

多因子组合的地质灾害易发性空间精度验证

解明礼, 巨能攀, 赵建军, 范强, 何朝阳

Evaluation on spatial accuracy and validation of geological hazard susceptibility based on a multi-factor combination

XIE Mingli, JU Nengpan, ZHAO Jianjun, FAN Qiang, and HE Chaoyang

在线阅读 View online: <https://doi.org/10.16031/j.cnki.issn.1003-8035.202302032>

您可能感兴趣的其他文章

Articles you may be interested in

基于遥感影像多尺度分割与地质因子评价的滑坡易发性区划

Landslide susceptibility assessment based on multi-scale segmentation of remote sensing and geological factor evaluation

李文娟, 邵海 中国地质灾害与防治学报. 2021, 32(2): 94-99

北京山区突发性地质灾害易发性评价

Assessment on the susceptibility of sudden geological hazards in mountainous areas of Beijing

罗守敬, 王珊珊, 付德荃 中国地质灾害与防治学报. 2021, 32(4): 126-133

香丽高速公路边坡地质灾害发育特征与易发性区划

Development characteristics and susceptibility zoning of slope geological hazards in Xiangli expressway

廖小平, 徐风光, 蔡旭东, 周文皎, 魏家旭 中国地质灾害与防治学报. 2021, 32(5): 121-129

基于遗传规划的泥石流多因子融合预测方法

Multi-factors fusion method of debris flow prediction based on genetic programming

翟淑花, 冒建, 南, 刘欢欢, 王云涛, 王强强, 熊春华, 王艳梅 中国地质灾害与防治学报. 2020, 31(6): 111-115

泥石流致灾因子敏感性分析——以四川都江堰龙溪河流域为例

Sensitivity analysis of debris flow to environmental factors: a case of Longxi River basin in Dujiangyan, Sichuan Province

李彩侠, 马煜, 何元勋 中国地质灾害与防治学报. 2020, 31(5): 32-39

基于快速聚类-信息量模型的汶川及周边两县滑坡易发性评价

Landslide susceptibility assessment based on K-means cluster information model in Wenchuan and two neighboring counties, China

周天伦, 曾超, 范晨, 毕鸿基, 龚恩慧, 刘晓 中国地质灾害与防治学报. 2021, 32(5): 137-150



关注微信公众号, 获得更多资讯信息

DOI: 10.16031/j.cnki.issn.1003-8035.202302032

解明礼, 巨能攀, 赵建军, 等. 多因子组合的地质灾害易发性空间精度验证[J]. 中国地质灾害与防治学报, 2023, 34(5): 10-19.

XIE Mingli, JU Nengpan, ZHAO Jianjun, et al. Evaluation on spatial accuracy and validation of geological hazard susceptibility based on a multi-factor combination[J]. The Chinese Journal of Geological Hazard and Control, 2023, 34(5): 10-19.

多因子组合的地质灾害易发性空间精度验证

解明礼, 巨能攀, 赵建军, 范 强, 何朝阳

(地质灾害防治与地质环境保护国家重点实验室(成都理工大学), 四川 成都 610059)

摘要: 地质灾害的发生是多种因素相互作用、制约和触发的结果。长期以来, 研究人员通过统计历史地质灾害所在区域的地质、地形、水文等环境因素预测未来地质灾害可能发生的位置、时间(或频率), 即地质灾害易发性评价。地质灾害易发性评价前提工作是进行影响因子选取, 而地质灾害发生的影响因子有数十种, 是否叠加因子越多模型评价精度就越高? 是否存在“最优因子数量”? 这一简单而又关键的问题值得探讨。文章以四川省汶川县为例, 选取 11 种广泛应用于地质灾害易发性评价的影响因子, 按照 4 种排列组合模型, 叠加 3 到 11 个因子信息量获得对应的地质灾害易发性指数分布图。运用成功率曲线确定线下面积值对各个结果进行预测精度评价。试验结果表明, 按照初步设定的 4 种排列组合模型叠加因子数量到 8 个时, 模型预测精度达到最大值; 但在因子叠加过程中发现各个因子对于易发性的控制性与个人经验确定的控制性存在一定差异, 按照实际因子控制性从大到小与从小到大排列组合后, 叠加多个关键因子后模型预测精度才会达到峰值。研究成果表明, 地质灾害易发性评价中叠加的因子数量越多, 模型预测精度越高, 叠加过程中如未加入关键因子, 模型预测精度将不会达到峰值, 说明地质灾害易发性评价不存在“最优因子数量”。

关键词: 地质灾害; 易发性; 影响因子; 最优因子数量; 关键因子

中图分类号: P642; TU43

文献标志码: A

文章编号: 1003-8035(2023)05-0010-10

Evaluation on spatial accuracy and validation of geological hazard susceptibility based on a multi-factor combination

XIE Mingli, JU Nengpan, ZHAO Jianjun, FAN Qiang, HE Chaoyang

(State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, Sichuan 610059, China)

Abstract: The occurrence of geological hazards is the result of the interaction, constraints, and triggers of various factors. For a long time, researchers have predicted the potential locations and time (or frequency) of future geological hazards in areas prone to historical geological disasters based on environmental factors such as geology, topography, and hydrology, known as geological hazard susceptibility assessment. A prerequisite for geological hazard susceptibility assessment is the selection of impact factors, and there are dozens of factors contributing to the occurrence of geological hazards. Does the accuracy of assessment model increase with the addition of more factors? Is there an “optimal number of factors”? This simple yet crucial question is worth exploring. Taking Wenchuan County, Sichuan Province, as an example, this study selects 11 commonly used influencing factors in geological hazard susceptibility assessment and arranges them into four different combination,

收稿日期: 2023-02-28; 修订日期: 2023-07-10

投稿网址: <https://www.zgdzhyfzxb.com/>

基金项目: 四川省科技计划项目(2022YFG0183); 地质灾害防治与地质环境保护国家重点实验室自主研究课题项目(SKLGP2020Z006)

第一作者: 解明礼(1992-), 男, 辽宁辽阳人, 土木工程专业, 博士, 主要从事地质灾害评价与预测研究工作。E-mail: 565725640@qq.com

通讯作者: 巨能攀(1973-), 男, 四川巴中人, 地质工程专业, 博士, 博导, 教授, 主要从事地质工程和岩土工程教学研究。

E-mail: jnp@cdut.edu.cn

superimposing information from 3 to 11 factors to obtain corresponding distribution maps of geological hazard susceptibility indexes. The area under curve (AUC) value was used to evaluate the predictive accuracy of the results. Experimental results show that the prediction accuracy of the model reaches its maximum value when the number of superposition factors added to the model reaches 8 according to the preliminary set four combinations. However, during the process of factor superposition, it is found that there are certain difference between the control of susceptibility by each factor and the control determined by individual experience. When the actual factor control is arranged from largest to smallest or from smallest to largest, the predictive accuracy of the model reaches its peak value after superposition of multiple key factors. The research findings indicates that the more factors added to the geological hazard susceptibility assessment, the higher the predictive accuracy of the model. If key factors are not included during the superimposition process, the predictive accuracy of the model will not reach its peak, indicating that there is no “optimal number of factors” in geological hazard susceptibility assessment.

