地区的地层或岩体放射性核素含量分布的普遍规律,需要在归并之后,做进一步统计。再统计时要用到按地层单元分组的数据和放射性特征参数归并、分组的数据。

### 2.6 最终统计结果输出

再统计的结果是按照图元名称来排序的,为了更加便于解释人员使用,我们希望能够按照地层时代顺序进行重新排序。排序方法是:通过输入地层时代顺序表,采用二叉树链表法按地层时代升序进行排序。

统计排序之后的结果,通常需要输出,如果将其直接生成 MS Word 文档形式,将有利于报告编写人员使用。根据这种想法,利用微软组件对象模型 (COM) 技术,采用 OLE Automation 机制,将 MS Word 程序作为后端服务组件,在应用系统中开发一个 Word 程序控制器,利用该控制器,将排序结果直接输出到 Word 表格中。该控制器包含的主要 Word 组件接口有: Application、Documents、Document、Tables 以及 Selection。

## 3 结语

航空伽马能谱参数的计算机自动统计方法的实现,得益于计算机硬件技术和地理信息系统的发展。

(1)使得大数据量、多参数的统计计算在短时间内完成,极大地降低了劳动繁琐程度,保证了数据统计的正确性,提高了工作效率。

(2)空间图形数据与计算结果的对应关系,可以直接了解每个地层单元的放射性数据分布情况,提高了可视化程度,对于地质填图有着积极的作用。

(3)MS Word 文档形式的输出,为报告编写人员带来了极大的方便。

参考文献：

[1] 王懋基,张文斌,欧介甫,等. 航空物探解释方法及应用[M]. 北京:地质出版社,1992.

[2] 邬伦,刘瑜,张晶,等. 地理信息系统—原理、方法和应用[M]. 北京:科学出版社,2001.

[3] 黄国瑜,叶乃菁. 数据结构(C语言版)[M]. 北京:清华大学出版社,2001.

[4] Kruglinski D J, Wingo S, Shepherd G. C++ 6.0 技术内幕[M]. 希望图书创作室译. 北京:北京希望电子出版社,1999.

[5] 熊盛青,范正国. IBM4341 计算机航空伽马能谱解释处理系统及应用[J]. 物探化探计算技术,1993,15(1).

[6] 张文斌,熊盛青. 一种有用的解释参数——航空伽马能谱变异系数[J]. 物探与化探,1990,14(4).

THE RADIOACTIVE PARAMETER STATISTICAL METHOD BASED ON PRIOR GEOLOGICAL DATA

HUANG Xu-zhao, FAN Zheng-guo, LIANG Yue-ming, XU Kun  
(China Aero Geophysical Survey and Remote Sensing Center for Land and Resources, Beijing 100083, China)

**Abstract:** Based on the mean values, variances and variation coefficients obtained from the statistical radioactive contents of various stratigraphic units, we can get to know distribution characteristics and migration regularity of radioactive elements, which serve as the basis for geological interpretation of radioactive fields and geological mapping. On the MAPGIS platform, the authors used the VC++ development language to design a computer automatic statistical software for airborne gamma power spectrum parameters on the basis of prior geological data. Such a database can not only make statistics of the mean values, variances and variation coefficients of radioactive elements in every stratigraphic unit but also check the rock naming or boundary accuracy of the stratigraphic units in available geological maps according to the variation coefficients.

**Key words:** mean value; variance; variation coefficient; geological mapping; correlation coefficient; spectrogram; GIS

作者简介: 黄旭钊(1962—),女,高级工程师。1984年毕业于长春地质学院物探系,现从事地理信息系统在航空物探领域的应用研究工作,已发表论文数篇。



上接 438 页

difference magnitude diffraction wave over weak reflection amplitude,so,the diffraction wave exact imaging and fidelity amplitude processing is the basic of carbonate karst reservoir study. The paper,by the real seismic data high fidelity processing with post-stack high precision migration arithmetic and true amplitude in Tahe oil filed , summed up a set of post-stack seismic data high fidelity amplitude processing technology to adapt the area Reservoir prediction.

**Key words:** Tahe oil field; seismic exploration; data processing; fidelity amplitude

作者简介: 王者顺(1955—),男,高级工程师,石油大学资源与信息学院博士,中国石油化工集团新星公司、中国石油化工股份有限公司教授级高工,一直从事研究工作,公开发表论著多篇。

加入台湾华艺思博网(CEPS)中文电子期刊服务声明

《物探与化探》期刊,將自 2004 年 8 月起,加入台湾中文电子期刊服务—思博网(CEPS)。中文电子期刊服务—思博网是目前台湾地区最大的期刊全文数据库,收录台湾地区 300 余种核心期刊的全文,其访问地址为: www.ceps.com.tw。读者可以通过这一网址检索《物探与化探》期刊 2004 年起各期的全文,在一段时期后,还可以回溯检索 1997~2003 年历年的全文。

由于《物探与化探》期刊被 CEPS 收录,故凡向本刊投稿者,均视为其文稿刊登后可供思博网(CEPS)收录、转载并上网发行;其作者文章著作权使用费与稿酬本刊一次付清,不再另付其它报酬。

请各位继续支持本刊,谢谢!