

DOI:10.19751/j.cnki.61-1149/p.2022.03.028

基于 RF - FR 模型的滑坡易发性评价 ——以略阳县为例

马啸¹, 王念秦^{2,*}, 李晓抗¹, 严冬³, 李嘉琳¹

(1. 陕西工程勘察研究院有限公司, 陕西 西安 710068; 2. 西安科技大学地质与环境学院,
陕西 西安 710054; 3. 陕西乾和实业有限公司, 陕西 西安 710000)

摘 要: 滑坡易发性评价是指导区域滑坡初步预警、预报的重要手段。为提高县域滑坡易发性评价的准确性, 以随机森林模型(RF)、频率比模型(FR)为基础模型, 结合 2 种模型的优越性, 建立随机森林-频率比模型(RF-FR), 进行滑坡易发性评价。以略阳县域为研究区, 选取高程、坡向、坡度、地层、地表粗糙度、距断层的距离、曲率、距道路的距离、地形湿度指数、距河流的距离及降雨量等 14 项影响因子建立数据库, 采用 Spearman 方法对各因子相关性进行分析, 剔除地形起伏度等 3 项相关性较高的评价因子, 并基于滑坡相对点密度(LRPD)进行评价因子分析。结果表明: ①滑坡灾害点与线状因子的距离呈负相关, 即距离越近, 灾害点越多。②FR、RF、RF-FR 模型预测率分别为 84.3%、90.1%、95.0%, RF-FR 模型较 FR、RF 模型预测精度分别提高了 10.7%、4.9%。③RF-FR 模型的滑坡灾害点在高、极高易发区的比例比 FR、RF 模型分别提高了 15.89%、5.29%。

关键词: 滑坡易发性; 频率比模型; 随机森林模型; RF-FR 模型; 滑坡相对点密度

中图分类号: P642.22 **文献标志码:** A **文章编号:** 1009-6248(2022)03-0335-10

Assessment of Landslide Susceptibility Based on RF - FR Model: Taking Lueyang County as an Example

MA Xiao¹, WANG Nianqin^{2,*}, LI Xiaokang¹, YAN Dong³, LI Jialin¹

(1. Shaanxi Engineering Investigation Research Institute Co. LTD, Xi'an, 710068, Shaanxi, China;
2. College of Geology and Environment, Xi'an University of Science and Technology, Xi'an, 710054,
Shaanxi, China; 3. Shaanxi Qianhe Industrial Co., LTD, Xi'an, 710000, Shaanxi, China)

Abstract: The assessment of landslide susceptibility is an important means to guide the preliminary early warning and forecast of regional landslide. In order to improve the accuracy of landslide susceptibility evaluation in county area, random forest model (RF) and frequency ratio model (FR) were used as the basic models, and RF - FR was established to evaluate landslide susceptibility combined with the advantages of the two models. In Lueyang county domain for the study area,

收稿日期: 2021-07-12; **修回日期:** 2022-01-11; **网络发表日期:** 2022-08-10; **责任编辑:** 李辉

基金项目: 国家自然科学基金面上项目“黄土山区城镇化行为过程与地质环境耦合机制”(41572287); 陕西省科技统筹创新工程计划项目“神府矿区矿山地质环境保护及其综合治理技术研究”(2016KTCL03-19)。

作者简介: 马啸(1995-), 男, 硕士, 研究方向为岩土体稳定与地质灾害防治。E-mail: 624540304@qq.com。

*** 通讯作者:** 王念秦(1964-), 男, 教授, 博士, 主要从事岩土体稳定与地质灾害防治教学和科研工作。E-mail: younglock@163.com。

selection of elevation, slope direction, slope, formation, surface roughness, the distance from the fault, curvature and the distance from the road, terrain humidity index, the distance from the river, a database of 14 factors such as rainfall, by adopting the method of Spearman correlation analysis of each factor, eliminate topographic relief degree of three high correlation of evaluation factors, and evaluate the relative point density (LRPD) based on landslide factor analysis. The results show that: ①there is a negative correlation between the distance between landslide disaster points and linear factors, that is, the closer the distance is, the more disaster points are. ②The prediction rates of FR, RF and RF - FR models are 84.3%, 90.1% and 95.0%, respectively. Compared with FR and RF models, the prediction accuracy of RF - FR model is 10.7% and 4.9% higher than that of FR and RF models. ③The proportion of landslide disaster points in high and extremely high-risk areas of 4RFmurFR model is 15.89% and 5.29% higher than that of FR and RF model, respectively.

Keywords: landslide susceptibility; frequency ratio model; random forests model; the RF - FR model; the relative point density

陕西是中国地质灾害最为严重的省份之一,据《陕西统计年鉴》:2012~2020 年陕西省因地质灾害伤亡 293 人,而滑坡是造成人员伤亡最为主要的灾害类型之一。近年来,滑坡易发性评价已经成为滑坡灾害风险管控和监测预警的重要手段,众多科研人员致力于此项研究,并取得众多成果。

滑坡灾害易发性评价方法主要分为定性评价、半定量评价和定量评价 3 种类型。其中,定量评价是滑坡易发性区划最为主要的方法。随着 GIS 技术的快速发展,将其应用于滑坡灾害易发性评价领域将更加优化了定量评价的准确性。易发性区划评价方法主要有:层次分析法、信息量模型、频率比、逻辑回归、随机森林、支持向量机和人工神经网络等。目前,将 2 种或多种模型进行对比与耦合来进行滑坡易发性评价已成为国内外学者研究的热点。Chen Wei 等(2018)以太白县为例,将随机森林与确定性因子进行耦合,得出 RF - CF 模型,结果表明,该模型的精度优于单一的模型;J. Cao 等(2019)采用逻辑回归和随机森林模型对川西山区的滑坡灾害易发性进行了评价,得出随机森林模型预测准确率高于逻辑回归模型。郭子正等(2020)基于加权频率比、逻辑回归和确定系数法,以万州区为例进行滑坡易发性评价,得出加权频率比的预测准确性更高。邱维蓉等(2020)通过 LR、SVM、BP、RF 4 种模型的对比与分析,得出 RF 模型优于其他 3 种模型。

