



DOI:10.12404/j.issn.1671-1815.2405487

引用格式:包峻帆,陈婕,杨文涛,等.利用可解释机器学习模型判别豫西巩义市康店镇黄土地质灾害易发性[J].科学技术与工程,2025,25(15):6200-6219.

Bao Junfan, Chen Jie, Yang Wentao, et al. Susceptibility identification of loess geological hazards in Kangdian Town, Gongyi City, Western Henan Province by using interpretable machine learning models[J]. Science Technology and Engineering, 2025, 25(15): 6200-6219.

利用可解释机器学习模型判别豫西巩义市康店镇黄土地质灾害易发性

包峻帆^{1,2}, 陈婕^{3*}, 杨文涛⁴, 杨泽强^{1,2}, 侯文青³, 陈恪^{1,2}, 袁野⁵, 杨明权^{1,2},
景斐媛^{1,2}, 刘森昕^{1,2}, 刘哲^{1,2}, 张媛媛^{1,2}, 黄灿^{1,2}

(1. 河南省地质矿产勘查开发局第三地质矿产调查院, 信阳 464000; 2. 河南省自然资源科技创新中心(信息感知技术应用研究), 信阳 464000; 3. 中国地质大学(武汉)地球物理与空间信息学院, 武汉 430074; 4. 河南理工大学资源环境学院, 焦作 454000; 5. 中国建筑材料工业地质勘查中心河南总队, 信阳 464000)

摘要 黄土丘陵区是地质灾害高发频发的地区之一,亟需采用合适的评价因子和训练模型开展地质灾害易发性评价研究。以郑州“7·20”特大暴雨期间受灾最严重的乡镇巩义市康店镇为研究区,基于卫星遥感解译、实地调查、无人机航拍及相关资料收集,构建覆盖黄土界面、人类工程活动、水动力作用3个主控因素13个影响因子的评价体系,采用CatBoost模型、XGBoost模型和LightGBM模型共3种机器学习算法,开展地质灾害易发性评价研究,基于性能最优的机器学习模型,运用SHAP(shapley additive explanations)算法完成特征全局解释与依赖性分析。结果表明:CatBoost模型的精度高于其他模型(XGBoost和LightGBM),在AUC(area under curve)值、SHAP准确度、精确率、召回率、 F_1 分数和野外验证中均表现最优,其极高、高、中、低、极低易发区域面积占比分别为3.19%、1.40%、2.04%、5.93%、87.44%,极高、高易发区域主要分布在人类活动强烈的冲沟两侧,切坡建房是地质灾害发生的重要诱因。本次研究旨在优化建模思路,对建模过程的不确定性和可解释性进行研究,对机器学习的易发性决策机理进行解释分析,为豫西黄土丘陵区地质灾害防治提供科学依据。

关键词 黄土丘陵区; 地质灾害易发性; 机器学习模型; SHAP; 模型解释

中图法分类号 P642.22 X43; 文献标志码 A

Susceptibility Identification of Loess Geological Hazards in Kangdian Town, Gongyi City, Western Henan Province by Using Interpretable Machine Learning Models

BAO Jun-fan^{1,2}, CHEN Jie^{3*}, YANG Wen-tao⁴, YANG Ze-qiang^{1,2}, HOU Wen-qing³,
CHEN Ke^{1,2}, YUAN Ye⁵, YANG Ming-quan^{1,2}, JING Fei-yuan^{1,2}, LIU Miao-xin^{1,2},
LIU Zhe^{1,2}, ZHANG Yuan-yuan^{1,2}, HUANG Can^{1,2}

(1. The Third Geological and Mineral Survey Institute of Henan Provincial Geological and Mineral Exploration and Development Bureau, Xinyang 464000, China; 2. Henan Provincial Natural Resources Science and Technology Innovation Center (Application Research of Information Perception Technology), Xinyang 464000, China; 3. School of Geophysics and Space Information, China University of Geosciences (Wuhan), Wuhan 430074, China; 4. School of Resources and Environment, Henan University of Technology, Jiaozuo 454000, China; 5. Henan Brigade of China Building Materials Industry Geological Exploration Center, Xinyang 464000, China)

[Abstract] The loess hilly area is one of the areas with a high incidence of geological disasters, and it is urgent to use appropriate evaluation factors and training models to conduct research on the susceptibility assessment of geological disasters. Kangdian Town, Gongyi City, the township hardest hit during the “7·20” extremely heavy rainstorm in Zhengzhou, was taken as the study area. Based on satellite remote sensing interpretation, field survey, UAV aerial photography and relevant data collection, an evaluation system covering 13 influencing factors of three main control factors, namely loess interface, human engineering activities and hydrodynamic effects, was constructed. CatBoost model, XGBoost model and LightGBM model were used to carry out the evaluation study

收稿日期: 2024-07-22 修订日期: 2024-10-29

基金项目: 河南省自然资源科研项目(2022-12)

第一作者: 包峻帆(1989—),男,汉族,河南信阳人,硕士,工程师。研究方向:地质矿产及地质灾害。E-mail:763323197@qq.com。

* 通信作者: 陈婕(1998—),女,汉族,湖北武汉人,博士研究生。研究方向:遥感及地质灾害。E-mail:chenjie@cug.edu.cn。

of geological disaster vulnerability. Based on the machine learning model with the best performance, SHAP (shapley additive explanations) algorithm was used to complete the global interpretation of characteristics and dependency analysis. The results show that the CatBoost model has higher accuracy than other models (XGBoost and LightGBM), and performs the best in AUC (area under curve) value, accuracy, precision, recall, F_1 score, and field validation. The proportion of areas with extremely high, high, medium, low, and extremely low susceptibility is 3.19%, 1.40%, 2.04%, 5.93%, and 87.44%, respectively. The extremely high and high susceptibility areas are mainly distributed on both sides of gullies with strong human activities, and slope cutting and building are important causes of geological disasters. The aim of this study is to optimize the modeling approach, investigate the uncertainty and interpretability of the modeling process, explain and analyze the decision-making mechanism of machine learning susceptibility, and provide scientific basis for geological disaster prevention and control in the loess hilly area of western Henan.

