

doi: 10.3969/j.issn.1005-3026.2017.04.024

# 基于 K-PSO 聚类算法和熵值法的滑坡敏感性

阮云凯<sup>1</sup>, 占洁伟<sup>1</sup>, 陈剑平<sup>1</sup>, 李严严<sup>2</sup>

(1. 吉林大学 建设工程学院, 吉林 长春 130026; 2. 中国科学院 地质与地球物理研究所, 北京 100029)

**摘要:** 引入 K-PSO 聚类算法和熵值法, 建立滑坡敏感性分析模型. 选取旭龙水电站库区 22 处典型滑坡作为研究对象, 确定 8 个主要影响因子: 岩体结构、斜坡结构、断层距离、变形迹象、坡体高度、平均坡度、诱发地震、淹没比例. 利用熵值法确定影响因子权重值分别为 0.152, 0.178, 0.035, 0.106, 0.106, 0.169, 0.193 和 0.061. 采用 K-PSO 算法对滑坡进行敏感性划分, 结果表明, 该库区 22 处滑坡有 8 处为轻度敏感, 9 处为中度敏感, 4 处为重度敏感和 1 处极度敏感. 将评价结果与现场实际调查情况对比分析知, 22 处滑坡的敏感度水平与现场实际发育情况具有较好的一致性, 该方法对旭龙水电站库区滑坡敏感性评价具有良好的指导作用.

**关键词:** 熵值法; 滑坡; K-PSO; 聚类模型; 敏感性

中图分类号: TU 473 文献标志码: A 文章编号: 1005-3026(2017)04-0571-05

## Landslide Sensitivity Based on K-PSO Clustering Algorithm and Entropy Method

RUAN Yun-kai<sup>1</sup>, ZHAN Jie-wei<sup>1</sup>, CHEN Jian-ping<sup>1</sup>, LI Yan-yan<sup>2</sup>

(1. College of Construction Engineering, Jilin University, Changchun 130026, China; 2. Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China. Corresponding author: CHEN Jian-ping, professor, E-mail: chenjpwq@126.com)

**Abstract:** The K-PSO clustering algorithm and entropy method were introduced to establish a sensitivity analysis model for landslide. The 22 typical landslides located in Xulong hydropower station reservoir area were investigated. Eight major factors including rock mass structure, slope structure, fault distance, signs of deformation, slope height, average gradient, induced earthquake and submerged ratio were determined for landslide sensitivity analysis. The weights of major factors determined by the entropy method are 0.152, 0.178, 0.035, 0.106, 0.106, 0.169, 0.193, 0.061, respectively. Sensitivity analysis results based on K-PSO clustering algorithm showed that among the 22 landslides, 8 landslides are evaluated as low sensitive, 9 as moderate, 4 as severely sensitive and one as extremely sensitive. Compared with the in-situ observations, the evaluated level of sensitivity of the 22 landslides agree very well with the actual development of the landslides. The proposed K-PSO method is effective for landslide sensitivity analysis in Xulong hydropower station reservoir area.

**Key words:** entropy method; landslide; K-PSO; clustering model; sensitivity

滑坡是最危险的地质灾害之一,其破坏力强、危害大,对人类的生活安全造成严重的威胁<sup>[1-2]</sup>. 随着西南各省水利水电的蓬勃发展,水电站坝址区产生的滑坡地质灾害对水电站的建设乃至运营都造成不可估量的危害和损失. 滑坡受到地质、地形和气象等因素的共同影响,而这些因素又具有

很大的不确定性,因此,研究滑坡的敏感性仍具有很大的挑战.

目前,许多学者开展了大量滑坡敏感性研究工作. 例如, Carrara<sup>[3]</sup> 建立 GIS 和统计分析模型, 进行滑坡敏感性分区研究; Biswajeet 等<sup>[4]</sup> 基于 RS 和 GIS 技术, 采用神经网络模型合理划分滑坡

收稿日期: 2015-04-06

基金项目: 国家自然科学基金重点资助项目(41330636); 吉林大学研究生创新基金资助项目(2016208).

作者简介: 阮云凯(1989-),男,福建建阳人,吉林大学博士研究生; 陈剑平(1957-),男,福建南平人,吉林大学教授,博士生导师.

