

文章编号: 1004- 4574(2009) 06- 0124- 05

# 泥石流危险度的改进集对分析模型

汪明武<sup>1</sup>, 李 丽<sup>2</sup>, 金菊良<sup>1</sup>

(1. 合肥工业大学 土木建筑工程学院, 安徽 合肥 230009; 2 合肥工业大学 资源与环境工程学院, 安徽 合肥 230009)

**摘 要:** 泥石流危险度的正确评估是泥石流灾害治理和工程措施的基础, 但泥石流危险度评价是一个极其复杂的不确定性问题。基于集对分析理论, 探讨了泥石流危险度的集对分析改进新方法, 即基于样本评价指标实测值与讨论泥石流标准危险度类别间的接近程度, 判别集对同异反, 进而应用模糊联系度的概念挖掘样本与讨论类别相邻类别的集对同异反有效信息, 以综合定量分析样本的危险度, 从而提高评价结果的可靠性和准确性。实例应用及与其他方法对比结果表明, 基于改进的集对分析方法来评价泥石流危险度是有效可行的, 可以取得理想的结果, 且操作简单。

**关键词:** 泥石流; 集对分析; 危险度; 模糊联系度

中图分类号: P642. 23

文献标识码: A

## An improved set pair analysis model for evaluation of dangerous degree of debris flow

WANG Ming-wu<sup>1</sup>, LI Li<sup>2</sup>, JIN Ju-liang<sup>1</sup>

(1. School of Civil Engineering Hefei University of Technology, Hefei 230009, China; 2 School of Resources and Environmental Engineering Hefei University of Technology Hefei 230009, China)

**Abstract** The dangerous degree evaluation of debris flow is a foundation for the disaster control and engineering measures. However, the evaluation of dangerous degree of debris flow is a very complicate and uncertainty problem. An improved model for the dangerous degree evaluation using the theory of set pair analysis (SPA) was discussed in this paper. The basic principle is that the identical- different- contrary is defined by an approaching degree between the measured index values of sample and corresponding index of standard sets of the dangerous degree. A model for determining discrepancy coefficient using the concept of the fuzzy degree of connection was investigated during the procedure of SPA. And the useful data was mined between the sample and adjacent discussed grade. Then the relative membership degree to each grade was calculated, and the dangerous degree of debris flow was predicted quantitatively. The results from the application to practical example and comparisons with other methods are fairly good. And they also show that the improved method of SPA used to assess dangerous degree of debris flow is feasible and effective, and easy to operate. The proposed novel method also can be used to evaluate the other uncertainty engineering problems.

**Key words** debris flow; set pair analysis; dangerous degree; fuzzy degree of connection

泥石流危险度评价是防治泥石流的基础。泥石流的发生、发展受各种不确定因素影响和控制, 故泥石流危险度区划是一个复杂的不确定问题。以往的泥石流危险度评价方法主要有泥石流活动的直接指标法和泥

收稿日期: 2008- 08- 23; 修订日期: 2009- 07- 21

基金项目: 国家自然科学基金资助项目 (70771035, 40702049); 安徽优秀青年科技基金资助项目

作者简介: 汪明武 (1972- ), 男, 教授, 博士, 主要从事智能岩土工程和防灾工程研究. E-mail: geotechnic@163.com

泥石流发育的间接指标法,如模糊信息方法、关联度分析法和模糊综合评判法等<sup>[1-5]</sup>,然而,它们应用中一般需人为设计主要评价因子及相应的权值,但泥石流是一种突发性的复杂自然灾变过程,专家确定的重要性权值或基于指标值确定的客观权值,很难全面反映评价指标的相互依赖或相互关联的关系和泥石流灾害演化的非线性特征。为此,人们引入神经网络方法来解决此难题<sup>[6]</sup>,但它受知识获取“瓶颈”问题的限制,存在应用局限性;近期出现的投影寻踪智能方法和可拓方法虽能全面反映泥石流的实际特征<sup>[7-8]</sup>,但也存在遗漏重要约束条件问题,而集对分析理论则可避免这些缺陷,也为泥石流危险度的系统分析提供了新的途径。本文以集对分析理论为基础,引入模糊联系度的思想对集对差异度系数充分进行数据挖掘,以建立可避免知识获取和指标权重确定等难题的、泥石流危险度集对分析新模型,以从新的角度探索泥石流危险度的评价方法。