Keywords: geological hazard; susceptibility; impact factors; optimal factor quantity; key factors

0 引言

地质灾害易发性评价是地质灾害风险评价的核心工作内容之一,是通过分析地质灾害位置与其相关因素之间的关系。利用统计、数据挖掘以及地理信息系统在空间上识别地质灾害易发区域,影响因子选取是易发性评价的前提工作^[1-3]。因子的正确选择取决于地质灾害的类型、机理、特征、案例区、分析的尺度、数据的可用性和使用的评价模型方法^[4-6]。地质灾害影响因子可划分为以下几类:(1)地质因子:岩性、断层距离(密度)、工程地质岩组、斜坡结构类型、向斜与背斜构造、褶皱轨迹等;(2)地形因子:坡度、坡高(地形起伏度)、高程、地形曲率(平面曲率、剖面曲率、标准曲率、坡形)、坡长、坡位、沟谷密度、地形表面纹理、地形位置指数(topographic position index, TRI)、地形耐用指数(terrain ruggedness index, TPI)、粗糙度指数、地表切割度等;(3)水文因子:河流距离(密度)、降雨、地形湿度指数(topographic wetness index, TWI)、水动力指数、地下水高度、有效补给率、侵蚀程度、输沙能力指数、河谷深度、流路长度、径流强度指数、河流流量等;(4)地表覆盖因子:土地利用、植被指数(NDVI)、土壤类型、土壤厚度、森林类型、土壤渗透率、地表太阳辐射强度等;(5)人类活动因子:道路距离(密度)、居民距离(密度)等;(6)地震因子:地震烈度、峰值加速度、地震密度、震中距离等^[7-10]。地形因子与地质因子可以表征主导滑坡发生的背景条件,地表覆盖因子、水文因子、人类活动因子与地震因子则反映附加因素加剧滑坡的可能。在已有的研究中,坡度是地质灾害易发性评价工作中最为常用的因素,岩性、高程、坡向、河流距离和断层距离等也是常用的评价因子,而其它因素的适用性因地制宜^[11-14]。

对于地质灾害影响因子的选择,目前还没有统一的标准,以往研究工作中地质灾害易发性评价选择的影响因子往往是根据经验选取地质因子、地形因子、水文因子中的部分参数,每项研究选取的因子类型存在一定差异,且因子数量不一致^[15-19]。可能存在以下问题:①选择因子较少,预测精度不足;②选择因子过多,叠加部分因子后预测精度可能达到峰值,叠加过多因子增加一定工作量。此外,是否在地质灾害易发性评价中叠加的因子数量越多,模型预测精度越高或者上下大幅度波动?易发性评价中是否存在“最优因子数量”这一概念?这些问题值得探讨。

基于以上提出的问题,本文在以往研究工作基础上,以四川省汶川县作为案例区,选取多个常用地质灾害影响因子,将所选取影响因子按照一定排列组合模式运用信息量法进行案例区地质灾害易发性评价,并采用成功率曲线验证叠加不同数量对地质灾害易发性评价精度的影响^[20]。

1 研究方法

本研究选取“5·12”汶川地震后地质灾害频发的汶川县为案例区,案例区共发育有地质灾害 690 处,包括崩塌 192 处、滑坡 351 处、泥石流 147 处(图 1)。基于所收集资料与已有研究基础^[16],选取地质灾害易发性评价常用影响因子 11 种:地貌类因子(高程、坡度、起伏度、坡形、起伏度、沟谷密度)、地质类因子(工程岩组、断裂)、水文类(河流)、人类工程活动(道路、植被指数),并对因子进行分级(表 1)。运用信息量模型(表 1),以案例区 70% 历史灾害点为训练样本,计算每个影响因子各个分级的信息量,信息量计算方法如式(1)——(2)所示。根据不同排列组合叠加 3 到 11 个因子信息量获得对应的地质灾害易发性指数分布图,易发性指数越高

