

基于可解释性机器学习的丘陵缓坡地区滑坡易发性区划研究*

张虹^{1,2}, 辜庆渝^{1,2}, 孙诚彬³, 孙德亮^{1,2}, 密长林⁴, 张凤太⁵

(1. 重庆师范大学 地理与旅游学院; 2. 重庆师范大学 GIS 应用研究重庆市高校重点实验室, 重庆 401331;
3. 临沂市公路事业发展中心莒南县中心, 山东 临沂 276000; 4. 山东省临沂市自然资源开发服务中心, 山东 临沂 276000;
5. 重庆理工大学 管理学院, 重庆 400054)

摘要:【目的】提出一种基于随机森林与 Permutation Importance、PDP 和 LIME 结合的可解释性模型,对滑坡易发性区划进行全局和局部解释,旨在为滑坡灾害精准防治与城市规划提供理论依据。【方法】以重庆市江津区为例,选取地形地貌、地质构造、气象水文、环境条件和人类活动共 5 个方面的 21 个因子,结合江津区 899 个历史滑坡点,建立 30 m×30 m 精度的栅格空间数据库,按照滑坡与非滑坡 1:1 的比例选取 899 个非滑坡点,利用随机森林算法构建滑坡易发性模型,将结果分为极低、低、中、高、极高等 5 个易发性等级,探讨了随机森林模型在三峡库区滑坡易发性区划中的普适性,最后通过 Permutation Importance、PDP、LIME 方法研究随机森林模型的可解释性。【结果】滑坡高-极高易发区内滑坡点数占历史总滑坡点的 71.3%,面积占区域总面积的 20.42%,混淆矩阵准确率为 0.968,全体数据集 AUC 值达 0.962。通过模型解释可知地形起伏度、年平均降雨量、坡度是滑坡易发性区划中最重要的因子,且地形起伏度、坡度为正影响,当年平均降雨量小于 1 300 mm 时,对滑坡的发生也产生正影响。【结论】基于可解释性机器学习的滑坡易发性区划模型预测精度高,对滑坡的精准防治有重要的实践意义。

关键词:随机森林;可解释性机器学习;江津区;滑坡易发性区划;三峡库区

中图分类号:X87

文献标志码:A

文章编号:1672-6693(2022)03-0078-15

滑坡是全球山区常见的一种自然灾害,具有分布广、频率高、速度快等特点,对人类的生命财产安全构成巨大的威胁^[1]。2019 年 7 月 23 日中国贵州省六盘水市水城县鸡场镇坪地村岔沟组发生特大山体滑坡,导致 21 幢房屋被埋,43 人遇难,9 人失联,1 600 多人受伤,直接经济损失达 1.9 亿元^[2]。从 2004 到 2016 年,世界各地共发生 4 862 起造成不同影响的滑坡灾害,导致 55 997 人死亡^[3]。因此,对区域进行滑坡易发性区划,对滑坡灾害的预警预报、防治管理等具有重要意义。

滑坡易发性区划是一种预防滑坡灾害的有效方法,它是根据当地地质环境因素来评估特定区域中滑坡发生概率的空间分布。滑坡易发性分析方法有很多,早期人们使用统计类方法预测滑坡发生的概率,该方法工作量大且评价结果不理想;基于专家经验的预测方法逐渐兴起,如层次分析法(AHP)、模糊逻辑、布尔逻辑^[4]等,此类方法以定性分析为主,主观性较强,且当指标过多时,权重难以确定。随着计算机技术的发展,越来越多学者将机器学习应用于滑坡灾害预测中。常用的机器学习算法有逻辑回归(LR)^[6]、支持向量机(SVM)^[8]、人工神经网络(ANN)^[10]等,这些算法实现了大量数据的快速分析,但具有建模复杂、抗噪声能力弱,易出现过拟合现象等缺点。相关研究显示,随机森林应用于滑坡易发性研究,比 LR,ANN,SVM 等模型性能更优^[11-16],该模型算法支持在大量数据集上有效运行,且有较高的预测精度,模型运行时不易产生过拟合现象,具有很好的抗噪声能力^[17]。

目前,滑坡易发性评估侧重于开发高精度的评价模型,而对模型的评估结果缺乏充分地解释。精度越高的机器学习算法可解释性越低^[18],限制了评价模型在滑坡灾害预测预警中的使用。为了提高机器学习的可解释性,建立用户与模型间的信任关系,近些年来学界兴起了可解释性机器学习的研究,并提出了一系列的可解释性方法。如基于 SHAP 的冠心病死亡率预测^[19],应用可解释性机器学习估计农作物产量^[20],基于 LIME 的交通安

* 收稿日期:2022-02-19 修回日期:2022-03-22 网络出版时间:2022-05-18 09:01

资助项目:国家自然科学基金(No. 42071217);国家重点研发计划(No. 2018YFC1505501);教育部人文社科规划项目(No. 20XJAZH002);重庆市自然科学基金(No. cstc2020jcyj-msxmX0841)

第一作者简介:张虹,女,副教授,博士,研究方向为人类活动扰动下的生态、地质环境风险,E-mail:zh_angh@cqu.edu.cn;通信作者:孙德亮,男,副教授,博士,E-mail:67130612@qq.com

网络出版地址:https://kns.cnki.net/kcms/detail/50.1165.N.20220517.1701.009.html

全的可解释性研究^[21]。

本文为了提高滑坡易发性模型的可解释性,以重庆市江津区为例,选取地形地貌、地质构造、气象水文、环境条件和人类活动共 5 个方面共 21 个评价指标,通过随机森林算法对江津区进行滑坡易发性分析,并通过 Permutation Importance,PDP,LIME 等算法研究模型的可解释性,这样在提高模型预测精度的同时,加强了模型的可解释性。

1 研究区概况与数据源

1.1 研究区概况

江津区位于长江中上游,地处重庆市西南部,面积约 3 218 km²。该区属四川盆地边缘,地形以山地和丘陵为主,其中低山占 31.8%,丘陵占 65.1%。全区地势南北高、中间低,长江自西南向东北横贯中北部,两侧地势均向江面倾斜。江津区气候属亚热带季风气候,降雨丰富,年平均降雨量为 1 086.5~1 312.5 mm,雨量北少南多。区域内地质岩性以砂岩、泥岩为主,岩体风化破碎。随着人类活动的加剧,加上降雨集中,该区地质灾害频发,其中尤以滑坡灾害最为突出,为江津区分布广、规模大、危害性最为严重的地质灾害。

1.2 数据来源

本研究选取了江津区多年来 899 个历史滑坡点数据,并收集地形地貌、地质构造、气象水文、环境条件和人类活动共 5 个方面的滑坡影响因子数据,具体数据来源如表 1 所示。

表 1 数据及来源
Tab.1 Data and their sources

数据名称	数据来源	类型	精度
历史滑坡	重庆地质监测站	数据表	
DEM	Aster 卫星	栅格	30 m×30 m
地质数据	国家地质资料数据中心 (www.ngac.org.cn)	栅格	1:200 000
土地利用(2017 年)	地理国情监测云平台 (www.dsac.cn)	矢量	1:100 000
行政区划	同上	矢量	1:100 000
年平均降雨量	同上	栅格	30 m×30 m
河网	重庆市水利局	矢量	1:100 000
道路	中国科学院资源环境科学与数据中心 (www.resdc.cn)	矢量	1:100 000
卫星影像(2017 年)	地理空间数据云平台	栅格	30 m×30 m
兴趣点(POI)	百度 API 爬虫	数据表	

2 地理空间数据库

2.1 影响因子

滑坡影响因子众多,但少有研究考虑到了人类活动对滑坡的影响。POI 核密度可反映城市功能区的集中程度,即 1 个 POI 可视为 1 个功能单元,泛指一切可以抽象为点的地理对象;POI 密度越大则该城市功能越集中,可间接反映人类活动强度。滑坡形成机理复杂,是内在因素与外在因素的共同作用的结果。其中,内在因素主要有斜坡地层的岩性、地质构造与内部结构;外在因素有地表水、地下水活动,地震及人类活动等。环境因子的客观性、存在性、继承性、显著性等是滑坡影响因素的选取原则。在此基础上,结合专家经验^[22-24],从研究区自身地理条件出发,选取高程、年平均降雨量、POI 核密度、距道路距离等 21 个影响因子,构建江津区滑坡易发性评价体系,各指标的详细描述见表 2。

