



引文格式: 李陈, 张志然, 张少华, 等. 基于机器学习算法 TabPFN 的滑坡易发性评价及可解释性分析[J]. 西北地质, 2026, 59(4): 144–154. DOI: 10.12401/j.nwg.2026007

Citation: LI Chen, ZHANG Zhiran, ZHANG Shaohua, et al. Landslide Susceptibility Assessment with TabPFN under Limited Sample Scenarios and Its Interpretability Analysis[J]. Northwestern Geology, 2026, 59(4): 144–154. DOI: 10.12401/j.nwg.2026007

基于机器学习算法 TabPFN 的滑坡易发性评价及可解释性分析

李陈^{1,2}, 张志然^{1,2,3,*}, 张少华^{1,2}, 庞正博^{1,2}, 朱良玉⁴, 赵益祯^{1,2,3}

(1. 西安石油大学地球科学与工程学院, 陕西 西安 710065; 2. 西安石油大学陕西省油气成藏地质学重点实验室, 陕西 西安 710065; 3. 油气资源智能地球物理探测陕西省高等学校重点实验室, 陕西 西安 710065; 4. 中国地震局第二监测中心, 陕西 西安 710054)

摘要: 滑坡易发性评价对于山区灾害防控具有重要意义, 但传统模型在小样本数据上的泛化能力往往受限, 且线性模型难以捕捉复杂的因子交互作用。TabPFN 机器学习模型具有免调参和小样本强泛化优势, 为破解上述建模难题提供了思路。本研究以黄山市为研究区, 综合选取地形地貌、地质构造、水文条件和人类活动等 4 类共 14 项评价因子构建滑坡易发性指标体系, 基于 736 个历史滑坡点, 在距滑坡点 1 km 外的区域选取等量非滑坡样本, 构建正负样本 1:1 的平衡数据集。在此基础上, 引入 TabPFN 模型开展滑坡易发性建模, 并结合 SHAP 方法对模型结果进行可解释性分析。结果表明, 与基线模型相比, TabPFN 在准确率(0.903)、精确率(0.965)、召回率(0.841)、F1 值(0.899)及 AUC(0.964)等指标上综合表现最优, 表现出更强的泛化能力与对复杂关系建模优势; SHAP 解释结果进一步揭示了距道路距离、NDVI 与岩性是研究区滑坡空间分布的主要驱动因子, 其中, 靠近道路区域滑坡易发性显著升高, 反映出人类工程活动对边坡稳定性的扰动作用突出, 较高的植被覆盖度对滑坡发生具有抑制作用, 不同岩性条件下滑坡敏感性差异显著, 体现了地质背景的基础控制作用。本研究结合 TabPFN 机器学习模型与 SHAP 方法, 实现了滑坡易发性的高精度预测与可解释性分析, 弥补了传统“黑箱”模型难以揭示成灾机理的不足, 为山地城市滑坡风险评估提供了兼具准确性与解释力的技术途径。

关键词: 地质灾害; 滑坡易发性; TabPFN; SHAP; 黄山市

中图分类号: P694; P280

文献标志码: A

文章编号: 1009-6248(2026)04-0144-11

Landslide Susceptibility Assessment with TabPFN under Limited Sample Scenarios and Its Interpretability Analysis

LI Chen^{1,2}, ZHANG Zhiran^{1,2,3,*}, ZHANG Shaohua^{1,2}, PANG Zhengbo^{1,2}, ZHU Liangyu⁴, ZHAO Yizhen^{1,2,3}

收稿日期: 2025-12-02; 修回日期: 2026-04-13; 责任编辑: 曹佰迪

基金项目: 国家自然科学基金项目“模型-数据联合驱动的兴趣区可解释智能推荐方法”(42371478), 陕西省自然科学基金基础研究计划项目“多源数据驱动的西安都市圈精细尺度活力评估及时空演变研究”(2025JC-YBQN-438)和西安市科技计划项目软科学一般项目“黄河流域生态保护与多情景土地利用模拟研究”(2025JH-KXYB-0213)联合资助。

作者简介: 李陈(2000-), 男, 硕士研究生, 主要研究方向为地质灾害风险评估。E-mail: lcliche@163.com。

*** 通讯作者:** 张志然(1991-), 女, 博士, 讲师, 主要研究方向为地学数据分析与挖掘。E-mail: zrzhang@xsyu.edu.cn。

© Editorial Office of *Northwestern Geology*. OA under CC BY-NC-ND

- (1. School of Earth Science and Engineering, Xi'an Petroleum University, Xi'an 710065, Shaanxi, China; 2. Shaanxi Key Laboratory of Petroleum Accumulation Geology, Xi'an Petroleum University, Xi'an 710065, Shaanxi, China; 3. Key Laboratory of Intelligent Geophysical Exploration for Oil and Gas Resources of Shaanxi Higher Education Institutions, Xi'an 710065, Shaanxi, China; 4. The Second Monitoring and Application Center, Xi'an 710054, Shaanxi, China)

Abstract: Landslide susceptibility assessment is crucial for the disaster mitigation in mountainous regions; however, traditional models often have limited generalization on small-sample datasets, and linear approaches struggle to capture the complex interactions among influencing factors. This study introduces the TabPFN machine learning model, which, with its hyperparameter-free nature and strong generalization advantage under small-sample conditions, provides a novel approach for overcoming the modeling bottlenecks. Taking Huangshan City as the study area, based on 736 landslide points, an equal number of non-landslide samples were selected from areas beyond 1 km of these points to construct a balanced dataset with a 1:1 ratio of positive to negative samples, this research establishes a landslide susceptibility indicator system by systematically selecting 14 evaluation factors across four categories: topography, geology, hydrology, and human activity. Building upon this foundation, The study introduces the TabPFN model for landslide susceptibility evaluation, and employs SHAP to analyze the underlying driving mechanisms. The results show that, compared with baseline models, TabPFN achieves the best overall performance in terms of accuracy (0.903), precision (0.965), recall (0.841), F1-score (0.899), and AUC (0.964), demonstrating stronger generalization capability and advantages in modeling complex relationships. Furthermore, SHAP analysis reveals that distance to roads, NDVI, and lithology are the primary driving factors influencing the spatial distribution of landslides, reflecting the fundamental pattern wherein human disturbance, vegetation cover, and geological conditions collectively govern landslide development. Specifically, landslide susceptibility increases significantly in areas proximal to roads, reflecting the prominent disturbance of human engineering activities on slope stability. Higher vegetation coverage exerts a notable inhibitory effect on landslide occurrence, while landslide sensitivity varies markedly across different lithological conditions, underscoring the fundamental controlling role of the geological setting. This study integrates the TabPFN machine learning model with SHAP-based interpretability to achieve both high-accuracy prediction and interpretability analysis of landslide susceptibility. It overcomes the shortcomings of traditional "black box" models in revealing the mechanisms of disaster formation, providing a technical pathway for landslide risk assessment in mountainous cities that ensures both precision and explanatory insight.