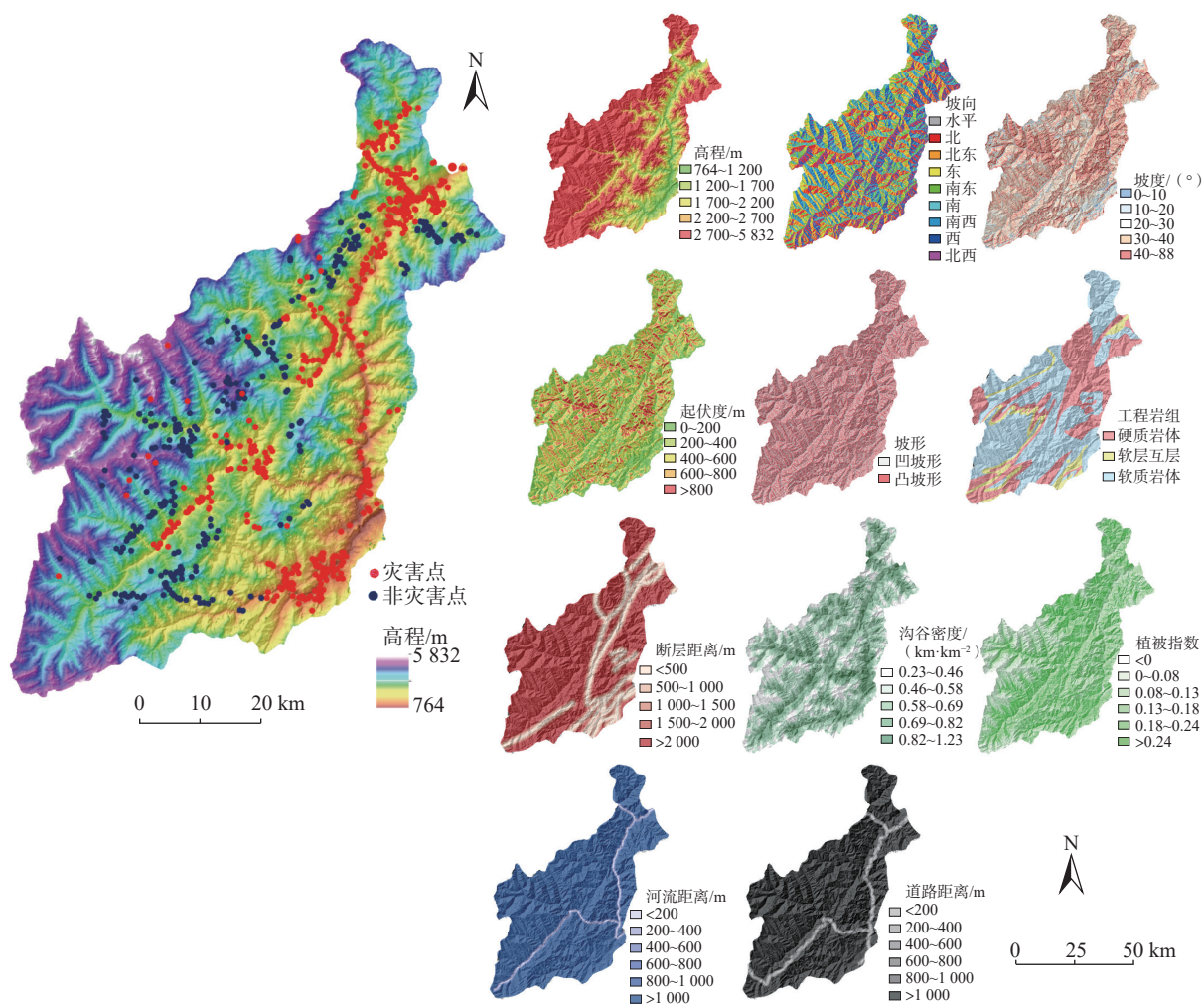


图 1 案例区基本信息

Fig. 1 Fundamental information of the case study area

代表地质灾害易发性越高。以 30% 历史灾害点和非灾害点为检验样本, 本文中非灾害点为历史地质灾害点数据 1 km 缓冲区范围外随机生成的点位数据, 运用成功率 (receiver operating characteristic, ROC) 曲线确定线下面积 (area under curve, AUC) 值对各个结果进行预测精度评价, 历史灾害点成功率曲线 AUC 值越趋近于 1 代表其评价精度越高, 而非灾害点成功率曲线 AUC 值越趋近于 0 代表其评价精度越高。对比各个组合结果, 分析叠加影响因子数量与地质灾害易发性评价精度的关系。研究思路如图 2 所示。

$$Y_i = \frac{N_i}{N} \cdot \left(\frac{S_i}{S} \right)^{-1} \quad (1)$$

$$I = \sum_{i=1}^n \lg(Y_i) \quad (2)$$

式中: I ——评价区某单元信息量预测值;

N_i ——分布在因素 X_i 内特定类别内的灾害点单元数;

N ——案例区含有灾害点分布的单元总数;

S_i ——案例区内含有评价因素 X_i 的面积;

S ——为案例区总面积;

Y_i ——致灾因子指标值。

2 因子排序组合

首先采用层次分析法确定每个因子的权重, 层次分析法是一种多指标分析评价方法, 具有精度高, 使用方便等特点。通过专家估计两两影响因子之间的关系构造矩阵对所有影响因子进行两两比较确定各个影响因子的权重, 这样避免了个别比较不合理而造成的结果偏差过大。

然而层次分析法带有一定的主观性, 为避免主观性, 选取 8 位从事工程地质研究工作学者对案例区 11

表 1 因子分级及信息量				
Table 1 Classification and information value of the factors				
因子	分级	灾害点比例/%	因子分级 面积比例/%	信息量
高程/m	[784, 1 200)	27.87	2.60	3.42
	[1 200, 1 700)	40.78	8.69	2.23
	[1 700, 2 200)	19.88	13.23	0.59
	[2 200, 2 700)	7.26	16.98	-1.23
	[2 700, 5 832]	4.21	58.50	-3.80
坡度/(°)	[0, 10)	19.88	3.21	2.63
	[0, 20)	26.85	13.30	1.01
	[20, 30)	34.40	32.75	0.07
	[30, 40)	14.95	37.66	-1.33
	[40, 88]	3.92	13.08	-1.74
地面起伏度/m	[0, 200)	15.09	4.49	1.75
	[200, 400)	68.36	42.07	0.70
	[400, 600)	15.38	45.22	-1.56
	[600, 800)	1.02	7.29	-2.84
	[800, ∞)	0.15	0.93	-2.69
沟谷密度 (/km·km ⁻²)	[0.23, 0.46)	1.30	13.65	-3.39
	[0.46, 0.58)	6.08	24.26	-2.00
	[0.58, 0.69)	31.84	33.00	-0.05
	[0.69, 0.82)	42.26	23.38	0.85
	[0.82, 1.23]	18.52	5.71	1.70
道路距离/m	[0, 200)	1.01	1.70	-0.75
	[200, 400)	2.32	1.68	0.46
	[400, 600)	2.32	1.66	0.48
	[600, 800)	2.32	1.66	0.48
	[800, 1 000)	2.32	1.64	0.50
断层距离/m	[1 000, ∞)	89.73	91.66	-0.03
	[0, 500)	28.94	9.96	1.54
	[500, 1 000)	25.90	8.80	1.56
	[1 000, 1 500)	10.27	7.27	0.50
	[1 500, 2 000)	7.96	6.07	0.39
工程岩组	[2 000, ∞)	26.92	67.90	-1.33
	硬质岩组	18.38	9.56	0.94
	软硬互层岩组	46.74	53.02	-0.18
	软质岩组	34.88	37.42	-0.10
河流距离/m	[0, 200)	8.10	1.59	2.35
	[200, 400)	10.27	1.59	2.69
	[400, 600)	11.29	1.59	2.82
	[600, 800)	5.79	1.57	1.88
	[800, 1 000)	3.47	1.57	1.14
坡向	[1 000, ∞)	61.07	92.07	-0.59
	北	6.34	11.30	-0.83
	北东	11.21	12.43	-0.15
	东	16.37	14.93	0.13
	南东	19.03	13.62	0.48
	南	7.96	11.82	-0.57
	南西	9.00	12.47	-0.47
	西	12.24	11.26	0.12
	北西	17.85	12.17	0.55
坡形	凹形坡	68.80	54.79	0.33
	凸形坡	31.20	45.21	-0.53

续表 1				
Table 1				
因子	分级	灾害点比例/%	因子分级 面积比例/%	信息量
植被指数	[-1, 0)	2.03	4.76	-1.23
	[0, 0.1)	18.43	22.12	-0.26
	[0.1, 0.25)	31.64	20.93	0.60
	[0.25, 0.4)	23.08	23.69	-0.04
	[0.4, 0.55)	21.04	18.31	0.20
	[0.55, 0.6]	3.77	10.18	-1.43