2.2 数据处理

借助 ArcGIS 对 DEM 数据处理后得到坡度、曲率、微地貌和地形粗糙度。重庆市 1:200 000 地质图矢量化得到岩性。基于欧氏距离计算距道路、距河流水系距离。基于 Landsat8 遥感数据获得归一化植被指数 (NDVI)。参照相关研究成果^[22-24],对所有滑坡影响因子进行分级,如表 3 所示。获得的滑坡影响因子专题图层

如图 1 和图 2 所示。

表 2 滑坡影响因素及相应描述
Tab. 2 Description of conditioning factors

类型	影响因子	描述
地形地貌	高程	直接影响降雨量、植被的生长情况、人类活动强度等
	坡度	反映了地表单元的陡缓程度,决定斜坡的稳定性
	坡向	影响坡体的可照时数,从而影响对植被的生长状况
	曲率	影响土壤侵蚀及地表径流在斜坡上的运动速度
	平面曲率	是垂直于最大坡度方向的斜率,影响地表径流,土壤侵蚀等
	剖面曲率	是高程变化的二次导数,反映地表曲面在垂直方向的弯曲变化情况,影响土壤沉积、水流速度
	坡位	是坡面所处的地貌部位,不同坡位的斜坡其地下水分布、岩层倾向、岩性等都存在差异
	微地貌	是最小地貌形态单元,影响边坡稳定性
	地形起伏度	区域地形特征的宏观指标
	地形粗糙度	是地球表面积与区域的投影面积之比,宏观描述地表形态
	地形湿度指数	是地形对水文过程控制的量化,定量地反映地形对径流流向和积蓄影响
地质构造	岩性	是滑坡发生的物质基础,反映岩石的软硬程度及成因
	倾坡类型	是坡向与岩层倾向的关系类型,影响边坡稳定性的重要因素
气象水文	距河流域距离	受到河流的侵蚀作用,斜坡稳定性下降,离水系越近,坡体含水量越高,软化越严重
	水流动力指数	是地表水侵蚀能力的定量表达,反映地表水的输送能力
	输沙指数	是地表水对泥沙输送能力的定量表达,值越大输送能力越强
	年平均降雨量	影响土壤侵蚀和植被长势
环境条件	NDVI	反映植被覆盖情况,选取 7—9 月份滑坡易发季节的 NDVI
	土地利用	反映人类对岩土层的改造程度
人类活动	距道路距离	反映边坡坡脚的稳定性
	POI 核密度	反映人类活动强度

3 方法

3.1 模型训练过程

整个训练过程如图 3 所示。首先构建数据库。选取地形地貌、地质构造、气象水文、环境条件、人类活动等 5 个方面共 21 个评价指标,所有指标重新采样为 30 m×30 m 的栅格数据;参照 899 个历史滑坡点,按 1 : 1 提取非滑坡点形成滑坡清单。其次,通过随机森林算法构建江津区滑坡易发性模型。利用 ROC 曲线和混淆矩阵进行模型评价。最后,运用 Permutation Importance, PDP, LIME 算法解释模型。

3.2 随机森林

随机森林最初是由 Breiman^[25]和 Cutler 等人^[26]提出,它是由多棵决策树组合形成的集成方法,可处理大数据,且具有较高的精度和抗噪声能力。它的基本思想是使用多棵树对样本进行训练与预测,对于每棵树使用的训练集均从总的训练集中采样得出。随机森林的结果是对多棵决策树的判断结果投票得到的。因此,使用随机森林模型时无需人为给定因子的权重,降低了人为主观因素对模型预测结果的影响,随机森林具体流程如图 4 所示。

随机森林的核心在于从原始数据中抽取 N 个与原始数据集容量相同的样本,然后对 N 个样本建立 N 颗互不关联的决策树模型 $[h(X, \theta N); N=1, \cdots]$,从而得到 N 种分类结果,并对结果进行投票分类决策。具体公式如下:

$$H(x) = \arg \max_Z \sum_{i=1}^k I(h_i(x) = Z)。$$

其中; $h_i(x)$ 表示某个单独的决策树, Z 表示输出变量(目标变量), $I(\cdot)$ 为示性函数。

表 3 滑坡影响因子分类

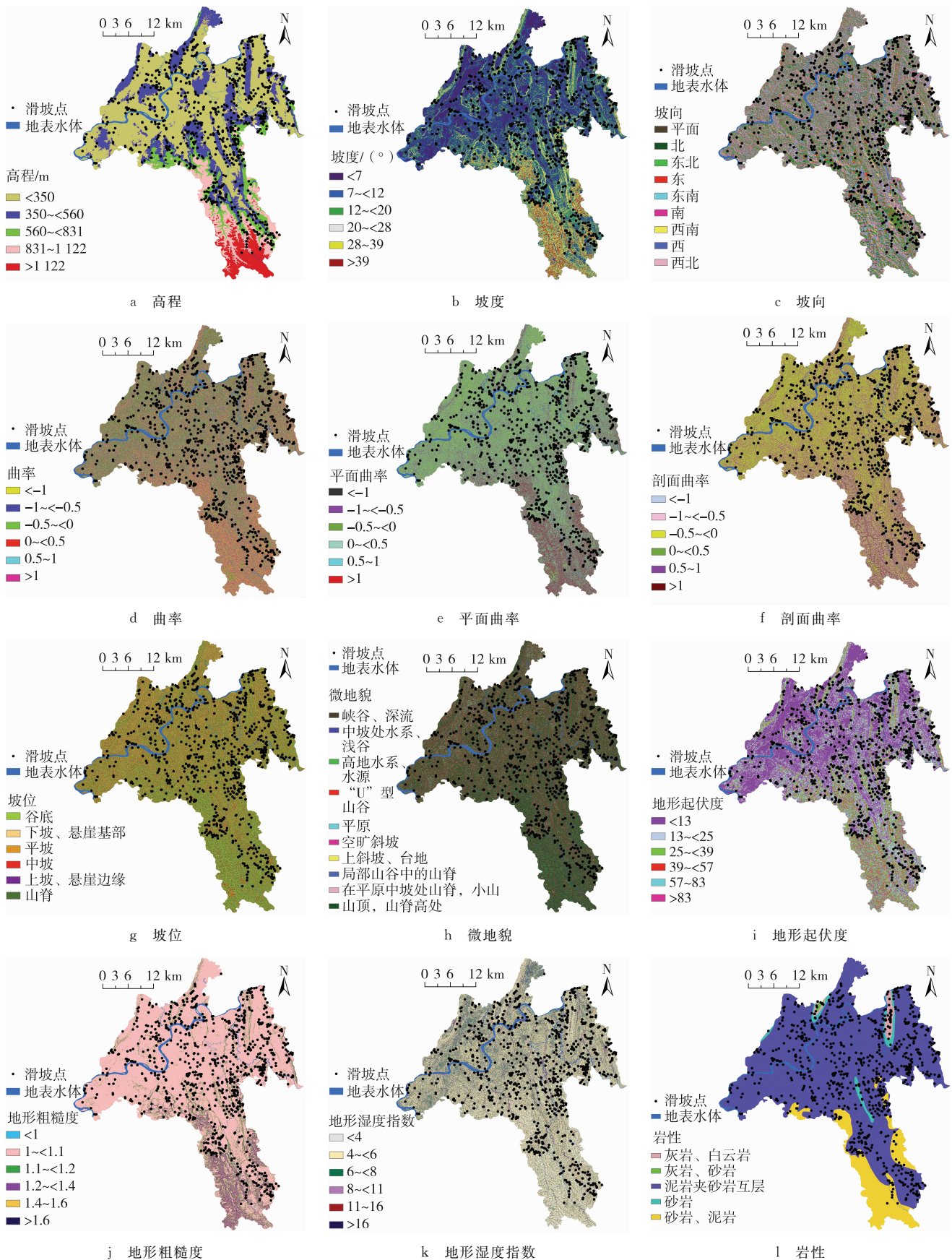
Tab. 3 Influencing factors categories of landslide