Keywords: geological hazards; landslide susceptibility; TabPFN; SHAP; Huangshan City

滑坡是全球范围内分布广泛、危害严重的地质灾害之一,造成大量人员伤亡和严重经济损失(Gutiérrez et al., 2015; 王朋伟等, 2023)。中国山地丘陵广布,部分地区滑坡灾害频发,且发生频率与危害程度均呈上升趋势(孙萍萍等, 2022)。滑坡不仅侵蚀土地资源,破坏居民建筑和工业设施安全(黄煜等, 2023),更对人民生命财产安全构成严重威胁,对区域经济和社会可持续发展产生不利影响(田媛等, 2022)。黄山市地处典型山地丘陵区,地质构造复杂、降雨充沛,滑坡事件频发,开展滑坡易发性评价对该区域的灾害防治和国土空间规划具有重要理论意义和现实需求。

准确预测滑坡可能发生的位置及其概率,对于降

低地质灾害风险与潜在损失具有重要意义(Pradhan, 2010)。随着遥感、地理信息系统和计算机技术的快速发展,多源时空数据的获取能力显著提升,推动了滑坡易发性评价由定性向定量的转变(Shen et al., 2019)。早期研究主要依赖逻辑回归、信息量法等统计模型(Reichenbach, 2018),该类方法便于计算,但通常基于因子独立性或线性假设,难以捕捉滑坡环境中复杂的非线性交互作用(Merghadi, 2020; 窦杰, 2023)。近年来,随机森林(RF)(刘坚等, 2018; 林荣福等, 2020; 马啸等, 2022)和支持向量机(SVM)(Zhu et al., 2022)等机器学习方法被广泛用于滑坡易发性评价中,显著提升了模型的分类精度。与此同时,卷积神经网络

(CNN)及注意力机制等深度学习模型,凭借强大的深层非线性特征提取能力,逐渐成为该领域的重要研究方向(Wang et al., 2019)。然而,传统深度学习网络结构复杂、参数量大,对大规模标注样本具有较强依赖性,在滑坡易发性评价中常面临样本规模有限和多源异构数据的情景时,易陷入过拟合。Hollmann et al.(2022)提出了一种基于 Transformer架构的表格先验数据拟合网络(TabPFN),该方法借助大规模预训练任务优化小样本建模能力,在捕捉复杂非线性关系方面展现出潜在优势(McElfresh et al.2023),为样本有限区域的滑坡易发性评价提供了新的思路。

尽管滑坡易发性评价方法不断发展,但在实际应用中仍面临若干挑战。首先,许多地区滑坡样本数量有限,且滑坡与非滑坡样本分布不均,导致部分统计模型与机器学习模型在小样本或样本不平衡条件下易出现过拟合或分类偏差,影响结果的稳定性与泛化能力(周晓亭等, 2022)。其次,滑坡孕育过程受地形、地质、气候和人类活动等多源因子共同影响,因子间存在复杂的非线性关系与空间依赖性,传统线性模型难以充分捕捉此类复杂结构,而高维变量输入又容易引入冗余信息,增加模型的不确定性(黄发明等, 2023)。此外,以随机森林、支持向量机和深度学习为代表的模型虽在预测精度方面有所提升,但其可解释性较弱,不利于识别关键控制因子及理解滑坡成灾机理(Fang et al, 2023; Zhang et al, 2023)。

基于上述背景,笔者以黄山市为研究区,整合多源数据构建滑坡数据库,提取包括高程、坡度在内的 14 项环境因子作为初始评价因子集,并借助皮尔逊相关性与信息增益率对因子进行筛选,以降低冗余、提升模型解释能力与计算效率。在模型构建方面,引入 TabPFN 方法,并选取随机森林与逻辑回归作为对比模型,系统评估不同方法在滑坡易发性评价中的性能差异与适用性。进一步利用 SHAP(Shapley Additive Explanations)方法开展模型可解释性分析,量化各因子对预测的贡献度,揭示黄山市滑坡易发性的主控因子及其作用机制。本文旨在构建一套精度高、泛化能力强且具备良好可解释性的滑坡易发性评价体系,为区域滑坡风险评估与关键因子识别提供理论支持与实证依据。

1 研究区概况

黄山市位于安徽省南部,地处 E 117°02'~118°55', N 29°24'~30°24', 总面积约为 9 678 km²;全市常住人口为 133.1 万人(第七次全国人口普查数据),主要分布于黟县盆地以及新安江河谷盆地等河谷平原区(图 1)。黄山市地处皖南山区核心地带,地貌类型多样,以山地(占总面积 51%)和丘陵(占 36.1%)为主。境内地表起伏显著,黄山山脉呈 EN-WS 走向横贯全市,构成了新安江与长江两大流域的分水岭,并发育有河流 600

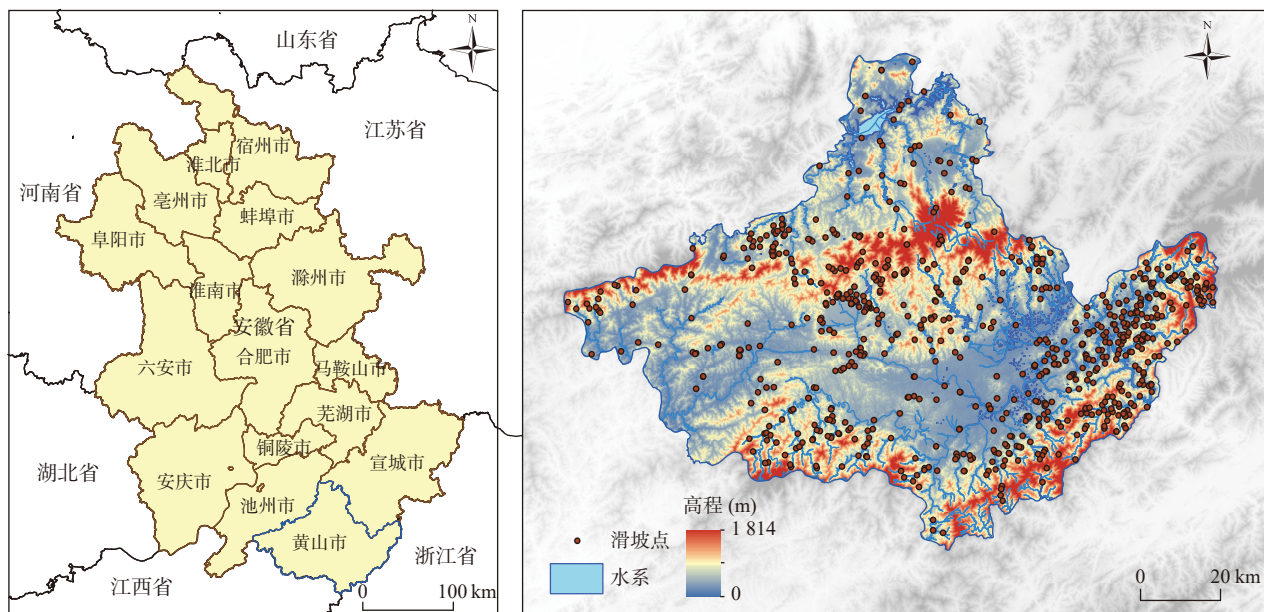


图1 研究区概况图

Fig. 1 Overview map of the study area

余条。气候方面,黄山市属中亚热带北缘季风湿润气候,四季分明,降水充沛。多年平均降雨量 1 816 mm,主要集中于 5~8 月,丰富的大气降水和密集的河网共同作用,使得该区地表水资源总量达 103.18 亿 m^3 ,人均水资源占有量约为全国平均水平的 4 倍。