敏感性等级,准确率高达 82.92%;Neaupane 等<sup>[5]</sup>建立了网络分析模型并应用于尼泊尔东部脆弱山区的滑坡灾害敏感性分区划分;唐川等<sup>[6]</sup>提出模糊综合评价模型并应用于云南滑坡敏感性分区划分;Xu 等<sup>[7]</sup>基于 GIS 和支持量机理论,建立地震诱发滑坡敏感性评价模型;李军霞等<sup>[8]</sup>利用组合赋权给因子赋权并基于未确知测度理论建立评价滑坡敏感性模型.传统的方法主观性较大,确定性模型则需要收集大量的水文地质和工程地质数据;而人工神经网络不确定性太强,验证参数困难;模糊数学评价模型则在滑坡影响因子太多时,权重分配会降低影响结果的精度.

数据挖掘在计算机领域中发挥重大作用,它从大量的数据中自动搜索隐藏于其中的有着特殊关系性信息的过程,进而获得最佳信息.将数据挖掘引入到滑坡评价<sup>[9]</sup>中,目的在于将滑坡与不滑坡的各种地质条件属性进行归类.聚类分析是数据挖掘中的一种有效算法,K-PSO 聚类算法是在 PSO 算法的基础上优化 K-Means 算法,它具有收敛稳定、计算效率高等特点.本文基于 K-PSO 聚类算法,建立了一种新的评价滑坡敏感性模型.

## 1 熵值法与 K-PSO 聚类算法的计算理论

### 1.1 熵值法

信息熵作为一种客观赋权的手段,用来确定影响因子的权重,它的评价过程完全依赖客观数据规律,使得权重具有绝对的客观性,从而在很大程度上避免人为因素的干扰,具体操作步骤如下.

待评价体系中有  $m$  个评价对象, $n$  个评价指标,那么第  $i$  个评价指标的熵值为<sup>[10]</sup>

$$H_i = -k \sum_{j=1}^m P_{ij} \ln P_{ij}. \quad (1)$$

式中: $H_i > 0$ ;  $k = 1/\ln m$ ;  $P_{ij} = u_{ij} / \sum_{j=1}^n u_{ij}$ , 当  $u_{ij} = 0$  时,  $u_{ij} \ln u_{ij} = 0$ ,  $i = 1, 2, \dots, n$ .

第  $i$  个指标的偏差度  $d_i = 1 - H_i$ , 第  $i$  个指标的权重为<sup>[10]</sup>

$$\theta_i = \frac{d_i}{\sum_{i=1}^n d_i} = \frac{1 - H_i}{n - \sum_{i=1}^n H_i}. \quad (2)$$

### 1.2 K 均值算法

K 均值聚类分析是对数据进行分类的一种算法,在数据挖掘等领域中广泛应用.步骤如下:

首先将待评价对象随机分为  $K$  类,根据初始聚类结果,由式(3)计算每一组的初始聚类中心  $V_j$ ,再计算每一个对象到各个聚类中心的距离  $d(V_j, X_i)$ ,并按距离最近重新分类,由式(3)更新聚类中心,然后由式(4)计算目标函数,通过迭代计算,使目标函数值不断减小,直到被评价对象不再分配到新的聚类中结束运算.

$$V_j = \frac{1}{n_j} \sum_{i=1}^k X_i. \quad (3)$$

式中, $n_j$  为第  $j$  类中被评价对象的个数.

$$f = \sum_{i=1}^k \sum d(V_j, X_i). \quad (4)$$

式中, $d(V_j, X_i)$  表示第  $j$  个评价对象与第  $i$  个聚类中心的距离.

### 1.3 PSO 算法介绍

Kennedy 等<sup>[11]</sup>提出粒子群算法(PSO),研究鸟群的觅食行为,它是一种有效的全局优化算法.

