

# 基于聚类分析的滑坡灾害危险性区划研究

桂 蕾,殷坤龙,王佳佳

(中国地质大学(武汉),武汉 430074)

**摘要:** 滑坡灾害危险性区划研究在城市规划决策方面具有重要的现实意义。聚类分析以统计学的形式将具有相似特征的数据进行归类,能够实现滑坡灾害危险性空间分布情况的定量评价。根据湖北省巴东县滑坡灾害统计资料,选择具有代表性的滑坡灾害影响因素作为危险性区划评价指标,采用熵权法和层次分析法相结合,综合评判各指标权重。并在此基础上,以 MapGIS 为操作平台,以 C#语言编程实现了快速聚类算法,对研究区 86216 个单元进行了滑坡灾害属性分类及危险性等级自动识别,预测结果较好。本研究将综合权重评判方法与聚类模型结合,同时克服了聚类结果不能自动排序的困难,对处理大批量,多属性数据具有一定的创新性和实用价值。

**关键词:** 滑坡;危险性区划;聚类分析;熵权法

**中图分类号:** P642. 22

**文献标识码:** A

**文章编号:** 1000-3665(2013)01-0100-06

滑坡灾害是一种常见的地质灾害,它不仅给人类生命安全带来威胁,而且对财产、环境、资源等具有破坏性。随着城市建设步伐的加快,土地需求日益增加,滑坡灾害危险性评价与区划在投资决策中具有重要意义。

滑坡灾害危险性区划研究经历着从定性到量的发展过程,目前,信息量模型、专家打分模型、破坏概率模型、神经网络模型等,在滑坡灾害危险性区划方面都有较好的应用<sup>[1]</sup>。但是,这些分析方法往往只拘泥于算法,忽视了样本数据的物理意义,预测结果有时难以解释。聚类分析是建立在具有实际物理意义的评判标准的基础上,以统计学的形式将具有相似特征的数据进行归类,预测算法含义清晰,能够对研究对象进行定量评价<sup>[2]</sup>。目前,聚类分析方法的应用较为广泛,如徐林荣等<sup>[3]</sup>利用模糊聚类方法确定了加筋土陡边坡状态评价量化指标;蔡美锋等<sup>[4]</sup>利用基于遗传算法的模糊 C 均值聚类方法对岩体结构面产状分类;蔡国军等<sup>[5]</sup>、张鹏等<sup>[6]</sup>利用聚类的方法进行了土层归并分类;李顺群等<sup>[7]</sup>由谱系聚类原理建立了粘土微结构几何属性的聚类方法,说明了成层地基的几何各向异性;

唐小松等<sup>[8]</sup>利用认知聚类分区方法进行了边坡可靠度分析;等。

湖北省巴东县是三峡库区受地质灾害影响严重的移民工程重点县之一,滑坡灾害严重制约了城市建设。本文在前人研究的基础上,以聚类分析方法为计算核心,结合层次分析法<sup>[9]</sup>和熵权法<sup>[10]</sup>综合评判样本属性指标的权重,以带权重的物理指标作为源数据,从空间分布的角度,开展巴东新城区及周边区域的滑坡危险性区划。同时,将分类等级目标层预设聚类算法中,解决了在前人研究中遇到的聚类结果不能自动排序的困难,研究思路在大批量数据处理上具有一定的创新性和实用价值。

## 1 研究区数据预处理

根据巴东县新城区已有滑坡的地质环境相关资料,统计分析选择具有代表性的滑坡灾害影响因素作为本次预测模型的数据源,主要包括岩组类型、地形坡度、斜坡结构、地质构造、水文地质条件及人类工程活动类型,其中水的作用又分为河流冲刷作用、地表汇水区分布情况和冲沟发育情况<sup>[11]</sup>。根据统计结果,按照滑坡发生频率越高、该影响因素类型的权重越高的原则对各影响因素进行分级,结果见表 1。

