

基于结果验证的信息量法地质灾害易发性评价

——以贵州省开阳县为例

范 强¹, 巨能攀¹, 向喜琼², 黄 健¹

(1. 成都理工大学 地质灾害防治与地质环境保护国家重点实验室, 四川 成都 610059; 2. 贵州大学 喀斯特重点实验室 贵州 贵阳 550025)

摘要:传统信息量法将所有评价因子的信息量值进行综合,未对评价因子作优化选择,选择带有主观性。鉴于此,采用结果验证对信息量法进行改进,选取工程岩组、断层、河流、道路、坡度、坡向、剖面曲率7个评价因子。首先利用地质灾害编录数据与评价因子进行空间叠加分析,计算单因子信息量,然后利用成功率曲线法评价单因子评价结果的模型拟合能力,按照模型拟合能力从高到低将单因子信息量排序,并依次进行总信息量试算,得到7组试算结果。最后对7组试算结果进行模型拟合能力评价,确定AUC为前4的因子组合为最优的因子组合。

关键词:信息量法; 结果验证; 地质灾害; 易发性评价; 开阳县

中图法分类号:P642

文献标志码:A

DOI:10.16232/j.cnki.1001-4179.2015.15.016

地质灾害作为一种自然现象,具有突发性、危害性大,且不易预测等特点,故对地质灾害的易发性评价成为工程界关注的课题。国内外不少学者对此进行了大量卓有成效的工作和研究,许多方法被用于地质灾害易发性评价,概括起来可分为两种类型^[1]:①定性方法;②定量方法。定性评价方法,也称为知识驱动方法,是以专家在野外调查的基础上形成的先验知识,并以此来判断哪些情况下有利于地质灾害的发生,进而对影响因素进行权重分配,最终将各影响因素综合而得地质灾害易发性图。定量方法也称为数据驱动方法,多为统计方法,是将滑坡灾害编录数据与相关评价因子进行空间关联分析,采用一定的数学模型,研究灾害发生与评价因子之间的关系,从而确定各评价因子的权重,最后叠加各个评价因子得到易发性图。信息量法作为一种基于统计学的定量方法,被广泛应用且被认为是一种有效的地质灾害易发性评价方法^[2-4]。但纵观该方法的应用,均是把所有评价因子的信息量进行综合,鲜有对评价因子进行优化选择,这就为评价因子的选择带来主观性。本文采用对单个评价因子评价结果进行验证,按照评价因子的模型拟合能力从高

到低进行总信息量试算,并选择最优组合,以期克服因子选择的主观性问题。本文选择西南山区贵州省开阳县为研究区,应用基于结果验证的信息量法进行地质灾害易发性评价。

1 研究区概况

开阳县位于贵州省中部,总面积2 026 km²,总人口43万,南距省会城市贵阳66 km²,北距历史名城遵义110 km²,位于连接贵阳与遵义两大城市的次中心区域,区位优势明显;地处云贵高原梯级斜坡地带,属苗岭山脉,地貌以山地、丘陵为主,山地、丘陵、盆地、河谷交错分布。地势总体上由西南向东北倾斜,南部多为侵蚀脊状中低山,北部多为峰丛谷地及河谷阶地,坡度较缓,平均海拔约1 080 m。研究区内地层除奥陶系中上统及部分石炭系地层外,从前震旦系到第四系均有出露。其中,寒武系地层分布最广,其厚度1 710 ~ 2 053 m,出露面积1 479 km²,占总面积的73.0%。研究区具有典型的喀斯特地貌特征,沟壑纵横,河流深切,岩石风化较为严重,局部地区有石漠化现象。区内构造活动强烈,断层方向以北东向居多。研究区属于

亚热带季风气候,多年平均气温 13℃,年平均降雨量在 1 258 mm 左右,每年的 4~10 月降雨集中,占年内降水量的 82.0%。同时研究区内矿产资源和水资源较为丰富,也从一定程度上加剧了地质灾害的发生。