# 1 泥石流危险度的改进集对分析模型

## 1.1 集对分析基本原理及评价流程

泥石流危险度集对分析基本原理是通过数据挖掘计算待估样本的评价指标实测值集合与泥石流危险度等级标准集合之间的集对联系度,进而定量评定泥石流危险度。待估样本的集对联系度同异反判别标准设计为:若评价指标值处于讨论等级内则为同一,处于讨论等级相邻等级内则为异,处于相隔等级内则为对立。

基于数据挖掘的泥石流危险度集对分析具体步骤如下:(1)首先基于样本实测值、泥石流危险度等级标准和同异反判别标准构建集对联系度表达式;(2)基于构建的表达式,计算样本中落入相邻等级评价指标与讨论泥石流危险度等级的接近程度,并应用模糊联系度进行差异度的数据挖掘,以求得最佳集对联系度;(3)计算样本对讨论等级的隶属度,并计算相应的特征值,以量化评定样本的泥石流危险度,具体分析流程如图1所示。

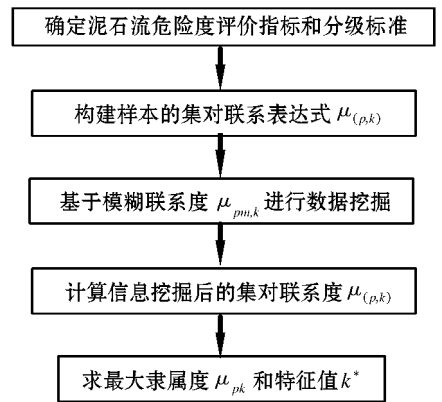


图1 改进集对分析流程图  
Fig 1 Flow chart of improved set pair analysis

## 1.2 集对联系度

集对分析方法是由我国学者赵克勤于1989年提出的一种关于确定不确定系统的同异反定量系统分析方法,该法能完整有效刻画确定不确定系统的对立统一关系,并在工程界得到应用<sup>[9]</sup>。它是将被研究客观事物的确定不确定性视为一个确定不确定系统,从同异反3方面研究客观事物之间的联系与转化,即以联系度来描述系统的各种不确定性。集对分析理论最基本的概念是集对和联系度。集对是指具有一定联系的两个集合组成的整体<sup>[9]</sup>。对给定集合A和B,则两个集合组成的集对为H=(A,B)。设集对有N个特性,其中共同具有的特性数为S,相互对立的特性数为P,既不共同也不相互对立的特性数为F,则相应的集对联系度表达式可构建如下,