个因子进行打分,8 位专家打分结果平均值作为因子最终权重值(表 2)。最终确定各个因子对地质灾害敏感度从高到低排序为:①断裂②岩性③坡度④河流⑤坡形⑥起伏度⑦沟谷⑧高程⑨公路⑩坡向⑪植被指数。结合汶川县地质灾害发育分布特征及每个因子的信息量综合分析,区内发育汶茂断裂与北川映秀断裂,地质灾害主要集中于河流两岸,受坡度控制明显,且区内地质灾害与构造活动有着高度耦合性,这一结论与已有研究成果是相同的^[16]。综上说明通过多位专家打分的汶川县各个因子对地质灾害敏感度排序结果合理性较高。

为了避免按照某种顺序叠加因子导致结果规律的偶然性,本文将各个因子按照不同排列组合成由 3 个因子至 11 个因子组成的评价模型,因子组合分为两类:顺序数组与随机数组。顺序数组涵盖两种组合:因子对地质灾害发生的敏感度从高至低排列与从低至高排列模式;随机数据由编程语言随机函数生成 1~11 中包含不同个数并且不重复的随机数组。因子排列组合如表 3 所示。

3 评价精度分析

将各个因子信息量按照表 3 中因子组合方式分别叠加,计算出各个组合的案例区地质灾害易发性指数图(图 3),运用成功率曲线验证和比较各个组合模型易发性精度。

如图 3 所示,为多个组合模型不同数量因子叠加的案例区地质灾害易发性指数图,结果表明当叠加因子数量 3~5 个时,易发性指数图受单个因子控制性较为明显,例如组合 1A(3)与组合 1B(4)中断层控制易发性指数图最为明显、组合 2A(3)中道路控制易发性指数图最为明显。而当叠加因子 6~7 个时,地质灾害易发性指数图受单个因子控制性不再明显,显现出了多个因子的叠加效应,但不同组合模型的地质灾害易发性指数图图面信息差异性较大,易发性高分布的区域和面积各不相同。当叠加至 8 个以上因子后,各个组合模型的易发性

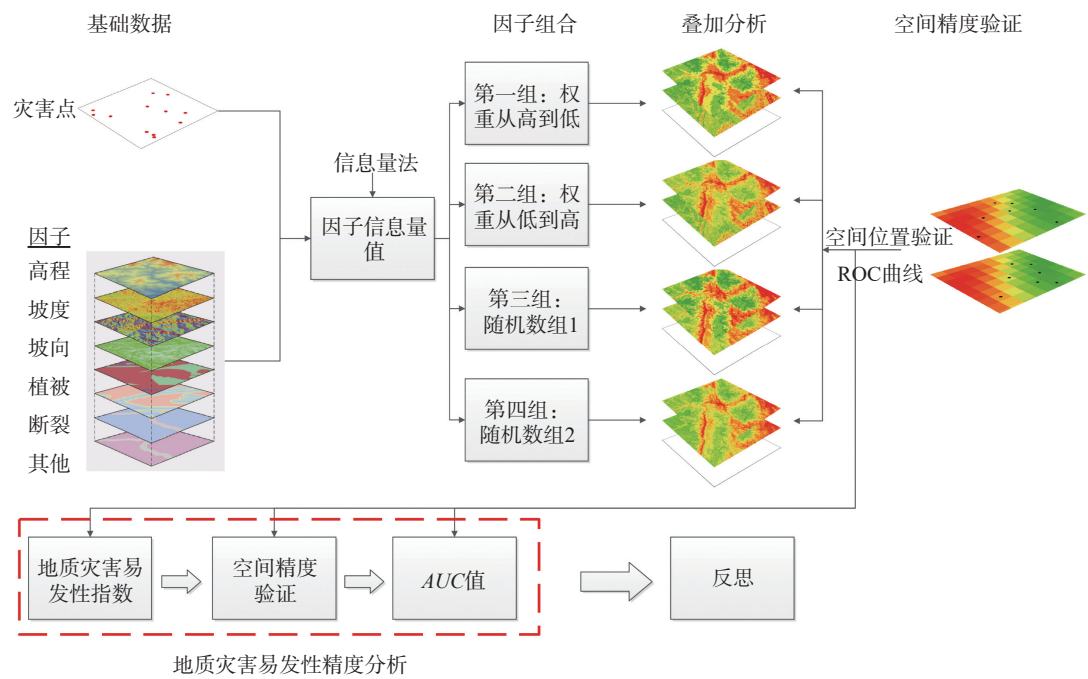


图 2 研究技术路线

Fig. 2 The research methodology flowchart

表 2 因子权重

Table 2 Factor weights table

专家因子		1	2	3	4	5	6	7	8	平均值
①	断层	0.055	0.269	0.193	0.223	0.138	0.182	0.209	0.135	0.176
②	岩性	0.023	0.133	0.182	0.124	0.168	0.106	0.182	0.143	0.133
③	高程	0.171	0.053	0.018	0.022	0.099	0.138	0.012	0.056	0.071
④	坡度	0.028	0.116	0.108	0.146	0.083	0.203	0.141	0.112	0.117
⑤	坡向	0.169	0.014	0.038	0.041	0.086	0.106	0.024	0.023	0.063
⑥	沟谷密度	0.063	0.064	0.082	0.055	0.082	0.106	0.096	0.073	0.078
⑦	坡形	0.128	0.105	0.046	0.100	0.073	0.046	0.105	0.090	0.087
⑧	河流	0.123	0.031	0.084	0.103	0.042	0.043	0.089	0.196	0.089
⑨	道路	0.128	0.042	0.078	0.064	0.057	0.021	0.057	0.075	0.065
⑩	植被指数	0.044	0.053	0.018	0.043	0.036	0.036	0.050	0.028	0.039
⑪	起伏度	0.069	0.119	0.153	0.079	0.138	0.014	0.035	0.067	0.084

表 3 因子排列组合

Table 3 Factor combination table

因子数量	顺序组合		随机组合	
	组合 1	组合 2	组合 3	组合 4
A(3)	①②③	⑪⑩⑨	③⑦⑪	②④⑥
B(4)	①②③④	⑪⑩⑨⑧	①③⑤⑩	②⑤⑨⑪
C(5)	①②③④⑤	⑪⑩⑨⑧⑦	②③⑤⑦⑨	⑤⑦⑨⑩⑪
D(6)	①②③④⑤⑥	⑪⑩⑨⑧⑦⑥	①③④⑤⑧⑩	②④⑥⑨⑩⑪
E(7)	①②③④⑤⑥⑦	⑪⑩⑨⑧⑦⑥⑤	①②④⑥⑦⑧⑪	①③④⑤⑧⑨⑩
F(8)	①②③④⑤⑥⑦⑧	⑪⑩⑨⑧⑦⑥⑤④	①④⑤⑥⑧⑨⑩⑪	①②③⑤⑦⑨⑩⑪
G(9)	①②③④⑤⑥⑦⑧⑨	⑪⑩⑨⑧⑦⑥⑤④③	①②④⑤⑥⑧⑨⑩⑪	②③④⑥⑦⑧⑨⑩⑪
H(10)	①②③④⑤⑥⑦⑧⑨⑩	⑪⑩⑨⑧⑦⑥⑤④③②	①②③⑤⑥⑦⑧⑨⑩⑪	①②③④⑤⑥⑦⑧⑨⑩
I(11)	①②③④⑤⑥⑦⑧⑨⑩⑪	⑪⑩⑨⑧⑦⑥⑤④③②①	①②③④⑤⑥⑦⑧⑨⑩⑪	①③④⑤⑥⑦⑧⑨⑩⑪