研究区以三叠系下统嘉陵江组( $T_{1j}^3$ )为基座,以三叠系中统巴东组第一段( $T_2b^1$ )为底层,以巴东组第二、三段( $T_2b^2$ 、 $T_2b^3$ )为主体,第四段( $T_2b^4$ )局部残留。该区地层岩性组合表现出软弱相间的特征,其中巴东组第二、三段最易发生滑坡,其原因为该段地层中存在泥岩、泥质粉砂岩及泥灰岩等软弱夹层,局部岩层还含

**收稿日期:** 2012-02-29; **修订日期:** 2012-04-06

**基金项目:** 国家自然科学基金项目(41002103);国家自然科学基金青年科学基金项目(41101515);三峡库区三期地质灾害防治重大科学研究项目(SXKY3-2-3-200704)

**作者简介:** 桂蕾(1986-),女,博士研究生,主要从事地质灾害预测预报及其风险评价方面的工作。

E-mail: ling200512@126.com

有亲水性较强的蒙脱石。而在嘉陵江组地层中滑坡难以发育,因为其岩性为中厚层白云岩,结构完整且坚硬。研究区地形坡度变化较大,从统计结果看出,滑坡及变形体多发生在斜坡中前部,坡度为  $10^{\circ} \sim 30^{\circ}$  时最利于滑坡发育。研究区近东西向的构造带形成了该区主要以顺向坡型为主的斜坡结构类型,加之研究区地层中存在软弱夹层,遇水易软化,斜坡结构类型为顺向坡的地区较反向坡的地区更利于滑坡的发育和发生。同时,研究区断层、节理、裂隙发育,根据统计结果,预测点距离断层作用位置 180m 以内时,对滑坡发生频率影响较大,距离越远,相关性越小。

表 1 滑坡频率统计及各影响因素分级表  
Table 1 Statistics of landslide frequency and factor classification

| 一级指标       | 二级指标                         | 滑坡频率   | 分级 |
|------------|------------------------------|--------|----|
| 岩组<br>类型   | $T_{1j}^3$                   | 0      | 1  |
|            | $T_2b^1$                     | 0.0345 | 2  |
|            | $T_2b^4$                     | 0.0345 | 2  |
|            | $T_2b^2$                     | 0.3500 | 3  |
|            | $T_2b^3$                     | 0.5863 | 4  |
|            | $<10^{\circ}$                | 0      | 1  |
| 地形<br>坡度   | $40^{\circ} \sim 50^{\circ}$ | 0.0476 | 2  |
|            | $>50^{\circ}$                | 0.0476 | 2  |
|            | $30^{\circ} \sim 40^{\circ}$ | 0.1429 | 3  |
|            | $10^{\circ} \sim 20^{\circ}$ | 0.2381 | 4  |
|            | $20^{\circ} \sim 30^{\circ}$ | 0.5238 | 5  |
| 斜坡<br>结构   | 反向坡                          | 0.0300 | 1  |
|            | 斜向坡                          | 0.3500 | 2  |
|            | 顺向坡                          | 0.6207 | 3  |
| 与断层<br>的关系 | $>300\text{m}$               | 0      | 1  |
|            | $180 \sim 300\text{m}$       | 0.0968 | 2  |
|            | $0 \sim 180\text{m}$         | 0.9032 | 3  |
|            | 无                            | 0      | 1  |
| 水的<br>作用   | 库水、汇水区、冲沟之一作用                | 0.3500 | 2  |
|            | 库水、汇水区、冲沟之二作用                | 0.5000 | 3  |
|            | 库水、汇水区、冲沟联合作用                | 0.1500 | 4  |
| 人类工<br>程活动 | 无                            | 0      | 1  |
|            | 公路或建筑                        | 0.4500 | 2  |
|            | 公路与建筑                        | 0.5500 | 3  |

水对滑坡的发生通常起到促进作用,本文从河流冲刷作用、地表汇水区分布情况及冲沟发育情况三个方面考虑水的作用对研究区滑坡灾害的影响。研究区河流冲刷作用主要表现在三峡库区库水位周期性涨落,将对岸坡进行反复冲刷,也将对斜坡地下水位产生周期性的影响,不利于滑坡的稳定;考虑地表汇水区的分布情况主要是因为汇水区分布情况可以反映降雨对坡体地下水水位的影响,从而可以间接体现降雨对坡体稳定性的影响;冲沟发育为坡体提供了良好的临空条