黄山市地处江南造山带东段与滨西太平洋构造域的复合部位。出露的岩层序列完整,包括新元古代浅变质火山碎屑岩(Pt3)、南华纪—中生代三叠纪海相沉积盖层(Nh-T)和侏罗纪—白垩纪陆相沉积盖层(J-K)。这一完整的层序为变质岩、碳酸盐岩和碎屑岩等多种异滑岩组及地貌景观的发育提供了物质基础(翁望飞等, 2020)。黄山市多山多丘的地形、集中而丰沛的降雨、复杂的岩土体结构,加之频繁的人类工程活动(如道路修建、城镇建设),共同构成了滑坡灾害发育的有利条件,导致区内滑坡灾害频发,已造成显著的经济损失与人员伤亡。因此,对该区域开展滑坡易发性评价研究具有重要的现实意义。

2 数据来源与评价方法

2.1 数据来源与预处理

本研究所使用的数据主要包括滑坡点数据、数字高程模型(DEM)、水系数据、道路数据、地质分布情况数据、归一化植被指数(NDVI)、土地利用和降雨数据。其中,滑坡点数据来源于地理遥感生态网平台(<https://www.gisrs.cn/>),包含灾害体名称、地理位置等信息,滑坡发生时间为 1980~2023 年,已剔除位置异常点;DEM 采用 2015 年欧洲航天局发布的哥白尼数字高程模型,分辨率 30 m;水系和道路网数据来源于 2024 年 OpenStreetMap(<https://www.openstreetmap.org/>);地质数据来源于全国地质资料馆 2017 版 1:100 万地质图(<http://www.cgs.gov.cn/>),包含断层、岩性等信息;NDVI 数据利用谷歌地球引擎(GEE)提取 2022~2025 年遥感影像的平均值;土地利用数据来自 Zenodo 平台(<http://doi.org/10.5281/zenodo.4417809>)提供的 2024 年土地覆盖数据集,包括耕地、林地、灌木、草地、水体和不透水面;降雨量来自 Zenodo 平台(<https://doi.org/10.5281/zenodo.3114194>),采用 2000~2023 年的平均降雨量。

为保证数据在空间基准上的一致性,对数据进行了预处理。对于矢量数据,将其统一转换至 CGCS 2000 地理坐标系,并从地质图中提取并重分类了岩性

信息。对于栅格数据,采用双线性插值法将其重采样至统一的空间分辨率 30 m,以消除因尺度差异引起的潜在误差。通过上述处理,构建了一个空间范围、坐标系及分辨率完全一致的多因子数据集,为后续滑坡易发性评价与可解释性分析提供了可靠的数据基础。

2.2 评价体系构建

滑坡易发性评价因子的筛选与构建是易发性评价研究的重要基础,针对性地筛选代表性因子是确保模型准确性的关键前提。本研究以黄山市 736 个历史滑坡点为样本基础,在系统分析区域滑坡分布规律、地形地貌特征、外部诱发条件及相关文献(Reichenbach et al., 2018),遵循代表性、可获取性及独立性原则,选出 4 大类 14 个评价因子(图 2)。这些因子涵盖了地形地貌、水文条件、人类活动与地质构造四大类影响机制:

(1)地形地貌因子:包括高程、坡度、坡向、曲率、地形起伏度及地形位置指数(TPI),用于表征边坡稳定性的基本地形条件。其中,坡度反映坡面倾斜程度;坡向指示坡面朝向;曲率刻画地表凹凸特征;地形起伏度表征区域高程差异;TPI 通过比较栅格单元与邻域平均高程描述相对地形位置。所有地形因子均基于 DEM 数据计算获得。

(2)水文条件因子:包括距河流距离、地形湿度指数(TWI)、降雨、NDVI,共同反映水文作用与植被固坡效应。TWI 基于坡度与汇流累积量计算,表征地表汇流能力与土壤含水状况;距河流距离通过欧氏距离计算,表征河流侵蚀与浸润影响;NDVI 采用 2022~2025 年多期遥感影像的年均值,反映植被覆盖状况及根系加固作用。

(3)人类活动因子:包括距道路距离与土地利用类型,表征人类工程活动对自然斜坡的扰动程度。距道路距离通过欧氏距离计算,反映道路开挖与工程活动强度;土地利用类型基于土地利用数据栅格化并重分类获得。

(4)地质构造因子:包括距断层距离与岩性,作为控制岩土体强度和结构面发育的内在条件。距断层距离通过欧氏距离计算,表征构造活动与断裂带影响;工程地质岩组由地质图中岩性单元重分类得到,综合反映岩土体强度、风化特征及结构面发育状况。