结合前人研究得出:将随机森林(RF)模型应用于滑坡易发性评价具有较高的准确性与实用性,而频率比(FR)模型更加能反映滑坡的发育规律。因

此,结合 2 种模型的优越性,将其进行耦合应用于滑坡易发性评价中可提高滑坡灾害易发性预测的准确性。笔者以略阳县域为研究区,首先将频率比(FR)、随机森林(RF)模型应用于滑坡灾害的易发性评价中,得出各模型县域滑坡易发性分区图及其准确性;其次以频率比替代各分级的评价因子数值,代入随机森林模型中进行 2 种方法的耦合,得出耦合模型下县域滑坡易发性分区图及其准确性。通过对比各模型准确性,得出研究结果。研究结果对滑坡易发性评价模型的优化具有借鉴意义,也可为相关部门进行滑坡防治提供理论依据。

1 研究区概况

略阳县位于陕西省汉中市西北部,地理坐标为东经 $105^{\circ}42'$ ~ $106^{\circ}31'$,北纬 $33^{\circ}07'$ ~ $33^{\circ}38'$ 。县域平面图略呈长方形,总面积为 $2\,831\text{ km}^2$ 。属暖温带湿润季风气候,多年平均降雨量为 860 mm ,且降雨量分配不均,呈由西北向东南递增之势,多年平均气温为 13.2°C 。该区高程分布范围较广,为 $559\sim 2\,399\text{ m}$,地势南北高、中间低。属于秦巴山区地貌类型。区内有嘉陵江和汉江 2 条水系,其中嘉陵江水系较大支流有 10 条,汉江水系较大支流有 8 条。区内构造断裂发育,主要有褒城断裂、茶店断裂、艾叶-官亭断裂、勉县-阳平关断裂。据统计,略阳县已发生地质灾害及隐患点共 482 处,其中,滑坡 445 处,占研究区内地质灾害总数 92%。研究区地理位置及滑坡灾害点分布见图 1。