件,将直接影响斜坡的稳定性。此外,随着移民工程的进行,巴东新城区人类工程活动的增加,必然对研究区地质环境造成不同程度的影响,是滑坡灾害频频发生的主要诱发因素之一。

在滑坡影响因素危险性等级划分的基础上,选择巴东新城区 20 个典型滑坡,如黄土坡滑坡、赵树岭滑坡、童家坪滑坡等作为权重评价样本源数据,每个滑坡的影响因素量化定级见表 1,部分滑坡各指标的评级结果见表 2。

表 2 巴东新城区典型滑坡影响因素分级表  
Table 2 Indicator level of factors of landslide in the new site of Badong County

| 编号 | 滑坡<br>名称 | 岩土<br>类型 | 地形<br>坡度 | 斜坡<br>结构 | 与断层<br>的关系 | 水的<br>作用 | 人类<br>活动 |
|----|----------|----------|----------|----------|------------|----------|----------|
| 1  | 黄土坡      | 4        | 4        | 3        | 3          | 4        | 3        |
| 2  | 岩湾桥      | 4        | 5        | 2        | 2          | 2        | 3        |
| 3  | 红石梁      | 3        | 4        | 2        | 3          | 2        | 2        |
| ⋮  | ⋮        | ⋮        | ⋮        | ⋮        | ⋮          | ⋮        | ⋮        |
| 18 | 赵树岭      | 4        | 4        | 3        | 2          | 4        | 3        |
| 19 | 童家坪      | 3        | 5        | 3        | 2          | 3        | 2        |
| 20 | 西五路      | 3        | 5        | 3        | 3          | 3        | 3        |

1.1 研究区滑坡影响因素权重计算

利用分级数据表,采用熵权法和层次分析法相结合,综合评判巴东新城区滑坡灾害影响因素权重。这种综合评判方法在求解指标权重时,结果较为可靠,因为熵权法是一种根据数据指标进行客观评价的方法,而层次分析法为主观评价方法,两者结合能够平衡主客观因素产生的误差。

① 熵权法:利用分级数据表建立 20 个滑坡、6 个评价指标的初始矩阵,用  $X = (x_{ij})_{m \times n} (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$  表示,其中  $m = 20, n = 6$ 。利用式(1)~(2)得到研究区滑坡影响因素权重值客观评价结果  $w_j'$ 。

② 层次分析法:根据影响因素之间的关系,建立重要性对比矩阵,求解矩阵最大特征值和特征向量,并进行一致性检验,得到研究区滑坡影响因素权重值主观评价结果,此主观评价结果与石菊松等<sup>[12]</sup>分析的巴东新城区滑坡敏感性指标权重较为一致,认为主观评价合理  $w_j$ 。

③ 综合评价:根据上述两步计算结果,利用式(3)得到最后研究区滑坡影响因素权重综合评判值(表 3)。

$$y_{ij} = \frac{x_{ij} - x_{ij, \min}}{x_{ij, \max} - x_{ij, \min}}$$

(1)

$$w_j' = \frac{1 - E_j}{n - \sum_{j=1}^n E_j} \quad (2)$$

其中,  $E_j = -\frac{1}{\ln m} \sum_{i=1}^m \frac{1 + y_{ij}}{y_i} \ln \frac{1 + y_{ij}}{y_i}$ ,  $y_i = \sum_{j=1}^m 1 + y_{ij}$

$$W_j = \alpha w_j + (1 - \alpha) w_j' \quad (3)$$

式中: $m$ ——样本个数;

$n$ ——评价指标个数;

$y_{ij}$ ——评价指标标准化值;

$E_j$ ——表征指标评价重要度的熵值;

$w_j$ ——评价指标层次分析法所得权重;

$w_j'$ ——评价指标熵权法所得权重;