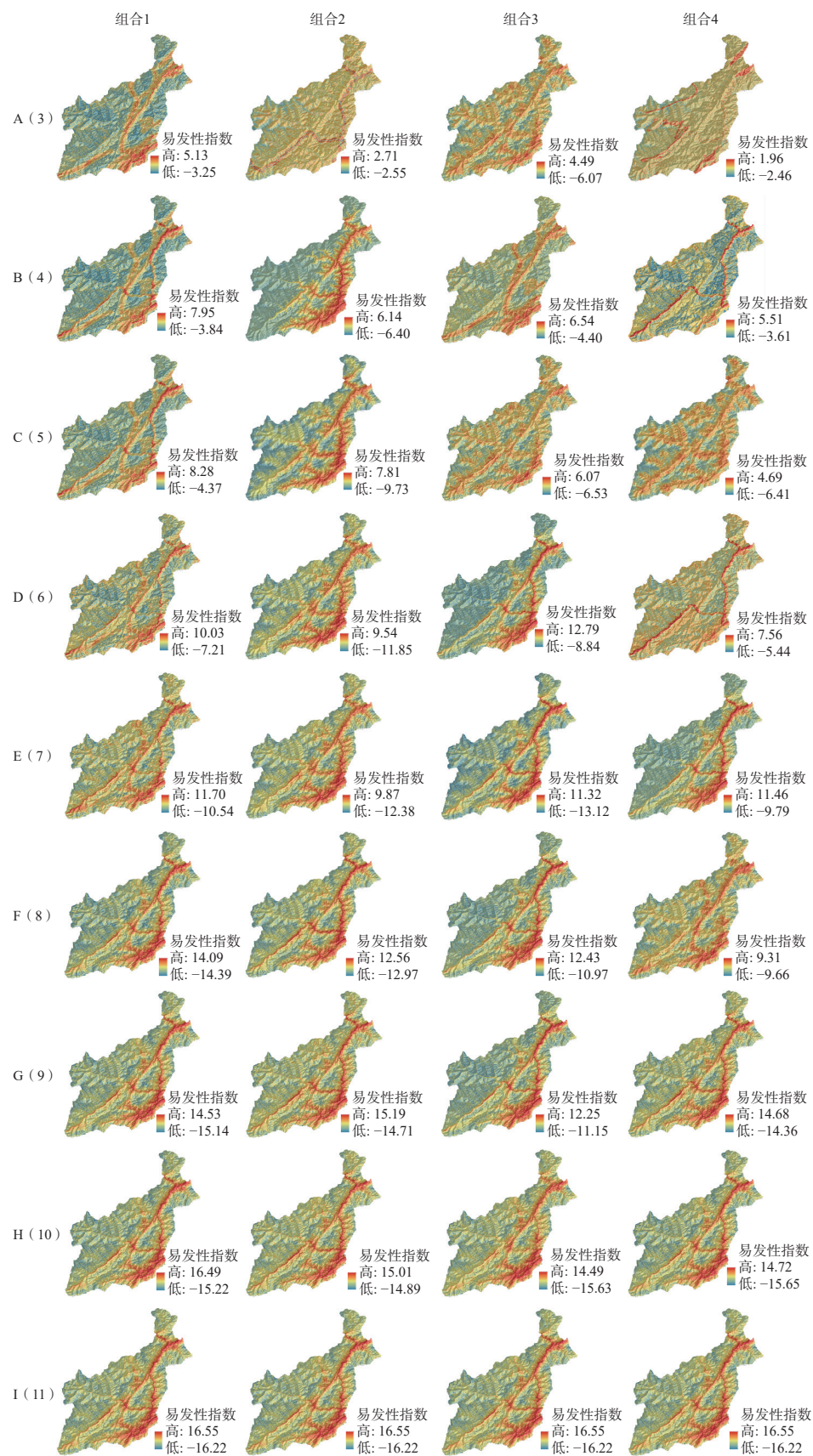


图 3 不同因子组合易发性指数图

Fig. 3 Geological hazard susceptibility index diagram for different quantitative factor combinations

指数图相似性较高, 显现出的高易发区与实际情况匹配度较高。

对比各类组合模型基于历史地质灾害点验证样本的成功率曲线(图 4), 统计出随因子数增多成功率曲线下面积(AUC)变化规律(图 5)。叠加 3 个因子预测精度较差, 组合 2 和组合 4 中 AUC 值仅在 0.65 左右, 而组

合 1 和组合 3 相对于组合 2 和组合 4 同等数量因子组合中叠加预测精度较高。其共同规律为: 4 种组合中 AUC 值随因子数增多而不断增高, 即随着叠加因子数量增多预测精度不断增高, 但叠加因子数至 8 个时, AUC 值不再明显上升与下降, 其值约为 0.9, 浮动幅度在 0.005 左右, 说明叠加 8 个以上因子时预测精度不再变化。