[Keywords] loess hilly region; geological disaster susceptibility; machine learning models; SHAP; model interpretation

黄土具有特殊的崩解性、水敏性、大孔隙性等特点,导致黄土区地质灾害频发^[1-4]。黄土地质灾害主要有崩塌、滑坡、泥石流、地面沉降、黄土湿陷、塌陷、地裂缝等7种类型,内部结构复杂、演化过程多样,具有隐蔽性深、影响范围广、破坏性强的特点^[5],呈现出点多面广、规模小、突发性强,降雨、冻融和人类工程扰动是诱发黄土地质灾害的主要因素^[6]。黄河流域黄土区目前记录了超过10 000多次地质灾害,经济损失超过67亿元^[7-8]。而通过建立地质灾害多类型因子评价体系,开展易发性评价是分析地质灾害潜在分布的重要手段^[9-11]。现有黄土地质灾害早期识别与易发性研究多针对陕西^[12-13]、甘肃^[14-16]、新疆^[17-18]、宁夏^[19-20]等地,鲜少学者对河南黄土地质灾害分布及易发性进行研究。康店镇北濒黄河,其地质灾害类型主要为崩塌,为力-水源型黄土地质灾害,相比与其他地区,具有地质背景复杂、节理裂隙发育、土体结构疏松多孔、地表覆盖参差不齐、干湿强度差异大、沟谷密度大、易受人类工程和降雨扰动等特点,为地质灾害高发频发的地区之一,受黄土界面、人类工程活动和水动力作用控制明显。据自然资源部门相关统计数据,仅2021年郑州“7·20”特大暴雨期间,巩义市康店镇黄土区发生地质灾害灾情点85处,主要为崩塌,造成直接经济损失3 909.1万元,2 972人受灾,分别占郑州地区灾情点的43.62%、直接经济损失的41.37%,受灾人数的57.42%,为豫西地区地质灾害受灾最严重的乡镇,亟需开展地质灾害易发性评价研究。

地质灾害易发性受直接和间接因素影响^[21]。以往研究大多通过分析地质灾害地形地貌、人类工程活动等多种因素,采用定性或定量方法来反映地质灾害发生的可能性及潜在空间分布^[22-23]。定性方法,如层次分析法^[24]等,依赖于专家经验,具有一定主观性。1980年以来,频率比法^[25]、信息量法等^[26]通过数理统计的定量方法很大程度上克服了定性方法的主观性,广泛且成熟地应用于地质灾害

易发性评价研究^[27-30]。然而,传统的统计分析方法容易忽视不同因子的非线性问题,需要大量先验知识,存在一定局限性^[31]。

随着人工智能的快速发展,逻辑回归^[32]、神经网络^[33]、随机森林^[34]、支持向量机^[35]、决策树^[36]等为代表的机器学习、深度学习方法能较好解决非线性问题,反映数据深层特征,并在地质灾害易发性评价中取得显著成果^[37]。不同算法模型存在不同的优点和不足^[38],算法模型直接影响评价结果的准确性及合理性^[39]。

而Boosting梯度提升新模型(CatBoost、XGBoost和LightGBM模型)很少在地质灾害易发性评价中被研究对比^[40]。不同外部环境中算法模型的精度有所不同,以往研究大都采用单一模型,缺乏对不同模型精度的检验和对比^[41]。

因此,针对康店镇黄土区地质灾害易发性评价研究少、Boosting梯度提升新模型(CatBoost、XGBoost和LightGBM模型)被对比研究少^[40]、不同模型精度的检验和对比被关注少^[41]的现状。现结合康店镇等豫西黄土地质灾害受黄土界面、人类工程活动和水动力作用控制明显的特点,以郑州“7·20”特大暴雨期间受灾最严重的乡镇巩义市康店镇为研究区,采用防止过拟合能力强的CatBoost模型、XGBoost模型和LightGBM模型,选取等高线密度、植被覆盖度、土地利用类型、距道路距离、距建筑物距离等13个能反映康店镇黄土地质灾害主要控制因素的因子,开展地质灾害易发性评价研究。相对于以往研究,本文中采用SHAP (shapley additive explanations)可解释算法,完成特征全局解释与依赖性分析,对机器学习的易发性决策机理进行解释分析,为豫西黄土区进一步开展地质灾害防治、国土空间规划管制等工作提供借鉴。

