

# 基于 MapGis 的区域 Cu、Pb、Zn 成矿预测

张鹏涛, 陈 艳

(中国地质大学资源学院遥感地质与数学地质研究所 湖北 武汉 430074)

**【摘要】:** 本文主要围绕云南三江南段局部 1:20 万区域土壤化探数据, 利用聚类分析中的 DBSCAN 方法和 MapGis 中的 DTM 分析对 Cu、Pb、Zn 元素含量数据进行处理, 把利用两种方法单独获得的预测区域进行区与区空间相交分析, 并结合已知矿点和断裂构造得出最终的成矿预测区域。

**【关键词】:** 成矿预测 聚类分析 DTM 分析

## 1、数据介绍

本文用到的数据是由云南省地调局提供的云南三江南段局部 1:20 万区域水系沉积物化探元素含量, 该区域位于东经 100 度 30 分至 101 度、北纬 23 度至 24 度之间, 共包含 1290 个采样点数据, 35 种金属和非金属元素数据的含量。除此之外还要用到该地区其他有关的地质矿产资料, 如已知矿点和断裂构造等。

## 2、数据预处理

### 2.1 剔除异常值

从化探元素数据的获得过程看, 由于样本在同一地点采集的差异性, 再加上元素含量化学分析的误差等因素, 最终导致某些样本中某一元素的含量远远大于正常值, 这一结果对于数据分析的影响极大。为了消除此影响, 首先要对整理好的原始数据进行剔除异常值处理。这里采用这样的方式, 把某一样本某一元素含量值与所有样本中同一元素的均值对比, 若含量值大于均值的十倍, 则用十倍均值替换该含量值。采用这种方式处理的好处是既消除了明显的噪声影响, 又尽可能的保证了数据的真实可靠性。

### 2.2 数据规格化

为了克服数据单位不同在接下来数据分析中的影响, 首先要对原始数据进行数据变换处理。所谓数据变换, 就是将原始数据矩阵中的每个元素, 按照某种特定的运算, 把它变为一个新值。这里进行数据预处理采用规格化变换方法。规格化变换又称为极差正规比变换。它以每一个原始数据减去该变量中的最小值, 再除以极差, 即得到规格化后的数据。经过规格化变换后, 每列的最大数据变为 1, 最小数据变为 0, 其余数据取值在 0~1 之间。数据规格化的计算公式如下:

$$X'_{ij} = \frac{X_{ij} - \min\{X_{ij}\}}{\max\{X_{ij}\} - \min\{X_{ij}\}} \quad (i=1,2,\dots,n; j=1,2,\dots,p)$$

## 3、聚类分析

经过数据预处理后的数据是一个 1290×37 (其中 35 列为元素含量, 2 列为平面 x、y 坐标) 的数据矩阵, 其中最大值为 1, 最小值为 0。根据研究区域已知的地质矿产资料可知, 该区 Cu 元素含量很高, 已知铜矿点就有 37 处, 因此, 选取 Cu 为主要研究对象有很好的代表性。另外选取 Sb、Zn 为辅助研究对象, 这是由于该区 Cu、Sb、Zn 伴生成矿现象比较明显。基于以上分析, 从 1290×37 矩阵中抽取出一个 1290×5 的子矩阵用于本文数据分析 (其中 3 列为 Cu、Sb、Zn 元素含量, 2 列为平面 x、y 坐标)。

之所以采用聚类分析方法处理数据, 是因为经过该方法处理后, 可以把 1290 个采样点分成有限个不同的区域类型, 这些不同的区域类型即为所分析元素含量分布的不同区域。在这里采用聚类分析中的基于密度的聚类算法 DBSCAN 处理所抽取出的 1290×5 子矩阵, 具体的 DBSCAN 算法思想不再详述。

先在 MapGis 中根据 1290×2 矩阵 (其中 2 列为平面 x、y 坐标) 做出 1290 个样本点分布的点图层。然后根据基于密度的聚类算法 DBSCAN 的思想, 在 VC++6.0 中编程加以实现, 利用实现好的算法针对 1290×3 矩阵 (其中 3 列为 Cu、Sb、Zn 元素含

量) 进行聚类分析。在聚类计算时, 为了接下来分析方便, 控制参数  $\epsilon$  和 MinPts, 使 1290 个样本被分为 5 类。把通过聚类算法获得的各样本点的类别号作为一系列属性输入样本点分布点图层, 根据此属性把 1290 个样本点分为五类, 把分类后的化探点图层与研究区已知矿点图层叠加显示。