$W_j$ ——评价指标综合权重。

表 3 滑坡影响因素综合评价权值表

Table 3 Weights of landslide factors with the entropy method

|       |      |        |        |        |
|-------|------|--------|--------|--------|
| 熵权法   | 影响因素 | 岩组类型   | 地形坡度   | 斜坡结构   |
|       | 权重   | 0.2614 | 0.1438 | 0.1202 |
|       | 影响因素 | 与断层的关系 | 水的作用   | 人类活动   |
|       | 权重   | 0.1152 | 0.1022 | 0.2572 |
| 层次分析法 | 影响因素 | 岩组类型   | 地形坡度   | 斜坡结构   |
|       | 权重   | 0.4523 | 0.2026 | 0.0850 |
|       | 影响因素 | 与断层的关系 | 水的作用   | 人类活动   |
|       | 权重   | 0.052  | 0.1339 | 0.0743 |
| 综合评判  | 影响因素 | 岩组类型   | 地形坡度   | 斜坡结构   |
|       | 权重   | 0.3759 | 0.1791 | 0.0991 |
|       | 影响因素 | 与断层的关系 | 水的作用   | 人类活动   |
|       | 权重   | 0.0773 | 0.1212 | 0.1475 |

\* 注: $\alpha$ 取 0.6。

## 1.2 研究区滑坡灾害预测单元数据汇总

结合巴东县新城区已有的地质环境数据,以 MAPGIS 为操作软件,整理绘制以下图件:地形坡度分级图、岩组类型分布图、斜坡结构类型分布图、断层分布图、库水位影响范围分布图、汇水区分布图、冲沟分布图及公路与居民用地分布图。同时,依据研究区地形特征进行不规则预测单元划分,再利用 GIS 空间分析中矢量叠加分析功能生成基本预测单元,本次研究共生成基本单元 86216 个(图 1)。

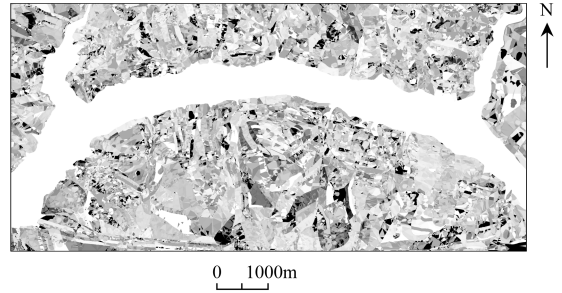


图 1 巴东县新城区空间预测单元划分图

Fig. 1 Spatial prediction units in new site of Badong County

根据所得研究区滑坡影响因素危险性分级表及权重综合评判结果,确定聚类预测模型的指标评价数据源,即每个预测单元所对应的影响因素量化值。例如,若 1 号预测单元其岩性类型危险性等级为 1,岩性类型权重为 0.3759,那么其计算量化值为  $1 \times 0.3759 = 0.3759$ 。部分预测单元量化值如表 4 所示。

表 4 预测单元指标量化值

Table 4 Quantitative value of forecasting units

| 单元号   | 岩组类型   | 地形坡度   | 斜坡类型   | 与断层的关系 | 水的作用   | 人类工程活动 |
|-------|--------|--------|--------|--------|--------|--------|
| 1     | 0.3759 | 0.3582 | 0.1982 | 0.0773 | 0.2424 | 0.1475 |
| 2     | 1.5037 | 0.5372 | 0.1982 | 0.0773 | 0.1212 | 0.1475 |
| 3     | 1.5037 | 0.5372 | 0.1982 | 0.0773 | 0.1212 | 0.1475 |
| ⋮     | ⋮      | ⋮      | ⋮      | ⋮      | ⋮      | ⋮      |
| 86214 | 0.3759 | 0.5372 | 0.1982 | 0.0773 | 0.1212 | 0.1475 |
| 86215 | 1.1278 | 0.5372 | 0.1982 | 0.2319 | 0.1212 | 0.1475 |
| 86216 | 0.7518 | 0.5372 | 0.1982 | 0.2319 | 0.1212 | 0.1475 |