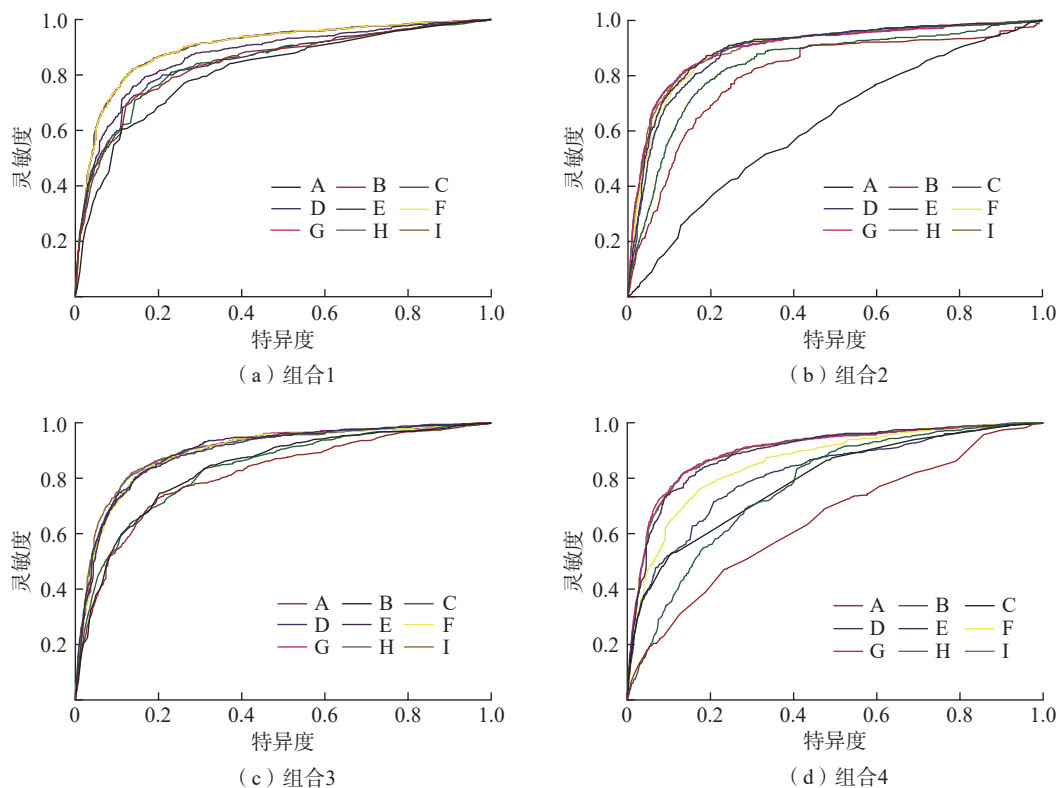


图 4 基于验证样本的不同因子组合成功率曲线

Fig. 4 Success rate curves of multifactor combination based on validation samples

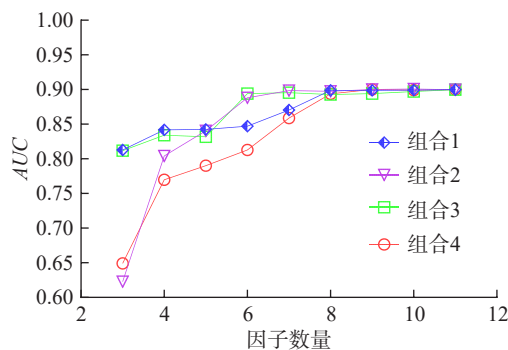


图 5 基于验证样本的 AUC 值统计

Fig. 5 AUC value statistics based on validation samples

对比各类组合基于非地质灾害点样本的成功率曲线, 统计出随因子数增多成功率曲线下面积(AUC)变化规律(图 6、图 7)。由图中可观察到叠加少于 8 个因子的组合随叠加因子数增多, AUC 值浮动较大, 且有着

随因子数增多而逐渐下降的趋势。叠加至 8 个以上因子的组合模型 AUC 值相对变化浮动较小, 稳定于 0.385 左右。

结合两种地质灾害易发性评价精度检验方法, 对比分析了按照不同组合方式叠加 3 至 11 个因子的 36 种组合模型, 分析结果发现随叠加因子数量增多, 组合模型精度不断提升, 但叠加至 8 个因子后, 模型精度不再变化, 精度值上下浮动较小, 历史灾害点验证样本的 AUC 值稳定于 0.9 左右, 非灾害点验证样本的 AUC 值稳定于 0.385 左右, 由于所选用的非地质灾害点为历史地质灾害点数据 1 km 缓冲区范围外随机生成的点位数据, 非地质灾害点又有可能在不久的将来成为新的地质灾害点, AUC 难以趋近于 0, 说明评价模型叠加至 8 个因子时模型精度已达到峰值, 叠加更多因子不会明显提升或降低其精度。

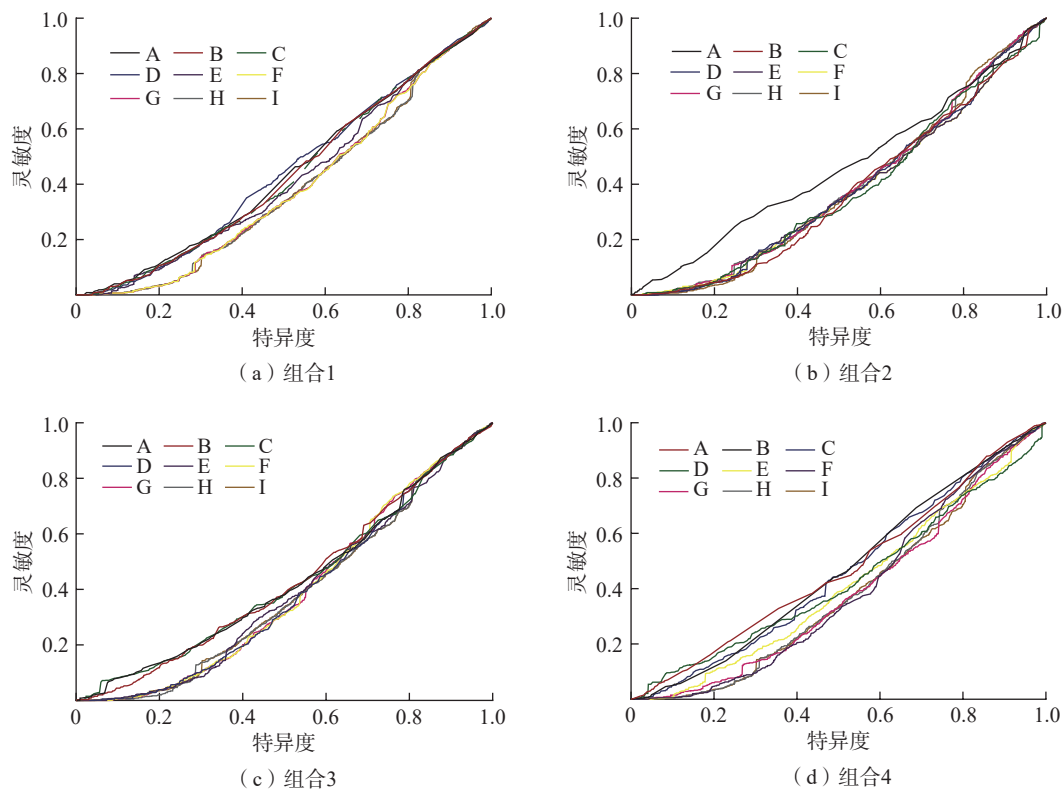


图 6 基于非灾害点验证样本的不同因子组合成功率曲线

Fig. 6 Success rate curves of multifactor combinations based on non-hazard validation samples