为了找出与已知矿点最接近的可能成矿区域分类, 尝试使用以下方法: 首先为五个分类各设一个类别积分变量, 分别为  $m_1, m_2, m_3, m_4, m_5$ , 初始值都是 0, 然后找出与各个已知矿点的中心坐标最近的分类, 假如某一矿点周围最近的采样点为五类中的第 1 类, 那么第 1 类的积分就加 1 (若为小型矿床, 则加 2), 即  $m_1 = m_1 + 1$  (或 2)。经过以上处理, 各个变量的最终值如下:

$$m_1=5; m_2=22; m_3=1; m_4=2; m_5=9。$$

由于第 2 类的积分最高, 所以第 2 类样本点所覆盖的区域就是最有可能的成矿区域。除去其它分类样本点, 只剩下第 2 类的样本点, 然后对所有的第 2 类样本点进行点缓冲区分析, 缓冲区半径为两点间距的  $\frac{\sqrt{2}}{2}$  倍, 选取这一半径是为了使区域内连线可以组成一个最小的正方形的四个采样点的缓冲区叠加在一起刚好可以填满它们所围成的正方形区域。

在矩阵中加入一列, 该列数据为 Cu、Sb、Zn 元素含量的均值。在 MapGis 中对 1290×3 的矩阵 (其中 2 列为平面 x、y 坐标, 另一列为 Cu、Sb、Zn 元素含量的均值) 进行 DTM 分析, 生成的 Cu、Sb、Zn 元素含量均值等值线图。在这里, 以均值与标准离差之和作为地球化学的异常下限值。经过计算, 新加列 (Cu、Sb、Zn 元素含量的均值) 的均值与标准离差之和为 0.86。提取等值线图中高于异常下限的区域, 然后与聚类分析所得的正方形样本点缓冲区进行区与区相交分析, 得出最有利于 Cu、Sb、Zn 的成矿预测区域如图一:

## 4、结果分析

从图一中挑选出面积最大的四个区域用 A、B、C、D 标示, 并把上图与该区已知矿点、断裂构造叠加显示如图二:

从图二可以看出, 在 A、C 区内都有已知铜矿点分布, 特别是 C 区, 有十几处已知矿点存在, 这正与本文的预测相吻合, 至于 B、D 区内没有已知矿点分布, 可能有两方面原因: 一是本文的分析讨论不够全面, 导致预测不够准确, 另一个原因可能是该区的矿产勘查工作还不够充分, 有待进一步完善。就以上四个成矿预测区域与区域断裂构造的关系看, B 和 C 为一个类型, A 和 D 为一个类型。由已知断裂构造可以看出, A 和 D 区域内仅有南北向的断裂, 且断裂构造不够复杂, 而 B 和 C 区域内不仅有南北向断裂, 还有近似东西向断裂, 且断裂构造相当复杂。当然, 这只是从已知断裂构造加以分析的, 并不排除 A 和 D 区内存在未发现东西向断裂构造的可能。因此, 四个预测区成矿的可能性都是存在的。

## 5、结论总结

本文通过对云南三江南段局部 1:20 万区域土壤化探数据中 Cu、Pb、Zn 数据的分析, 得出了该区内 Cu、Pb、Zn 的成矿预测区域。把预测结果与已知矿点对比, 有一定的吻合性。通过分析, 本人就该研究区域获得了以下认识: (下转第 144 页)

参数变量、变量个数、条件、被选取的概率、产生式后件,并将这些因子存入结构体 ProductionJX。对于每条规则都需要执行此项操作。

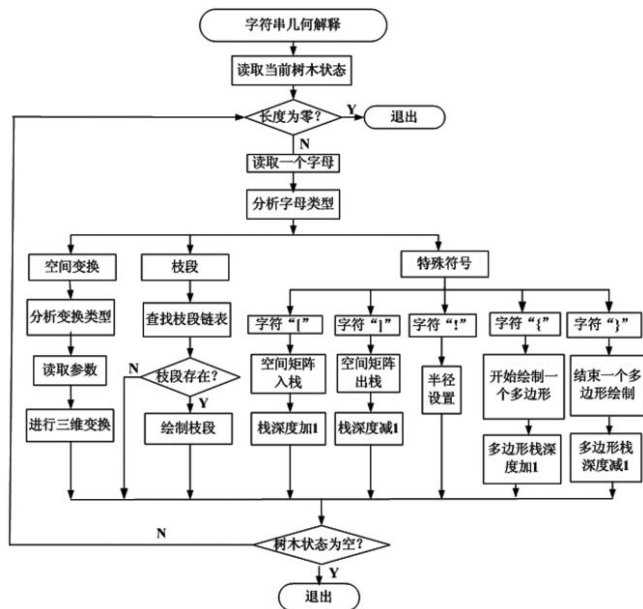


图1 芒果树花期的三维可视化编译流程图

```

struct ProductionJX // 产生式结构体.
{
    char symbol; // 符号.
    float probability; // 产生式被选取的概率.
    char production[1024]; // 产生式后件.
    char condition[128]; // 产生式被选取的条件.
    int numparam; // 参数的个数.
    char param[MAXPARAM][256]; // 参数.
};
  