1 研究区概况及数据来源

1.1 研究区概况

巩义市位于河南省中部,属郑州市。康店镇,

隶属河南省郑州市巩义市,地处巩义市西部,如图1所示,地理坐标为 $112^{\circ}47'56'' \sim 112^{\circ}58'32''E$, $34^{\circ}43'19'' \sim 34^{\circ}51'15''N$,面积 109.15 km^2 。地形地貌以黄土丘陵为主,属邙岭山系,东依伊洛河,北濒黄河。总体地形为南高北低。平均气温 $15.33 \text{ }^{\circ}\text{C}$,年降水量 390 mm ,全年无霜期 232 d 。黄土丘陵区地层岩性为中更新统(Qp_2^{el+dl})、上更新统(Qp_3^{al})黄土,具有中-微湿陷性。地形标高 $107 \sim 280 \text{ m}$,相对高差 $20 \sim 75 \text{ m}$,冲沟十分发育,地面沟壑纵横,地形支离破碎,沟谷多呈“U”字形,壁陡谷深,坡角近于直立状。北部黄河右岸和东部伊洛河左岸构成康店镇冲洪积倾斜平原区,全新统(Qh)黄土状亚砂土、亚黏土及粉细砂、中粗砂等为其主要岩性,湿陷等级中等。人类工程活动有切坡建房、筑路、农作物种植等,区内分布85处崩塌地质灾害点。

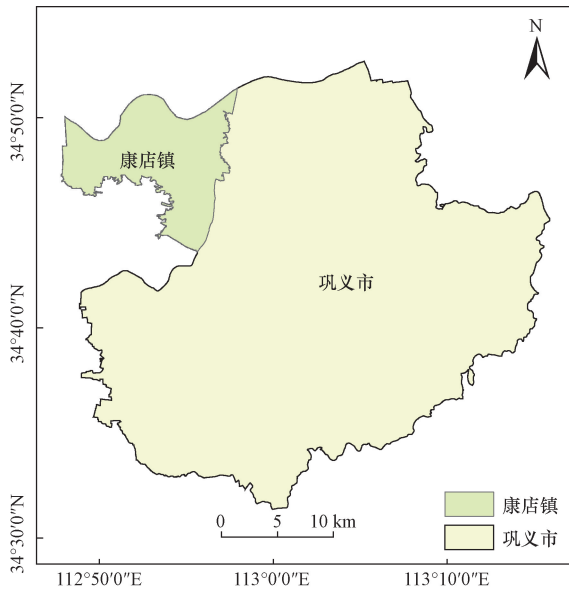


图1 研究区位置图

Fig. 1 Location map of the study area

1.2 数据来源

通过卫星遥感解译、实地调查、无人机航拍及相关资料收集,获取巩义市康店镇地质灾害点共85处,均为崩塌,面密度 $0.77 \text{ 处}/\text{km}^2$,分布在黄土丘陵区,分布相对集中。

数字高程模型(digital elevation model, DEM)数据通过收集到的巩义市1:10 000地形图构建,共收集2017年5月最新测制覆盖巩义市康店镇全域 109.15 km^2 的1:10 000地形图19幅。

DEM生成采用由等高线构建不规则三角网(triangulated irregular network, TIN)法,由等高线生成TIN,使用栅格转换方法由TIN进行内插快速生成格网DEM。

坡度、地形起伏度、剖面曲率通过ArcGIS空间

分析功能基于DEM生成;等高线密度基于1:10 000地形图等高线计算得到;植被覆盖度通过地理空间数据云(<http://www.gscloud.cn>)下载的Landsat 8 OLI数据计算得到;地层和构造数据来源于1:50 000地质图;土地利用、道路和水系数据来源于全国第三次土地调查数据。

为确保各因子间具有一致的空间分辨率,本文采用网格单元作为评价单元,其具有划分简单、易于操作和速度快的优点^[42],以DEM数据为基准,设置网格大小为 $5 \text{ m} \times 5 \text{ m}$ 。

2 评价方法

2.1 XGBoost 模型

XGBoost模型是在梯度提升决策树(gradient boosting decision tree, GBDT)的基础上进行了工程优化和算法优化。XGBoost的目标函数主要由损失函数和L2正则项共同构成,目标函数为

$$\text{Obj} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (1)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (2)$$

式中: $\sum_{i=1}^n l(y_i, \hat{y}_i)$ 为损失函数,由预测值 \hat{y}_i 与真实值 y_i 表示; $\sum_{k=1}^K \Omega(f_k)$ 为L2正则项,是全部决策树叶子节点数和节点权重形成的向量决定; T 为叶子数,叶子节点越少模型越简单; γ 和 λ 为正则项参数; ω 为叶子节点权重。

2.2 LightGBM 模型

LightGBM是新的梯度提升算法^[43],在GBDT算法基础上,采用直方图分裂决策树算法、带深度限制节点展开法、单边梯度采样及互斥特征捆绑算法,每次选择一个最合适的叶子节点进行分裂,显著提高计算效率和模型精度。单边梯度采样算法可保留对计算信息增益有利的高梯度样本,方差增益为

$$V_j(d) = \frac{1}{n} \left\{ \frac{\left[\sum_{x_i \in A_{l,1}} G_i + \left(\frac{1-a}{c} \right) \sum_{x_i \in A_{l,2}} G_i \right]^2}{n_l^j(d)} + \frac{\left[\sum_{x_i \in A_{r,1}} G_i + \left(\frac{1-a}{c} \right) \sum_{x_i \in A_{r,2}} G_i \right]^2}{n_r^j(d)} \right\} \quad (3)$$

式(3)中: n 为样本数; j 为分裂特征; d 为分裂点; G 为样本梯度; a 和 c 分别为大梯度和小梯度样本的采样率; A 为所分裂节点的梯度样本; l, r 分别表示左、右子节点; 1 和 2 分别代表大梯度和小梯度样本。

