

文章编号:1673-0836(2006)08-1451-04

基于逻辑回归法的滑坡危险度区划研究^{*}

王志旺^{1,2}, 廖勇龙², 李端有²

(1. 中国地质大学 地质过程与矿产资源国家重点实验室, 武汉 430074;

2 长江科学院 水利部岩土力学与工程重点实验室, 武汉 430010)

摘要:滑坡危险度区划能降低由于滑坡所造成的危害,越来越受到重视。结合长江三峡库区秭归-巴东一带滑坡灾害发生的环境和基本特征,选取地层岩性、断裂构造、坡度、坡向、距离河流远近、沟壑密度、高程、植被条件等作为研究区滑坡灾害危险度分区评价因子,利用GIS的空间分析功能将各因素图件栅格化为400 m×400 m的单元共7461个。采用逻辑回归方法对研究区滑坡灾害进行了危险度分区定量评价,评价结果与该区现有滑坡的分布情况比较吻合。

关键词:滑坡;危险度区划;地理信息系统;逻辑回归

中图分类号:TV698.1

文献标识码:A

Zonation of Landslide Hazards Based on Logistic Regression

WANG Zhi-wang^{1,2}, LIAO Yong-long², LI Duan-you²

(1. State Key Laboratory of Geological Processes and Mineral Resources,
China University of Geosciences, Wuhan 430074, China;

2. Key Laboratory of Geotechnical Mechanics and Engineering of Ministry of Water Resources,
Yangtze River Scientific Research Institute, Wuhan 430010, China)

Abstract: The study on hazard area zoning is a very useful method to reduce the harm of landslide. Based on the environment and characters of landslides in the region between Zigui and Badong in the TGP's reservoir area, the main factors including lithology, fault, slope angle, slope aspect, elevation, drainage network, distance to river and distribution of plant are derived for analyzing zonation of the landslide hazards in the studied area. Using the spatial analysis technology of GIS, each map of factors is divided into 7461 units measured at 400 m×400 m. The statistic method called logistic regression has been used to create a landslide probability prediction map for the studied area. The research result is very similar to the actual field conditions.

Keywords: landslide; hazard zoning; GIS; logistic regression

1 引言

滑坡危险度区划研究属于滑坡空间预测的一个分支,它是将一定区域内产生滑坡的基础,引起滑坡发生的外界因素,以及滑坡所造成的灾害综合统计,定量或半定量地评价区域内滑坡危害性等级

程度即滑坡危险度分区。滑坡危险区划理论的研究开始于20世纪60年代,在美国以及一些西欧国家为了合理的开发和利用斜坡土地资源,利用敏感因素对斜坡的危险性进行分带。日本采用地震、坡度以及降雨等因素对滑坡进行空间预测。由于滑坡危险度区划能降低滑坡所造成的危害,减少人们

• 收稿日期:2006-08-20(修改稿)

作者简介:王志旺(1971-),男,湖北新洲人,高级工程师,在职博士研究生,主要从事滑坡地质灾害监测技术研究工作。E-mail:wzwhbwh@126.com。

基金项目:科技部社会公益研究专项资金面上项目(2002DIB10044)

生命财产所受到的威胁,并指导国民生产,因此越来越受到各国政府的重视,并逐步将其纳入法规范畴。

我国滑坡灾害分布广泛,发生频繁,且灾情严重。从我国国情出发,不可能同时对所有滑坡灾害进行工程治理,滑坡的治理只能针对危害对象分级分期进行。因此,加强滑坡灾害的前期科学论证工作,科学地划分滑坡危险度,是防灾减灾工作的重要环节。

滑坡危险度区划的数学模型已有很多,如应用较多的有信息量模型、专家评分模型等。本文采用了逻辑回归方法,对长江三峡库区秭归-巴东一带进行了滑坡危险度区划研究。

2 逻辑回归

线性回归模型是比较常用的定量分析的统计方法,然而在许多情况下,线性回归会受到限制,例如,当因变量是一个分类变量而不是一个连续变量时,线性回归就不适用了。对于分类变量,经常采用逻辑回归分析方法。

逻辑回归分析是对定性变量的回归分析,其逻辑函数的形式为:

$$f(x) = \frac{e^x}{1 + e^x} \quad (1)$$

根据离散型随机变量期望值的定义,用 P 代表自变量为 x 时 $y=l$ 的概率,可得:

$$f(p) = \frac{e^p}{1 + e^p} = \frac{e^{(a+b_1x_1+b_2x_2+\dots+b_nx_n)}}{1 + e^{(a+b_1x_1+b_2x_2+\dots+b_nx_n)}} \quad (2)$$