$$\mu_{(A,B)} = \frac{S}{N} + \frac{F}{N}i + \frac{P}{N}j \tag{1}$$

式中,j为对立度系数,常取j=-1;i为差异度系数,i∈[-1,1],是宏观与微观两个层次相互结合的参量,它是集对分析的关键。

## 1.3 模糊联系度及计算

为确保泥石流危险度评价结果的准确性和可靠性,且能体现对不确定性程度的刻画,在集对分析过程中,不仅仅是简单确定被评价样本的指标集合与评价标准集合间的同异反联系度,还必须对差异度系数信息进一步挖掘,以挖掘处于相邻等级中的评价指标与所讨论等级间的联系度。本文应用模糊联系度来对差异度系数进行数据信息的充分挖掘和定量分析。而实现集对能完整描述客观对象和主观认识的确定不确定性,对某一特性上的同异反模糊刻画,公式(1)应该定义为信息意义上的确定性与不确定性模糊结构表达式。设模糊子集A,A<sup>c</sup>为论域X中确定性信息和不确定性信息,则它们是一种对立描述,相应定义信息x关于A,A<sup>c</sup>的确定度和差异度分别为μ<sub>A</sub>(x)=a+c,μ<sub>A<sup>c</sup></sub>(x)=b且μ<sub>A</sub>(x)+μ<sub>A<sup>c</sup></sub>(x)=1其中a,c为x对确定信息的同一度和对立度。若令a=SN,b=FN,c=PN,a+b+c=1,则待评样本p的泥石流危险度属于k级(k

= 1 2 ... K)的模糊联系度定义为,

$$\mu_{p,k} = a_{p,k} + b_{p,k}i + c_{p,k}j \tag{2}$$

式中,  $a_{p,k}$ ,  $b_{p,k}$ ,  $c_{p,k}$  分别是待估样本  $p$  对危险度等级  $k$  的同一度、差异度和对立度。由上分析可知, 差异度的合理确定是基于信息意义集对分析的关键, 而差异度的数据挖掘过程实质是  $b_{p,k}$  的分解过程。当样本评价指标值随危险性等级增加而增大时, 设样本  $p$  实测的评价指标  $x_{pm}$  落在相邻等级  $n$  的指标标准内共有  $M$  个, 以接近程度来描述  $x_{pm}$  与  $k$  级危险度的模糊联系度。当  $x_{pm}$  落在相邻等级  $k+1$  内时,  $x_{pm}$  与危险度  $k$  级的接近程度为  $S_{k+1,m}/x_{pm}$ ,  $x_{pm}/S_{k+2,m}$ , 并将它们归一化后的数值作为接近程度的肯定和否定, 即作为集对  $(x_{pm}, B_k)$  的同一度与对立度, 则相应的差异度为

$$b = 1 - \frac{S_{k+1,m}S_{k+2,m}}{(S_{k+1,m} + S_{k+2,m})x_{pm}} - \frac{x_{pm}}{S_{k+1,m} + S_{k+2,m}} = \left( \frac{x_{pm} - S_{k+1,m}}{(S_{k+1,m} + S_{k+2,m})x_{pm}} \right) i + \frac{x_{pm}}{S_{k+1,m} + S_{k+2,m}} j \tag{3}$$

式中,  $S_{k+1,m}$ ,  $S_{k+2,m}$  为相邻等级  $m$  指标的界限值。则落在相邻等级内的评价指标  $m$  对  $k$  等级相应指标分级标准的模糊联系度表达式如下:

$$u_{pm,k} = \frac{S_{k+1,m}S_{k+2,m}}{(S_{k+1,m} + S_{k+2,m})x_{pm}} + \left( \frac{x_{pm} - S_{k+1,m}}{(S_{k+1,m} + S_{k+2,m})x_{pm}} \right) i + \frac{x_{pm}}{S_{k+1,m} + S_{k+2,m}} j \tag{4}$$

而当  $x_{pm}$  落在相邻等级  $k-1$  ( $k-1 \neq 0$ ) 时, 相应的模糊联系度表达式则为

$$u_{pm,k} = \frac{x_{pm}}{S_{k-1,m} + S_{k,m}} + \left( \frac{x_{pm} - S_{k-1,m}}{(S_{k-1,m} + S_{k,m})x_{pm}} \right) i + \frac{S_{k-1,m}S_{k,m}}{(S_{k-1,m} + S_{k,m})x_{pm}} j \tag{5}$$

对  $M$  个指标对等级  $k$  分别进行数据挖掘分析, 并取平均值作为相应的综合模糊联系度, 即

$$\bar{a}_{p,k} = \frac{1}{M} \sum_m a_{pm,k}, \bar{b}_{p,k} = \frac{1}{M} \sum_m b_{pm,k}, \bar{c}_{p,k} = \frac{1}{M} \sum_m c_{pm,k} \tag{6}$$