类型	评价因子	分级/个	分级标准
地形地貌	高程/m	5	1) <350;2) 350~<560;3) 560~<831;4) 831~1 122;5) >1 122;
	坡度/(°)	6	1) <7;2) 7~<12;3) 12~<20;4) 20~<28;5) 28~39;6) >39;
	坡向	9	1) 平面;2) 北;3) 东北;4) 东;5) 东南;6) 南;7) 西南;8) 西;9) 西北
	曲率	6	1) <-1;2) -1~<-0.5;3) -0.5~<0;4) 0~<0.5;5) 0.5~1;6) >1
	平面曲率	6	1) <-1;2) -1~<-0.5;3) -0.5~<0;4) 0~<0.5;5) 0.5~1;6) >1
	剖面曲率	6	1) <-1;2) -1~<-0.5;3) -0.5~<0;4) 0~<0.5;5) 0.5~1;6) >1
	坡位	6	1) 谷底;2) 下坡、悬崖基部;3) 平坡;4) 中坡;5) 上坡、悬崖边缘;6) 山脊
	微地貌	10	1) 峡谷、深流;2) 中坡处水系、浅谷;3) 高地水系、水源;4) “U”型山谷;5) 平原;6) 空旷斜坡;7) 上斜坡、台地;8) 局部山谷中的山脊;9) 在平原中坡处山脊,小山;10) 山顶,山脊高处
	地形起伏度	6	1) <13;2) 13~<25;3) 25~<39;4) 39~<57;5) 57~83;6) >83
	地形粗糙度	5	1) <1;2) 1~<1.1;3) 1.1~<1.2;4) 1.2~<1.4;5) 1.4~1.6;6) >1.6
	地形湿度指数	6	1) <4;2) 4~<6;3) 6~<8;4) 8~<11;5) 11~16;6) >16
地质构造	岩性	5	1) 灰岩、白云岩;2) 灰岩、砂岩;3) 泥岩夹砂岩互层;4) 砂岩;5) 砂岩、泥岩
	倾坡类型	7	1) 第Ⅰ型顺向坡/外倾坡;2) 第Ⅱ型顺向坡/外倾坡;3) 斜交外倾坡;4) 斜交坡;5) 横交坡/切向坡;6) 逆向坡/内倾坡;7) 平叠坡
气象水文	距河流水系距离/m	11	1) <100;2) 100~<200;3) 200~<300;4) 300~<400;5) 400~<500;6) 500~<600;7) 600~<700;8) 701~<800;9) 800~<900;10) 900~1 000;11) >1 000
	水流动力指数	7	1) <15;2) 15~<30;3) 30~<45;4) 45~<60;5) 60~<100;6) 100~1 000;7) >1 000
	输沙指数	6	1) <20;2) 20~<40;3) 40~<70;4) 7~<100;5) 100~200;6) >200
	年平均降雨量/mm	11	1) <1 114;2) 1 114~<1 133;3) 1 133~<1 156;4) 1 156~<1 186;5) 1 186~<1 219;6) 1 219~<1 251;7) 1 251~<1 282;8) 1 282~<1 311;9) 1 311~<1 340;10) 1 340~1 371;11) >1 371
环境条件	NDVI	7	1) <-0.6;2) -0.6~<-0.2;3) -0.2~<0.1;4) 0.1~<0.35;5) 0.35~<0.55;6) 0.55~0.8;7) >0.8
	土地利用	8	1) 旱地;2) 林地;3) 水域;4) 建设用地;5) 草地;6) 裸岩;7) 园地;8) 水田
人类活动	POI 核密度	10	1) <1;2) 1~<2;3) 2~<3;4) 3~<4;5) 4~<5;6) 5~<100;7) 100~<500;8) 500~<1 000;9) 1 000~2 000;10) >2 000
	距道路距离/m	7	1) <100;2) 100~<200;3) 200~<300;4) 300~<400;5) 400~<500;6) 500~600;7) >600

3.3 滑坡易发性区划的解释模型

无法解释的滑坡易发性模型会带来一定风险,因为随机森林模型类似于黑盒,无法控制模型内部的运行,只能在不同的参数和随机种子之间进行尝试。故研究引入 Permutation Importance、PDP 与 LIME 解释模型。

3.3.1 Permutation Importance Permutation Importance 是与模型无关的衡量特征重要性的有效方法,通过改变某个特征的特征值大小,证明该值的重要性,变化越大,重要性越大。具体步骤如下:1) 模型训练完成后选择一个特征 X_1 ;2) 对该特征置随机数,并计算新的预测结果;3) 将新旧两个预测结果进行对比即可得到 X_1 对模型预测的影响。



注:底图数据来源于国家地球系统科学数据中心(<http://www.geodata.cn>),审图号:渝 S(2017)032 号,下同

图 1 滑坡影响因子专题图层(I)

Fig. 1 The thematic layers of landslide influencing factor (I)

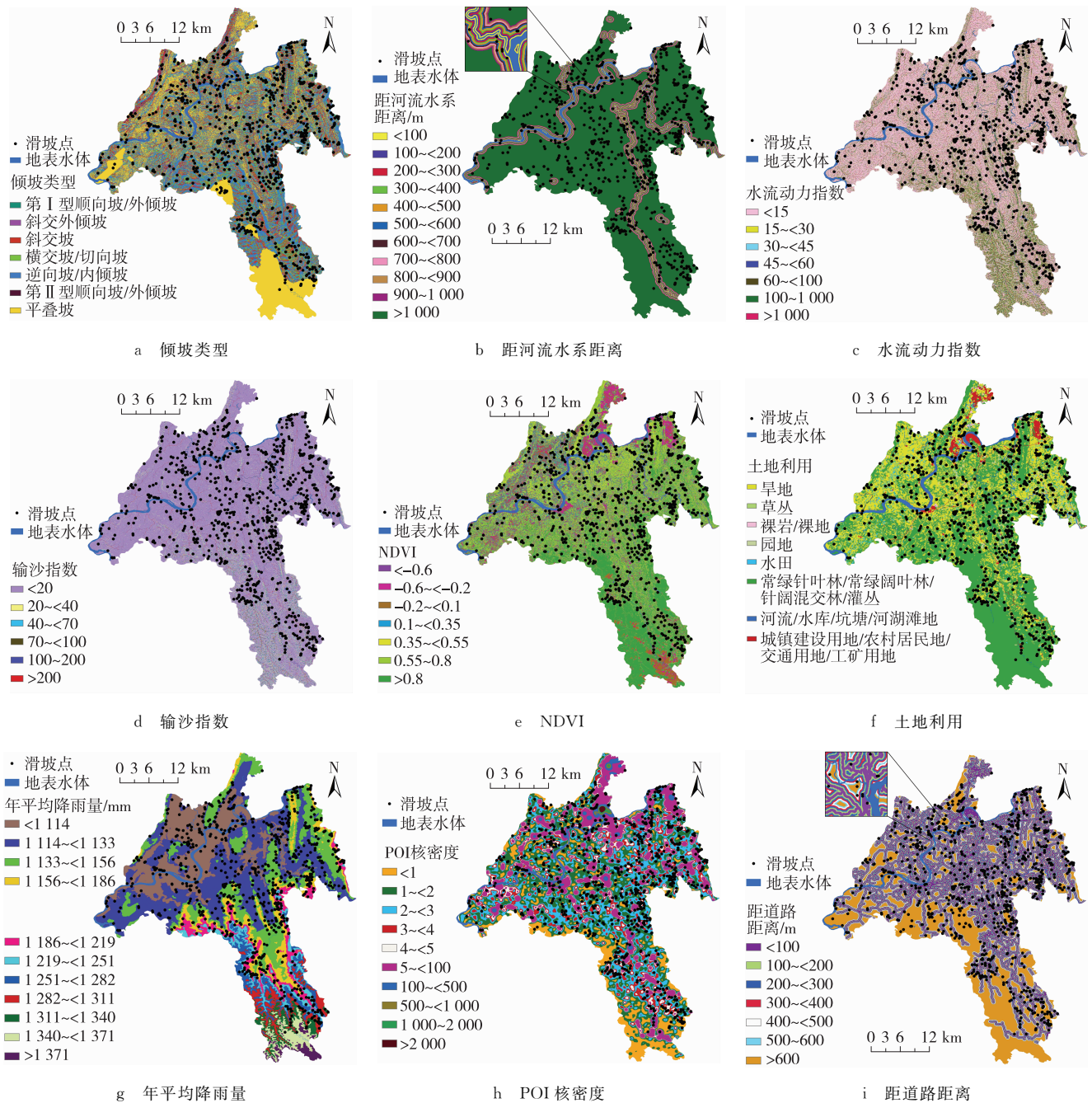


图2 滑坡影响因子专题图层(II)

Fig. 2 The thematic layers of landslide influencing factor (II)

3.3.2 PDP 该方法基本思想是针对一个机器学习模型的 PDP, 改变重点关注的特征, 控制其他特征不变, 观察该特征对预测结果的影响^[27]。同时, PDP 可以说明特征变量和预测结果之间的关系^[28]。具体公式如下:

$$\hat{f}_s(x_s) = E_{x_c}[\hat{f}(x_s, x_c)] = \int \hat{f}(x_s, x_c) dP(x_c),$$