在  $D$  维空间中,单个粒子按照自身位置  $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$  和速度  $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$  来表示.每个粒子通过调整自身的位置进行搜索,而位置的变化是由速度的改变来实现的.记第  $i$  个粒子群最优位置  $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$  和粒子在整个群的最优位置  $P_g = (p_{g1}, p_{g2}, \dots, p_{gD})$ .粒子速度和位置的改变由式(5)和式(6)计算求解<sup>[12]</sup>:

$$v_{id} = wv_{id} + \mu_1 r_1 (p_{id} - x_{id}) + \mu_2 r_2 (p_{gd} - x_{id}), \quad (5)$$

$$x_{id} = x_{id} + v_{id}. \quad (6)$$

式中: $d$  为搜索方向( $d = 1, 2, \dots, D$ );  $w$  ( $0.4 \leq w \leq 0.9$ ) 为控制粒子飞行速度的惯性权重; $\mu_1$  和  $\mu_2$  为学习因子( $\mu_1 = \mu_2 = 2$ );  $r_1$  和  $r_2$  为取值在  $[0, 1]$  之间的随机变量<sup>[12]</sup>.

### 1.4 K-PSO 算法介绍

K 均值算法比 PSO 算法收敛速度要快,但 K 均值算法聚类不够精确,为了提高 PSO 算法的性能,需要采用 K 均值算法重新初始化群. K-PSO 算法流程图见图 1<sup>[12]</sup>,具体操作过程如下:

- 1) 运行 K 均值算法一次,其结果为  $N_c$  个聚类中心向量;将这些中心向量作为粒子群中的一个粒子,粒子群中其余数据随机初始化;
- 2) 运行 PSO 算法的流程.

## 2 研究区概况及滑坡影响因子

### 2.1 研究区概况

旭龙水电站地处四川省得荣县和云南省德钦县交界的金沙江干流上游河段,工程区范围上游

坡地质灾害,分布如图 2 所示. 滑坡发育地层主要为金沙江蛇绿岩群(DTJ).

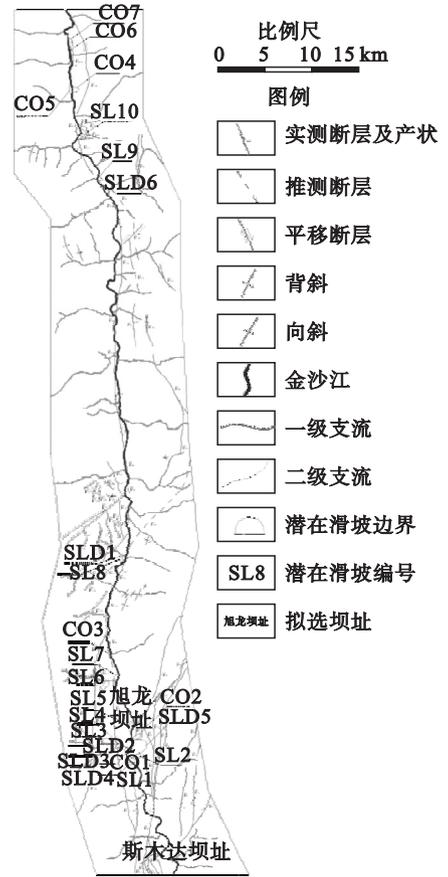


图 2 滑坡空间分布

Fig. 2 Landslide spatial distribution

### 2.2 评价因子选取

本文按文献[13]中的方法,选取滑坡的岩体结构  $C_1$ 、斜坡结构  $C_2$ 、断层距离  $C_3$  (m)、变形迹象  $C_4$ 、坡体高度  $C_5$  (m)、平均坡度  $C_6$  ( $^\circ$ )、诱发地震  $C_7$  和淹没比例  $C_8$  (%) 共 8 个因子作为滑坡敏感性评价因子,并据此建立滑坡敏感性评价指标体系,如图 3 所示。