2 评价方法与数据

2.1 方法介绍

信息量法在本质上属于统计分析方法^[4],是一种数据驱动的定量方法。信息量法在矿产资源预测^[5-6]、斜坡稳定性预测^[4,7]、地质灾害危险性评价^[8-11]等方面均有应用。信息量法应用于地质灾害危险性评价的主要思路为:首先由专家根据经验知识选取地质灾害的影响因子(坡度、地层岩性等),再将已经发生的灾害点与各个影响因子叠加分析,计算各个影响因子与地质灾害之间的关系,并将其转化为影响区域斜坡稳定性的信息量,最后将所有影响因子的信息量求和,得到每个评价单元总的信息量值。

GIS 支持下,信息量值的计算过程如下。

(1) 计算评价因子信息量。将研究区分成 S 个大小相等的评价单元,经过专家知识选取了 m 个评价因子(F),每个评价因子有 n 个等级。第 i 个评价因子的第 j 个等级(F_i^j)的信息量计算公式为

$$I(F_i^j) = \ln \frac{N_i^j/N}{S_i^j/S} \quad (1)$$

式中, N_i^j 为分布在影响因素 F_i^j 内的灾害单元数, N 为研究区内地质灾害分布的单元总数, S_i^j 为研究区内含有影响因素 F_i^j 的单元数, S 为研究区所划分的评价单元总数。

(2) 计算评价单元总的信息量。将所有评价因子的信息量进行空间叠加,即得每个单元总的信息量(式 2)。该值用以表示对地质灾害产生贡献率大小,信息量值越大,表示越容易发生地质灾害,反之亦然。

$$I = \sum_{i=0}^n I(F_i) \quad (2)$$

信息量法采用概率分析,计算评价因子等级对地质灾害的贡献大小,客观地揭示了评价因子和地质灾害之间的关系,最终将所有评价因子的信息量进行综合,作为地质灾害易发程度的表征。传统的应用未对根据专家经验选取的评价因子做出优化选择,难免带有一定的主观性。并且地质灾害具有很强的地域特性,一个地方的经验知识不一定适合另一个区域的实际情况。因此,对评价因子进行优化选择,可以弥补评价因子选取的主观性缺陷。为选取最合适的因子组合,本文采取对单因子评价结果逐个验证的方法。首

先采用成功率曲线法对每个单因子的模型拟合效率进行评价,即将每个单因子的信息量计算结果作为最终的信息量,然后和已知灾害点进行叠加分析,得到每个灾害点所处单元的信息量。根据信息量值从高到低累积面积比和相应信息量值所处单元中的灾害点的百分比作曲线图,曲线下方的面积(AUC)越大,说明该评价因子的模型拟合能力越好。该方法的实质是认为已发生灾害点的评价单元,计算出来的信息量越大越好。根据每个评价因子计算的 AUC ,将所有评价因子排序(F_1, F_2, \dots, F_n),按照 AUC 从高到低依次选择评价因子,得到 N 个因子组合。然后对 N 个因子组合进行总信息量的计算。最后再次采用成功率曲线法,选择最优组合,完成评价因子的选取。

2.2 数据及预处理

采用的数据包括:1:5 万地质图;1:5 万地形图;Aster GDEM;TM 图像;从地质图中提取的地层岩性图层、断层图层;从地形图中提取的道路图层、水系图层;从 DEM 数据生成坡度、坡向、剖面曲率。