由于函数 $f(p)$ 对 x 的变化在 $f(p)=0$ 或 $f(p)=1$ 的附近是不敏感的、缓慢的,并且非线性程度较高,因此引入 $f(p)$ 的 Logistic 变换,即

$$\begin{aligned} \text{logit}(p) &= \text{logit}\left(\frac{p}{1-p}\right) \\ &= (a + b_1x_1 + b_2x_2 + \dots + b_nx_n) \end{aligned} \quad (3)$$

此时,可计算 P 的概率:

$$p = \frac{e^{(a+b_1x_1+b_2x_2+\dots+b_nx_n)}}{1 + e^{(a+b_1x_1+b_2x_2+\dots+b_nx_n)}} \quad (4)$$

由于滑坡作用的影响因素中有些地质因素是可以量化的,有些是难以量化的而宜采用定性方式进行表达,并且滑坡发生与否可以用 0 或 1 表示,故需要采用逻辑回归分析方法进行分析研究。

3 基于逻辑回归法的滑坡危险度区划

3.1 研究区概况

研究区为三峡库区秭归-巴东一带(面积为

1188.9 km²),区内滑坡地质灾害较为发育,有新滩、黄腊石等大型滑坡体。滑坡主要发生在沿河流两岸,在软硬岩石互层的分布地段,地形较陡,经河流冲刷、山洪冲蚀以及雨水的侵蚀剥蚀,常常造成岩层间的滑动而形成。

区内地层出露齐全,发育良好,除缺失泥盆系下统、石炭系上统、白垩系的一部分和第三系外,自前震旦系至第四系均有出露,从东向西依次出露前震旦纪、震旦纪、寒武纪、奥陶纪、志留纪、泥盆纪、石炭纪、二叠纪和三叠纪地层;白垩纪地层仅在周坪一带出露,第四系堆积物零星覆盖于河流阶地、剥夷面上及斜坡地带。

研究区处于上扬子台褶带,除北北东向的黄陵背斜和秭归向斜外,断裂较为发育。与库岸稳定性有关的主要有九湾溪、仙女山、天阳坪和新华断裂等。

采用 400 m × 400 m 的单元对研究区进行单元划分,共划分 7461 个单元。

3.2 滑坡危险度分区评价因子选择与分级

滑坡是各种内外影响因素综合作用的产物,因此滑坡影响因素的选取是危险度区划研究的基础。结合研究区自然地理特征,在参阅以往研究和实地考察的基础上,选取地层岩性、断裂构造、坡度、坡向、距离河流远近(河流缓冲区)、水系发育程度(沟壑密度)、高程、植被条件(植被指数)等作为研究区滑坡灾害危险度分区评价因子。滑坡危险度分区评价因子提取框架图见图 1。各种评价因子经过矢量化形成各种图层后,还要重新分级,分成若干个区段,才能应用逻辑回归分析方法进行滑坡危险度分区研究。各种图层分组情况见表 1。

3.2.1 地层岩性

地层岩性是发生滑坡的物质基础。根据前人的研究表明,研究区与滑坡相关性较大的地层主要为侏罗系上统遂宁组(J_s)、侏罗系中统聂家山组(J_n)、侏罗系中统沙溪庙组二段(J_s2)、侏罗系下统桐竹园组(J_t)、三叠系中统巴东组(T_b)、三叠系下统嘉陵江组(T_js)。此外,二叠系与石炭、泥盆、志留系的接触面附近滑坡也较为发育。因此,地层岩性主要考虑上述对滑坡发育影响较大的地层,并对二叠系与石炭、泥盆、志留系的分界线取 500 m 的缓冲区。

地层岩性因子的提取主要通过将地质图进行扫描,然后进行几何校正,建立地理座标系统。再对各种地层边界线进行数字化(矢量化),并分别对

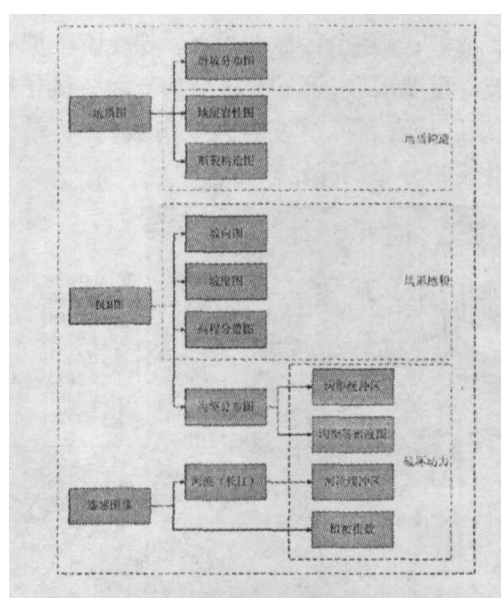


图 1 滑坡危险度分区评价因子提取框架图

Fig. 1 The flow chart of picking up factors on analyzing zonation of landslide hazards