上述因子在空间上系统覆盖了黄山市滑坡发育的主要控制条件,为后续易发性精确建模与成灾机制解析提供了可靠的数据支撑。

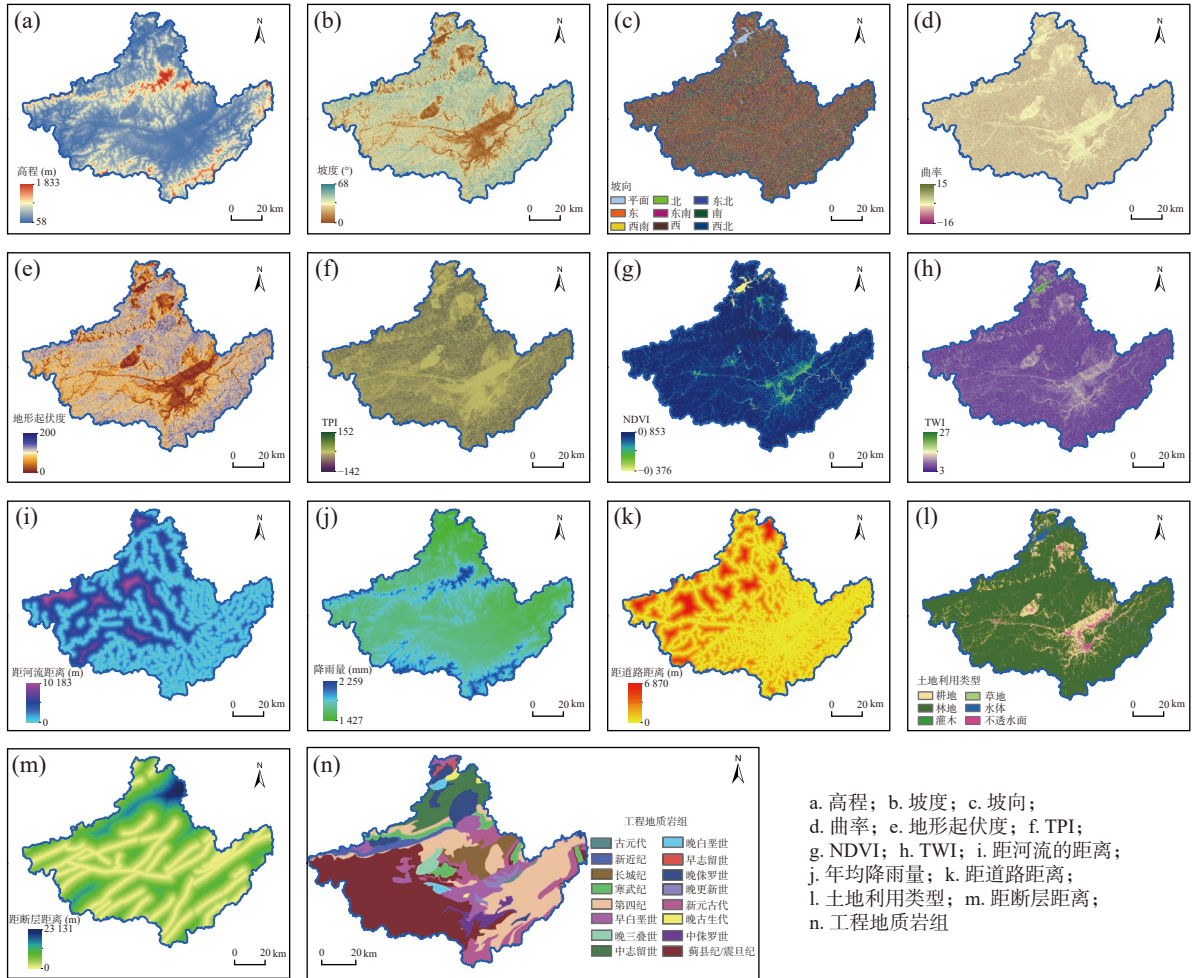


图2 评价因子分布图

Fig. 2 Spatial distribution of evaluation factors

2.3 研究方法

本研究基于 TabPFN 模型与 SHAP 可解释性分析方法开展滑坡易发性评价。基于构建的评估指标体系,通过皮尔逊相关性分析和信息增益率去除剔除冗余因子,保留高价值因子;引入 TabPFN 模型开展滑坡易发性建模,并与逻辑回归(LR)、RF、SVM 及 CNN 进行性能对比;利用 SHAP 模型进行模型结果的可解释性分析。

2.3.1 因子选择

构建高精度的滑坡易发性评价模型,所选因子既要能有效表征致灾机理,又需避免因子间的高度相关性所导致的模型偏差。为此,本研究采用皮尔逊相关性分析与信息增益率对初选因子进行筛选与优化,皮尔逊相关性分析用于剔除冗余因子,信息增益率用于筛选高价值因子,共同确定最终输入模型的评价因子组合。

在滑坡易发性建模中,并非所有因子都具有同等

贡献,部分因子可能包含噪声(付泉等, 2024)。信息增益率可以定量评估单个因子的预测能力。信息增益率值越高,表明该因子对滑坡易发性的预测能力越强(Chen et al., 2017);若其值等于或接近 0,则意味着该因子缺乏有效信息,应予剔除。

给定训练数据集 D , 其信息熵 $Entropy(D)$ 的计算公式为:

$$Entropy(D) = - \sum_{k=1}^y p_k \log_2(p_k) \quad (1)$$

式中: y 为类别数, 本研究为滑坡与非滑坡; p_k 表示数据集中第 k 类样本所占比例。

针对某一特定因子 s , s 的取值将数据集 D 划分为 m 个子集 $D_v, v = 1, 2, \dots, m$, 计算该因子带来的信息增益 $Gain(D, s)$:

$$Gain(D, s) = Entropy(D) - \sum_{v=1}^m \frac{|D_v|}{D} Entropy(D_v) \quad (2)$$

为消除因子取值数目多少带来的内在偏差, 计算分裂信息 $IV(s)$:

$$IV(s) = - \sum_{v=1}^m \frac{|D_m|}{|D|} \log_2 \frac{|D_m|}{|D|} \quad (3)$$

最终, 因子的信息增益率 $IGR(s)$ 为:

$$IGR(s) = - \frac{Gain(D, s)}{IV(s)} \quad (4)$$

2.3.2 TabPFN 模型

TabPFN 是一种基于 Transformer 架构的表格数据分类模型, 其核心思想是将监督学习问题转化为近似贝叶斯推断过程(Hollmann et al., 2022)。TabPFN 通过在大规模合成数据集上进行预训练, 结合元学习与贝叶斯推断思想, 能够在极少训练样本下完成推理, 并直接输出概率形式的预测结果, 适用于样本稀缺的滑坡易发性评价任务。

TabPFN 的建模过程包括预训练与推断两个阶段。在预训练阶段, 由先验分布生成大量合成数据, 模型在这些任务上进行训练, 学习如何从有限样本中提取特征, 并近似贝叶斯后验分布。通过这一过程, 模型捕捉数据中的结构信息和变量关系。在推断阶段, 模型将训练样本与待预测样本作为输入序列, 通过 Transformer 的自注意力机制建模样本间的关系, 通过一次前向传播输出每个评估单元的滑坡发生概率。在具体实施时, 本研究调用 TabPFN 官方开源库构建模型。由于 TabPFN 将超参数选择与训练过程内化于预训练阶段, 无需额外超参数调优, 因此采用其默认参数配置(Vaswani et al, 2017)。

2.3.3 SHAP 算法

SHAP 方法基于博弈论中的 Shapley 值, 能够公平量化每个特征对模型预测结果的贡献程度(Lundberg et al., 2017), 其兼具全局与局部解释能力, 既可从数据集层面识别关键特征, 又能针对单个预测展示各特征如何通过交互作用影响输出结果, 有效增强“黑箱”模型的决策透明度。由于滑坡灾害是地形、地质、水文及人类活动等因子复杂耦合的结果, 而随机森林、神经网络等机器学习模型虽能捕获这些复杂关系并获得较高预测精度, 但其不具有可解释性, 在滑坡易发性分析中引入该方法(Huang et al., 2025)具有重要价值。SHAP 通过将复杂预测拆解为可理解的特征贡献, 不仅能验证模型可靠性、发掘潜在致灾规律, 更能揭示传统地学分析中难以识别的因子耦合机制。

对于特征集合 F 中的第 i 个特征, 其 SHAP 值的计

算公式为:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \quad (5)$$

式中: S 为不包含特征 i 的特征子集; $f(S)$ 表示仅使用子集 S 中的特征进行模型预测的输出值; $f(S \cup \{i\})$ 表示在子集 S 中加入特征 i 后的模型预测值; 组合系数 $\frac{|S|!(|F|-|S|-1)!}{|F|!}$ 确保所有特征子集的贡献被公平分配。