2.3 CatBoost 模型

CatBoost 在传统 GBDT 基础上,直接处理类别特征并平滑处理避免过拟合,采用顺序建树算法避免信息泄露,使用对称树提高计算速度,采用动态学习率以加速收敛。CatBoost 算法具有处理类别特征的能力、高效性、灵活性、鲁棒性和可解释性。

2.4 基于 Optuna 框架的参数优化方法

Optuna 是一款适用于机器学习的自动超参数优化框架,可以动态地构造超参数搜索空间,对比传统网格搜索法和贝叶斯调参法具有运行速度快、占用计算资源少的优点。基于 Optuna 框架对 3 种模型进行自动超参数确定,参数调节结果如表 1 所示。

表 1 参数调节结果

Table 1 Parameter adjustment results

模型	参数调节结果
CatBoost	iterations = 859, depth = 9, learning_rate = 0.07, random_strength = 0.78, bagging_temperature = 0.32, od_type = lter, od_wait = 24
XGBoost	n_estimators = 650, max_depth = 5, min_child_weight = 2, gamma = 0.86, subsample = 0.79, colsample_bytree = 0.63, learning_rate = 0.06, reg_alpha = 0.94
LightGBM	n_estimators = 684, max_depth = 6, learning_rate = 0.09, num_leaves = 171, min_child_samples = 11, min_child_weight = 0.01, subsample = 0.97, colsample_bytree = 0.80, reg_alpha = 0.40, reg_lambda = 0.06

2.5 SHAP 可解释性算法

SHAP (shapley additive explanations) 算法是一种机器学习解释方法,采用博弈论的 Shapley 值特征输入模型的每个特征对输出的贡献程度。特征 i 的 SHAP 值可表示为

$$\varphi_i(f, x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f_x(S \cup \{i\}) - f_x(S)] \quad (4)$$

式(4)中: N 为特征集合; S 为不包含特征 i 的任何子集; $|S|$ 和 $|N|$ 分别为集合 S 的大小和所有特征的总数; $f_x(S \cup \{i\})$ 和 $f_x(S)$ 表示包含和不包含特征 i 的模型预测结果。

3 地质灾害易发性划分与评价

3.1 评价因子选取与分析

选择适用、专业的评价因子是开展地质灾害易发性评价的重要工作^[34]。康店镇等豫西黄土丘陵区地质灾害类型主要为崩塌,为力-水源型黄土地质灾害,其演化模式为:在切坡建房等人类活动的作用下,坡体界面扩张开裂,稳定性降低,再叠加水

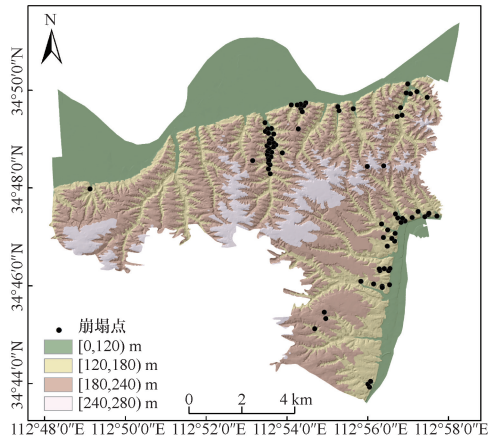
动力的作用,最终变形失稳形成崩塌^[5],黄土界面、人类工程活动、水动力作用是黄土灾害形成的主要控制因素。黄土界面隔开了其两侧的物质,影响着黄土的力学行为和坡体形态^[5]。人类活动会改变斜坡的应力状态与原有的岩土结构^[44]。水是黄土本身的三相之一,也是黄土地质灾害的重要驱动力^[5],地形地貌控制斜坡的水分和应力分布^[45-46]。野外调查是评价因子选取的最优方法^[47],在详细实地调查所有灾害点的前提下,结合研究区地质灾害特征规律与影响因子、黄土地质灾害演化机理及前人研究成果,选取覆盖黄土界面、人类工程活动、水动力作用 3 个主控因素的 13 个与研究区地质灾害密切相关的评价因子:高程、坡度、地形起伏度、地貌、地层、距构造距离、坡形、植被覆盖度、土地利用类型、距道路距离、距水系距离、距建筑物距离、等高线密度,评价因子图及评价因子分级面积百分比、灾害面积百分比与信息量值见图 2 和图 3。

(1) 高程。高程影响黄土区植被分布、人类活动、斜坡稳定、河流切割作用等,进而影响地质灾害分布^[48-49]。康店镇海拔高程 107 ~ 280 m,如图 2(a) 和图 3(a) 所示,将其划分为 4 个区间,其中,高程在 120 ~ 180 m 区间信息量值最高,最易发生地质灾害。

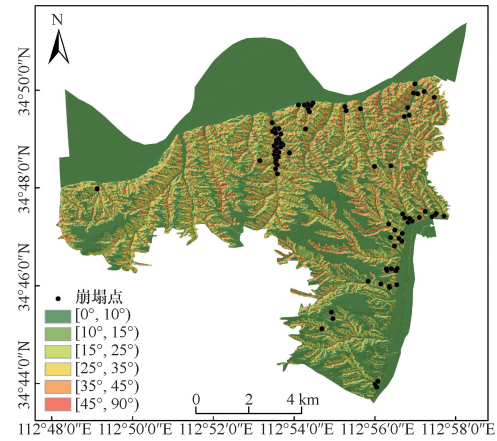
(2) 坡度。坡度决定黄土区斜坡应力分布,进而影响斜坡稳定、地表水动力作用^[50-52]、堆积物厚度及分布^[53]等,为地质灾害的发生提供了地形及水源条件^[54]。康店镇斜坡坡度 $0^\circ \sim 90^\circ$,如图 2(b) 和图 3(b) 所示,将其划分为 6 个区间,大部分地质灾害分布在坡度 $15^\circ \sim 90^\circ$,特别是坡度在 $35^\circ \sim 45^\circ$ 区间信息量值最高,最易发生地质灾害。