其中: x_s 为某一特征变量, 通常集合 s 中仅有 1 个到两个特征; $\hat{f}_s(x_s)$ 为在 x_s 每个取值下对应的标签估计值; x_c 为除 x_s 以外的其他特征变量。E 表示数学期望, 它是求当特征变量 x_s 固定时, x_c 在它对应边际分布变化时的平均值; $P(x_c) = \int p(x_{all}) dx_s$, 式中 x_{all} 表示所有特征变量, $\int p(x_{all})$ 是模型中所有特征变量的联合概率密度。

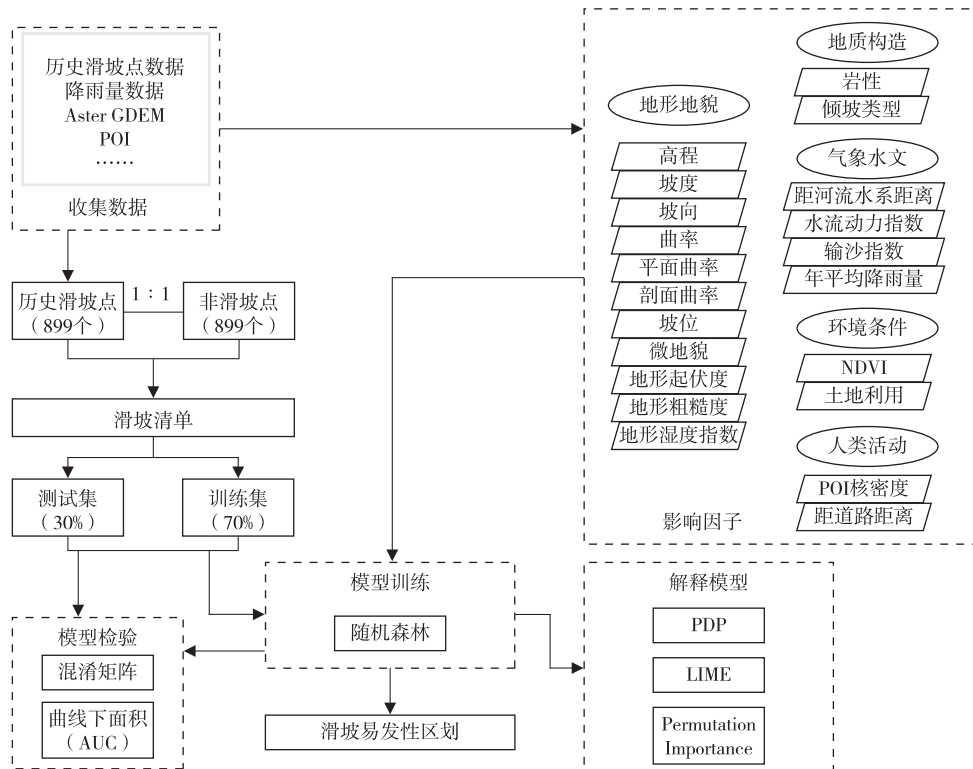


图 3 方法流程图

Fig. 3 Flowchart of the method

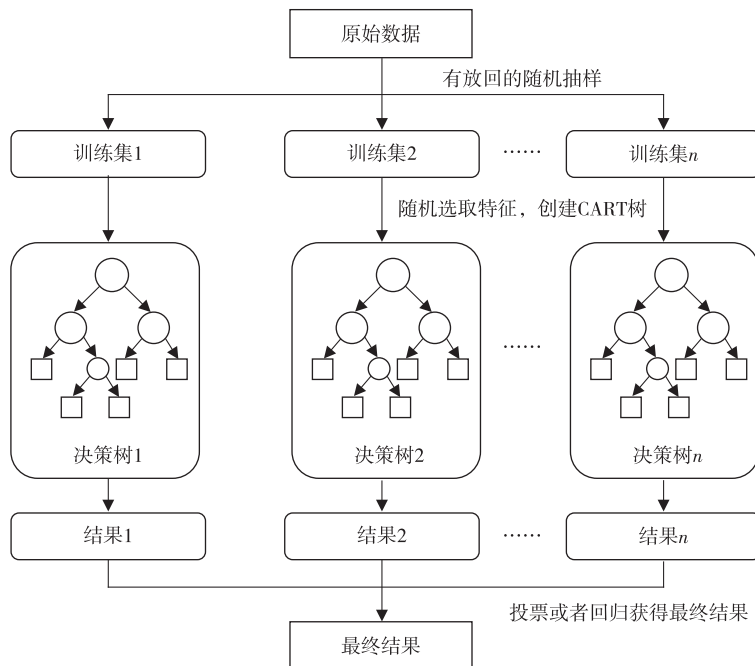


图 4 随机森林流程图

Fig. 4 Flowchart of random forest

3.3.3 LIME LIME 是一种与模型无关的局部可解释性算法,由 Marco 等人^[29]于 2016 年提出。LIME 能真实地反应分类器在预测样本时的行为,它针对单条样本,假定在局部模型是简单的线性模型,对局部数据点做出解释。LIME 将输入值在局部点周围做微小的扰动,观察模型的预测行为,并根据扰动点与原始数据间距离分配权重,得到一个可解释的模型和预测结果。具体公式如下:

$$M_{\text{explanation}}(x) = \arg \min_{g \in G} L(f, g, \pi_x) + \Omega(g) \quad (1)$$

对于实例 x 的解释模型 g , 通过最小化损失函数来比较模型 g 和原模型 f 的近似性, (1) 式中: $\Omega_{(g)}$ 代表了解释模型 g 的模型复杂度, G 表示所有可能的解释模型, π_x 定义了 x 的邻域, 通过最小化 L 使得模型变得可解释。

4 结果分析

4.1 混淆矩阵

滑坡为典型的二分类问题, 本文将 0.5 作为阈值, 对结果进行二分类, 大于 0.5 表示会发生滑坡, 小于 0.5 表示不会发生滑坡。本文使用混淆矩阵检验模型精度, 结果如表 4 所示。模型准确率为 0.968, 召回率为 0.967, 全体数据集 AUC 值为 0.982, 训练集构建的模型较为精确。

4.2 受试者工作特征(ROC)曲线

ROC 曲线是衡量模型准确性的一种常用指标。在 ROC 曲线图中, 横坐标表示假阳性率, 纵坐标表示真阳性率, 越靠近左上角的点对应模型精度越高。AUC 值为 ROC 曲线所覆盖的区域面积, AUC 值越大, 模型精度越高。图 5 为随机森林模型的 ROC 曲线, 其中训练集、测试集、全体数据集的 AUC 值分别为 0.975, 0.794, 0.962。

表 4 随机森林模型混淆矩阵
Tab. 4 Confusion matrix of random forest

随机森林 预测值	真实值	
	滑坡	非滑坡
滑坡	874	33
非滑坡	25	866

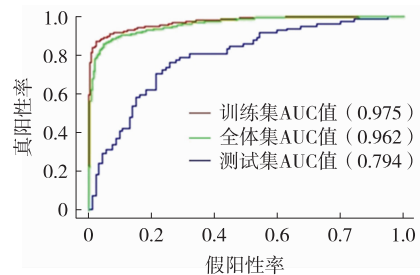


图 5 ROC 曲线及 AUC 值

Fig. 5 ROC curve and AUC value

4.3 滑坡易发性区划

将训练完成的随机森林模型代入 GIS 软件中, 得到每个栅格范围的滑坡发生概率 p , 并按专家经验划分为 5 个易发性等级, 分别是极低易发性 ($p < 0.037$)、低易发性区划 (p 为 $0.037 \sim < 0.100$)、中易发性 (p 为 $0.101 \sim < 0.175$)、高易发性 (p 为 $0.176 \sim 0.340$)、极高易发性 ($p > 0.340$)。低-中易发区主要分布于浅丘平坝区, 地形起伏小具有明显的山地性; 高-极高易发区主要分布于长江沿岸、低山山麓带、暴雨区及高强度人类活动区。图 6 显示: 1) 长江及支流沿岸地区受河水侵蚀冲刷, 岸坡稳定性被破坏, 属于滑坡高易发区; 2) 各背斜翼部低山山麓带区域切割强烈, 相对高差大, 滑坡易发; 3) 多雨和暴雨区, 如年平均降雨量在 1 300 mm 以上的珞璜镇、四面山镇等地滑坡易发; 4) 四面山高速、绕城高速周围等人类活动较强的地区, 滑坡易发。

分别统计各等级区域的栅格数以及滑坡数, 并计算滑坡密度, 结果如表 5 所示。极低-低易发区面积占全区总面积的 60.4%, 而滑坡点数仅占全区历史滑坡点数的 15.68%; 而高-极高易发区面积占比为 20.43%, 滑坡点数占比达 71.3%。因此, 随着滑坡易发性等级的增加, 区域面积减小, 滑坡密度增大, 历史滑坡与易发性等级呈明显的正相关关系^[31], 说明该滑坡易发性区划模型与实际情况契合较好。