式中,  $a_{pm,k}$ ,  $b_{pm,k}$ ,  $c_{pm,k}$  分别为样本  $p$  落在相邻等级内的指标对  $k$  级危险度的同一度、差异度和对立度。将上分解值叠加入式 (2), 即可得数据充分挖掘后的集对联系度,

$$\mu_{p,k} = (\bar{a}_{p,k} + \bar{a}_{p,k}b_{p,k}) + \bar{b}_{p,k}b_{p,k}i + (\bar{c}_{p,k} + \bar{c}_{p,k} + \bar{c}_{p,k}b_{p,k})j \tag{7}$$

可见, 模糊联系度的数据挖掘过程是对不确定性的进一步分析过程, 它是在对不确定性系统深入认识过程中获到的有效信息, 从而可提高评价结果的准确性和可靠性。

### 1 4 级别特征值

基于数据挖掘后的集对联系度, 计算样本对讨论等级的隶属度  $\mu_{pk}$ 。对于泥石流危险度评价问题, 经数据挖掘后构成的集对联系度已能满足实际工程的准确度要求, 即取  $i=0$   $j=-1$  来计算样本对讨论等级的隶属度, 再按最大隶属度对应的等级作为评价结果的准则来判定泥石流危险度等级。为避免应用最大隶属度原则造成判定失真, 以及实现定量描述经数据挖掘后的集对分析结果和表征样本向另一类别转换的趋势, 本文采用特征值  $k^*$  来定量化评定结果, 计算公式如下,

$$k^* = \frac{\sum_{k=1}^K k \cdot \bar{\mu}_{pk}}{\sum_{k=1}^K \bar{\mu}_{pk}} \tag{8}$$

$$\bar{\mu}_{pk} = \frac{\mu_{pk} - m \ln \mu_{pk}}{m \max(\mu_{pk}) - m \ln(\mu_{pk})} \tag{9}$$

式中,  $k^*$  为待评样本的泥石流危险度特征值;  $\mu_{pk}$  为样本对泥石流危险度  $k$  的隶属度。

## 2 应用实例

为验证构建模型的正确性和有效性, 选用文献 [1 8] 数据作为实例, 实例中选用了选用泥石流规模、泥石流发生频率、流域面积、主沟长度、流域最大相对高差、流域切割密度、主沟床弯曲系数、泥砂补给段长度比、24 h 最大降雨量、人口密度等 10 个指标作为泥石流危险度的评价指标, 将泥石流危险度划分为轻度危险、中度危险、高度危险和极度危险 4 个等级, 则泥石流危险度评价问题的集对  $H(P, B_k)$  中的集合  $P$ , 由待评样本  $p$  的评价指标  $m$  ( $m = 1 2 \dots 10$ ) 实测值构成, 集合  $B_k$  ( $k = 1 2 \dots 4$ ) 为泥石流危险度对应的评价指标等级构成。样本指标实测值和危险度分级标准见表 1 和表 2。基于数据挖掘的泥石流危险度集对分析结果及