(3) 地形起伏度。地形起伏度反映地形的最大高差^[55],是黄土区地质灾害物源分布和储量的重要影响因素^[56],进而影响研究区地质灾害发育。康店镇地形起伏度 0 ~ 114 m,如图 2(c) 和图 3(c) 所示,将其划分为 5 个区间,由因子面积占比可知,康店镇地形起伏度主要分布在 0 ~ 20 m,其中,地形起伏度在 10 ~ 20 m 区间信息量值最高,最易发生地质灾害。

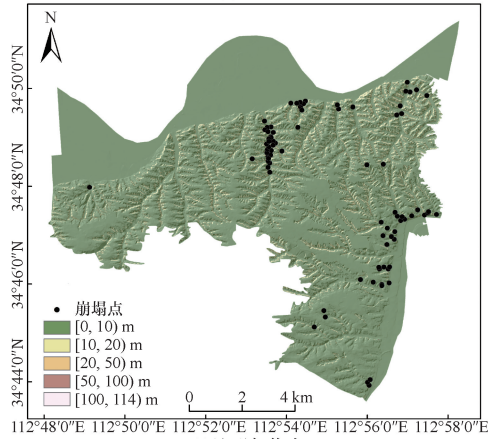
(4) 地貌。地貌类型是黄土区地表岩土体及其地质作用的综合形态^[57],控制斜坡的水分和应力分布从而影响地质灾害发育条件^[58-59]。康店镇地貌类型主要为冲洪积倾斜平原和黄土丘陵两类,如图 2(d) 和图 3(d) 所示,黄土丘陵地貌信息量值远高于冲洪积倾斜平原地貌,最易发生地质灾害。