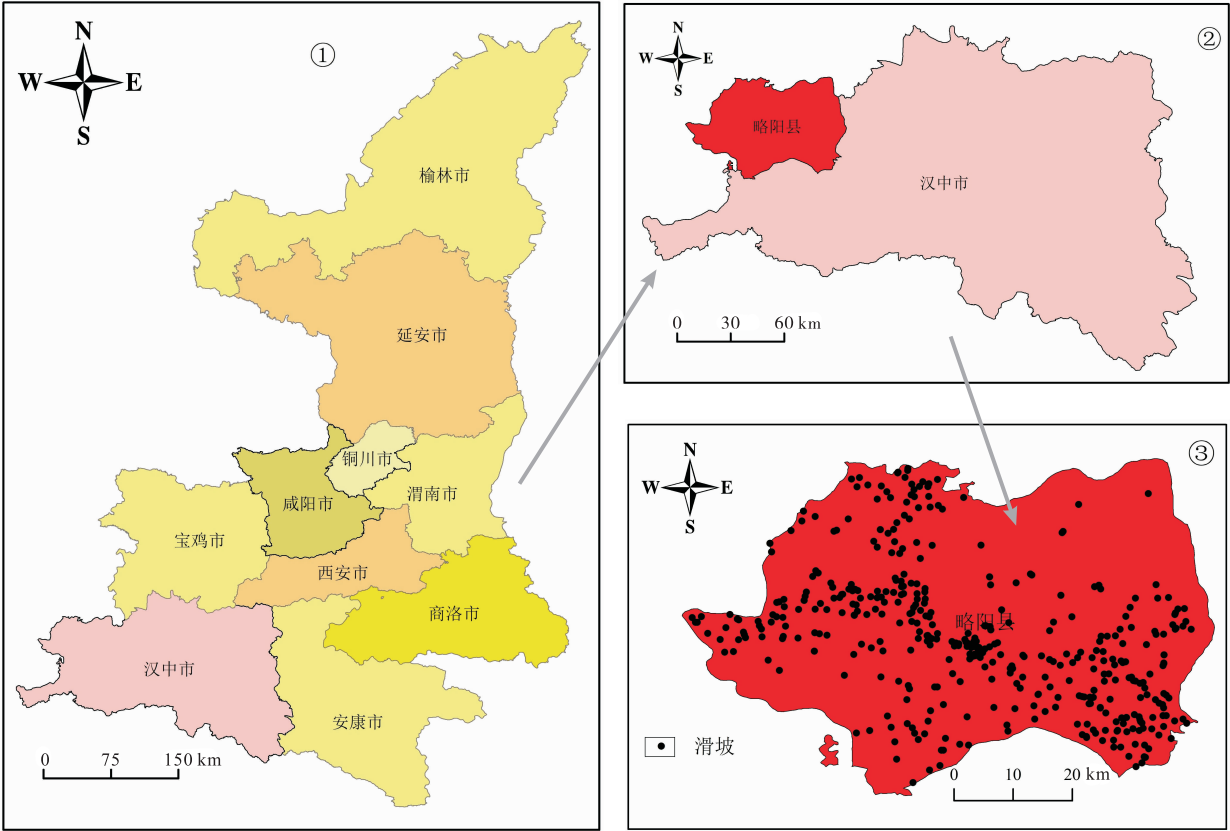


图 1 研究区地理位置及滑坡灾害点分布图

Fig. 1 Geographical location and landslide points distribution in research area

2 评价模型简介

2.1 随机森林模型

随机森林 (Random Forest Model, 简称 RF) 是将多棵决策树进行集成的算法。随机森林算法能有效解决二分类问题, 准确率高, 使大数据、高维特征数据能够有效地运行, 并且能够得到各评价因子的权重。其方法是: 从数据集 (N) 中随机选择 k 个特征 (列), 共 m 个特征 (其中 k 小于等于 m); 然后根据这 k 个特征建立决策树; 重复 n 次, 这 k 个特性经过不同随机组合建立起来 n 棵决策树, 对每个决策树都传递随机变量来预测结果, 得到高票数的预测目标作为随机森林算法的最终预测。

2.2 频率比模型

频率比模型 (frequency ratio model, 简称 FR) 是一种基于数理统计的预测方法, FR 值的大小表征各级评价因子对滑坡灾害发生的重要程度, 即

$FR > 1$ 表明该分级区间有利于滑坡发生, $FR < 1$ 表现该分级区间对滑坡的发生不利。其公式如下:

$$FR = \frac{\frac{n_i}{N}}{\frac{S_i}{S}}$$

(1)

式中: n_i 为各分级区间内滑坡面积, N 为研究区滑坡总面积, s_i 为分类面积, S 为研究区总面积。

3 略阳县域滑坡评价因子构建

3.1 建立滑坡数据库

基础数据获取方式为: ①略阳县 $30\text{ m} \times 30\text{ m}$ 的数字高程模型 (DEM), 从国家地理空间数据云下载 (<http://www.gscloud.cn>), 并提取坡向、曲率、坡度等评价因子。②略阳县气象站点降水量数据。③ 1: 50 000 地质图、路网图、河流分布图在 91 卫图下载。④陕西省地质灾害点编录数据库。

3.2 评价因子的选取

影响滑坡形成的因素很多,根据研究区内滑坡分布特征及形成条件,选取高程(A)、坡向(B)、坡度(C)、地层(D)、地表粗糙度(E)、距断层距离(F)、曲率(G)、距道路距离(H)、地形湿度指数(I)、距河流距离(J)、降雨量(K)、地形起伏度(L)、高程变异系数(M)和地表切割深度(N)14 个因素作为评价因子。

借助 ArcGIS 软件,将研究区划分为 3 122 182 个 30 m×30 m 的栅格。各评价指标的获取方法为:①坡向、坡度、曲率通过 ArcGIS 工具箱中的表面分析从 DEM 中提取。②地形起伏度、地表粗糙度、地形湿度指数、地表切割深度和高程变异系数通过空间分析里的栅格计算器对 DEM 处理得出。③河流通过对河流分布图矢量化获得。④道路通过对路网图进行矢量化,使用欧氏距离分析工具得到。⑤断层、地层岩性通过对 1:50 000 地质图矢量化得到。

3.3 评价因子的相关性分析

进行滑坡评价模型预测时,要求各评价因子间

相关性小,因而有必要进行评价因子的相关性分析。采取 Spearma 法得出各因子间的相关性系数(R),R 值能反应各评价指标的相关性。通常将相关性分为微弱相关、低度相关、显著相关、高度相关和线性相关,所对应的 R 取值范围分别为 $|R|<0.3$ 、 $0.3\leq|R|<0.5$ 、 $0.5\leq|R|<0.8$ 、 $0.8\leq|R|<1$ 、 $|R|=1$ 。其公式如下:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

将研究区的各评价因子的数据导入到 SPSS 软件中的相关性分析工具中,得到各因子相关系数(表 1)。将相关性系数大于 0.3 的评价因子剔除掉,根据结果可知地形起伏度、高程变异系数、地表切割深度与坡度、地表粗糙度相关性高。因此,将这 3 项评价指标剔除,最终得到 11 项评价因子进行易发性评价。

表 1 因子相关性系数表
Tab.1 Correlation coefficient between factors

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	1.000													
B	-0.022	1.000												
C	0.076	-0.009	1.000											
D	0.044	-0.017	-0.028	1.000										
E	0.076	-0.009	1.000	-0.028	1.000									
F	0.144	0.023	0.033	-0.241	0.033	1.000								
G	0.138	-0.031	0.205	0.018	0.105	0.090	1.000							
H	0.181	-0.022	0.208	0.002	0.208	0.108	0.170	1.000						
I	-0.249	0.028	-0.263	0.036	-0.263	-0.152	-0.219	-0.238	1.000					
J	0.715	-0.055	0.048	-0.008	0.048	0.123	0.084	0.258	-0.148	1.000				
K	-0.197	0.004	-0.042	0.272	-0.042	-0.190	-0.108	-0.079	0.123	0.134	1.000			
L	0.077	-0.011	0.976	-0.027	0.976	0.017	0.300	0.213	-0.572	0.046	-0.047	1.000		
M	-0.309	0.007	0.822	-0.004	0.822	-0.055	0.183	0.007	-0.362	0.235	0.028	0.810	1.000	
N	0.123	-0.012	0.962	-0.011	0.962	0.026	0.388	0.249	-0.605	0.091	-0.063	0.981	0.780	1.000