地层岩性影响了斜坡的破坏模式,不同岩组的岩性不同,受风化程度各异,对产生地质灾害的贡献也不同,本文根据前人经验,结合贵州省岩土分类标准,按照岩土体的工程特性将从地质图中提取的地层单元归类为硬质岩体、较硬质岩体、软硬互层岩体和软质岩体。断层构造对地质灾害的影响主要有两个方面:其一是断层活动使其周围岩石较为破碎,在一定程度降低了岩土体的自身强度,同时也有利于雨水入渗,进一步降低坡体的稳定性。其二断层活动是有利于形成地质灾害发生的地形条件。同时,断层对地质灾害的影响具有距离效应,即距离断层越近,所受影响越大。故本文将根据断层线图层生成距离断层、距离图层,每个评价单元的值距它最近的断层的距离值。道路修建会产生切坡及填坡,改变坡体内部的原始应力状态,使坡体趋于不稳定。流水对河道产生侵蚀作用,尤其是在山区,河流对岸坡的不断掏蚀,形成众多的临空面,不断改变坡体内部应力状态,增加了岸坡失稳的可能性。道路和河流对地质灾害的影响也具有距离效应,同样据道路线图层、河流图层生成距道路距离、距河流距离。坡度是影响斜坡稳定的重要因素之一,坡度越陡,发生斜坡失稳的几率就越大,坡向决定了坡体日照时点及时长,在一定程度上影响了斜坡岩体的风化速度,同时不同坡向的斜坡可能形成不一样的局部小气候,进而影响了地质灾害的发生,本文采用的坡度及坡向来自于 DEM 数据,应用 ArcGIS 中 Slope、Aspect 工具生成。剖面曲率是对斜坡沿最大坡降方向地面高程变化率的度量,反映了斜坡的形态,不同形态的斜坡其

稳定性不同,一般认为凸形坡较凹形坡更容易发生地质灾害,本文所用剖面曲率数据应用 ArcGIS 中 Curvature 工具生成。各评价因子图层的具体分级见表 1。

表 1 评价因子信息量计算结果

项目	指标	面积比	地质灾害比	信息量
工程岩组	硬质岩体	0.43412	0.106383	-1.40628
	较硬质岩体	0.115884	0.276596	0.869964
	软硬互层岩体	0.38169	0.537234	0.341824
	软质岩体	0.068305	0.079787	0.012277
距断层距离(m)	0~200	0.188033	0.212766	0.123573
	200~400	0.149835	0.191489	0.245294
	400~600	0.124512	0.074468	-0.51403
	600~800	0.096615	0.06383	-0.41451
	800~1000	0.081293	0.037234	-0.78084
	>1000	0.359712	0.420213	0.155458
距河流距离(m)	0~200	0.172134	0.196809	0.133956
	200~400	0.156596	0.239362	0.424307
	400~600	0.144661	0.154255	0.064216
	600~800	0.120827	0.143617	0.172788
	800~1000	0.103355	0.085106	-0.19427
	>1000	0.302426	0.180851	-0.51416
距道路距离(m)	0~200	0.125363	0.223404	0.577767
	200~400	0.115043	0.180851	0.452365
	400~600	0.10833	0.132979	0.20501
	600~800	0.095995	0.111702	0.151543
	800~1000	0.088175	0.053191	-0.50542
	>1000	0.467094	0.297872	-0.44987
坡度(°)	<10	0.408338	0.191489	-0.75726
	10~20	0.39693	0.37766	-0.04977
	20~30	0.140801	0.287234	0.712949
	30~40	0.040828	0.12234	1.097435
	>40	0.013103	0.021277	0.4848
剖面曲率	<-0.5	0.031226	0.015957	-0.67134
	-0.5~-0.1	0.269713	0.234043	-0.14185
	-0.1~0.1	0.346541	0.393617	0.127378
	0.1~0.5	0.316476	0.308511	-0.02549
	>0.5	0.036044	0.047872	0.283799
坡向	N	0.131978	0.074468	-0.57226
	NE	0.135384	0.143617	0.059036
	E	0.137173	0.159574	0.151268
	SE	0.128657	0.095745	-0.29546
	S	0.112457	0.095745	-0.16088
	SW	0.107758	0.095745	-0.1182
	W	0.117118	0.143617	0.203968
	NW	0.129476	0.191489	0.391339

3 易发性评价

3.1 单因子评价结果

根据公式(1),分别对 7 个评价因子计算信息量,计算结果见表 1。对于工程岩组,其中硬质岩体的信息量值为负值,说明在硬质岩体地层中,不易发生地质灾害,这和人研究成果相符合。在容易发生地质灾害的工程岩组中,较硬质岩体比软硬互层岩体更容易