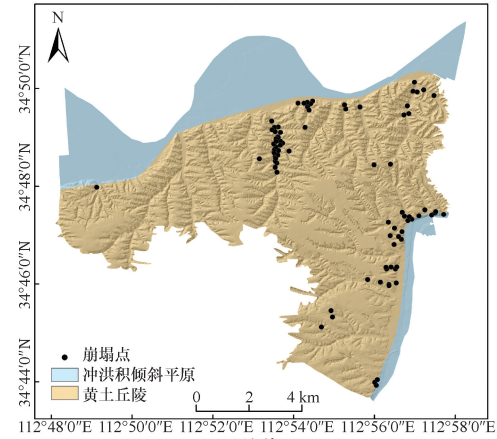
(a)高程



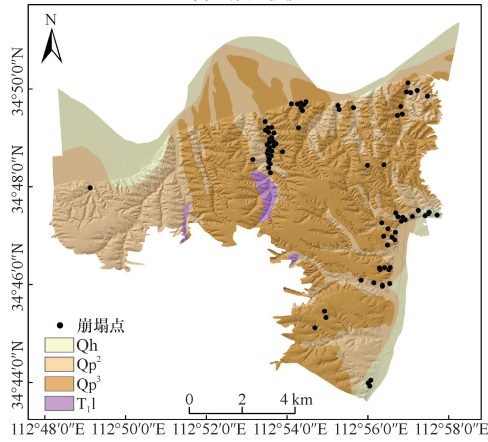
(b)坡度



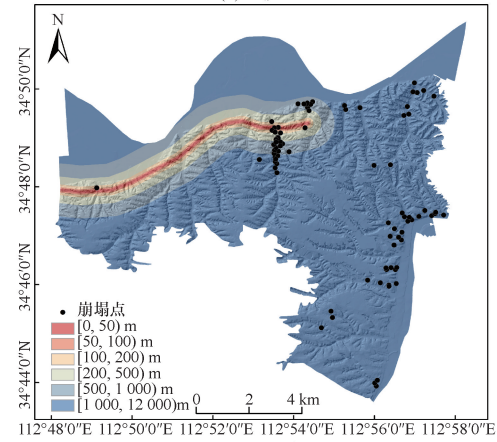
(c)地形起伏度



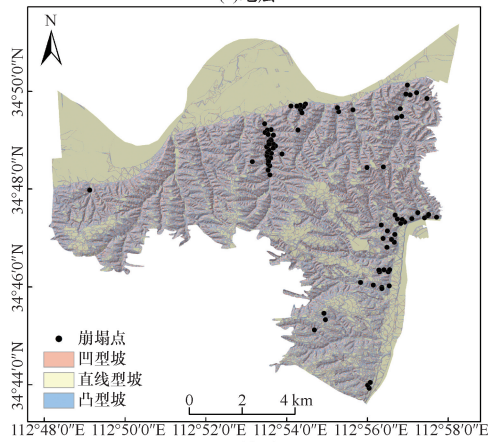
(d)地貌



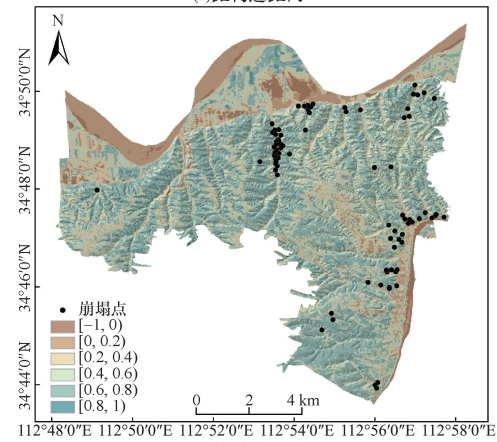
(e)地层



(f)距构造距离



(g)坡形



(h)植被覆盖度

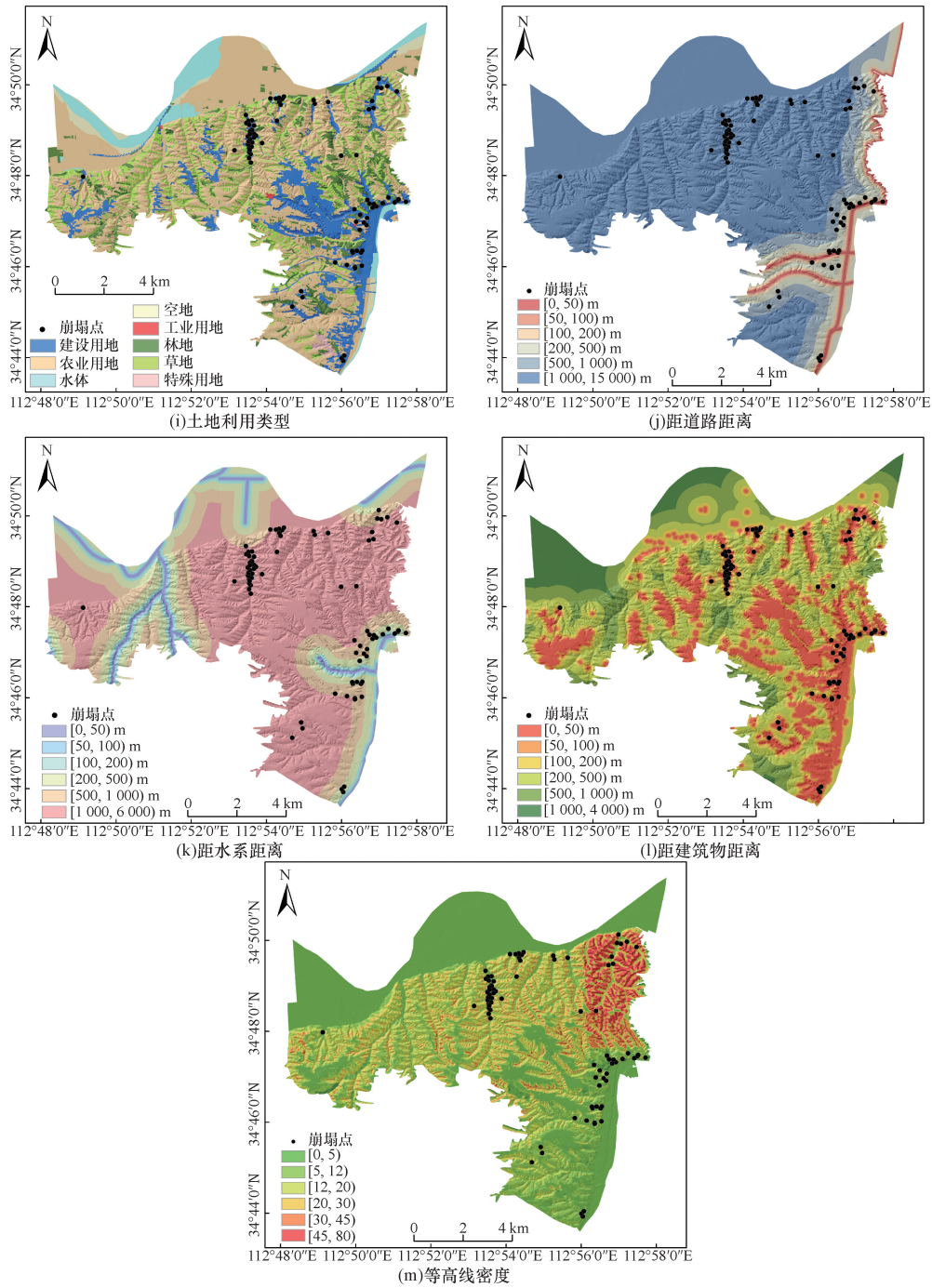


图2 地质灾害评价因子图

Fig. 2 Geological hazard assessment factors

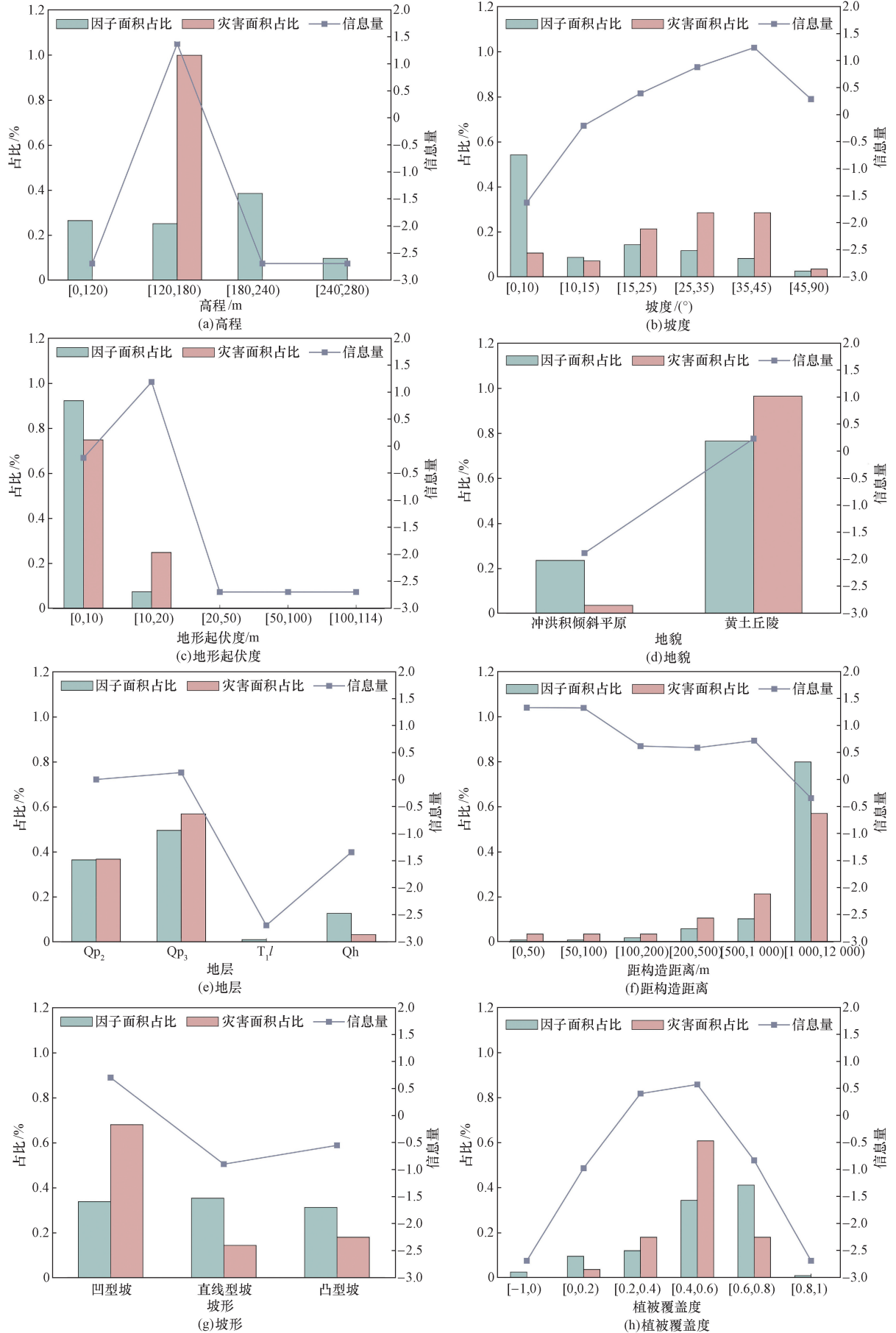
(5) 地层。地层岩性决定黄土区地质灾害物源结构、力学性质及失稳方式^[60], 是黄土区地质灾害的重要影响因素。康店镇地层岩性分为4类: 中生界三叠系下统刘家沟组 (T_1l)、第四系中更新统 (Qp_2^{el+dl})、上更新统 (Qp_3^{al})、全新统 (Qh^{al}), 如图2(e)和图3(e)所示, 上更新统 (Qp_3^{al}) 的信息量值最大, 最易发生地质灾害。

(6) 距构造距离。地质构造运动促使黄土区岩土体挤压破碎, 形成有利于地质灾害发育的地形地

貌^[61]。康店镇距构造距离0~12 000 m, 如图2(f)和图3(f)所示, 将其划分为6个区间, 其中, 距构造距离在0~100 m范围信息量值最高, 最易发生地质灾害。

(7) 坡形。坡形反映坡体在内外力作用下的演变过程^[62], 控制黄土区地质灾害发育状态。坡形分为凸型坡、直线型坡和凹型坡, 如图2(g)和图3(g)所示, 凹型坡信息量值最高, 最易发生地质灾害。