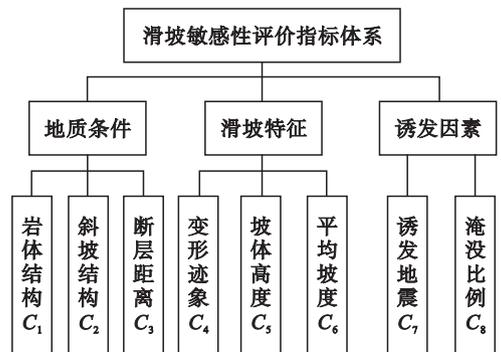


图 3 滑坡敏感性评价指标体系

Fig. 3 Landslide sensitivity evaluation system

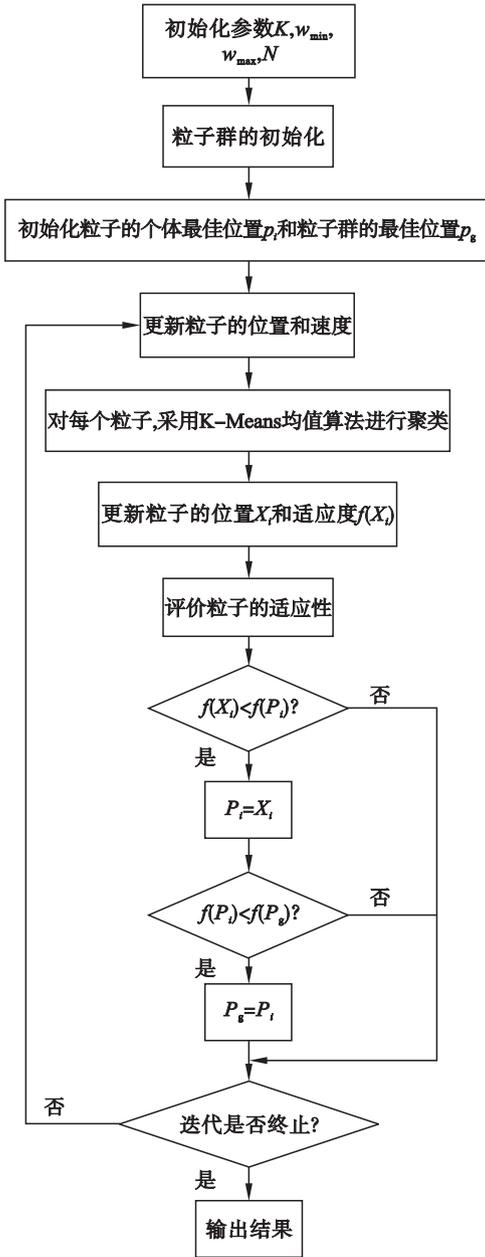


图 1 K-PSO 算法流程图

Fig. 1 Flow chart of K-PSO method

距昌波坝址 75.5 km, 下游距奔子栏镇 72.8 km. 库坝区地处横断山脉北端,属于西南部青藏高原地貌区域. 水电站正常蓄水水位 2 302 m,库容为  $8.19 \times 10^8 \text{ m}^3$ , 正常蓄水时水库面积约为  $16.5 \text{ km}^2$ . 库坝区地下水可分为松散堆积层孔隙水、碎屑岩孔隙裂隙间水、基岩裂隙水三大类型. 研究区岩层主要发育有中元古界、古生界、中生界、新生界,同时也发育有岩浆岩、变质岩. 区域构造基本特征为断裂构造:发育德登—老君山断裂、德格—乡城断裂、德钦—中甸—大具断裂、江达—施坝断裂、巴塘断裂. 近库区内历史上未记录破坏性的地震发生. 现场野外调查表明,综合确定了坝址区沿金沙江两岸分布的 22 处典型潜在滑

### 3 基于熵值法和 K-PSO 的评价模型

#### 3.1 权重的确定及聚类分析结果

用式(1)和式(2)计算评价因子的权重:

$\theta_i = (\theta_1, \theta_2, \dots, \theta_8) = (0.152, 0.178, 0.035, 0.106, 0.106, 0.169, 0.193, 0.061)$ . 将研究区滑坡敏感性分为轻度、中度、重度、极度敏感 4 类. 研究区滑坡各评价因子指标值和 K-PSO 聚类结果如表 1 所示.