3 结果与讨论

3.1 数据集划分与因子分析

为降低多重共线性对模型稳定性的干扰, 首先对初选的 14 个因子进行皮尔逊相关性分析(图 3)。结果显示, 高程与降雨量呈高度正相关(相关系数 0.96), 地形起伏度与坡度亦高度相关(相关系数 0.95), 均超过一般认为的高度相关阈值 0.7(Gu et al., 2023), 为避免冗余特征对模型训练产生不利影响, 须对四个因子进行取舍。

进一步采用信息增益率对因子重要性进行排序(图 4), 评估因子的预测能力, 结果显示降雨量的信息增益率优于高程, 地形起伏度的信息增益率优于坡度。同时, 高度集中的强降雨是诱发滑坡的直接外部水动力条件, 地形起伏度能在宏观尺度上表征区域地形的切割深度与相对高差。因此, 本研究从初选的 14 项因子中剔除了高程和坡度两个因子, 保留对预测更有贡献的地形起伏度和降雨量。

3.2 模型性能对比分析

本研究基于 736 个已知滑坡点构建正样本。为了确保样本的代表性和平衡性, 以已知滑坡点为中心建立 1 000 m 的缓冲区(Fu et al., 2025), 并在该缓冲区之外随机生成等量的非滑坡点。由此, 构建了正负样本比例为 1:1 的平衡数据集, 以供后续模型训练与评估使用, 该策略有效缓解了样本偏差, 为模型训练提供了可靠的数据基础。同时, 为了确保模型的可靠性和泛化能力, 将数据集按照 7:3 的比例划分为训练集和测试集, 训练集用于模型的训练和参数优化, 测试集用于评估模型的预测性能。

为客观评估模型性能, 将 TabPFN 模型与 LR (Budimir et al., 2015)、RF(Pham et al., 2018)、SVM (Huang et al., 2025) 和 CNN(Wang et al., 2019) 等基线模型进行对比分析。在评估参数上, 采用准确率、精确率、召回率、F1 分数和 AUC 值指标, 在测试集上对

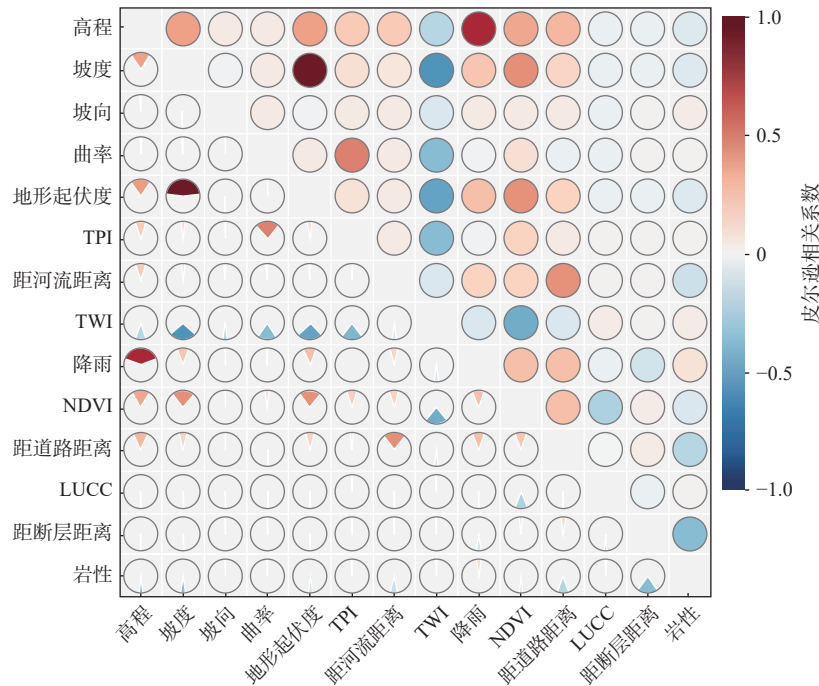


图3 评价因子相关性分析

Fig. 3 Correlation analysis of evaluation factors

5 种模型进行对比。准确率表示模型预测正确的样本占总样本的比例;精确率表示预测为滑坡的样本中实际为滑坡的比例;召回率表示实际为滑坡的样本中被预测为滑坡的比例;F1 分数是准确率和精确率的调和平均值,用于衡量模型的整体性能;AUC 用于衡量模型区分滑坡和非滑坡能力的大小。

与多种经典机器学习和深度学习方法相比,TabPFN 在滑坡易发性评估中展现出明显优势。TabPFN 在综合评价指标上均优于 LR、RF、SVM 和 CNN 模型(图 5)。TabPFN 准确率为 0.903, AUC 值为 0.964, 准确率分别比 LR、RF、SVM 和 CNN 提高 9.2%、1.7%、4.8% 和 2.7%, 精确率分别提升 12.57%、3.29%、4.67% 和 7.76%, 表明其在预测滑坡样本时具有更低的误报率,

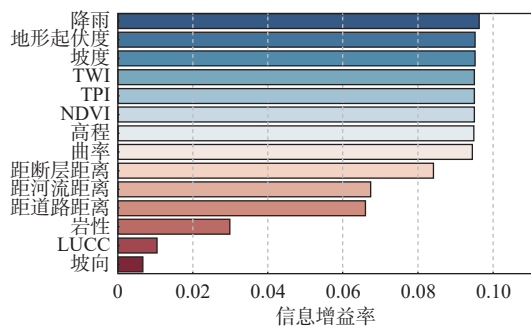


图4 评价因子信息增益率

Fig. 4 Information gain ratio of evaluation factors

有助于提升灾害防控的精准性。此外, TabPFN 在 F1 分数较 LR、RF、SVM 和 CNN 提高 11.22%、1.63%、5.08%、1.88%, 在 AUC 上分别提高了 9.6%、1.8%、4.7% 和 2.8%, 进一步说明其在不同类别样本间的区分能力更强。

为检验模型性能差异的统计学显著性,本研究将所有模型独立运行 20 次,采用 DeLong 检验(DeLong et al., 1988)比较各模型在 AUC 指标上的差异,利用 McNemar 检验(McNemar, 1947)评估分类结果的一致性差异。DeLong 与 McNemar 检验结果显示, TabPFN 与 LR、SVM 及 CNN 之间的差异均达到显著水平($p < 0.001$)。与 RF 模型相比, TabPFN 的性能提升并非在所有重复实验中都达到统计显著。综合来看, TabPFN 在该滑坡数据集上整体表现出最优的预测性

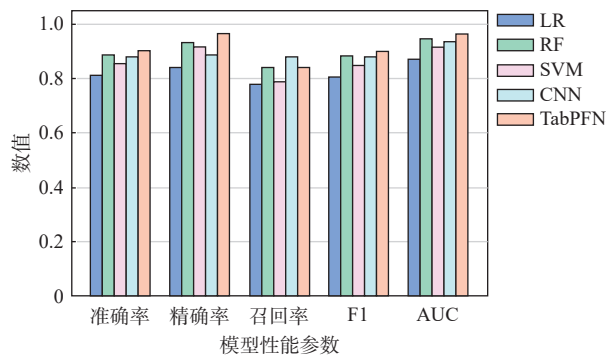


图5 模型性能对比

Fig. 5 Comparison of model performance

能,体现了其在处理小样本数据方面的优势,RF 可作为实际应用中的备选模型。

从建模机制来看,随机森林、逻辑回归和支持向量机等机器学习方法在有限样本上通常依赖反复训练与参数调优,易受样本量少和类别不平衡影响,导致过拟合或欠拟合。卷积神经网络作为典型的深度学习模型,具备较强的特征提取能力,但高度依赖大规模训练数据,在有限样本场景下易发生过拟合,且可解释性相对较弱。TabPFN 基于元学习与贝叶斯推断框架,借助在大规模合成任务上预训练的 Transformer 结构,使模型具备从少量样本中学习复杂非线性关系的能力,从而在本研究场景中展现出更稳健的预测性能与概率输出能力。