各种地层赋以相应的属性,建立地层的矢量化图层。

3.2.2 断裂构造

滑坡的发育一般都与断裂构造密切相关,尤其是在区域性断裂构造的交叉复合部位,由于岩石较为破碎,常常形成有利于滑坡形成和发育的构造条件。从研究区滑坡分布图和断层分布图可以看出,

在 NNE 向水田坝断层 - 黄荆树坪断层与 NNW 向仙女山断层的交汇部位以及 NNE 向刘家坪 - 观垭断层与 NNW 向张家河断层的交汇部位附近,滑坡较为发育。因此,除考虑所以断裂构造的 1m 的缓冲区外,还分别建立 NNW 和 NNE 这两组断裂交汇部位的 1 m、2 km、3 km、4 km、5 km 的缓冲区。

断裂构造指标的提取也是通过将地质图进行扫描,然后进行几何校正,建立地理座标系统。再对各种断裂构造线进行矢量化,并分别对各种断裂构造线赋以相应的属性,建立断裂构造的矢量化图层。

3.2.3 坡度

坡度是滑坡发生的一个主要因素,它与土层厚薄、气候条件、水文条件、岩性条件等许多因素密切相关。绝大多数滑坡危险性评估都利用坡度因素作为几个重要的自变量之一。一方面,随着坡度的增加,包括重力在内的剪切力增大,相应的滑坡发生的概率也会增大;另一方面,坡度较大的地方往往崩塌会为滑坡的发生提供大量的物质来源,从而为滑坡发生创造了条件。通过数字高程模型 DEM,提取研究示范区的坡度图,并按表 1 进行重新分级。

表 1 各图层分组情况表

Table 1 Reclassification of the analyzing factors

因子图层	类别及其重新分类后的对应值							
地层岩性 *	<i>J_s</i>	<i>J_n</i>	<i>J_{s2}</i>	<i>J_t</i>	<i>T_b</i>	<i>T_{js}</i>	<i>P_{cds}</i>	
	1	2	3	4	5	6	7	
重要断裂交汇	0~1	1~2	2~3	3~4	4~5			
部位缓冲区(km)	1	2	3	4	5			
到长江的距离(km)	0~0.5	0.5~1	1~1.5	1.5~2	2~2.5	2.5~3		
	1	2	3	4	5	6		
海拔高度	0~100	100~500	500~1000	1000~1500	>1500			
(m)	1	2	3	4	5			
沟壑密度	5~262	262~519	519~776	776~1033	1033~1291	1291~1548	1548~1805	1805~2320
	1	2	3	4	5	6	7	8
坡向	0~45	45~90	90~135	135~180	180~225	225~270	270~315	315~360
(°)	1	2	3	4	5	5	6	8
坡度	0~5	5~20	20~35	35~50	50~65	65~80	80~90	
(°)	1	2	3	4	5	6	7	
植被指数	105~160	160~210	210~260					
	1	2	3					

* J_s 、 J_n 、 J_{s2} 、 J_t 、 T_b 、 T_{js} 分别对应于表 1 中的地层 (遂宁组、聂家山组、沙溪庙组二段、桐竹园组、巴东组、嘉陵江组), P_{cds} 为二叠系与石炭、泥盆、志留系的分界线 500 m 的缓冲区。

3.2.4 坡向坡向指山坡朝向。由于朝向不同,山坡的小气候和水热比情况有规律性差异,导致不同朝向坡面上自然地理诸要素的规律性分异,也往往

存在滑坡分布的规律性分异。通过数字高程模型 DEM,提取研究示范区的坡向图,并按表达进行重新分类分级。

3.2.5 距离河流远近(河流缓冲区)

河流的侵蚀作用是岸坡破坏、演变的主要外力之一,因此有必要考虑距离河流远近这一影响因素。

对 SPOT 卫星影像进行遥感分类,提取水体(河流)的信息,将水体(河流)栅格分类数据转换为矢量数据。由于本研究区的河流侵蚀作用主要表现为主要河流(长江)的侵蚀作用,因此,这里主要考虑距离长江的远近指标,在水体(河流)矢量数据的基础上,分别建立主要河流(长江)的 500 m、1 000 m、1 500 m、2 000 m、2 500 m、3 000 m 缓冲区。

3.2.6 水系发育程度(沟壑密度)