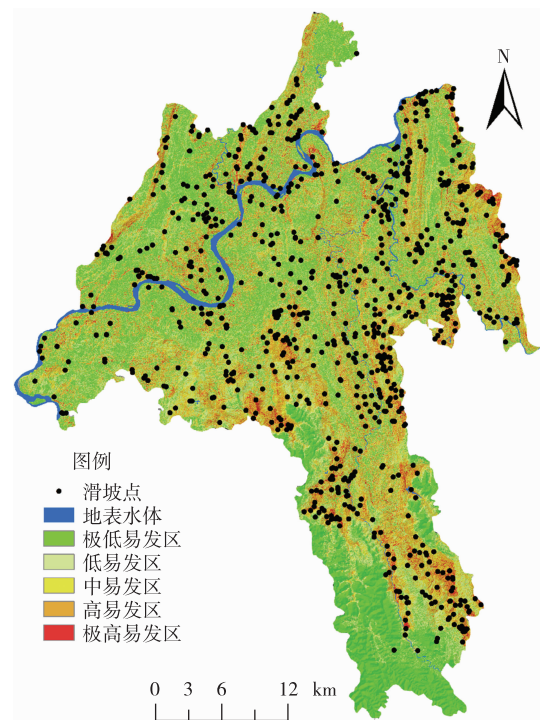


图 6 滑坡易发性区划图

Fig. 6 Landslide susceptibility zoning map

表 5 滑坡易发性分级统计

Tab. 5 Classification statistics of landslide susceptibility

易发性等级	滑坡发生概率 p	面积占比	滑坡点数/个	密度/(个·km ⁻²)
极低易发区	<0.037	32.66%	41	0.039 4
低易发区	0.037~<0.100	27.74%	100	0.113 3
中易发区	0.100~<0.175	19.17%	117	0.191 7
高易发区	0.175~0.340	16.85%	219	0.408 6
极高易发区	>0.340	3.58%	422	3.702 8

5 模型的可解释性研究

5.1 因子的重要性

不同因子对滑坡的影响程度不同,本研究从全局角度来确定滑坡因子的影响大小。即计算每个因子的重要性,重要性的值越高,因子对预测结果的影响越大(表 6)。

表 6 各因子重要性

Tab. 6 Importance of each factor

因子	重要性	因子	重要性	因子	重要性
地形起伏度	0.039 2±0.018 2	POI 核密度	0.008 5±0.015 3	剖面曲率	0.000 0±0.007 5
年平均降雨量	0.037 0±0.015 1	平面曲率	0.007 1±0.015 8	坡向	-0.000 9±0.013 4
坡度	0.034 3±0.014 3	地形湿度指数	0.003 6±0.008 3	输沙指数	-0.000 9±0.004 5
高程	0.022 3±0.010 5	微地貌	0.001 8±0.008 2	NDVI	-0.001 3±0.010 4
距河流水系距离	0.016 0±0.013 6	曲率	0.000 9±0.007 6	倾坡类型	-0.001 3±0.016 3
距道路距离	0.012 5±0.021 6	岩性	0.000 4±0.005 9	水流动力指数	-0.001 8±0.005 9
土地利用	0.008 5±0.011 8	地形粗糙度	0.000 4±0.001 8	坡位	-0.002 2±0.008 5

5.2 PDP

PDP 可以探测每个因子内部对滑坡预测概率的影响。本文选取重要性排名前 4 个的因子进行分析,部分依赖图如图 7 所示。图中横坐标为因子的取值范围,纵坐标为预测结果,灰色区域为置信区,即可能发生波动的范围。由图 7 可知:当地形起伏度小于 39 m 时,该因子对预测结果呈正影响(即对滑坡的发生有促进作用),且随着地形起伏度的增大,对预测值的影响能力也增大;当地形起伏度大于 39 m 时,它对预测值的影响开始缓慢下降。当坡度小于 20°时对预测值成正影响,且随着坡度的增大对预测结果的影响也增大,当坡度大于 20°时对预测值的影响开始缓慢下降。年平均降雨量在 1 133~<1 219 mm 时,与预测结果成正相关关系,随后下降。高程小于 831 m 时,对预测结果产生正影响,当高程大于 831 m 时,对预测结果产生负影响。

5.3 LIME 的模型解释

通过 LIME 分别取预测结果为发生滑坡与不发生滑坡的样本各两个(表 7),并对它们进行解释(图 8 和图 9),结果显示每个因子对该样本预测结果的贡献大小。图 8 和图 9 中正数代表该因子对结果产生正影响,负数相反。由两图可知,地形起伏度、坡度、距河流水系距离、距道路距离是模型识别滑坡、非滑坡的最重要因子,其中:距河流水系距离、距道路距离对滑坡预测结果产生负影响;起伏度、坡度对非滑坡预测结果产生负影响。

6 讨论

6.1 模型的可解释性

前人对滑坡易发性模型的讨论仅停留在因子重要性与滑坡易发性区域的空间分布上,未就每个因子内部对滑坡预测概率的影响进行探讨。本文利用 PDP 来解释因子内部不同取值区间对滑坡预测概率的影响。结果显示:研究区地形起伏度与滑坡预测概率成正相关关系;年平均降雨量小于 1 300 mm 时对滑坡的发生产生正影响,大于 1 300 mm 后对滑坡的发生产生负影响。

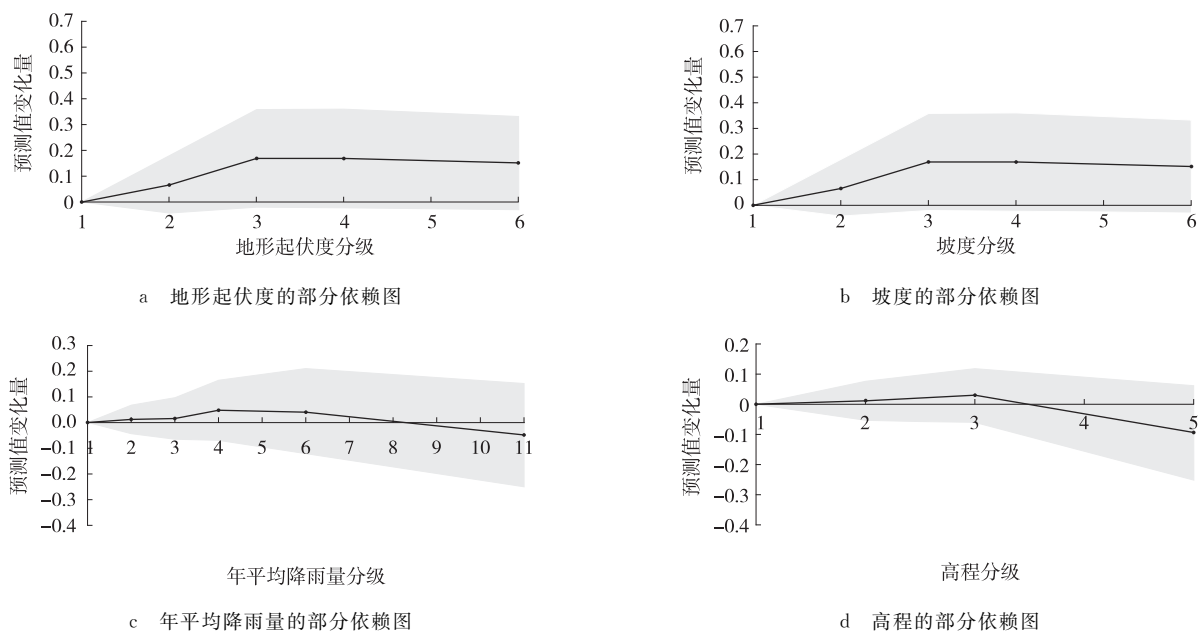


图 7 重要性排名前 4 的因子的部分依赖图
Fig. 7 PDP for the top 4 factors of importomce

表 7 LIME 解释的四条数据
Tab. 7 Four numbers explained by LIME

因子	研究对象编号				因子	研究对象编号			
	339	939	519	1025		339	939	519	1025
地形起伏度	2	2	5	2	坡位	1	1	1	2
坡度	3	3	6	2	曲率	1	4	1	3
距河流水系距离	6	9	10	11	平面曲率	2	4	3	3
距道路距离	2	7	7	2	地形湿度指数	3	2	1	3
岩性	6	6	10	6	高程	1	3	4	1
水流动力指数	4	5	3	2	坡向	5	6	2	7
土地利用	2	1	2	1	NDVI	5	6	3	5
倾坡类型	3	2	7	5	剖面曲率	5	4	6	4
年平均降雨量	3	5	7	2	输沙指数	1	2	1	1
POI 核密度	6	1	2	2	地形粗糙度	2	2	4	2
微地貌	10	7	1	1					