3.3 滑坡易发性评价结果分析

基于上述模型生成了研究区的滑坡易发性区划图(图6),并采用自然断点法将研究区划分为极低易发区、低易发区、中易发区、高易发区和极高易发区五个级别。从空间分布来看,黄山市滑坡易发性呈现沿河谷与山地带状分布的特征。极高与高易发区主要分布于主要河流及其支流的河谷沟谷地带,并延伸至中东部和南部地形起伏较大的丘陵山区。这些区域在地形因子方面表现出明显差异,如高程、坡度、地形起伏度及 TPI 值均与周边环境形成较大对比。同时,该区域水系发育密集, TWI 值较高,反映地表汇流条件良好,土体含水量较大;距河流与道路的距离

较近,说明受河流切割与人类工程活动影响。此外,极高与高易发区多位于植被覆盖较低或 NDVI 低值带,并与岩性破碎、风化强烈、工程地质条件较差的岩组分布高度吻合,揭示不良地质条件与强烈外营力共同控制滑坡的发育背景。

相比之下,低与极低易发区主要分布于市域西北部及部分盆地等地形平缓区域。这些地区坡度较缓,曲率与 TPI 变化较小,水系稀疏或位于分水岭区域, TWI 及降雨汇集效应较弱,距离主要道路与河流较远。工程地质岩组整体性较好,抗风化能力强,区域稳定性高。实地滑坡点主要分布于中、高和极高易发区,而在低与极低易发区鲜有分布,验证了易发性分区与主控因子空间分布的一致性与地质合理性。

各模型在易发等级的面积分配上存在明显差异(表1)。LR 模型各等级面积均在 18%~21%,占比较为均衡,整体分区结果平滑保守;RF、SVM、CNN 和 TabPFN 模型呈现“两极分化”特征,极低与极高易发区面积占比较大,中、低与高易发区占比小。其中, CNN 的极低与极高易发区面积占比分别为 32.87% 与 33.77%, TabPFN 模型的极低与极高易发区面积占比分别达 32.54% 与 31.41%,这两类模型在识别稳定区与高风险区时更为明确,空间区分能力更强。

在空间细节的刻画上,各模型表现出显著差异。RF、LR 与 SVM 模型能识别研究区内的主要高风险区域,但局部预测精度与边界细节的过渡上略显粗糙。

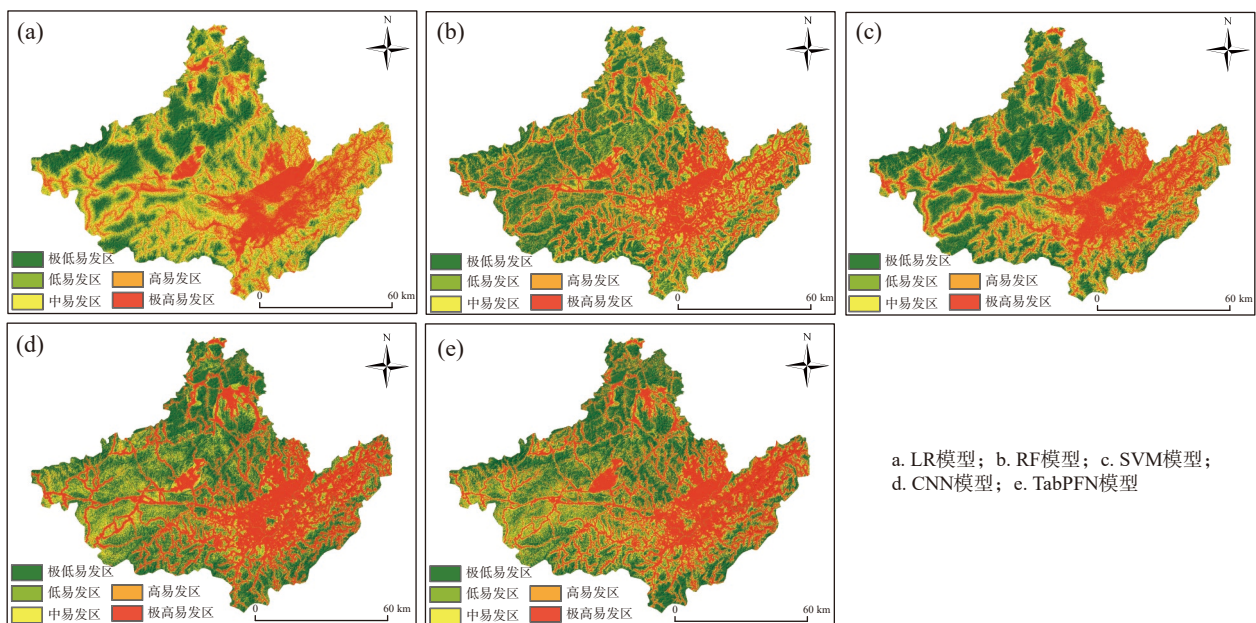


图6 黄山市滑坡易发性分区图

Fig. 6 Landslide susceptibility zoning maps

表 1 五类易发区的面积占比(%)

Tab. 1 Area proportions of the five susceptibility zones(%)

模型	极低易发区占比	低易发区占比	中易发区占比	高易发区占比	极高易发区占比
LR	18.09	19.78	20.66	20.39	21.08
RF	30.66	17.53	13.37	13.56	24.88
SVM	30.18	15.31	12.45	14.11	27.95
CNN	32.87	16.65	9.62	7.09	33.77
TabPFN	32.54	16.05	10.93	9.07	31.41

TabPFN 易发性图在空间细节上表现最优, 不仅准确刻画了断裂带周边的高风险区, 还识别出传统模型未能反映的潜在滑坡区域, 说明 TabPFN 能更有效捕捉致灾因子与滑坡发生之间的复杂非线性关系, 具备更强的地质环境适应性与预测识别能力。

3.4 基于 SHAP 模型的可解释性分析

为揭示 TabPFN 模型的决策机制并量化各因子的影响, 采用 SHAP 方法进行可解释性分析。蜂群图直观展示了各特征对单个样本预测的影响方向与强度(图 7), SHAP 特征重要性排序图进一步量化各因子的贡献(图 8)。每个点代表一个特征值的 SHAP 值(图 7)显示, 点在水平轴上的位置表示该特征值对模型预测结果的影响程度, 正值表示增加滑坡易发性的概率, 负值表示降低滑坡易发性的概率; 点越靠右, 说明该特征对模型输出有较大的正向影响; 点越靠左, 则说明其负向影响更大。点的颜色表示特征值的相对大小, 从低值(蓝色)到高值(红色)变化。