3.4 基于滑坡相对点密度(LRPD)的评价因子分析

滑坡相对点密度(LRPD)能反应灾害点与各评价因子的空间分布关系,得到不同分级状态的因子对滑坡发育的影响。LRPD 值越大,则表明该分级区间是滑坡发育的有利因素,各评价因子分级的 LRPD 值见表 2。其计算公式为:

$$LRPD = \frac{\frac{m_i}{M}}{\frac{n_i}{N}} \quad (3)$$

式(3)中, m_i 表示因子某一分级内滑坡点个数; M 表示研究区滑坡点总数; n_i 表示因子对应分级所占面积; N 表示研究区总面积。

表 2 各评价因子分级标准、频率比值及 *LRPD* 值表
Tab. 2 Grading standard, frequency ratio and *LRPD* value of each evaluation factor

评价因子	分级	灾害点 (个)	栅格数 (个)	FR 值	<i>LRPD</i> (个/km ²)	评价因子	分级	灾害点 (个)	栅格数 (个)	FR 值	<i>LRPD</i> (个/km ²)
A	559~924 m	214	620 099	2.421	0.383	E	1.16~1.28	50	564 713	0.621	0.098
	924~1 143 m	157	963 750	1.143	0.181		1.28~1.48	7	183 317	0.268	0.042
	1 143~1 375 m	62	885 122	0.491	0.078		1.48~3.72	2	30 784	0.456	0.072
	1 375~1 695 m	9	433 622	0.146	0.023	F	<1000m	193	1 541 574	0.878	0.139
	1 695~2 399 m	3	219 589	0.096	0.015		1 000~2 000 m	95	440 982	1.512	0.239
B	-1°~0°	0	178	0.000	0.000		2 000~3 000 m	59	410 785	1.008	0.160
	337.5°~22.5°	45	381 495	0.828	0.131	G	3 000~4 000 m	41	242 749	1.185	0.188
	22.5°~67.5°	47	395 367	0.834	0.132		>4 000 m	57	486 092	0.823	0.130
	67.5°~112.5°	38	364 225	0.732	0.116		<0	321	1 970 574	1.143	0.181
	112.5°~157.5°	62	413 917	1.051	0.166	H	0~1	100	901 622	0.803	0.123
	157.5°~202.5°	95	401 553	1.660	0.263		>1	24	349 986	0.497	0.076
	202.5°~247.5°	66	387 784	1.194	0.189		<500 m	235	898 275	1.836	0.291
	247.5°~292.5°	55	385 476	1.001	0.159	I	500~1 000 m	86	750 091	0.804	0.127
C	292.5°~337.5°	37	392 187	0.662	0.105		1 000~1 500 m	46	341 375	0.945	0.150
	0°~10°	94	401 949	1.641	0.260		1 500~2 000 m	45	566 299	0.558	0.088
	10°~20°	129	765 279	1.183	0.187	J	>2 000 m	33	566 142	0.409	0.065
	20°~30°	136	916 899	1.041	0.165		2.0~4.9	113	966 987	0.820	0.130
	30°~40°	75	740 158	0.711	0.113		4.9~6.4	165	1 092 785	1.059	0.168
D	>40°	11	297 897	0.259	0.041	K	6.4~8.6	114	852 716	0.938	0.149
	1-第四纪	0	63 928	0.000	0.000		8.6~12.4	36	151 008	1.673	0.265
	2-侏罗纪	6	237 870	0.177	0.028		12.4~33.9	17	58 686	2.032	0.322
	3-二叠纪	12	214 531	0.392	0.062	L	<500 m	194	833 630	1.633	0.259
	4-三叠纪	28	232 187	0.846	0.134		500~1 000 m	98	551 062	1.248	0.198
	5-石炭纪	10	813 04	0.863	0.137		1 000~1 500 m	54	565 020	0.671	0.106
	6-泥盆纪	139	643 924	1.515	0.240	M	1 500~2 000 m	43	567 780	0.531	0.084
	7-志留纪	138	1 291 160	0.750	0.119		>2 000 m	56	604 690	0.650	0.103
	8-奥陶纪	19	50 974	2.615	0.414		705~743 mm	44	313 031	0.978	0.156
	9-震旦纪	74	248 266	2.091	0.331	N	743~773 mm	115	966 779	0.828	0.132
	10-元古代	19	58 038	2.297	0.364		773~801 mm	87	884 335	0.685	0.109
E	1.00~1.07	221	1 415 032	1.096	0.174		801~828 mm	104	606 854	1.193	0.190
	1.07~1.16	165	928 336	1.247	0.197		828~867 mm	95	351 183	1.883	0.301

注:A. 高程;B. 坡向;C. 坡度;D. 地层;E. 地表粗糙度;F. 距断层距离;G. 曲率;H. 距道路距离;I. 地形湿度指数;J. 距河流距离;K. 降雨量

