

## GIS 中多模型预测自动综合的原理与算法

陈永良<sup>1</sup>, 刘大有<sup>1</sup>, 王世称<sup>2</sup>

(1. 吉林大学, 长春 130023; 2. 长春科技大学, 长春 130026)

[摘要]提出了在 GIS 矿产资源靶区预测研究中自动综合多模型统计结果的两种方案,概括了用两种方案进行自动综合的基本原理,研制了在 GIS 软件平台中用这两种方案实现多模型统计结果自动综合的计算机算法。

[关键词]矿产资源 找矿靶区 GIS 自动综合

[中图分类号]P628 [文献标识码]A [文章编号]10495-5331(2000)06-0055-03

### 技术方法

在用传统方法开展的矿产资源统计预测研究中,为了提高统计预测结果的可靠性,研究者们常常使用相同的观测数据应用多种统计预测模型来预测矿产资源靶区,由于不同的预测模型着眼点和理论依据不同,不同模型的统计结果之间常常存在一定的差异,在这种情况下,须要按照一定的原则对几种统计结果做进一步的“综合”,以便确定出优选找矿靶区的数量标准。基于定量类比的方法算得的统计结果,是描述各个地质统计单元之间相似程度的一种统计量,更确切地说,是表征统计单元与模型(由建模单元组成的集合)之间相似程度的数量指标,单元联系度越高单元与模型之间的相似性就越强,其成矿有利性也越大,因此,单元联系度也可以称做成矿有利度。多模型预测结果的综合,就是将不同预测模型计算的单元联系度“综合”成一种反映单元成矿有利程度的数量指标,以便实施预测。

在 GIS 矿产资源靶区预测研究中,为了增强预测结果的可靠性,也有必要同时使用多种统计模型预测矿产资源靶区,并对统计结果作进一步的综合。借助于一定的综合原则,在 GIS 软件平台内以“人机交互”的方式来实现几种预测结果的统计综合,是比较困难的。因此,设计出合理的自动综合原则,研制相应的计算机算法,并编制算法的计算机程序,在 GIS 软件平台中实现多模型统计预测结果的自动综合是非常必要的。

### 1 自动综合的基本原理

在 GIS 软件平台中,自动综合多种统计数学模

型的预测结果,是借助于 GIS 软件平台能够处理空间数据和属性数据的功能以及 GIS 平台具有“二次开发”环境的特点,将在传统方法中只能用“手工方式”来完成的多模型预测结果的综合问题,通过计算机算法在 GIS 平台中由计算机来自动完成。

根据几种统计模型的计算结果,可以在 GIS 中自动综合出一种找矿靶区圈定的数量标准。不同预测模型统计结果的自动综合可以采用以下两种原则:

1)“求同”原则 即确定出不同统计结果之间相一致的部分作为找矿靶区圈定的依据。将每一种预测模型的统计结果都看成是对地质统计单元所属的找矿靶区级别的一种判定,那么,可以用“求同”的原则确定最终的靶区分级结果。如果地质统计单元的几种预测结果都属于某一级找矿靶区,则该地质统计单元即被归属于这一级找矿靶区;如果地质统计单元用不同模型预测的结果属于不同级别的找矿靶区,则该地质统计单元应归属于上述几种模型预测结果中最低一级靶区。

2)“折中”原则 即计算出几种统计量平均意义下的统计量作为找矿靶区圈定的数量标准。将每一种模型的统计结果看成是对地质统计单元成矿有利程度的一种“排序”,有几种统计结果就有几种“排序”,我们可以以几种排序为基础,用“折中”的原则确定地质统计单元成矿有利程度的综合排序,可以根据具体情况确定折中算法。

上述两种综合原则各据优缺点,求同原则强调几种统计结果的一致性,即只有几种统计结果相一致时,才能作为找矿靶区圈定出来,因此,用该原则圈定的较高级别的找矿靶区相对较少,找矿目标比

[收稿日期]1999-08-20; [修定日期]1999-10-01; [责任编辑]张启芳。

[基金项目]吉林省科委资助项目“计算机综合信息矿产预测系统”研究成果的一部分。

较集中。由于该原则只注重靶区分级结果,忽略了同一级靶区中不同单元之间的细微差异,因此,不能进行同一级靶区中不同统计单元之间的成矿有利性对比。折中原则强调最终结果应尽可能地与几种统计结果相似,因此,最终的结果实质是几种统计结果的平均值,其最终圈定的找矿靶区可能不如如同原则那样目标集中,但是,用该原则综合后的数量指标与原来几种模型计算结果的意义相同,可以对比任意两个统计单元之间的成矿有利性差异。