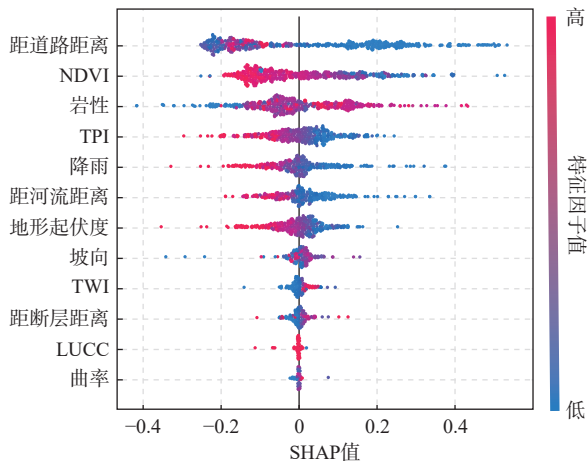


图7 SHAP 蜂群图

Fig. 7 SHAP beeswarm plot of feature importance

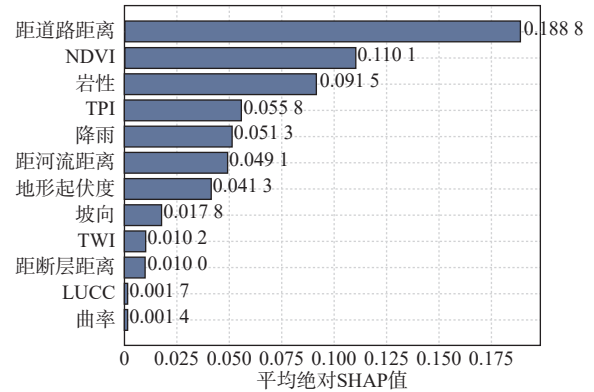


图8 SHAP 特征重要性柱状图

Fig. 8 SHAP bar plot of feature importance

距道路距离和 NDVI 是最为显著的两个驱动因子。距道路距离因子具有最高的平均绝对 SHAP 值 0.1888, 表明其对模型输出贡献最大。蜂巢图点分布范围广泛且密集, 特别是在负值区域有大量红色和蓝色点聚集, 表明其对模型预测结果具有重要影响; 较低的距道路距离值(蓝点)倾向于产生正的 SHAP 值, 显著增加滑坡概率。这是因为道路建设往往伴随着切坡、填方及车辆震动等工程扰动, 破坏了斜坡原有的应力平衡, 从而降低了坡体稳定性; 反之, 远离道路的高距离值(红点)其 SHAP 值主要分布在负值区间, 表明人为扰动较小, 滑坡风险相对降低。

NDVI 点分布较为分散, 表现出复杂的非线性关系。中等偏低的 NDVI 值往往对应较高的滑坡概率, 可能与次生植被、裸土或植被扰动有关, 而极高的 NDVI(如成熟、密集森林)通常对滑坡存在负贡献, 反映了成熟植被的固坡效应。该非线性关系提示植被类型与生长阶段在稳坡作用中扮演不同角色。同时, 岩性和地形位置指数 TPI 也显示出明确的影响模式, 揭示了地质背景与地貌部位对滑坡的基础控制作用。相比之下, 曲率等某些指标的平均绝对 SHAP 值较小 0.0014, 表明其在本次模型中的相对影响有限。

总体来看, SHAP 分析不仅验证了模型输出的物理合理性, 还为定量识别主控因子与制定有针对性的治理措施提供了科学依据。应优先对靠近道路且植被破坏的河谷坡地开展工程或生态修复干预, 以降低滑坡发生概率。

4 结论

(1)以黄山市为研究对象, 选取 14 个因子, 并通

过皮尔逊相关性分析和信息增益率对因子进行筛选, 建立滑坡易发性评价体系。基于 TabPFN 模型进行滑坡易发性评估, 其在各项性能指标上均优于 LR、RF、SVM 和 CNN 模型, 准确率分别比 LR、RF、SVM 和 CNN 提升 9.2%、1.7%、4.8% 和 2.7%, 精确率分别提升 12.57%、3.29%、4.67% 和 7.76%, F1 分数分别提升 11.22%、1.63%、5.08% 和 1.88%, AUC 分别提升 9.6%、1.8%、4.7% 和 2.8%, 表明预测精度有提升。

(2) 利用 SHAP 算法分析各评价因子对模型预测结果的影响方向与强度, 结果显示距道路距离 ($\phi = 0.1888$)、NDVI ($\phi = 0.1101$)、岩性 ($\phi = 0.0915$)、TPI ($\phi = 0.0558$) 和降雨 ($\phi = 0.0513$) 对模型结果贡献较高, 是主导性预测因子, 可解释性分析不仅量化了不同环境因子在模型决策中的权重, 也凸显了关键致灾因子的控制作用。

(3) 基于 TabPFN 模型的滑坡易发性评价结果表明, TabPFN 展现出更强的空间区分和细节表现能力, 将 32.54% 与 31.41% 的区域分别划定为极低与极高易发区, 压缩了模糊的中间过渡带, 更能果断地界定稳定区与高危区; 精细勾勒断裂带沿线的高险区, 能更深层地捕捉致灾因子间复杂的非线性关系, 是进行精细化滑坡风险评估的优选工具。

(4) 尽管 TabPFN 在本研究中表现优异, 但仍存在一定局限性。由于滑坡环境因子普遍存在空间自相关性, 随机划分训练集和测试集的方式可能导致模型评价指标被高估; 未来将引入空间交叉验证机制, 在空间上划分相互独立的训练集与测试集, 更准确地评估模型的真实泛化能力。模型在地形复杂或数据稀缺地区的推广效果有待验证, 后续将探索模型在不同地理区域的适用性。此外, 当前评价体系偏重静态因子, 对动态降雨过程及深部孕灾信息的表征不足, 将整合合成孔径雷达干涉 (InSAR) 数据与高分辨率降雨数据, 进一步提升模型在时空尺度上的预测能力。

利益冲突声明 (Conflict of Interests):

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

参考文献 (References):

窦杰, 向子林, 许强, 等. 机器学习在滑坡智能防灾减灾中的应用与发展趋势[J]. 地球科学, 2023, 48(5): 1657–1674.

DOU Jie, XIANG Zilin, XU Qiang, et al. Applications and development trends of machine learning in intelligent landslide disaster prevention and reduction[J]. Earth Science, 2023, 48(5): 1657–1674.

付泉, 党光普, 李致博, 等. 基于分形维数耦合支持向量机和熵权模型的滑坡易发性研究[J]. 西北地质, 2024, 57(6): 255–267.

FU Quan, DANG Guangpu, LI Zhibo, et al. Study of landslide susceptibility mapping based on fractal dimension integrating support vector machine with index of entropy model[J]. Northwestern Geology, 2024, 57(6): 255–267.