```

(2) 对公理进行解析操作:①无条件约束下,根据符号及变量个数,选取相应的产生式,并给对应的变量赋值,赋值好的后件存入字符缓冲区。②条件约束下,对公理的操作主要是根据条件中的变量进行算术或者布尔运算后,如果条件为真,再根据符号及变量个数,选取相应的产生式,并给对应的变量赋值,赋值后的产生式后件存入字符缓冲区 ocb。

(3) 按照步骤 (2), 将字符缓冲区 ocb 中的内容执行解析操作,直至迭代完成为止。

#### 4.3 显示结果

按照本文 3.2 章节的规则,改变花序生长信号 m 值,实现了芒果树从初花期、盛花期、末花期至终花期四个阶段的开花生长过程的模拟,如图 2 所示。

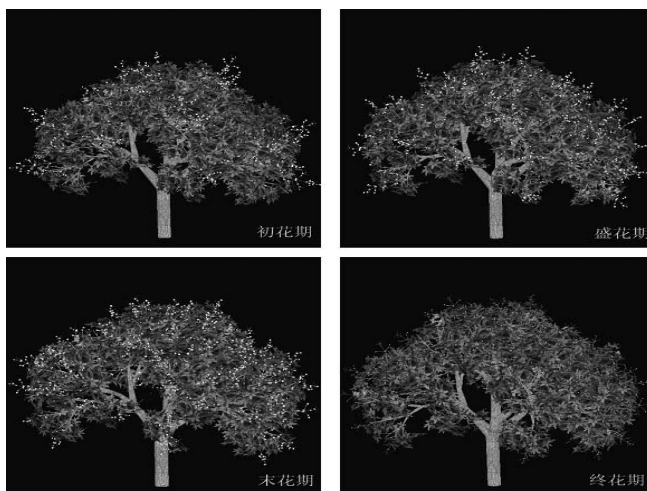


图2 成年芒果树的花芽分化模型及三维生长可视化

#### 5、结论

L 系统是经典的基于植物学规则的树木建模方法之一。在模拟花序方面真实感较强。本文以开花时期的芒果树为研究对象,探讨了 L 系统理论在果树花期时期生长模拟中的具体应用,实现了芒果树花期模型生长的三维可视化,取得较为满意的建模效果。

#### 参考文献:

- 1.赵星,Philippe de Reffye,熊范纶等.基于双尺度自动机模型的植物花序模拟[J].计算机学报,2002,vol.25:116~124
- 2.Prusinkiewicz.P. and Lindenmayer.A. The algorithmic beauty of plants [M]. Springer-Verlag, New York, 1990
- 3.Prusinkiewicz P, Lindenmayer A, Hanan J. Developmental models of herbaceous plants for computer imagery purpose [J]. Computer Graphics, 1998, 22(4): 141~150
- 4.Frijters D. Mechanisms of developmental integration of *Aster novae angliae* L. And *Hieracium murorum* L. Annals of Botany. 1978, 42: 561~575
- 5.Janssen J M, Lindenmayer A. Models for control of branch positions and flowering sequences of capitula in *Mycelis muralis* (L.) Dumont( Compositae). New Phytologist, 1987, 105: 191~220

(上接第 129 页)

- 1 研究区内的 A、B、C、D 四个区域为 Cu、Pb、Zn 找矿远景区域;
- 2 断裂构造为 B 和 D 区内 Cu、Pb、Zn 成矿的主要控矿因素;

#### 参考文献:

- 1.矫东风,吕新彪,基于 GIS 空间分析的成矿预测,地质找矿论丛,Vol.18 No.4 Dec.2003.
- 2.刘治国,池顺都,周顺平,成矿预测中应用 GIS 的主要步骤,地质找矿论丛,Vol.17 No.2 Jun.2002.
- 3.杨岳清,杨建民,徐德才,杨建华,云南澜沧江南段火山岩演化及其铜多金属矿床的成矿特点,矿床地质,Vol.25 No.4 Aug.2006.
- 4.鲍光淑,姚锐,戴塔根,谢文兵,地学信息系统在矿产预测中的应用,中南工业大学学报,Vol.33 No.5 Oct.2002.
- 5.吴堃虹,利用 GIS 编制矿产预测图,地质与勘探,Vol.36 No.3 May. 2000.