3.4.1 地形地貌因子

选取的地形地貌因子为高程、坡度、坡向、曲率和地表粗糙度。采用自然间断点分类方法将高程划分为 5 级(图 2a)(表 2), 在 559~924 m 内 *LRPD* 值最大, 为 0.383, 而在高程 >1 375 m 范围内滑坡发育较少, 仅占全部灾害点的 2.47%, 说明滑坡主要发生在高程中低区间的边坡上, 原因是在此高程范围内的人工活动强烈、多进行切坡建房和道路建设; 加上地表水及降雨的侵蚀作用, 易形成地质灾害。按照方位角等间隔将坡向划分为 9 级(图 2b)(表 2), 其中 -1~0 表示水平坡, 即平地, 在 157.5~202.5 m(正南)区间内灾害点最多, *LRPD* 值为 0.262, 反映了正南朝向的边坡更容易受到光照、降雨等自然作用的影响, 进而形成滑坡灾害。以相等间隔法将坡度分为 5 级(图 2c)(表 2), 在坡度 0°~30°内滑坡集中发育, 坡度在 30°以上的灾害点仅 11 处, 在 0°~10°*LRPD* 值最大, 为 0.260, 说明中低坡度的边坡对滑坡发育影响最大。其原因为在中低坡度处人工活动强烈, 人为切坡建房及公路修建导致岩土体局部地区坡度较陡, 而图层所提取出来的坡度为 30×30 m 地区的整体平均坡度, 导致坡度值变小, 但能反映出在中-低坡度为形成滑坡灾害的因素, 而在平均坡度较大的地区, 很少有人居住, 仅在人工活动的地区构成灾害, 因此, 在坡度较大的区域滑坡灾害较少。曲率的大小代表边坡的类型, 当曲率小于 0 时为凸型坡, 曲率大于 1 时为凹型坡, 曲率等于 1 时为直线坡(图 2d)(表 2)。滑坡灾害点多发生于凸坡处, 占总灾害点的 73.3%, *LRPD* 值为 0.181。地表粗糙度表示该地区抗侵蚀能力, 指数越大, 则抗侵蚀能力越强(图 2e)(表 2)。滑坡多发生在地表粗糙度指数较小的地貌类型处, 在 1~1.07 分级处, *LRPD* 值最大, 表明该分级状态为滑坡的发育有利状态。

3.4.2 地质因子

地质因子包括地层岩性、距断层距离。地层岩性作为内在因子对滑坡发育起重要作用, 是由于不同的岩性物理力学参数差异大, 直接影响边坡的稳定性(图 2f)(表 2)。研究区共 10 套地层, 而滑坡灾害点主要分布在泥盆纪、志留纪地层, 占灾害点总数的 62.2%。在此 2 处地层的 *LRPD* 值分别为 0.24、0.119, 表明上述 2 套地层对滑坡发育有利。以相等

间隔法将距断层距离以 1 000 m 划分为 5 级(图 2g)(表 1)。滑坡灾害点主要分布在 0~1 000 m 内, *LRPD* 值为 0.139, 表明滑坡常发育在断层构造活动带和岩土体结构软弱处。

3.4.3 道路因子

略阳县进行了大量的道路建设, 破坏了原始地形、地貌, 进而促进了滑坡发育。以等间隔法进行分级, 将距道路的距离以 500 m 等分(图 2h)(表 2)。灾害点主要分布在 0~500 m 内, *LRPD* 值为 0.291。表明距离道路越近, 人类活动对滑坡的发育影响越大。

3.4.4 水文因子

水文因素是滑坡形成、致灾的重要影响因素之一。水文因子包括降雨量、距河流的距离和地形湿度指数。降雨能降低岩土体的物理力学性质, 使剪切强度下降, 摩擦系数变小, 使岩土体结构面软化, 是影响滑坡发生的最重要的因子, 依据自然间断点将其分为 5 级(图 2i)(表 2)。降雨量在 828~867 mm 时, *LRPD* 值最大, 为 0.156; 河流的下蚀和侧蚀作用容易侵蚀其周围的岩土体, 破坏边坡的稳定性, 进而为滑坡灾害的形成提供了条件; 依据 500 m 等间隔法将距河流的距离分为 5 级(图 2j)(表 2)。灾害点分布数量与距河流距离成反比, *LRPD* 值随距河流距离的增加而递减。地形湿度指数表示在某一流域内的土壤干湿状态(图 2k)(表 2); 依据自然间断点法将评价因子分为 5 级, *LRPD* 值在各分级中依次递增, 表明地形湿度指数越大越有利于滑坡灾害的发生。

4 评价结果

4.1 基于 RF 模型评价结果

滑坡预测可看作一个二元分类的过程, 在进行滑坡的易发性评级时, 需要建立等量的正数据集(滑坡)、负数据集(非滑坡)。借助 R 软件中的随机选取功能, 将灾害点数据集和研究区数据集导入; 从 445 个滑坡灾害点随机选取 312(70%)个正数据集与从研究区非滑坡灾害点随机选取 312 个负数据集作为训练集; 将滑坡灾害点剩余的 133(30%)个正数据集和研究区非滑坡灾害点数据集随机选取的 133 个负数据集作为测试集。将训练集代入 RF 模型进行训练, 再用测试集对 RF 模型进行预测, 得到