黄发明, 陈彬, 毛达雄, 等. 基于自筛选深度学习的滑坡易发性预测建模及其可解释性[J]. 地球科学, 2023, 48(5): 1696–1710.

HUANG Faming, CHEN Bin, MAO Daxiong, et al. Landslide susceptibility prediction modeling and interpretability based on self-screening deep learning model[J]. Earth Science, 2023, 48(5): 1696–1710.

黄煜, 谢婉丽, 刘琦琦, 等. 基于 GIS 与 MaxEnt 模型的滑坡易发性评价——以铜川市中部城区为例[J]. 西北地质, 2023, 56(1): 266–275.

HUANG Yu, XIE Wanli, LIU Qiqi, et al. Landslide susceptibility assessment based on GIS and MaxEnt model: example from central districts in Tongchuan City[J]. Northwestern Geology, 2023, 56(1): 266–275.

林荣福, 刘纪平, 徐胜华, 等. 随机森林赋权信息量的滑坡易发性评价方法[J]. 测绘科学, 2020, 45(12): 131–138.

LIN Rongfu, LIU Jiping, XU Shenghua, et al. Evaluation method of landslide susceptibility based on random forest weighted information[J]. Science of Surveying and Mapping, 2020, 45(12): 131–138.

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 Jian, LI Shulin, CHEN Tao. Landslide susceptibility assessment based on optimized random forest model[J]. Geomatics and Information Science of Wuhan University, 2018, 43(7): 1085–1091.

马啸, 王念秦, 李晓抗, 等. 基于 RF-FR 模型的滑坡易发性评价——以略阳县为例[J]. 西北地质, 2022, 55(3): 335–344.

MA Xiao, WANG Nianqin, LI Xiaokang, et al. Assessment of landslide susceptibility based on RF: taking Lueyang County as an example[J]. Northwestern Geology, 2022, 55(3): 335–344.

孙萍萍, 张茂省, 贾俊, 等. 中国西部黄土区地质灾害调查研究进展[J]. 西北地质, 2022, 55(3): 96–107.

SUN Pingping, ZHANG Maosheng, JIA Jun, et al. Geo-hazards Research and Investigation in the Loess Regions of Western China[J]. Northwestern Geology, 2022, 55(3): 96–107.

田媛, 巨能攀, 解明礼, 等. 滑坡编录表达模式对易发性评价结果的影响[J]. 成都理工大学学报 (自然科学版), 2022, 49(5): 606–615.

- TIAN Yuan, JU Nengpan, XIE Mingli, et al. Analysis of the impact of landslide cataloguing expression patterns on the evaluation results of landslide susceptibility[J]. Journal of Chengdu University of Technology (Science & Technology Edition), 2022, 49(5): 606–615.
- 王朋伟, 安玉科. 滑坡匀速变形阶段快速诊断方法研究[J]. 西北地质, 2023, 56(5): 197–203.
- WANG Pengwei, AN Yuke. Research on rapid diagnosis method of landslide's uniform deformation stage[J]. Northwestern Geology, 2023, 56(5): 197–203.
- 翁望飞, 李维, 徐生发, 等. 黄山市地质遗迹资源类型、特征及成因[J]. 华东地质, 2020, 41(3): 215–228.
- WENG Wangfei, LI Wei, XU Shengfa, et al. Types, features and genesis of geological heritage resources in Huangshan City[J]. East China Geology, 2020, 41(3): 215–228.
- 周晓亭, 黄发明, 吴伟成, 等. 基于耦合信息量法选择负样本的区域滑坡易发性预测[J]. 工程科学与技术, 2022, 54(3): 25–35.
- ZHOU Xiaoting, HUANG Faming, WU Weicheng, et al. Regional landslide susceptibility prediction based on negative sample selected by coupling information value method[J]. Advanced Engineering Sciences, 2022, 54(3): 25–35.
- Budimir M, Atkinson P, Lewis H. A systematic review of landslide probability mapping using logistic regression[J]. Landslides, 2015, 12(3): 419–436.
- Chen W, Xie X, Wang J, et al. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility[J]. Catena, 2017, 151: 147–160.
- DeLong E R, DeLong D M, Clarke-Pearson D L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach[J]. Biometrics, 1988, 44(3): 837–845.
- Fang H, Shao Y, Xie C, et al. A New Approach to Spatial Landslide Susceptibility Prediction in Karst Mining Areas Based on Explainable Artificial Intelligence[J]. Sustainability, 2023, 15(4): 3094.
- Fu Y, Fan Z, Li X, et al. The Influence of Non-Landslide Sample Selection Methods on Landslide Susceptibility Prediction[J]. Land, 2025, 14(4): 722.
- Gu T, Li J, Wang M, et al. Study on landslide susceptibility mapping with different factor screening methods and random forest models[J]. PLoSone, 2023, 18(10): e0292897.
- Gutiérrez F, Linares R, Roqué C, et al. Large landslides associated with a diapiric fold in Canelles Reservoir (Spanish Pyrenees): detailed geological–geomorphological mapping, trenching and electrical resistivity imaging[J]. Geomorphology, 2015, 241: 224–242.
- Hollmann N, Müller S, Eggersperger K, et al. TabPFN: A transformer that solves small tabular classification problems in a second[J]. arXiv, 2022: 220701848.
- Huang F, Liu K, Jiang S, et al. Optimization method of conditioning factors selection and combination for landslide susceptibility prediction[J]. Journal of Rock Mechanics and Geotechnical Engineering, 2025, 17(2): 722–746.
- Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Advances in Neural Information Processing Systems. Long Beach, CA, 2017.
- Merghadi A, Yunus A P, Dou J, et al. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance[J]. Earth-Science Reviews, 2020, 207: 103225.
- McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages[J]. Psychometrika, 1947, 12(2): 153–157.
- McElfresh D, Khandagale S, Valverde J, et al. When do neural nets outperform boosted trees on tabular data?[C]//Advances in Neural Information Processing Systems. 2023, 36.
- Pham B T, Shirzadi A, Bui D T, et al. A hybrid machine learning ensemble approach based on a radial basis function neural network and rotation forest for landslide susceptibility modeling: A case study in the Himalayan area, India[J]. International Journal of Sediment Research, 2018, 33(2): 157–170.
- Pradhan B. Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches[J]. Journal of the Indian Society of Remote Sensing, 2010, 38(2): 301–320.
- 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.
- Shen C, Feng Z, Xie C, et al. Refinement of landslide susceptibility map using persistent scatterer interferometry in areas of intense mining activities in the karst region of southwest China[J]. Remote Sensing, 2019, 11(23): 2821.
- Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems, 2017: 5998–6008.
- Wang Y, Fang Z, Hong H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China[J]. Science of The Total Environment, 2019, 666: 975–993.
- Zhang J, Ma X, Zhang J, et al. Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model[J]. Journal of Environmental Management, 2023, 332: 117357.
- Zhu Z, Gan S, Yuan X, et al. Landslide susceptibility mapping with integrated SBAS-InSAR technique: a case study of Dongchuan District, Yunnan (China)[J]. Sensors, 2022, 22(15): 5587.