## 2 自动综合的计算机算法

在 GIS 中自动综合多模型统计结果,需要合理地存储与地质统计单元和地质变量相关的信息。在矢量结构 GIS(以 MapInfo 为例)中,可以将地质统计单元存放在一个矢量图层(每一个地质统计单元都是一个多边形实体)中,而将地质变量存放在该图层的属性表中。假设有  $n$  个地质统计单元、 $m$  个地质变量,可以将统计单元属性表结构定义为  $n \times (m + 1)$  的浏览表,浏览表的第一列存放统计单元编号,第二列至第  $m + 1$  列存放  $m$  个地质变量的观测值。在 GIS 中自动生成地质统计单元分布图层、自动选择地质变量并对其赋值后(另有文章介绍),就可以用多种统计数学模型计算统计单元联系度。所得的联系度数据作为统计单元的新增属性附加在用于存放地质变量数据的属性段之后,则有几单元联系度计算结果就需要增加几个属性段。单元联系度数据产生后,就可以对其进行自动综合。

### 2.1 求同原则下的自动综合算法

该算法大致内容为:将找矿靶区的不同级别看作有序尺度变量<sup>[1,2]</sup>的不同状态,用离散型数据来表示。用阈值法将单元联系度数据离散化为靶区分级数据,则有几统计数学模型即可得到几种靶区分级数据。用逻辑组合运算将数种靶区分级数据综合成一种靶区分级数据,并用综合后的靶区分级数据圈定找矿靶区。

设有  $G$  种单元联系度数据,预测靶区共分  $K$  级。我们可以分别用数字  $1, 2, \dots, K, K + 1$  来表示一级靶区、二级靶区、...、 $K$  级靶区和非找矿靶区。那么,  $G$  种预测结果的自动综合由以下三步来完成:

步骤 1:把每一种单元联系度数据均离散化为靶区分级数据。

如果  $N$  个模型单元的第  $L$  ( $L = 1, 2, \dots, G$ ) 种单元联系度数据的变化范围为  $\min_L \sim \max_L$ , 令  $y_L = (\max_L - \min_L) / (K + 1)$ , 那么,把第  $L$  种单元联系度

数据离散化为靶区分级数据的  $K + 1$  个阈值分别为  $\min_L, \min_L + y_L, \min_L + 2 y_L, \dots, \min_L + K y_L$ 。

对于某个地质统计单元的第  $L$  种单元联系度数据,当其取值小于或等于  $\min_L$  时,该单元为非找矿靶区,取值转化为  $K + 1$ ;当其取值大于  $\min_L$  而小于或等于  $\min_L + y_L$  时,该单元为  $K$  级找矿靶区,取值转化为  $K$ ;当其取值大于  $\min_L + y_L$  而小于或等于  $\min_L + 2 y_L$  时,该单元为  $K - 1$  级找矿靶区,取值转化为  $K - 1$ ;...;当其取值大于  $\min_L + K y_L$  时,该单元为一级找矿靶区,取值转化为 1。将所有地质统计单元的每一种单元联系度数据均转化成靶区分级数据为止。

步骤 2:按照取大( )的原则将  $G$  种靶区分级数据综合成最终的靶区分级数据。假设某一个地质统计单元的  $G$  个靶区分级数据分别为  $B_1, B_2, \dots, B_G$ , 那么,综合后的靶区分级数据为  $B_1, B_2, \dots, B_G$ , 即综合后的靶区分级数据应该是  $B_1, B_2, \dots, B_G$  中数值最大者(靶区分级数据越大,靶区级别越低)。将所有的地质统计单元的靶区分级数据均综合完为止。

步骤 3:根据最终的靶区分级数据对统计单元图层做专题渲染,生成靶区预测图。

### 2.2 折中原则下的自动综合算法

该算法的主要内容为:由于不同统计模型的计算结果数量级不同,必须将每一种统计结果正规化,使他们的数值变化范围均在  $0 \sim 1$  之间,再进行综合。折中法综合可以用几种正规化的单元联系度的简单算术平均值作为综合后的单元联系度。最后,用综合的单元联系度生成靶区预测图。

设有  $G$  种单元联系度数据,我们可以通过以下步骤实现方案二的自动综合。

步骤 1:将  $G$  种统计单元联系度数据作正规化处理,计算公式为:

$$y_l = \frac{y_l - \min_{1 \leq i \leq n} \{y_i\}}{\max_{1 \leq i \leq n} \{y_i\} - \min_{1 \leq i \leq n} \{y_i\}} \quad (l = 1, 2, \dots, n)$$