与其他方法的对比分析结果见表 3

表 1 样本评价指标实测值

Table 1 Measured values of evaluation indexes of samples

样本 $p$	规模 / $10^4 \text{m}^3$ ( $m=1$ )	发生频率 / % ( $m=2$ )	流域面积 / $\text{km}^2$ ( $m=3$ )	主沟长度 / km ( $m=4$ )	流域最大相 对高差 / km ( $m=5$ )	流域切割 密度 / ( $\text{km} \cdot \text{km}^{-2}$ ) ( $m=6$ )	主沟床弯 曲系数 ( $m=7$ )	泥砂补给 段长度比 ( $m=8$ )	24 h最大降 雨量 /mm ( $m=9$ )	人口密度 / ( $\text{人} \cdot \text{km}^{-2}$ ) ( $m=10$ )
1	195.1	1500	47.10	12.00	2.19	23.8	1.45	0.80	102.0	260
2	105.0	450	53.10	18.35	2.92	21.2	1.28	0.62	97.0	210
3	7.8	20	10.61	4.61	1.66	12.8	1.12	0.40	100.4	290
4	10.0	30	14.10	8.00	1.94	17.8	1.10	0.35	100.4	310
5	1.8	80	4.20	2.10	1.19	12.3	1.11	0.42	100.4	318
6	82.0	1200	18.05	18.05	1.66	22.8	1.39	0.72	100.4	0
7	9.0	50	28.32	28.32	2.80	15.6	1.22	0.45	100.4	260
8	5.0	50	3.28	3.28	1.13	22.0	1.28	0.78	98.0	276

表 2 泥石流危险度分级标准

Table 2 Grade standard of dangerous degree of debris flow

级别	指标 $m$									
	1	2	3	4	5	6	7	8	9	10
轻度 ( $k=1$ )	0~1	0~10	0~0.5	0~1	0~0.2	0~50	0~1.1	0~0.1	0~25	0~50
中度 ( $k=2$ )	1~10	10~50	0.5~10	1~5	0.2~0.5	5~10	1.1~1.25	0.1~0.3	25~50	50~150
高度 ( $k=3$ )	10~100	50~100	10~35	5~10	0.5~1.0	10~20	1.25~1.4	0.3~0.6	50~100	150~250
极度 ( $k=4$ )	100~200	100~1600	35~55	10~25	1.0~3.5	20~30	1.4~2.0	0.6~1.0	100~150	250~350

表 3 集对分析方法评价结果及对比

Table 3 Results from SPA and comparisons

样本 $p$	$\mu_{p1}$	$\mu_{p2}$	$\mu_{p3}$	$\mu_{p4}$	本文方法		可拓方法 <sup>[8]</sup>	模糊数学 <sup>[1]</sup>
					$k$	$k^*$		
1	-1.0000	-1.0000	-0.0630	1.0000	4	3.68	极度危险	极度危险
2	-1.0000	-0.7349	0.2884	0.7349	4	3.45	极度危险	极度危险
3	-0.7051	0.1647	0.4366	-0.1647	3	2.87	中度危险	中度危险
4	-0.7040	0.0085	0.4160	-0.0085	3	2.99	高度危险	高度危险
5	-0.5844	0.1025	0.3357	-0.1025	3	2.90	中度危险	中度危险
6	-0.8000	-0.6654	0.2078	0.5654	4	3.49	极度危险	极度危险
7	-0.8423	-0.0738	0.5403	0.0738	3	3.05	高度危险	高度危险
8	-0.7128	-0.0280	0.3732	0.0280	3	3.02	中度危险	中度危险

现以样本实测值分布等级较广的样本 4 为例, 说明基于数据挖掘的泥石流危险度集对分析过程。基于样本 4 指标实测值和分级标准构建的联系度表达式如下,  $\mu_{(p_4 B_1)} = \frac{1}{10} + \frac{2}{10}i + \frac{7}{10}j$ ,  $\mu_{(p_4 B_2)} = \frac{2}{10} + \frac{2}{10}i + \frac{7}{10}j$ ,  $\mu_{(p_4 B_3)} = \frac{4}{10} + \frac{5}{10}i + \frac{3}{10}j$ ,  $\mu_{(p_4 B_4)} = \frac{3}{10} + \frac{5}{10}i + \frac{1}{10}j$ 。样本 4 中  $\mu_{(p_4 B_1)} = \frac{1}{10} + \frac{2}{10}i + \frac{7}{10}j$  集对差异度的数据挖掘过程: 由  $\mu_{(p_4 B_1)}$  集对联系度表达式可知, 指标实测值落在中度危险范围内的指标有泥石流规模指标 ( $m=1$ ) 和发生频率指标 ( $m=2$ ) 共 2 个。据实测值可知泥石流规模指标值落在轻度危险等级 ( $n=1$ ) 的右侧, 界限值为  $S_2=1$ ,  $S_3=10$  将实测值和界限值代入式 (4) 即可求得  $\mu_{411} = 0.0909 + 0.0000i + 0.9091j$  同理可求得发生频率指标 ( $m=2$ ) 相应的模糊联系度, 即  $\mu_{421} = 0.2778 + 0.2222i + 0.5000j$  再按式 (6) 和式 (7) 计算数据挖掘后的最终模糊联系度, 可得  $\mu_{41} = 0.1843 + 0.1111i + 0.7046j$ ;  $\mu_{(41)} = 0.1369 + 0.0222i + 0.8409j$  取  $i=0, j=$