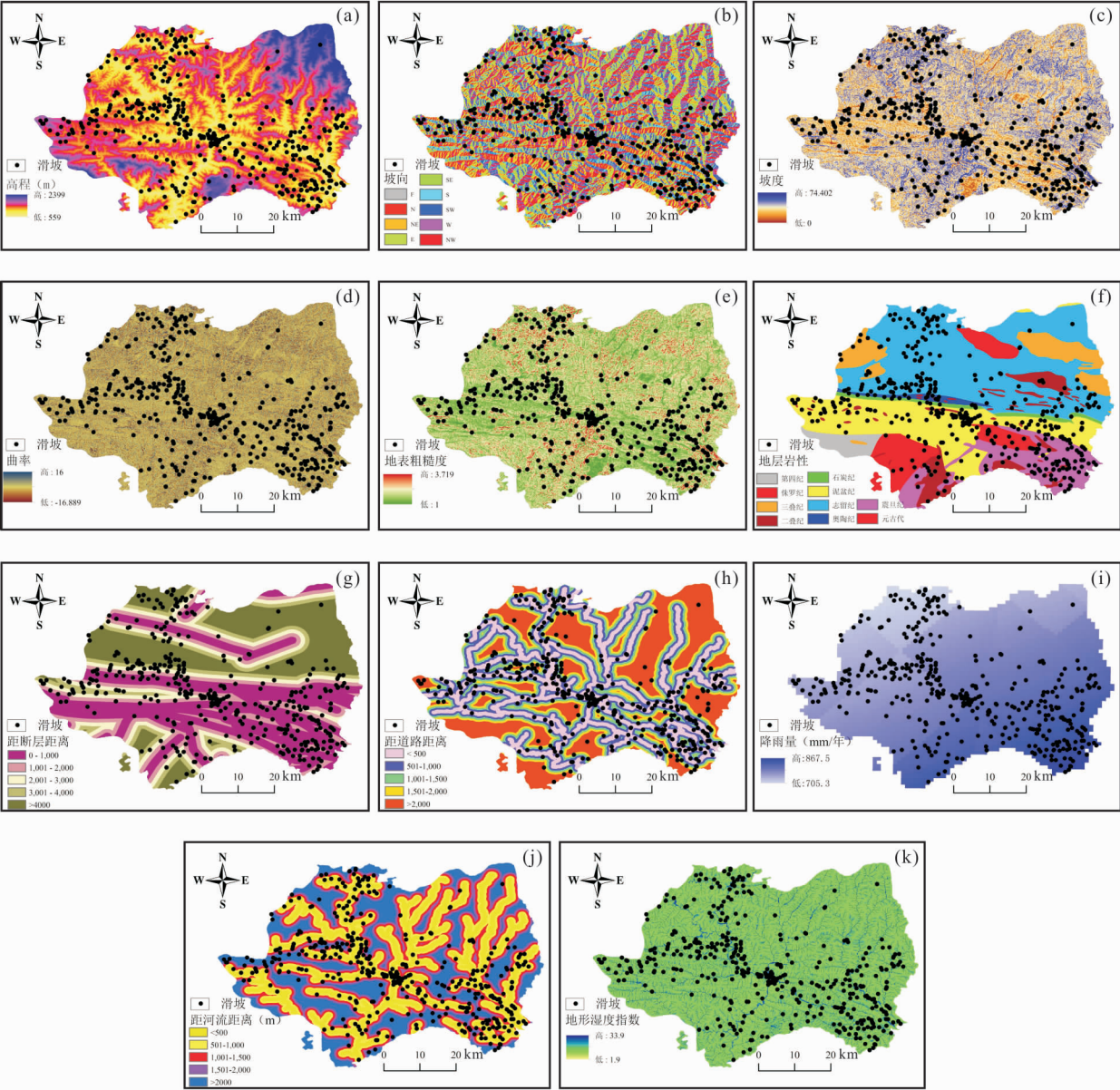


图 2 评价因子分层图

Fig. 2 Evaluation factor classification

模型的准确率为 85.36%；然后将训练模型用于整个研究区,得到全区的易发性指数值,易发性指数值越高,则表明越容易发生滑坡灾害。自然断点法的原理为:通过聚类计算方式使得各分组内的相似性保持最大化,组外的相异性也最大化。将其应用于滑坡的易发性分级,通过国内外学者的研究,分类 5 级有较为优良的效果,能表现出滑坡易发性分组之间的差异性与组内的相似性。因此,将滑坡的易发性等级分为极低易发区、低易发区、中易发区、高易

发区和极高易发区 5 类(图 3)。

4.2 基于 FR 模型评价结果

根据 FR 模型方法,得出各级评价因子的频率比值,计算结果见表 2。因子频率比值越大,表明滑坡易发性程度越高;最后运用 ArcGIS 栅格计算器工具将所有评价因子进行叠加,按照自然间断点法将其分为极低、低、中、高和极高易发区 5 类,得到基于 FR 模型的略阳县滑坡易发性评价区划图(图 4)。