## 1.3 研究区滑坡灾害危险性区划分级

结合滑坡频率统计及各影响因素分级表(表 1),将滑坡灾害危险性区划等级定为极高危险区、高危险区、中危险区、低危险区和极低危险区五级(表 5)。再依据滑坡影响因素综合评价权值表(表 2),对表 5 中各指标进行量化定值,并将危险性等级从高到低用 5、4、3、2、1 标记(表 6)。

表 5 滑坡危险性评价指标分级

Table 5 Rating of landslide risk assessment index

| 危险性 | 岩组类型                | 地形坡度(°)                  | 斜坡结构 | 与断层的关系          | 水的作用                 | 人类工程活动 |
|-----|---------------------|--------------------------|------|-----------------|----------------------|--------|
| 极高  | $T_2b^3$            | $20^\circ \sim 30^\circ$ | 顺向坡  | $\leq 180m$     | 库水作用、汇水区、冲沟发育三种联合作用  | 公路与建筑  |
| 高   | $T_2b^2$            | $10^\circ \sim 20^\circ$ | 斜向坡  | $180 \sim 300m$ | 库水作用、汇水区、冲沟发育有其中两种作用 | 公路或建筑  |
| 中   | $T_2b^2$            | $30^\circ \sim 40^\circ$ | 斜向坡  | $180 \sim 300m$ | 库水作用、汇水区、冲沟发育有其中两种作用 | 公路或建筑  |
| 低   | $T_2b^1$ 、 $T_2b^4$ | $> 40^\circ$             | 反向坡  | $> 300m$        | 库水作用、汇水区、冲沟发育有其中一种作用 | 无      |
| 极低  | $T_{1j}^3$          | $< 10^\circ$             | 反向坡  | $> 300m$        | 无                    | 无      |

以表 5 作为聚类分析的目标层,以表 6 作为目标层数据的计算形式,为聚类分析模型的分类结果进行等级排序提供依据,最终得到研究区的滑坡空间危险性等级分布图。

表 6 目标层指标量化值  
Table 6 Quantitative value of target layer

| 危险性等级 | 岩组类型   | 地形坡度   | 斜坡类型   | 与断层的关系 | 水的作用   | 人类工程活动 |
|-------|--------|--------|--------|--------|--------|--------|
| 1     | 0.3759 | 0.1791 | 0.0991 | 0.0773 | 0.1212 | 0.1475 |
| 2     | 0.7518 | 0.3582 | 0.0991 | 0.0773 | 0.2424 | 0.1475 |
| 3     | 1.1278 | 0.5372 | 0.1982 | 0.1546 | 0.3637 | 0.2949 |
| 4     | 1.1278 | 0.7163 | 0.1982 | 0.1546 | 0.3637 | 0.2949 |
| 5     | 1.5037 | 0.8954 | 0.2972 | 0.2319 | 0.4849 | 0.4424 |

2 聚类分析模型

聚类分析的方法较多,针对本文样本规模的情况,选择快速聚类法<sup>[13]</sup>进行空间危险性区划,其基本思想是:待聚类的样本比较多时,先给出一个大致得初始分类,然后用某种原则进行修改,直到分类结果比较合理为止。本文以 Visual studio 2005 为计算平台,利用 C# 语言编辑计算模型,具体流程如图 2 所示。

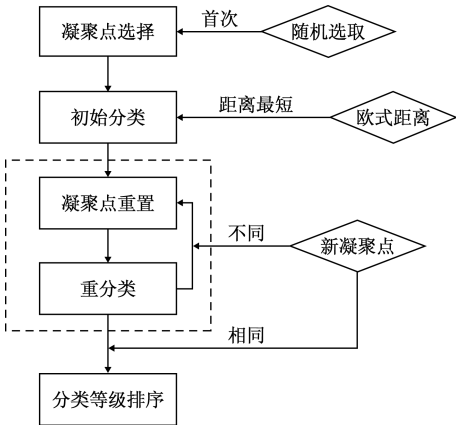


图 2 聚类模型计算流程图  
Fig. 2 Flow chart of the clustering model

首先通过 rand() 函数,在所有 86216 个预测单元里随机选择 5 个单元作为初始凝聚点。利用欧式距离公式(4)计算各样本与初始凝聚点之间的距离。

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^p (x_{ki} - x_{kj})^2}$$
 (4)