发生地质灾害。断层、河流、道路对地质灾害发生的影响,均表现出一定的距离效应,和前人研究成果一致。坡度小于 20°不易发生地质灾害,坡度在 20°~40°之间,是地质灾害主要的发生地带。当坡度大于 40°时,信息量值呈负值,说明较大的坡度反而不易发生地质灾害,而一般认为坡度越大,越利于发生地质灾害,笔者分析产生这一结果的原因有二:① 地质灾害编录数据的精度不够。在坡度很大的地方,往往人迹罕至,存在发生地质灾害但没有被发现的情况。② 地质灾害信息提取精度不够准确。由于地质灾害数据精度限制,以一个点表示地质灾害所在位置,然后以该点所在评价单元的坡度作为该地质灾害的坡度,存在点的位置位于地质灾害体上较为平缓的部位的情况。坡向为 NE、E、W、NW 的斜坡更容易发生地质灾害。

根据表 1 的计算结果,将各评价因子重新分类,其栅格值为对应的信息量值。将地质灾害点数据与各单因子信息量图层叠加,并绘制各单因子易发性图的成功率曲线,见图 1。

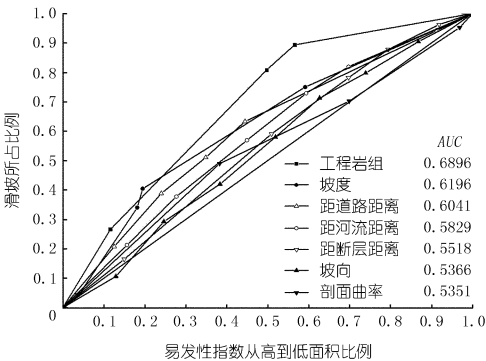


图 1 单因子易发性指数成功率曲线

3.2 最佳因子组合评价

根据图 1 可知,AUC 值从高到低依次为:工程岩组、坡度、距道路距离、距河流距离、距断层距离、坡向、剖面曲率。按照 AUC 值从高到低选择因子组合,共有 7 种组合,依据公式(2)对 7 种组合分别计算总信息量,并与地质灾害点数据叠加,绘制总信息量成功率曲线图如图 2。

从图中可以看出,参与最终信息量值计算的评价因子小于 4 个时,每增加一个评价因子,其模型拟合能力有所提高,当前 5 个评价因子的参阅总信息量计算时,其模型拟合能力开始降低。因此,认为将单因子 AUC 从高到低前 4 的评价因子参与总信息量计算为最优选择。根据公式(2),将工程岩组、坡度、距道路距离、距河流距离作为最终评价因子,计算得到研究区的地质灾害易发性图(见图 3)。