注:339 号和 939 号研究对象的二分类结果为 1, 519 号和 1025 号研究对象的二分类结果为 0

另外,针对以往滑坡易发性模型欠缺模型局部解释的问题,本研究引入了 LIME 算法,对模型进行局部解释,揭示滑坡发生概率的空间分异性。结果显示:地形起伏度、坡度对预测为滑坡起促进作用;距河流水系距离、距道路距离对预测为滑坡起抑制作用。

以往基于随机森林的滑坡易发性研究中普遍用平均基尼指数反应特征的重要性^[31-32],但平均基尼指数与随机森林算法本身相关,方法不具通用性,本文采用的 Permutation Importance 与模型算法本身无关,能更客观地解释滑坡易发性模型的因子重要性。

6.2 滑坡的制约因素及重要性

不同因子对滑坡的影响程度不同,本文从全局角度计算了各因子的重要性(图 10)。各因子的重要性由大到小依次为地形起伏度、年平均降雨量、坡度、高程、距河流水系距离、距道路距离、土地利用、POI 核密度、平面曲

率、地形湿度指数、微地貌、岩性、地形粗糙度、剖面曲率、坡向、输沙指数、NDVI、倾坡、水流动力指数、坡位。

通过部分依赖图可知,地形起伏度是描述区域地形特征的宏观指标^[34],是对滑坡的发生最重要的影响因子,且地形起伏度与区域的地质环境关系密切,江津区南部四面山区系云贵高原过渡到四川盆地的梯形地带,地质结构稳定性较差,易发生滑坡。对滑坡发生第二重要的因子是年平均降雨量。随着年平均降雨量的增加,滑坡密度呈先增大后减小的趋势,当降雨量在 1 186~1 219 mm 时,滑坡最易发生。降雨形成的地表径流将坡面上不稳定的岩土颗粒冲刷带走,导致坡面侵蚀,同时,降雨量影响植被发育,进而影响滑坡的发生^[12]。坡度影响斜坡的稳定性,从模型解释结果看,当坡度 $13^{\circ}\sim 28^{\circ}$ 时,对滑坡影响最大,之后略微下降;当坡度大于 40° 时,斜坡表面不容易积累滑坡物质,且坡度较大的地区人类活动影响较小。高程不仅直接影响滑坡发生,也间接体现人类活动对滑坡的影响。高程小于 831 m 时,对预测结果产生正影响;在低海拔地区,人口集聚,人类活动强烈,易改变当地的地质环境,滑坡发生概率高^[35]。

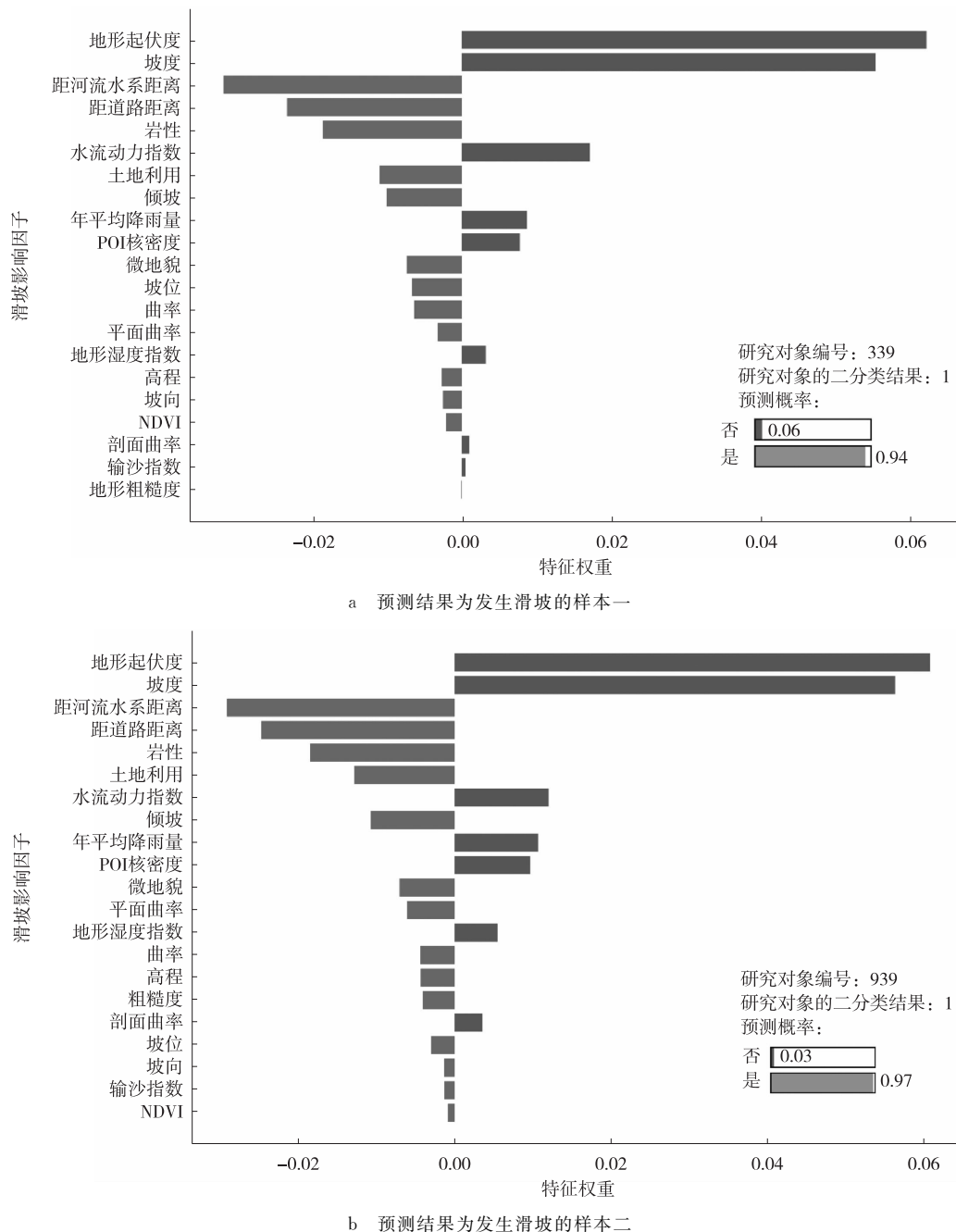


图 8 LIME 对被预测为滑坡的数据的解释

Fig. 8 LIME's interpretation of data predicted to be a landslide

6.3 随机森林的普适性

已有学者用随机森林模型对三峡库区中段中山地区进行滑坡易发性区划,并取得了良好的预测效果。孙德亮等人^[22]利用超参数优化后的随机森林模型对奉节县进行滑坡易发性区划,证明了随机森林模型有较强的预测能力。肖婷等人^[23]用随机森林模型对万州区进行滑坡易发性区划,训练集和测试集 AUC 值分别为 0.975 和 0.794,极高易发区的滑坡频率比为 7.304,表明模型可靠度高。而本文以三峡库区西南段江津区为研究对象,区域多为低山、丘陵地貌,地形起伏较小,结果表明随机森林在低山缓坡区的滑坡易发性区划中仍能达到良好的预测精度,全体数据集 AUC 值为 0.962。区域内高-极高滑坡易发区包含了 71.3% 的滑坡点,滑坡贡献率随易发性等级的增高而增高,可见随机森林模型在丘陵缓坡地区也具有很好的适应性与预测能力。