式中: $x_i$ ——样本;  
 $x_j$ ——凝聚点;  
 $p$ ——样本指标属性个数,即滑坡灾害影响因素个数。

根据距离最短的原则<sup>[14]</sup>进行初始分类,分类结果记为  $G^{(0)} = \{G_1^{(0)}, G_2^{(0)}, \dots, G_k^{(0)}\}$ 。利用重心公式(5)确定新的 5 个凝聚点  $x_j^{(1)}$ 。

$$x_j^{(1)} = \frac{1}{n_{i \in G_j^{(0)}}} \sum x_i$$
 (5)

接着,再对样本进行重分类,得到  $G^{(1)}$ ,直至分类到凝聚点相同,分类结果相同时,即  $x_j^{(m)} = x_j^{(m+1)}$ ,  $G^{(m)} = G^{(m+1)}$ ,  $m$  表示重复计算次数,聚类分析完成。

最后,将预设的目标层作为分类排序的比较对象,用已分好类的预测单元与目标层比较,依然依据距离最短的原则,得出每一类分类结果所对应的危险性等级。

表 7 中列出了研究区部分预测单元危险性预测值。并利用 MAPGIS 软件属性管理中的属性连接功能对每个单元赋值,形成巴东县新城区滑坡灾害危险性区划图(图 3)。

表 7 预测单元危险性等级  
Table 7 Risk level of prediction units

| 单元号   | 危险等级 |
|-------|------|
| 1     | 1    |
| 2     | 3    |
| 3     | 3    |
| ⋮     | ⋮    |
| 86214 | 1    |
| 86215 | 3    |
| 86216 | 1    |

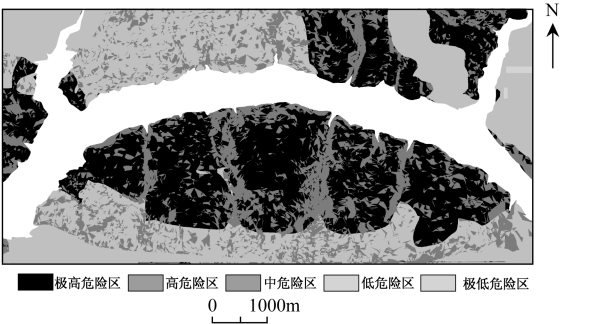


图 3 巴东县新城区滑坡灾害危险性空间预测图  
Fig. 3 Spatial prediction of landslide hazard map in the new site of Badong County

3 预测结果可靠性分析

为了验证权重分析和聚类分析结合所得滑坡灾害危险性区划分布图的可靠性,将本文预测结果与已知滑坡的空间位置(图 4)进行对比,可见,如赵树岭滑坡、黄土坡滑坡、太矶头滑坡等滑坡的分布位置与预测

结果危险性等级为极高的区域基本吻合。

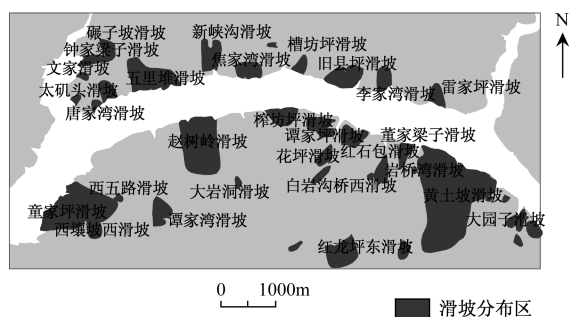


图4 巴东县新城区已知滑坡灾害分布图

Fig. 4 Distribution of landslide disaster in the new site of Badong County

同时,查询 86216 个预测单元中危险性等级为极高和高的单元,将其与已知滑坡区域进行比较,若重合,则将此等级区域合并,求其面积之和,再与所对应的已知滑坡区域总面积求比值,精度计算结果为 56.6%。