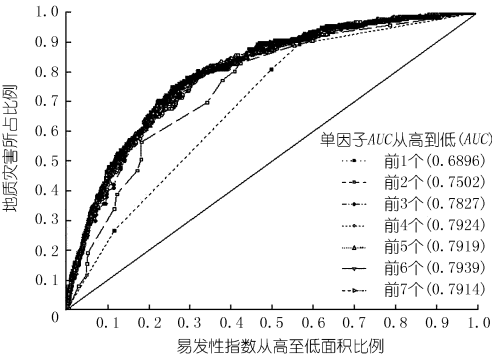


图2 总信息量试算成功率曲线

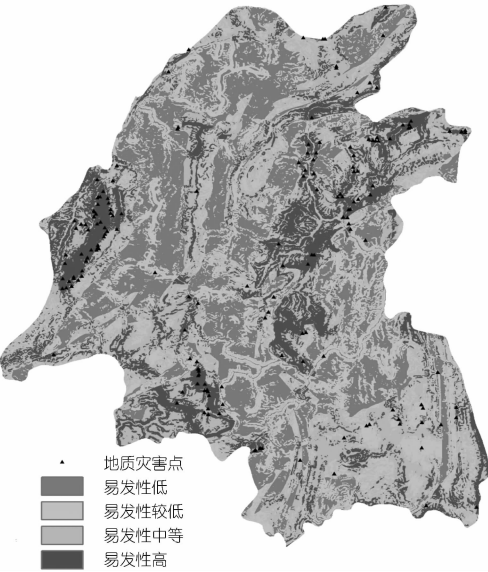


图3 研究区地质灾害易发性图

4 结 论

本次评价过程中,获得以下两点认识:① 地质灾

害易发性评价的评价因子的数量不是越多越好,如本文采用 7 个评价因子的模型拟合能力不及采用 4 个评价因子。因此,对于初步选择的评价因子一定要进行优化选择。② 采用对单因子评价结果进行验证的方法是可行的,且是较有成效的,该方法可以推广到其它的应用中。

参考文献:

[1] 谢洪斌,杨雪,谭德军,等.模糊证据权法在地震滑坡危险度区划中的应用[J].中国安全科学学报,2011,(8):164-170.

[2] 阮沈勇,黄润秋.基于GIS的信息量法模型在地质灾害危险性区划中的应用[J].成都理工大学学报,2001,(1):89-92.

[3] 吴柏清,何政伟,刘严松.基于GIS的信息量法在九龙县地质灾害危险性评价中的应用[J].测绘科学,2008,(4):146-147,131.

[4] 汤连生,周萃英.基于信息量法的斜坡破坏概率预测与制图[J].中山大学学报:自然科学版,1995,(4):100-105.

[5] 陈广洲,徐晓春,汪家权.一种信息量和专家证据权重法耦合的成矿预测方法[J].测绘科学,2013,(1):146-149.

[6] 魏冠军,党亚民,章传银,等.GIS的信息量法在澜沧老厂成矿预测中的应用[J].测绘科学,2010,(6):217-218,246.

[7] 方国柱,张正禄,章传银.工程高边坡稳定性评价的信息量法[J].武汉测绘科技大学学报,1996,(4):38-43.

[8] 光磊.基于GIS的信息量法在滑坡危险性评价中的应用[J].地质与资源,2005,(3):231-233.

[9] 邓辉,何政伟,陈晔.信息量模型在山地环境地质灾害危险性评价中的应用[J].自然灾害学报,2014,23(2):67-75.

[10] 罗真富,蒲达成,谢洪斌,等.基于GIS和信息量法的泥石流流域滑坡危险性评价[J].中国安全科学学报,2011,(11):144-150.

[11] 杜娟,殷坤龙,陈丽霞.基于GIS的巴东县新城区滑坡灾害危险性区划[J].自然灾害学报,2011,(1):149-155.

(编辑:超凤超)

Geohazard susceptibility assessment by using information quantity model with result validation: a case study in Kaiyang County, Guizhou Province

FAN Qiang¹, JU Nengpan¹, XIANG Xiqiong², HUANG Jian¹

(1. State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China; 2. Key Laboratory of Karst Environment and Geohazard Prevention, Ministry of Education, Guizhou University, Guiyang 550025, China)

Abstract: The traditional information quantity model synthesizes all the evaluation factors without optimization, which is subjective. Therefore, the results validation is adopted to improve the information quantity model by selecting seven evaluation factors, namely lithology, fault, stream, road, slope ratio, slope direction and profile curvature. Firstly, the geological logging data and the evaluation factors are employed to carry out an overlay analysis, so the information value of single factor is obtained. Secondly, the success ratio curve method is used to evaluate the model fitting capacity of the single factor; the information quantity of evaluation factors are ranked according to the model fitting capacity, and seven sets of total information quantity are obtained by trial calculation. Finally, the calculation results are evaluated so as to provide an assessment to the model fitting capacity, and the combination of top 4 AUC factors is determined as the best.

Key words: information quantity model; results validation; geohazard; susceptibility assessment; Kaiyang County