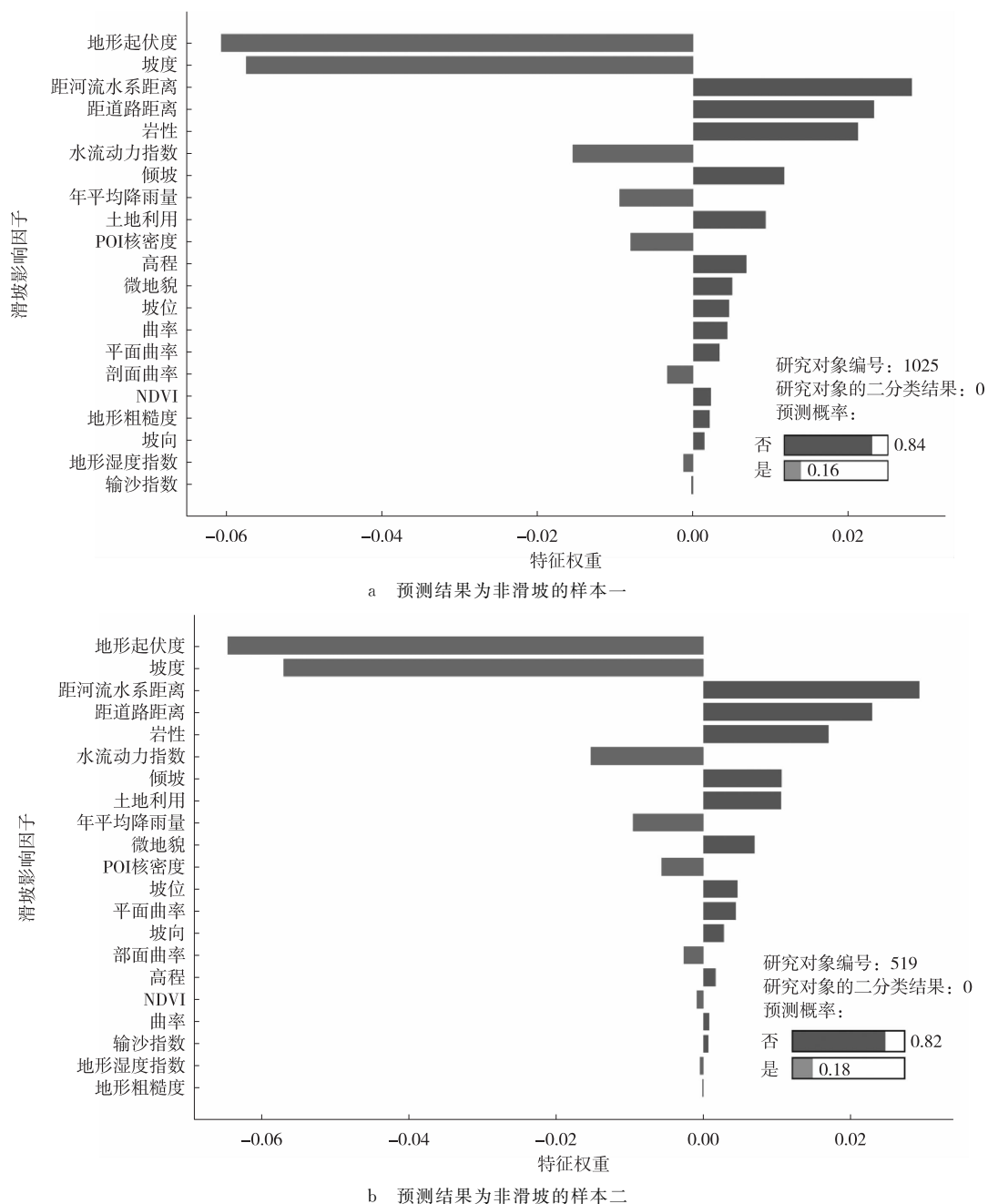


图9 LIME对被预测为非滑坡的数据的解释

Fig.9 LIME's interpretation of data predicted to be a non-landslide

7 结论

本文以江津区为对象,选取了高程、倾坡类型、岩性、距河流水系距离等 21 个因子作为滑坡易发性影响因

子,利用随机森林模型,进行丘陵缓坡地区滑坡易发性区划,得到较高的预测精度。主要结论有:1) 随机森林适用于丘陵缓坡地区的滑坡易发性区划,江津区滑坡易发区区划混淆矩阵准确率为 0.968,全体数据集 AUC 值为 0.962。2) 江津区滑坡高-极高易发区面积占比为 20.43%,但滑坡点数占全区历史滑坡的 71.3%。受地形地貌、气象水文、人类活动等影响,高-极高易发区主要分布在长江及支流沿岸、各背斜翼部低山山麓带及人类活动强烈的地区。3) 从全局解释角度出发,地形起伏度、坡度、年平均降雨量这 3 个因素对丘陵缓坡地区滑坡易发性影响最大。4) 地形起伏度、坡度对滑坡的发生产生正影响,年平均降雨量小于 1 300 mm 时,对滑坡的发生产生正影响,大于 1 300 mm 后对滑坡的发生产生负影响。

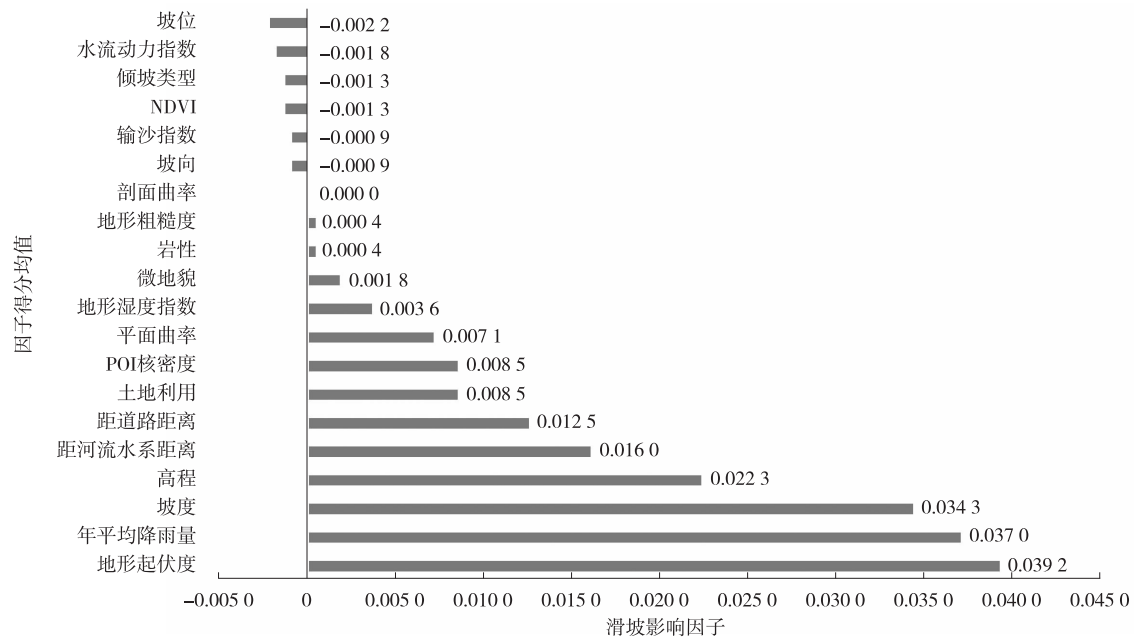


图 10 滑坡的影响因子重要性排序

Fig. 10 Importance ranking of influencing parameters of landslides

参考文献:

- [1] SUN D L, WEN H J, ZHANG Y, et al. An optimal sample selection-based logistic regression model of slope physical resistance against rainfall-induced landslide[J]. *Natural Hazards*, 2020; 1-25.
- [2] 郑光, 许强, 刘秀伟, 等. 2019 年 7 月 23 日贵州水城县鸡场镇滑坡-碎屑流特征与成因机理研究[J]. *工程地质学报*, 2020, 28(3): 541-556.
ZHENG G, XU Q, LIU X W, et al. The Jichang landslide on July 23, 2019 in Shuicheng, Guizhou: characteristics and failure mechanism[J]. *Journal of Engineering Geology*, 2020, 28(3): 541-556.
- [3] FROUDE M J, PETLEY D N. Global fatal landslide occurrence from 2004 to 2016[J]. *Natural Hazards & Earth System Sciences*, 2018; 2161-2181.
- [4] NEAUPANE K M, PIANTANAKULCHAI M. Analytic network process model for landslide hazard zonation[J]. *Engineering Geology*, 2006, 85(3/4): 281-294.
- [5] OTHMAN A N, WAN M N, NORAINI W M, et al. GIS based multi-criteria decision making for landslide hazard zonation[J]. *Procedia-Social and Behavioral Sciences*, 2012, 35(35): 595-602.
- [6] BAI S B, WANG J, LÜ G N, et al. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China[J]. *Geomorphology*, 2010, 115(1/2): 23-31.
- [7] COLKESEN I, SAHIN E K, KAVZOGLU T. Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression[J]. *Journal of African Earth Sciences*, 2016, 118: 53-64.
- [8] MARJANOVI M, KOVAEVI M, BAJAT B, et al. Landslide susceptibility assessment using SVM machine learning algorithm[J]. *Engineering Geology*, 2011, 123(3): 225-234.
- [9] HONG H Y, LIU J Z, ZHU A X, et al. A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China)[J]. *Environmental Earth Sciences*, 2017, 76:

652.