指深切河谷(支流)在采样区内的分布密度。沟壑是由地表水的侵蚀、搬运和堆积作用形成的地貌景观,也是所在地区内外地质营力作用的产物。沟壑的形成和发育与气候、地层岩性、地质构造等诸因素有关。沟壑密度越大,为滑坡创造的空间条件越多。

通过数字高程模型 DEM,提取示范研究区的水系图,然后对研究示范区按照一定的网格大小进行网格化,分别按照每个网格单元的水系长度计算水系等密度图,即为沟壑密度。在此基础上,按照表 1 进行分级。

3.2.7 高程

根据文献的统计研究表明,地层岩性及河谷切割深度控制着谷坡的高度,从而控制滑坡的后缘高程。在三峡河谷段,滑坡的后缘高程在 500 ~ 1000 m 范围内,大部分在 600 m 上下;川江河谷段,该高程逐渐下降至 300 ~ 400 m 之间;三峡与川江河谷交接地段,滑坡的高程在 400 ~ 500 m 之间。因此,滑坡与高程分布的关系是十分密切的。

通过数字高程模型 DEM,按照表 1 及大于 1500 m 的分级提取研究区的高程分带图。(8) 植被条件(植被指数)

研究表明,滑坡的发育常常与地下水及植被的发育程度有关,而植被指数是对地表附近含水量及植被发育程度的综合反映,因此,植被指数与滑坡的发育和分布是密切相关的。利用 SPOT 卫星影像数据,提取归一化植被指数 NDVI,并进行分级。

3.3 滑坡危险度区划结果及评价

图 2 为采用逻辑回归法的滑坡危险度区划结果。从图中可以看出,在研究区内主要有三个区域发生滑坡的危险性较大。一个区域为泄滩 - 黄腊石沿长江一带,该区域有黄腊石滑坡和范家坪滑坡

等重要滑坡;另一个区域为香溪一带,该区域有香溪滑坡等重要滑坡;第三个区域为新滩 - 链子崖一带,在该区域有新滩滑坡等大型滑坡;分析结果与现有滑坡的分布情况比较吻合。

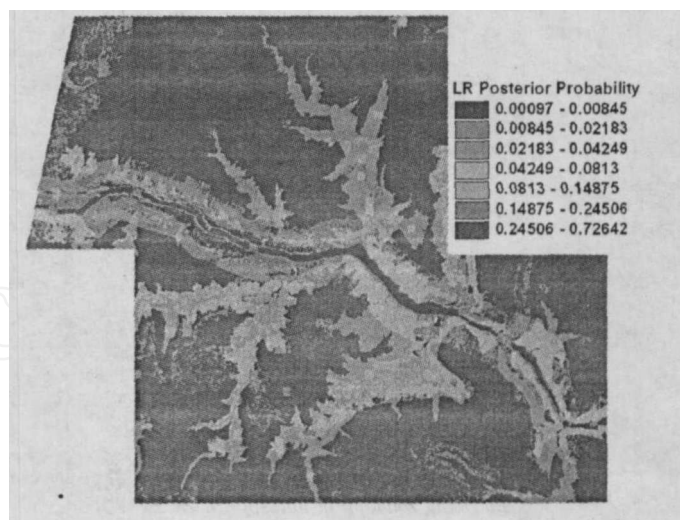


图 2 滑坡危险度区划结果

Fig. 2 Zonation of landslide hazards in the studied area

4 结语

采用逻辑回归方法对长江三峡库区秭归 - 巴东一带进行了滑坡危险度区划分析研究,分析结果与现有滑坡的分布情况比较吻合。研究表明,逻辑回归法作为一种数据驱动方法,可以较好地避免滑坡危险度区划研究中评价因子选择及权重赋值的主观性,因此具有较高的评价精度。

参考文献:

- [1] 杜榕恒,刘新民,袁建模,李天池. 长江三峡工程库区滑坡与泥石流研究[M]. 成都:四川科学技术出版社,1991.4
- [2] F. P. Agterberg, G. F. Bonham - Carter, Q. Cheng and D. F. Wright Weights of Evidence Modeling and Weighted Logistic Regression for Mineral Potential Mapping. From Davis, J. C. and Herzfeld, U. C. (eds.), 1993, Computers in Geology, 25 Years of Progress, Oxford University Press, Oxford, pp. 13 - 32
- [3] 王济川,郭志刚. Logistic 回归模型—方法与应用[M]. 北京:高等教育出版社,2001.9
- [4] 李端有,王志旺. 滑坡危险度区划研究初探[J]. 长江科学院院报, 2005, 22(4): 41 - 44