表 1 各影响因子指标值  
Table 1 Index value of each influence factor

滑坡名称	滑坡类型	体积 $\times 10^{-4}/\text{m}^3$	$C_1$	$C_2$	$C_3/\text{m}$	$C_4$	$C_5/\text{m}$	$C_6/(\text{°})$	$C_7$	$C_8/\%$	等级	现场结果
SL10	堆积体滑坡	1 294	极破碎	松散体	<0	不稳体、裂缝	220	29	3	0	轻度	轻度 4
SL9	顺层滑坡	639	较破碎	顺向坡	141	局部崩塌	309	30	3	1.3	中度	中度 1
SLD6	顺层滑坡	67.8	较完整	顺向坡	71	破碎带、变形体	154	35	3	6	轻度	轻度 4
SLD5	堆积体滑坡	11 146	极破碎	碎裂体	<0	无	590	33	3	22	中度	中度 1
SL2	顺层滑坡	722	较破碎	顺向坡	<0	危岩体	254	15	3	54	重度	重度
SLD1	切层滑坡	34.5	破碎	横向坡	298	表层流水迹象	145	27	4	55	重度	中度
SL8	切层滑坡	396.1	较破碎	逆向坡	175	小型褶曲	220	23	4	50	重度	中度
SL7	切层滑坡	570.6	破碎	逆向坡	58	拉张裂隙	149	26	3	100	极度	极度
SL6	切层滑坡	3 370	较破碎	逆向坡	<0	挤压带、危岩体	633	30	3	2	轻度	轻度
SL5	切层滑坡	1 470	较完整	逆向坡	356	缓倾裂隙	200	28	4	100	轻度	轻度
SL4	顺层滑坡	1 130	破碎	斜向坡	<0	冲沟、流水	230	29	3	22	轻度	轻度
SL3	顺层滑坡	28 900	破碎	斜向坡	260	冲沟、流水	580	20	3	0	轻度	轻度
SLD2	顺层滑坡	2 482	较完整	顺向坡	<0	隆起、流水	365	19	3	0	中度	轻度
SLD4	切层滑坡	5.0	较破碎	横向坡	174	无	70	21	3	0	中度	中度
SL1	堆积层滑坡	16.7	极破碎	松散体	313	无	177	27	3	100	重度	重度
CO7	堆积体滑坡	430	破碎	碎裂体	<0	无	120	33	3	0	轻度	轻度
CO6	堆积体滑坡	74.5	破碎	碎裂体	119	无	60	26	3	0	轻度	轻度
CO4	堆积体滑坡	127	破碎	碎裂体	847	无	254	32	3	0	中度	中度
CO2	堆积体滑坡	1 490	较破碎	逆向坡	<0	表层崩塌	303	22	3	40	中度	中度
CO5	堆积体滑坡	276.4	破碎	碎裂体	1 397	无	121	29	3	0	中度	中度
CO3	堆积体滑坡	156	较完整	顺向坡	<0	危岩体	82	42	4	100	中度	中度
CO1	堆积体滑坡	190.1	较破碎	顺向坡	739	危岩体	123	17	3	100	中度	中度

#### 3.2 评价结果分析

从表 1 可以看出,研究区的 22 处大型滑坡基本处于稳定性较高的状态,以轻度、中度敏感为主.其中处于轻度敏感 8 处,中度敏感 9 处,重度敏感 4 处,极度敏感 1 处. SL7 滑坡(图 4)岩体破碎,蓄水后完全淹没,聚类分析结果为极度敏感.

SL2 滑坡等 4 处岩体较破碎,变形特征比较明显,蓄水后前缘淹没,稳定性较差. SL9 滑坡等 9 处岩体较破碎,有变形特征,蓄水后浸润坡脚,分析结果为中度敏感. SL3 滑坡(图 5)等 8 处岩体较完整,变形特征不明显,聚类分析结果为轻度敏感.

本文评价结果与现场的实际调查基本吻合,



图 4 SL7 滑坡全貌

Fig. 4 Full view of landslide SL7

仅 SLD1, SL8, SLD2 三处的敏感性等级高于现场调查, 结合现场情况, 这三处滑坡岩体都比较破碎, 且变形现象较明显, 表层有流水、隆起等现象, 但对于滑坡防治灾害而言, 结果偏于安全, 有利于防灾观念的提高。