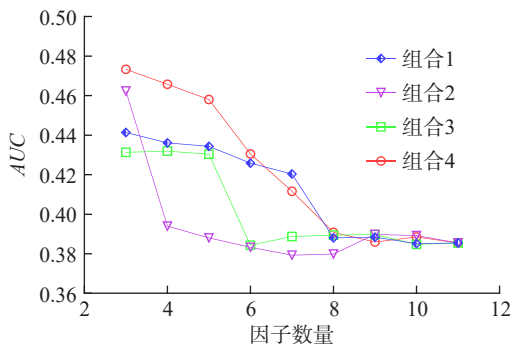


图 7 基于非灾害点验证样本的 AUC 值统计

Fig. 7 AUC value statistics based on non-hazard validation samples

4 讨论

根据不同因子组合方式可发现,各个因子对于案例区的地质灾害易发性影响存在较大的差异性。由于汶川地区受到 2008 年“5·12” $M_s8.0$ 地震的震裂影响,在断裂带区域地质灾害分布较为密集,断层缓冲区因子对于案例区的地质灾害易发性控制性最强。综合图 5 和图 7 中叠加各个因子后 AUC 值的变化幅度,重新梳理各个影响因子的控制性排序为:断层>河流>道路>岩性>高程>起伏度>坡度>沟谷密度>坡形>坡向>植被指数。这一排序结果与前文通过专家打分确定因子重要性等

级排序存在一定差异,分析原因为专家打分存在一定主观性,尽管采用了多个专家打分的平均值,但还是难以消去其主观性。

对比前文四种因子组合模型,它们有着共同的特点,运用验证样本成功率曲线检验各种组合模型评价精度时:模型精度随着叠加因子数增多而提高,叠加至 8 个因子时模型精度不再变化,趋于平稳状态,AUC 稳定于 0.9 左右,上下浮动约 0.005;运用非灾害点样本成功率曲线检验各种组合模型评价精度时:模型 AUC 值随叠加因子数增多而下降,即模型精度随着叠加因子数增多而上升,同样在叠加至 8 个因子时模型精度趋于稳定,AUC 稳定于 0.385。根据这一结果,可以确定当模型选取 8 个以上因子时,模型精度将不会改变,8 个因子可能是地质灾害易发性评价叠加最佳因子数。但这一结论是否正正确值得再次证明与讨论。

按照前文多次叠加后确定的因子实际控制性从高到低与从低到高两种组合模型再次检验随着因子叠加数量增多易发性评价精度的变化规律。如图 8—9 所示,按照因子实际控制性从高到低排列组合,叠加断层、河流、道路 3 个因子后 AUC 值已经接近峰值,为 0.889,其后再次叠加其它因子,AUC 值上下浮动约 0.02。

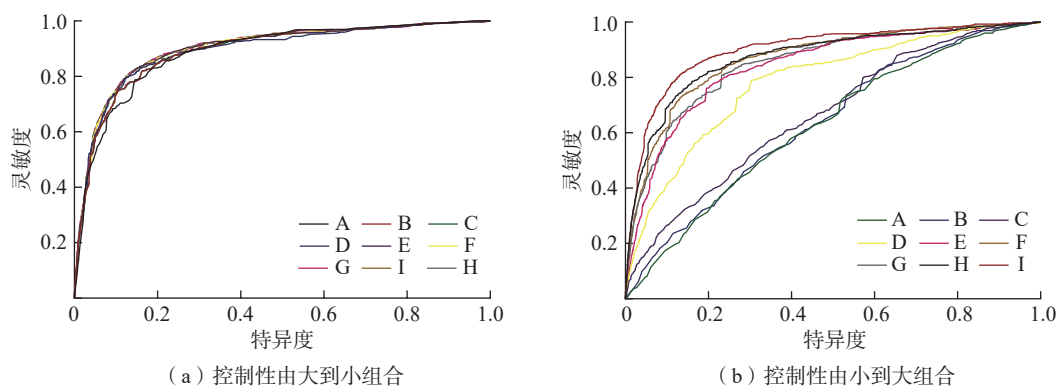
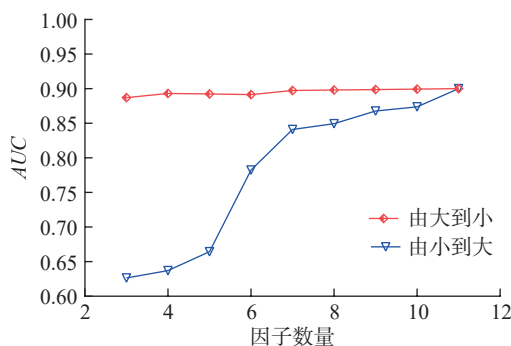


图 8 两种模型成功率曲线

Fig. 8 Success rate curves of two models

而按照实际因子控制性从低到高排列组合,当叠加到最后一个因子(断层)时 AUC 值才达到峰值。

图 9 两种模型 AUC 值统计Fig. 9 AUC value statistics of two models

结合前面的试验研究与后面的验证结果综合分析,造成叠加至 8 个因子时易发性指数的 AUC 值最大的原因在于叠加过程中存在一定偶然性:前期四种组合模型在叠加因子时仅靠个人经验或随机组合,未将关键因子优先组合, AUC 值无法快速达到峰值。而当叠加至 7~8 个因子时已经包含了这类关键因子(例如断层、河流、道路),此时达到了评价结果精度的峰值,其 AUC 值在 0.9 上下以 0.005 浮动。

综上试验研究表明,开展某地区地质灾害易发性评价时,最先开展的工作应是确定出该区域地质灾害的主控因素,例如构造、水文、岩性、地形等因素,即需要开展的是孕灾条件分析。且对于大区域,例如省级地质灾害易发性评价,应根据地质环境条件与地质灾害发育特征,对研究区开展综合分区,找出各个分区的主控因素,进行分区评价。可采用反演分析模式,综合运用信息量模型与 ROC 曲线法,将逐个因子不同等级所对应的信息量作为检验变量,利用 ROC 曲线法进行单因子分析,

根据 AUC 值确定各个因子对研究区地质灾害敏感度重要程度排序。

5 结论

本文以四川省汶川县为案例区,选取广泛应用的 11 种地质灾害影响因子进行不同排列组合,验证“是否在地质灾害易发性评价中叠加的因子数量越多,模型预测精度越高或者上下波动。”这一问题。经对比试验研究,得出以下结论:

(1)地质灾害影响因子进行随机组合时,叠加因子数量越多,地质灾害易发性评价结果精度越高,但叠加至一定数量因子后评价精度达到峰值,叠加更多因子不会明显提升或降低精度。

(2)地质灾害的发生在不同区域有着不同的主控因子,因子选取原则不仅仅根据个人经验,更应该计算出来每个因子独立的控制性,可采用单因子信息量与 ROC 曲线组合模型评价结果确定出主控因子,优先叠加控制性较强的因子,能够快速达到易发性评价精度的最高值。