分析预测结果,精度计算等级不高的原因主要有两个方面:(1)已有资料中滑坡分布区域大多数为古滑坡,圈定面积大,但是并不是整体滑动,在精度计算时,分母为整体面积,导致计算结果偏小;(2)从预测图上看到,长江北岸临江的部分地段预测结果为低危险区,这应与滑坡影响因素指标权重的确定有关。本文在确定指标权重的时候,参考统计资料和相关文献,认为岩土类型影响因素较大,水的作用相对较小,对预测结果有一定的影响。从总体上综合考虑,认为此次滑坡预测结果仍是可靠的,对该研究区的城市规划建设具有参考价值。

## 4 结论

通过对巴东新城区滑坡灾害地质资料的分析和统计,选择有代表性的影响因素作为本次滑坡灾害危险性区划指标,利用熵权法和层次分析法相结合,综合评价了该区滑坡灾害影响因素的权重,并采用快速聚类分析方法确定了滑坡灾害预测单位危险性等级值。本研究主要结论如下:

(1)熵权法与层次分析法相结合,综合评判权重,虽然能够较好的较小主客观因素引起的评价误差,但是,权重计算仍然建立在大量数据统计的基础上,需要进行校正。

(2)快速聚类分析方法对样本规模较大的情况适用性较好,运算速度较快。本文在此基础上,预设目标层,克服了聚类分析方法分类结果不能自动排序的困

难,对大规模数据的分类计算有参考意义。

(3)权重分析和聚类分析相结合的方法,在本次滑坡灾害危险性区划研究中得到了较好的应用,认为预测方法意义明确,预测结果真实可靠,对区域灾害研究及城市规划等其他领域研究具有一定的参考价值。

(4)滑坡灾害区划研究是建立在资料收集整理以及详细的野外调查的基础之上,GIS 区划结果严格受到样本质量和数量控制,必须加强数据整理和优化。

## 参考文献:

- [1] 殷坤龙,滑坡灾害预测预报研究[M]. 武汉:中国地质大学出版社,2004. [YIN K L. Landslide hazard prediction and evaluation [M]. Wuhan: China University of Geosciences Press, 2004. (in Chinese)]
- [2] EVERITT B. Cluster analysis [M]. New York: Halsted-Wiley, 1974.
- [3] 徐林荣,华祖焜,杨灿文. 加筋土陡边坡状态评定的模糊聚类分析[J]. 岩土工程学报, 1999, 21(4): 475-480. [XU L R, HUA Z Y, YANG C W. Investigation and analysis of performance of reinforced steep slope using fuzzy cluster [J]. Chinese Journal of Geotechnical Engineering, 1999, 21(4): 475-480. (in Chinese)]
- [4] 蔡美峰,王鹏,赵奎,等. 基于遗传算法的岩体结构面的模糊 C 均值聚类方法[J]. 岩石力学与工程学报, 2005, 24(3): 371-376. [CAI M F, WANG P, ZHAO K, et al. Fuzzy C-Means cluster analysis based on genetic algorithm for automatic identification of joint sets [J]. Chinese Journal of Rock Mechanics and Engineering, 2005, 24(3): 371-376. (in Chinese)]
- [5] 蔡国军,刘松玉,童立元,等. 基于聚类分析理论的 CPTU 土分类方法研究[J]. 岩土工程学报, 2009, 31(3): 416-424. [CAI G J, LIU S Y, TONG L Y, et al. Soil classification using CPTU data based upon cluster analysis theory [J]. Chinese Journal of Geotechnical Engineering, 2009, 31(3): 416-424. (in Chinese)]
- [6] 张鹏,张在明,杨宇友,等. 模糊聚类在地层分析中的应用[J]. 岩土力学, 2009, 30(8): 2348-2352. [ZHANG P, ZHANG Z M, YANG Y Y, et al. Application of fuzzy cluster to geotechnical stratum analysis [J]. Rock and Soil Mechanics, 2009, 30(8): 2348-2352. (in Chinese)]
- [7] 李顺群,郑刚,崔春义,等. 粘土微结构各向异性评