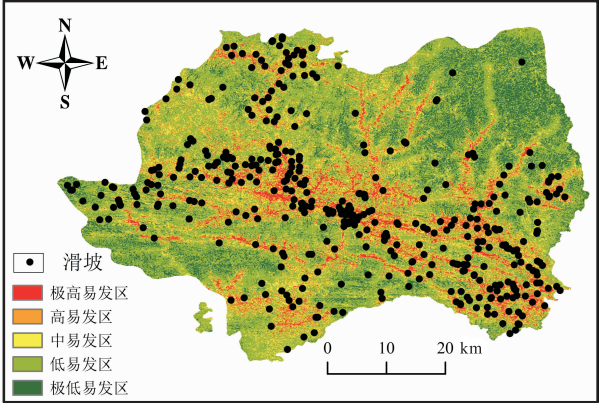


图 3 基于 RF 模型的滑坡易发性区划图

Fig. 3 Landslide susceptibility map based on RF

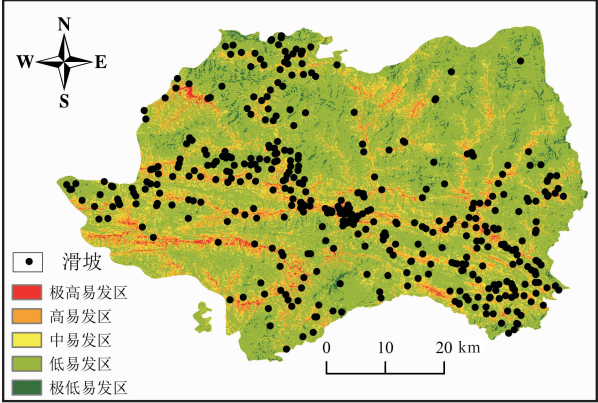


图 5 基于 FR-RF 模型的滑坡易发性区划图

Fig. 5 Landslide susceptibility map based on FR-RF

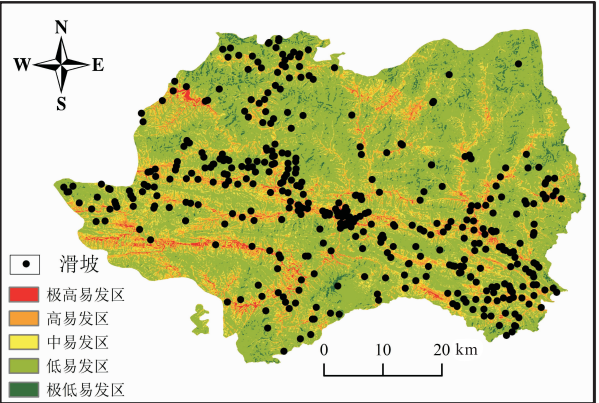


图 4 基于 FR 模型的滑坡易发性区划图

Fig. 4 Landslide susceptibility map based on FR

4.3 基于 FR-RF 模型评价结果

在滑坡易发性评价过程中,由于各级评价因子数据量大,而且数据类型各异,进而影响评价模型的准确性。先将各级评价因子数值采用各分级的频率比替代,则能提高数据分布的规律性;其次把评价因子的数值替换为在该分级的频率比,代入 RF 模型进行计算,得到基于 FR-RF 模型的易发性指数,模型的准确率为 92.83%;将易发性指数导入 ArcGIS 中,得到易发性分区图(图 5),按照自然间断点法将其分为极低易发区、低易发区、中易发区、高易发区和极高易发区 5 类。

5 各评价模型的检验

将 3 种模型的测试样本集的易发性指数值及对

应的真实类标签导入 SPSS 软件 ROC 曲线分析工具中进行分析,得到 3 种评价模型的预测率曲线(图 6),由图 6 可知,FR、RF、RF-RF 3 种模型下训练样本的成功率分别为 0.843、0.901、0.950。其中,RF-RF 模型较其他 2 种模型更好,预测率比 FR、RF 模型分别高出 0.107、0.049,可见 RF-RF 模型更适合研究区滑坡易发性评价。

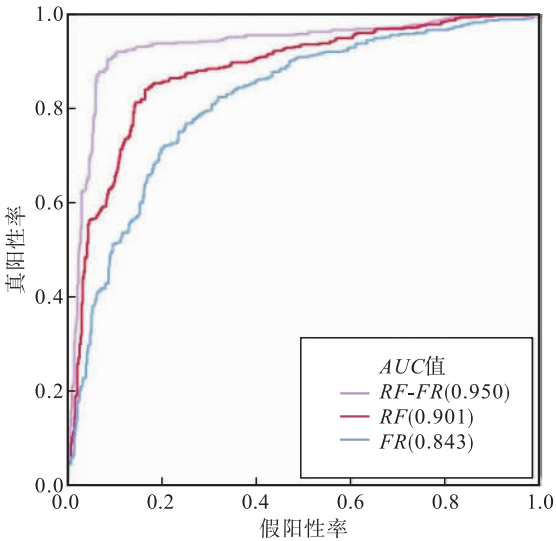


图 6 预测率曲线图

Fig. 6 Graph of success rate

采用 ArcGIS 多值提取致点工具提取各模型灾害点的易发性分级,得到灾害点在各分级内的比例(表 3)。结果表明,RF-RF 模型滑坡灾害点落在极高易发区、高易发区的比例为 87.45%;RF、FR 模

型滑坡灾害点落在极高易发区、高易发区的比例分别为 82.16%、71.56%。说明用 RF - FR 模型进行易发性评价有更佳的效果。

表 3 各易发区灾害点比例表
Tab.3 Percentages of landslide susceptibility

易发性	灾害点比例(%)		
	FR	RF	RF - FR
极低易发性	2.07	1.23	1.16
低易发性	9.83	4.38	2.74
中易发性	16.54	12.23	8.65
高易发性	19.87	15.68	11.84
极高易发性	51.69	66.48	75.61

随机选取略阳县 70 个滑坡点进行实地调研,结果表明,极高、高和中易发区内分布滑坡点个数分别为 53、10 和 7,滑坡点在低易发区与极低易发区无分布。通过与 FR - RF 模型中灾害点比例对比,现场调查得出的各易发区灾害点比例与其相近,验证了 FR - RF 模型预测结果的准确性。

6 结论

(1)根据 Pearson 相关系数的统计分析,地形起伏度、高程变异系数、地表切割深度与坡度的相关性系数绝对值分别为 0.972、0.822、0.962,呈强相关性。剔除地形起伏度、高程变异系数和地表切割深度 3 项因子。

(2)基于滑坡相对点密度(LRPD)对评价因子进行分析,得出滑坡灾害点与线状因子的距离呈负相关,即距离越近,灾害点越多;反之,灾害点越少。

(3)将 RF、FR 2 种模型进行耦合,得到 RF - FR 模型。以 ROC 曲线对 3 种模型精度进行检验,结果表明,RF - FR 模型测试样本的预测率比 RF、FR 模型分别提高了 10.7%、4.9%,显示耦合模型优于单一模型,且 RF - FR 模型用于地质灾害易发性评价具较高的准确性与实用性。

(4)基于 FR、RF 和 RF - FR 模型,滑坡灾害点落在高易发区一极高易发区所占的比例分别为 71.56%、82.16%和 87.45%,得出滑坡在高一极高

易发区发育较为集中,空间分布呈树枝状发育,主要集中在河流、道路所在区。

参考文献(References):

孙果梅,况明生,曲华.陕西秦巴山区地质灾害研究[J].水土保持研究,2005,21(05):244-247.