(3)根据本文有限的多次测试结果表明,地质灾害易发性评价中叠加的因子数量越多,模型预测精度越高,叠加过程中如未加入关键因子,模型预测精度将不会达到峰值,说明地质灾害易发性评价存在关键因子,但不存在“最优因子数量”。

参考文献(References):

- [1] 黄发明. 基于 3S 和人工智能的滑坡位移预测与易发性评价 [D]. 武汉: 中国地质大学, 2017. [HUANG Faming. Landslide displacement prediction and vulnerability evaluation based on 3S and artificial intelligence [D]. Wuhan: China University of Geosciences, 2017. (in Chinese with English abstract)]

- [2] CÁRDENAS N Y, MERA E E. Landslide susceptibility analysis using remote sensing and GIS in the western Ecuadorian Andes [J]. *Natural Hazards*, 2016, 81(3): 1829 – 1859.
- [3] POURGHASEMI H R, YANSARI Z T, PANAGOS P, et al. Analysis and evaluation of landslide susceptibility: A review on articles published during 2005-2016 (periods of 2005-2012 and 2013-2016) [J]. *Arabian Journal of Geosciences*, 2018, 11(9): 193.
- [4] ERCANOGLU M. Landslide susceptibility assessment of SE Bartın (West Black Sea region, Turkey) by artificial neural networks [J]. *Natural Hazards and Earth System Sciences*, 2005, 5(6): 979 – 992.
- [5] 张俊,殷坤龙,王佳佳,等.三峡库区万州区滑坡灾害易发性评价研究[J].岩石力学与工程学报,2016,35(2): 284 – 296. [ZHANG Jun, YIN Kunlong, WANG Jiajia, et al. Evaluation of landslide susceptibility for Wanzhou District of Three Gorges Reservoir [J]. *Chinese Journal of Rock Mechanics and Engineering*, 2016, 35(2): 284 – 296. (in Chinese with English abstract)]
- [6] MANZO G, TOFANI V, SEGONI S, et al. GIS techniques for regional-scale landslide susceptibility assessment: The Sicily (Italy) case study [J]. *International Journal of Geographical Information Science*, 2013, 27(7): 1433 – 1452.
- [7] ABBASZADEH SHAHRI A, SPROSS J, JOHANSSON F, et al. Landslide susceptibility hazard map in southwest Sweden using artificial neural network [J]. *CATENA*, 2019, 183: 104225.
- [8] REICHENBACH P, ROSSI M, MALAMUD B D, et al. A review of statistically-based landslide susceptibility models [J]. *Earth-Science Reviews*, 2018, 180: 60 – 91.
- [9] ZHAO Yu, WANG Rui, JIANG Yuanjun, et al. GIS-based logistic regression for rainfall-induced landslide susceptibility mapping under different grid sizes in Yueqing, Southeastern China [J]. *Engineering Geology*, 2019, 259: 105147.
- [10] TRIGILA A, IADANZA C, ESPOSITO C, et al. Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy) [J]. *Geomorphology*, 2015, 249: 119 – 136.
- [11] HE Qingfeng, SHAHABI H, SHIRZADI A, et al. Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF Classifier, and RBF Network machine learning algorithms [J]. *Science of the Total Environment*, 2019, 663: 1 – 15.
- [12] REGMI N R, GIARDINO J R, VITEK J D. Modeling susceptibility to landslides using the weight of evidence approach: western Colorado, USA [J]. *Geomorphology*, 2010, 115(1/2): 172 – 187.
- [13] GUZZETTI F, CARRARA A, CARDINALI M, et al. Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy [J]. *Geomorphology*, 1999, 31(1/2/3/4): 181 – 216.
- [14] 黄发明,汪洋,董志良,等.基于灰色关联度模型的区域滑坡敏感性评价[J].地球科学,2019,44(2): 664 – 676. [HUANG Faming, WANG Yang, DONG Zhiliang, et al. Regional landslide susceptibility mapping based on grey relational degree model [J]. *Earth Science*, 2019, 44(2): 664 – 676. (in Chinese with English abstract)]
- [15] 张纪恺,凌斯祥,李晓宁,等.九寨沟县滑坡灾害易发性快速评估模型对比研究[J].岩石力学与工程学报,2020,39(8): 1595 – 1610. [ZHANG Qikai, LING Sixiang, LI Xiaoning, et al. Comparison of landslide susceptibility mapping rapid assessment models in Jiuzhaigou County, Sichuan Province, China [J]. *Chinese Journal of Rock Mechanics and Engineering*, 2020, 39(8): 1595 – 1610. (in Chinese with English abstract)]
- [16] 樊芷吟,苟晓峰,秦明月,等.基于信息量模型与 Logistic 回归模型耦合的地质灾害易发性评价[J].工程地质学报,2018,26(2): 340 – 347. [FAN Zhiyin, GOU Xiaofeng, QIN Mingyue, et al. Information and logistic regression models based coupling analysis for susceptibility of geological hazards [J]. *Journal of Engineering Geology*, 2018, 26(2): 340 – 347. (in Chinese with English abstract)]
- [17] 武雪玲,沈少青,牛瑞卿. GIS 支持下应用 PSO-SVM 模型预测滑坡易发性[J].武汉大学学报(信息科学版),2016,41(5): 665 – 671. [WU Xueling, SHEN Shaoqing, NIU Ruiqing. Landslide susceptibility prediction using GIS and PSO-SVM [J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(5): 665 – 671. (in Chinese with English abstract)]
- [18] 陈涛,钟子颖,牛瑞卿,等.利用深度信念网络进行滑坡易发性评价[J].武汉大学学报(信息科学版),2020,45(11): 1809 – 1817. [CHEN Tao, ZHONG Ziying, NIU Ruiqing, et al. Mapping landslide susceptibility based on deep belief network [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(11): 1809 – 1817. (in Chinese with English abstract)]
- [19] 范强,巨能攀,向喜琼,等.证据权法在区域滑坡危险性评价中的应用——以贵州省为例[J].工程地质学报,2014,22(3): 474 – 481. [FAN Qiang, JU Nengpan, XIANG Xiqiong, et al. Landslides hazards assessment with weights of evidence: A case study in Guizhou, China [J]. *Journal of Engineering Geology*, 2014, 22(3): 474 – 481. (in Chinese with English abstract)]
- [20] 许冲,戴福初,姚鑫,等.基于 GIS 的汶川地震滑坡灾害影响因子确定性系数分析[J].岩石力学与工程学报,2010,29(增刊1): 2972 – 2981. [XU Chong, DAI Fuchu, YAO Xin, et al. GIS based certainty factor analysis of landslide triggering factors in Wenchuan earthquake [J]. *Chinese Journal of Rock Mechanics and Engineering*, 2010, 29(Sup 1): 2972 – 2981. (in Chinese with English abstract)]