- [10] YILMAZ I. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat—Turkey)[J]. *Computers & Geosciences*, 2009, 35(6): 1125-1138.
- [11] KAWABATA D, BANDIBAS J. Landslide susceptibility mapping using geological data, a DEM from ASTER images and an Artificial Neural Network (ANN)[J]. *Geomorphology*, 2009, 113(1/2): 97-109.
- [12] SUN D L, WEN H J, WANG D Z, et al. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm[J]. *Geomorphology*, 2020, 362: 107201.
- [13] WANG Y, SUN D, WEN H, et al. Comparison of random forest model and frequency ratio model for landslide susceptibility mapping (LSM) in Yunyang county (Chongqing, China) [J]. *International Journal of Environmental Research and Public Health*, 2020, 17(12): 4206.
- [14] SUN D, XU J, WEN H, et al. Assessment of landslide susceptibility mapping based on Bayesian hyperparameter optimization: a comparison between logistic regression and random forest[J]. *Engineering Geology*, 2020, 281: 105972.
- [15] SEVGEN E, KOCAMAN S, NEFESLIOGLU A H, et al. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest[J]. *Sensors*, 2019, 19(18): 3940.
- [16] ADA M, SAN B T. Comparison of machine-learning techniques for landslide susceptibility mapping using two-level random sampling (2LRS) in Alakir catchment area, Antalya, Turkey[J]. *Natural Hazards*, 2018, 90: 237-263.
- [17] SVETNIK V, LIAW A, TONG C. Random forest: a classification and regression tool for compound classification and QSAR modeling[J]. *Journal of Chemical Information & Computer Sciences*, 2003, 43(6): 1947-1958.
- [18] IBRAHI M, LOUIE M, MODARRES C, et al. Global explanations of neural networks: mapping the landscape of prediction[EB/OL]. (2019-02-06)[2022-03-25]. <https://arxiv.org/abs/1902.02384>.
- [19] WANG K, TIAN J, ZHENG C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP[J]. *Computers in Biology and Medicine*, 2021, 137(5): 104813.
- [20] MATEO-SANCHIS A, PILES M, AMORÓS-LÓPEZ J, et al. Learning main drivers of crop progress and failure in Europe with interpretable machine learning[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2020, 104: 102574.
- [21] DAS S, DATTA S, ZUBAIDI H A, et al. Applying interpretable machine learning to classify tree and utility pole related crash injury types[J]. *IATSS Research*, 2021, 45(3): 310-316.
- [22] 孙德亮. 基于机器学习的滑坡易发性区划与降雨诱发滑坡预报预警研究[D]. 上海: 华东师范大学, 2019.
- SUN D L. Mapping landslide susceptibility based on machine learning and forecast warning of landslide induced by rainfall[D]. Shanghai: East China Normal University, 2019.
- [23] 肖婷. 三峡库区万州区及重点库岸段滑坡灾害风险评价[D]. 武汉: 中国地质大学, 2020.
- XIAO T. Landslide susceptibility assessment based on statistical and machine learning method in Wanzhou district[D]. Wuhan: China University of Geosciences, 2020.
- [24] SUN D, XU J, WEN H, et al. An optimized random forest model and its generalization ability in landslide susceptibility mapping: application in two areas of Three Gorges reservoir, China[J]. *Journal of Earth Science*, 2020, 31(6): 1068-1086.
- [25] BREIMAN L. Bagging predictors [J]. *Machine Learn*, 1996, 24(2): 123-140.
- [26] CUTLER A, CUTLER D R, STEVENS J R. Random forests[J]. *Machine Learn*, 2004, 45(1): 157-176.
- [27] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. *Annals of Statistics*, 2001, 29(5): 1189-1232.
- [28] JEREMY P, SHUANG D, WALTER N, et al. Opening the black box: the promise and limitations of explainable machine learning in cardiology, Canadian Journal of Cardiology[J]. *Canadian Journal of Cardiology*, 2021, 38(2): 204-213.
- [29] RIBEIRO M T, SINGH S, GUESTRIN C. “Why should I trust you?”: Explaining the predictions of any classifier[C]// *Proceeding of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 1135-1144.
- [30] 许嘉慧, 孙德亮, 王月, 等. 基于 GIS 与改进层次分析法的奉节县滑坡易发性区划[J]. *重庆师范大学学报(自然科学版)*, 2020, 37(2): 36-44.
- XU J H, SUN D L, WANG Y, et al. Landslide susceptibility mapping of Fengjie county based on GIS and improved analytic hierarchy process[J]. *Journal of Chongqing Normal University(Natural Science)*, 2020, 37(2): 36-44.
- [31] 孙德亮, 马祥龙, 唐小娅, 等. 基于不同因子分级的滑坡易发性区划对比: 以万州区为例[J]. *重庆师范大学学报(自然科学版)*, 2021, 38(5): 43-54.

- SUN D L, MA X L, TANG X Y, et al. Comparison of landslide susceptibility mapping based on different factor classifications: taking Wanzhou district as an example[J]. Journal of Chongqing Normal University (Natural Science), 2021, 38(5): 43-54.
- [32] 刘坚, 李树林, 陈涛. 基于优化随机森林模型的滑坡易发性评价[J]. 武汉大学学报: 信息科学版, 2018, 43(7): 1085-1091.
- LIU J, LI S L, CHEN T. Landslide susceptibility mapping by random forest model[J]. Geomatics and Information Science of Wuhan University, 2018, 43(7): 1085-1091.
- [33] GHIMIRE M. Geo-hydrological hazard and risk zonation of Banganga watershed using GIS and remote sensing[J]. Journal of Nepal Geological Society, 2001, 23: 99-110.
- [34] 吴森, 李虎杰, 张魁, 等. 高程因素对平昌县滑坡灾害发育的敏感性分析[J]. 三峡大学学报(自然科学版), 2015, 37(4): 62-66.
- WU S, LI H J, ZHANG K, et al. Sensitivity analysis of elevation to landslide growth in Pingchang county[J]. Journal of China Three Gorges University (Natural Sciences), 2015, 37(4): 62-66.

Resources, Environment and Ecology in Three Gorges Area

Landslide Susceptibility Mapping in Hilly and Gentle Slope Region Based on Interpretable Machine Learning

ZHANG Hong^{1,2}, GU Qingyu^{1,2}, SUN Chengbin³, SUN Deliang^{1,2}, MI Changlin⁴, ZHANG Fengtai⁵

(1. School of Geography and Tourism, Chongqing Normal University, Chongqing Normal University, Chongqing 401331;

2. Key Laboratory of GIS Application in Chongqing University, Chongqing 401331;

3. Linyi Highway Development Center Junan County Center, Linyi Shandong 276000;

4. Linyi Natural Resources Development Service Center, Linyi, Shandong 276000;

5. School of Management, Chongqing University of Technology, Chongqing 400054, China)

Abstract: [Purposes] An interpretable model based on random forest and factor importance, PDP and LIME was proposed to give global and local explanations for landslide susceptibility zoning, in order to provide theoretical basis for landslide disaster prevention and management and urban planning in Jiangjin district. [Methods] A 30 m × 30 m grid spatial database is established based on 899 historical landslide sites in Jiangjin district with 21 factors including topography, geomorphology, geological structure, meteorological and hydrological conditions, environmental conditions and human activities, and selected 8990 non-landslide points, landslide and non-landslide sample points into the random forest model for training and building landslide susceptibility model. The results are divided into five grades: very low, low, medium, high and very high. Finally, the universality of the random forest model in the Three Gorges reservoir area is discussed. Finally, the interpretability of random forest model was studied by Permutation Importance, PDP and LIME interpretation method. [Findings] The number of landslide points in high-extremely high landslide-prone area accounts for 71.3% of the total historical landslide points, the area accounts for 20.42% of the total area of the area, the accuracy of confusion matrix is 0.968, and the AUC of the whole data set is 0.962. According to the explanatory nature of the model, it can be seen that the fluctuation, annual average rainfall and slope are the most important factors, among which, the fluctuation and slope have a positive impact on the occurrence of landslides, and when the annual average rainfall is less than 1300 mm, the occurrence of landslides has a positive impact. [Conclusions] The landslide susceptibility zoning model based on interpretable machine learning has high prediction accuracy and has important practical significance for landslide prevention and control.

Keywords: random forest; machine learning interpretable; Jiangjin district landslide; landslide susceptibility mapping; Three Gorges reservoir region

(责任编辑 许 甲)