- 1 即可计算的样本 4 对轻度危险的隶属度  $\mu_{41} = 0.7040$ 。同理可求得  $\mu_{42} = 0.0085$   $\mu_{43} = 0.4160$   $\mu_{44} = 0.0085$  再按式 (8) 和式 (9) 可求得级别特征值为 2.99 按等级判定准则可判定样本 4 的危险度为高度。

从表 3 可知, 基于数据挖掘的集对分析结果与文献 [1], [8] 方法的评价结果基本吻合。下面对样本 5 和 8 的评价等级与可拓及模糊数学方法结果存在一定差别, 由样本 5 和 8 的评价指标可知, 样本 5 和 8 的实测评价指标值中超过中度危险等级的都存在 6 个, 处于中度等级为 4 个, 但基于可拓学和模糊数学方法的评价结果为中度危险, 评价方法的风险性和可靠性则有待深入探讨。可见, 基于数据挖掘的集对分析评价过程不需确定权值, 定量结果也更接近样本特征的实际情况, 故应用集对分析方法来评价泥石流危险度是有效可行的, 而且操作过程相对简单。

### 3 结 语

泥石流危险度评价问题是一个复杂的不确定问题, 本文应用集对分析理论和模糊联系度概念, 建立了基于数据挖掘的泥石流危险度集对分析模型, 实例应用及与其他方法对比结果表明, 应用集对分析方法来评价泥石流危险度是有效可行的, 且可深入挖掘样本的评价信息, 该方法具有简便、定量严密的特点, 也为其他不确定工程问题评价提供了新的参考, 但模型的应用可靠性有待深入研究。

### 参考文献:

- [1] 刘希林, 唐川. 泥石流危险度评价 [M]. 北京: 科学出版社, 1995: 1-93.
- [2] 刘希林, 张松林, 唐川. 沟谷泥石流危险度评价研究 [J]. 水土保持学报, 1993, 7(3): 20-25.
- [3] 刘丽, 王士革. 滑坡、泥石流区域危险度二级模糊综合评判初探 [J]. 自然灾害学报, 1996, 5(3): 51-59.
- [4] 魏永明, 谢又予, 伍永秋. 关联度分析法和模糊综合评判法在泥石流沟谷危险度划分中的应用—以北京市北部山区怀柔、密云两县为例 [J]. 自然灾害学报, 1998, 7(2): 109-117.
- [5] 刘涌江, 胡厚田, 白志勇. 泥石流危险度评价 [J]. 水土保持学报, 2000, 14(2): 84-87.
- [6] 汪明武. 基于神经网络的泥石流危险度区划 [J]. 水文地质工程地质, 2000, 27(2): 18-19.
- [7] 汪明武, 金菊良, 李丽. 投影寻踪方法在泥石流危险度评价中的应用 [J]. 水土保持学报, 2002, 16(6): 79-81.
- [8] 匡乐红, 徐林荣, 刘宝琛. 基于可拓方法的泥石流危险性评价 [J]. 中国铁道科学, 2006, 27(5): 1-5.
- [9] 赵克勤. 集对分析及其初步应用 [M]. 杭州: 浙江科学技术出版社, 2000.