SUN Guomei,KUANG Mingsheng,QU Hua. Reserch of Geological Disaster in Qinling-Bashan Mountains[J]. Reserch of Soil and Water Conservation,2005,21(05):244-247.

殷坤龙,张桂荣.地质灾害风险区划与综合防治对策[J].安全与环境工程,2003,10(01):32-35.

YIN Kunlong,ZHANG Guirong. Risk Zonation of Geo-hazards and Its Comprehensive Control[J]. Safety and Environmental Engineering,2003,10(01):32-35.

王高峰,郭宁,邓兵,等.不同组合模型区域滑坡易发性及精度分析[J].西北地质,2021,54(02):259-272.

WANG Gaofeng,GUO Ning,DENG Bing,et al. Analysis of Landslide Susceptibility and Accuracy in Different Combination Models [J]. Northwestern Geology, 2021, 54(02):259-272.

胡涛,樊鑫,王硕,等.基于逻辑回归模型和 3S 技术的思南县滑坡易发性评价[J].地质科技通报,2020,39(02):113-121.

HU Tao,FAN Xin,WANG Shuo,et al. Landslide Susceptibility Evaluation of Geological Sinan County Using Logistics Regression Model and 3Stechology[J] Bulletin of Geological Science and Technology, 2020, 39(02):113-121.

许冲,戴福初,姚鑫,等. GIS 支持下基于层次分析法的汶川地震区滑坡易发性评价[J].岩石力学与工程学报,2009,28(S2):3978-3985.

XU Chong,DAI Fuchu,YAO Xin,et al. GIS-Based Landslide Susceptibility Assessment Using Analytical Hierarchy Process in Wenchuan Earthquake Region[J]. Chinese Journal of Rock Mechanics and Engineering, 2009, 28(S2):3978-3985.

高克昌,崔鹏,赵纯勇,等.基于地理信息系统和信息量模型的滑坡危险性评价——以重庆万州为例[J].岩石力学与工程学报,2006,25(05):991-996.

GAO Kechang,CUI Peng,ZHAO Chunyong,et al. Langslide Hazard Evaluation of Wanzhou Based on GIS Informa-

- tion Value Method in the Three Gorges Reservoir[J]. Chinese Journal of Rock Mechanics and Engineering, 2006, 25(05): 991-996.
- 齐信, 黄波林, 刘广宁, 等. 基于 GIS 技术和频率比模型的三峡地区秭归向斜盆地滑坡敏感性评价[J]. 地质力学学报, 2017, 23(01): 97-104.
- QI Xin, HUANG Bolin, LIU Guangning, et al. Landslide Susceptibility Assessment in The Three Gorges Area, China, Zigui Synclinal Basin, Using GIS Technology and Frequency Ratio Model[J]. Journal of Geomechanics, 2017, 23(01): 97-104.
- 刘月, 王宁涛, 周超, 等. 基于 ROC 曲线与确定性系数法集成模型的三峡库区奉节县滑坡易发性评价[J]. 安全与环境工程, 2020, 27(04): 61-70.
- LIU Yue, WANG Ningtao, ZHOU Chao, et al. Evaluation of Landslide Susceptibility Based on ROC and Certainty Factor Method in Fengjie County, Three Gorges Reservoir[J]. Safety and Environmental Engineering, 2020, 27(04): 61-70.
- 王念秦, 郭有金, 刘铁铭, 等. 基于支持向量机模型的滑坡危险性评价[J]. 科学技术与工程, 2019, 19(35): 70-78.
- WANG Nianqin, GUO Youjin, LIU Tieming, et al. Landslide susceptibility assessment based on support vector machine model[J]. Science Technology and Engineering, 2019, 19(35): 70-78.
- 郭子正, 殷坤龙, 黄发明, 等. 基于滑坡分类和加权频率比模型的滑坡易发性评价[J]. 岩石力学与工程学报, 2019, 38(02): 287-300.
- GUO Zizheng, YIN Shenlong, HUANG Faming, et al. Evaluation of Landslide Susceptibility Based on Landslide Classification and Weighted Frequency Ratio Model[J]. Chinese Journal of Rock Mechanics and Engineering, 2019, 38(02): 287-300.
- 邱维蓉, 吴帮玉, 潘学树, 等. 几种聚类优化的机器学习方法在灵台县滑坡易发性评价中的应用[J]. 西北地质, 2020, 53(01): 222-233.
- QIU Weirong, WU Bangyu, PAN Xueshu, et al. Application of Several Cluster-optimization-based Machine Learning Methods in Evaluation of Landslide Susceptibility in Lingtai County [J]. Northwestern Geology, 2020, 53(01): 222-233.
- 吴常润, 赵冬梅, 刘澄静, 等. 基于 GIS 和信息量模型的陇川县滑坡易发性评价[J]. 西北地质, 2020, 53(02): 308-320.
- WU Changrun, ZHAO Dongmei, LIU Chengjing, et al. Landslide Susceptibility Assessment of Longchuan County Based on GIS and Information Value Model[J]. Northwestern Geology, 2020, 53(02): 308-320.
- Cheng Wei, Xie Xiaoshen, Peng Jianbing, et al. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method[J]. Catena, 2018, 164(01): 135-14.
- Cao J, Zhang Z, Wang C Z, et al. Susceptibility assessment of landslides triggered by earthquakes in the Western Sichuan Plateau[J]. Catena, 2019, 175(02): 63-76.