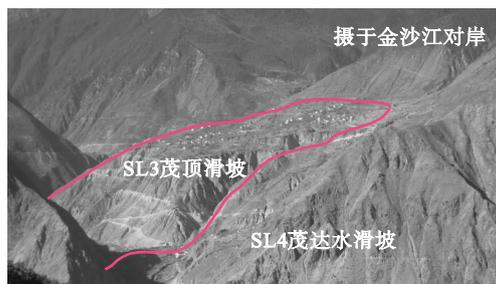


图5 SL3 滑坡全貌

Fig. 5 Full view of landslide SL3

## 4 结 论

1) 基于熵值法和 K-PSO 聚类算法进行滑坡敏感性评价, 可知研究区滑坡处于轻度敏感 8 处, 中度敏感 9 处, 重度敏感 4 处, 极度敏感 1 处。

2) 评价结果与现场调查情况基本一致, 仅 SLD1, SL8 和 SLD2 三处滑坡的敏感性高于现场, 结果偏于保守, 但更加吻合现场调查情况, 验证了该方法的准确性。

3) 影响滑坡敏感性的因子具有不确定、非线性和模糊性, 利用聚类模型来评价滑坡敏感性可以有效地解决因子之间不确定问题, 尤其是在地质资料缺少的条件下, 该方法可以快速合理地分析滑坡的敏感性。

## 参考文献:

[1] Baeza C, Corominas J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques [J]. *Earth Surface Processes and Landforms*, 2001, 26: 1251-1263.

[2] Jeong G C, Kim K S, Choo C O, et al. Characteristics of landslides induced by a debris flow at different geology with emphasis on clay mineralogy in South Korea [J]. *Natural Hazards*, 2011, 59: 347-365.

[3] Carrara A. GIS techniques and statistical models in evaluating landslide hazard [J]. *Earth Surface Process and Landform*, 1991, 16: 427-445.

[4] Biswajeet P, Lee S. Utilization of optical remote sensing data and GIS tools for regional landslide hazard analysis using artificial neural network model [J]. *Earth Science Frontiers*, 2007, 14(6): 143-151.

[5] Neaupane K M, Piantanakulchai M. Analytic network process model for landslide hazard zonation [J]. *Engineering Geology*, 2006, 85(3/4): 281-294.

[6] 唐川, 周钜, 朱静. 云南崩塌滑坡危险度分区的模糊综合解析法 [J]. *水土保持学报*, 1994, 8(4): 48-54. (Tang Chuan, Zhou Ju, Zhu Jing. The method of fuzzy comprehensive analysis of the risk zoning of rockfall and landslide in Yunnan [J]. *Journal of Soil and Water Conservation*, 1994, 8(4): 48-54.)

[7] Xu C, Dai F C, Xu X W, et al. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang river watershed, China [J]. *Geomorphology*, 2012, 145/146: 70-80.

[8] 李军霞, 王常明, 王钢城. 基于组合赋权-未确知测度理论的滑坡危险性评价 [J]. *岩土力学*, 2013, 34(2): 468-474. (Li Jun-xia, Wang Chang-ming, Wang Gang-cheng. Landslide risk assessment based on combination weighting-uncertain measure theory [J]. *Rock and Soil Mechanics*, 2013, 34(2): 468-474.)

[9] Wan S, Lei T C, Chou T Y. A landslide expert system: image classification through integration of data mining approaches for multicategory analysis [J]. *International Journal of Geographical Information Sciences*, 2012, 26(4): 747-770.

[10] 张晨, 王清, 陈剑平, 等. 金沙江流域泥石流的组合赋权法危险性评价 [J]. *岩土力学*, 2011, 32(3): 831-836. (Zhang Chen, Wang Qing, Chen Jian-ping, et al. Evaluation of debris flow risk in Jinsha River based on combination weighting process [J]. *Rock and Soil Mechanics*, 2011, 32(3): 831-836.)

[11] Kennedy J, Eberhart R C. Particle swarm optimization [C]// Proceedings of the IEEE International Conference on Neural Networks. Perth, 1995: 1942-1948.

[12] Li Y Y, Wang Q, Chen J P, et al. K-means algorithm based on particle swarm optimization for the identification of rock discontinuity sets [J]. *Rock Mechanics and Rock Engineering*, 2015, 48: 375-385.

[13] 周福军. 日冕水电站库区滑坡稳定性早期智能判别及危害模糊综合预测研究 [D]. 长春: 吉林大学, 2013. (Zhou Fu-jun. Early intelligence identification methods of stability and fuzzy synthetic prediction of hazard for landslides in Rimian hydropower station reservoir [D]. Changchun: Jilin University, 2013.)