式中,  $y_l$  为正规化后的单元联系度数据,  $y_l$  为原始的单元联系度数据。正规化的目的是使所有单元联系度数据具有相同的数量级,数值变化范围均在  $0 \sim 1$  之间。

具体做法是:在统计单元属性表中,找到第  $j$  ( $j = 1, 2, \dots, G$ ) 种单元联系度数据所在的列(Column),再找到该列中数据的最大值和最小值,将该列中所有数据均减去最小值,再除以最大值与最小

值之差,即得该种单元联系度的正规化数据。重复上述操作直到所有单元联系度数据均被正规化为止。

步骤 2:求  $G$  种正规化的单元联系度的平均值,作为“折中”的单元联系度数值。设第  $m+1, m+2, \dots, m+G$  列(前  $m$  列为地质变量数据)为正规化的单元联系度数据,第  $i(i=1, 2, \dots, n)$  个统计单元的“折中”的联系度为

$$y_i = \frac{1}{G_{j=m+1}^{m+G}} y_{ij} \quad (i = 1, 2, \dots, n)$$

式中,  $y_i$  为第  $i$  个统计单元的“折中”的单元联系度,  $y_{ij}$  为正规化单元联系度。重复  $n$  次步骤 2 中的操作,可以计算出所有统计单元的“折中”联系度。

步骤 3:将折中联系度数据转化成靶区分级数据,算法与方案一的步骤 1 相同(该步可以省略)。

步骤 4:用折中的单元联系度数据或靶区分级数据对统计单元图层进行专题渲染,生成靶区预测图。

### 3 结论及讨论

多模型统计预测结果的自动综合是 GIS 矿产资源预测研究必须解决的问题之一。本文主要讨论可用于自动综合的两个基本原则及其计算机算法。基于上述的两个原则,笔者已经在 MapInfo 软件平台中

用 MapBasic 语言开发出了相应的计算机程序,并嵌入计算机综合信息矿产预测系统中。使用该系统进行矿产资源靶区预测时,用户可以使用特征分析<sup>[3]</sup>、秩特征分析<sup>[4]</sup>、序分析<sup>[5]</sup>、独立权法<sup>[6]</sup>、定向独立权法<sup>[7]</sup>和秩相关权法<sup>[8]</sup>进行地质统计单元单元联系度计算,然后将上述 6 种单元联系度计算结果自动综合成求同的靶区分级数据或折中的单元联系度数据。

#### [参考文献]

- [1] 周光亚,夏立显. 非定量数据分析及其应用[M]. 北京:科学出版社,1993.
- [2] 周光亚,董文泉,夏立显. 数量化理论及其应用[M]. 长春:吉林人民出版社,1979.
- [3] Botbol J M. An application of characteristic analysis to mineral exploration[J]. Proc gth Int, 1971, (2):92~99.
- [4] 陈永良,李学斌,许亚光. 秩特征分析方法在矿产资源预测中的应用[J]. 地质与勘探,1998,(2):37~42.
- [5] 成秋明,范继璋. 序分析方法及其在地质学中的应用[J]. 长春地质学院学报,矿产资源评价专集,1989:64~70.
- [6] 陈永良,杜德文,刘安洲. 独立权法及其在矿产预测中的应用[J]. 长春地质学院学报,第三十届国际地质大会专集,1996:117~121.
- [7] 陈永良,刘大有,王世称. 定向独立权模型及其在矿产资源预测中的应用[J]. 地球科学,1999,(4).
- [8] 陈永良,许亚光,于大伟. 秩相关权法及其在矿产预测中的应用[J]. 长春地质学院学报,1997,(1):92~97.

## PRINCIPLE AND ALGORITHM FOR AUTOMATIC INTEGRATING MULTI-MODEL RESULTS IN GIS

CHEN Yong-liang, LIU Da-you, WANG Shi-cheng

**Abstract:** The two plans and basic principles for automatic integrating multi-model statistical results of mineral resource target prediction in GIS are studied. The corresponding algorithms of the two plans in GIS software platform are introduced.

**Key words:** mineral resources, ore-hunting target, GIS, automatic integrating

#### 第一作者简介:

陈永良(1965年-),男,1992年毕业于长春地质学院数学地质专业,获硕士学位。1998年获工学博士学位。现任长春科技大学综合信息矿产预测研究所副研究员,吉林大学计算机科学系博士后。主要从事综合信息矿产预测、数学地质方法、地理信息系统(GIS)及计算机应用等方面的研究工作。

通讯地址:吉林省长春市解放大路83号 吉林大学计算机科学系 邮政编码:130023