- 估的谱系聚类方法[J]. 岩土工程学报, 2010, 32(1): 109–114. [LI S Q, ZHENG G, CUI C Y, *et al.* Pedigree cluster method to evaluate geometrical anisotropy of soil micro-structure fuzzy cluster[J]. Chinese Journal of Geotechnical Engineering, 2010, 32(1): 109–114. (in Chinese)]
- [8] 唐小松, 李典庆, 周创兵. 基于认知聚类分区方法的边坡可靠度分析[J]. 岩土力学, 2011, 32(2): 571–578. [TANG X S, LI D Q, ZHOU C B. Knowledge-based clustered partitioning method for reliability analysis of slope stability[J]. Rock and Soil Mechanics, 2011, 32(2): 571–578. (in Chinese)]
- [9] 黄锦林, 赵吉国, 张婷, 等. 基于层次分析法的库岸高速滑坡危险度评价[J]. 自然灾害学报, 2011, 20(5): 95–99. [HUANG J L, ZHAO J G, ZHANG T, *et al.* AHP-based hazard degree assessment of high-speed landslide of reservoir bank[J]. Journal and Natural Disasters, 2011, 20(5): 95–99. (in Chinese)]
- [10] 闫文周, 赵彬. 基于熵权法的生态城市建设平价[J]. 统计与决策, 2009(2): 66–68. [YAN W Z, ZHAO B. Entropy-based method evaluation of eco-city construction[J]. Statistics and Decision-making, 2009(2): 66–68. (in Chinese)]
- [11] 刘长春, 殷坤龙, 杜鹃. 湖北省巴东县新城区滑坡灾害空间预测[J]. 中国地质灾害与防治学报, 2009, 20(4): 13–18. [LIU C C, YIN K L, DU J. Spatial Prediction of Landslide Hazards in New Site of Badong County, Hubei Province[J]. The Chinese Journal of Geological Hazard and Control, 2009, 20(4): 13–18. (in Chinese)]
- [12] 石菊松, 张永双, 董诚, 等. 基于GIS技术的巴东新城区滑坡灾害危险性区划[J]. 地球学报, 2005, 26(3): 275–282. [SHI J S, ZHANG Y S, DONG C, *et al.* GIS-based Landslide Hazard Zonation of the New Badong County Site [J]. ACTA GEOSCIENTICA SINICA, 2005, 26(3): 275–282. (in Chinese)]
- [13] 向进东. 实用多元统计[M]. 武汉: 中国地质大学出版社, 2005. [XIANG J D. Applied multivariate statistical analysis[M]. Wuhan: China University of Geosciences Press, 2005. (in Chinese)]
- [14] MILLIGAN G W. Clustering validation: results and implications for applied analysis[M]. New Jersey: World Scientific, 1996.

## Landslide hazard zonation based on cluster analysis

GUI Lei, YIN Kun-long, WANG Jia-jia

(Department of Engineering Faculty, China University of Geosciences, Wuhan 430074, China)

**Abstract:** Researches in landslide hazard zonation are of important practical significance in urban planning decision-making. With the statistical form, cluster analysis can be used to classify the data with similar characteristics, and the quantitative assessment of the spatial distribution of landslide hazard may be realized. Based on landslide hazard data in Badong County, representative landslide hazard risk factors have been chosen as index in the zoning evaluation. Entropy method and AHP method have been combined to give a comprehensive evaluation of the index weight. On the basis of the above results, MapGIS is used as a platform and C# programming language is used to achieve a faster clustering algorithm. 86216 units in this study area are classified according to landslide hazard properties and are identified automatically to different risk levels. The results of this research are reliable and the method based on weight analysis and cluster model has innovative and practical value in processing high-volume and multi-attribute data in field studies.

**Key words:** landslide; hazard zonation; cluster analysis; entropy method

责任编辑: 汪美华