(8) 植被覆盖度。植被会对斜坡稳定性产生水



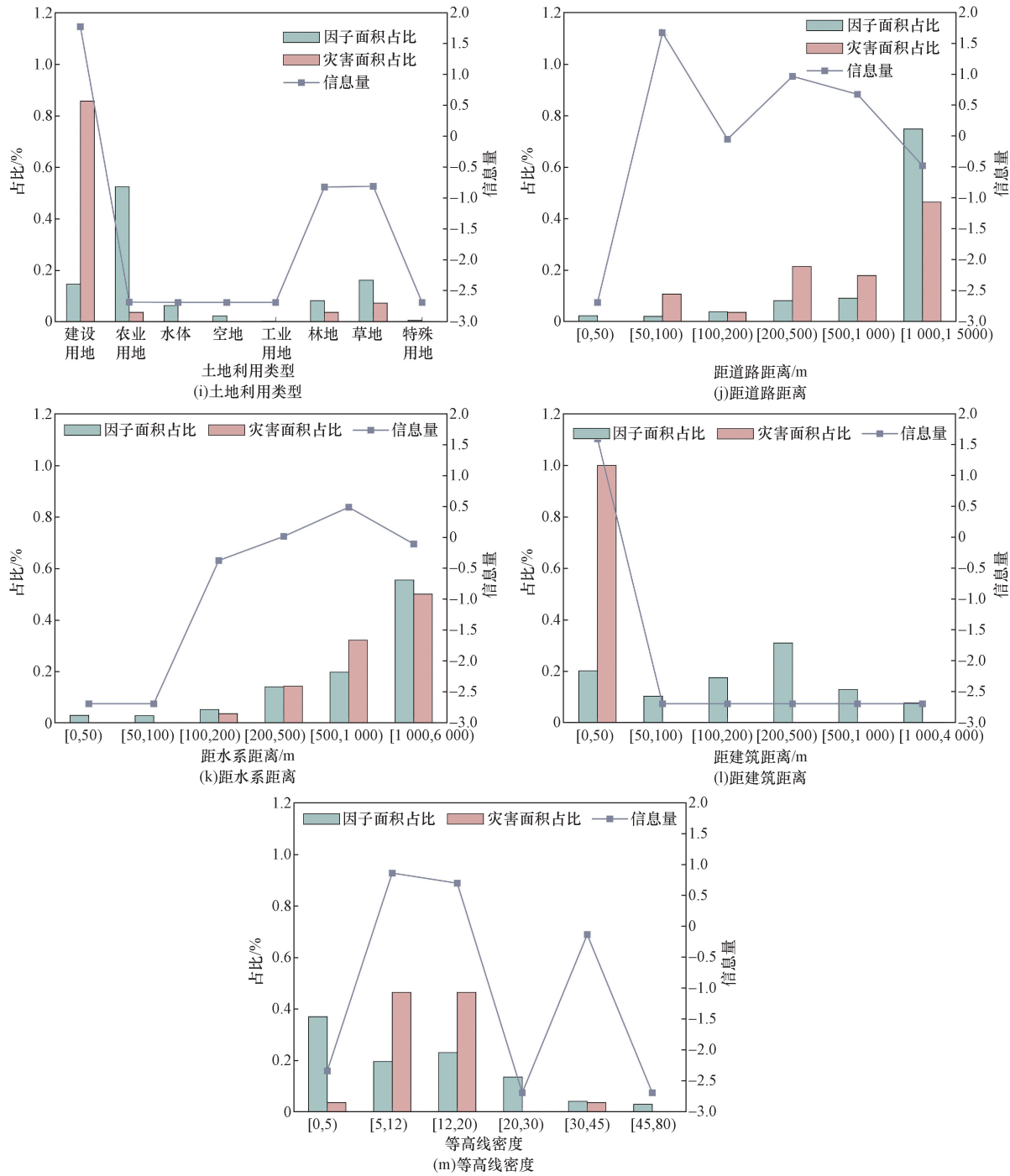


图3 评价因子分级面积百分比、灾害面积百分比与信息量值

Fig. 3 Evaluation factor graded area percentage, disaster area percentage, and information value

文效应和力学效应^[63],能增强水土保持能力,进而降低地质灾害发生概率^[64]。康店镇地表覆盖参差不齐,很大程度上影响了斜坡稳定,地表植被覆盖裸露的地区极易遭受降雨的侵蚀作用从而引发地质灾害^[65]。归一化植被指数(normalized difference vegetation index, NDVI)体现植被覆盖程度,间接反映地质灾害发育情况^[66],NDVI值范围-1~1,如图2(h)和图3(h)所示,将其划分为6个区间,其

中,NDVI值在0.4~0.6范围信息量值最高,最易发生地质灾害。

(9)土地利用类型。土地利用类型是体现人类工程活动的重要指标^[67],也是黄土区地质灾害的重要影响因素。康店镇土地利用类型分为建设用地、农业用地、水体、空地、工业用地、林地、草地、特殊用地8类,如图2(i)和图3(i)所示,建设用地的信息量值最高,最易发生地质灾害。

(10)距道路距离。切坡修路开挖卸荷,改变斜坡原有应力平衡,使坡体变形,降低植被覆盖率^[68],加之汽车震动、降雨等外部因素扰动加速了黄土区斜坡失稳变形,进而诱发地质灾害,距道路距离不同会对黄土区地质灾害有不同影响^[69]。康店镇距道路距离0~15 000 m,如图2(j)和图3(j)所示,将其划分为6个区间,距道路距离50~100 m区间信息量值最高,最易发生地质灾害。

(11)距水系距离。河流为地质灾害发育提供了有利临空面及水动力条件^[70],对坡脚的侵蚀作用加速了黄土区斜坡失稳变形^[71],进而诱发地质灾害发生。康店镇距水系距离0~6 000 m,如图2(k)和图3(k)所示,将其划分为6个区间,距水系距离500~1 000 m区间信息量值最高,最易发生地质灾害。

(12)距建筑物距离。人类大规模开挖边坡、爆破以及修建房屋形成的高陡边坡会改变黄土区斜坡的应力状态与原有的岩土结构,降低斜坡的稳定性,叠加降雨等外部因素扰动,极易诱发地质灾害^[72]。康店镇距建筑物距离0~4 000 m,如图2(l)和图3(l)所示,将其划分为6个区间,距建筑物距离越小,信息量有增大趋势,距建筑物距离0~50 m区间信息量值最高,最易发生地质灾害,表明建筑物两侧为黄土区地质灾害的高发区,切坡建房等人类工程活动是影响研究区地质灾害发育的重要因素。

(13)等高线密度。等高线密度体现了地形坡度的缓陡程度,一定程度上控制着黄土区地表水动力作用和斜坡应力分布,是黄土区地质灾害的重要影响因素。康店镇等高线密度0~80,如图2(m)和图3(m)所示,将其划分为6个区间,等高线密度5~20区间因子面积占比最大,分布最广,信息量值最高,最易发生地质灾害。

3.2 康店镇崩塌易发性评价

3.2.1 因子共线性检验

文献[73-75]显示,地质灾害易发性评价因子的高度相关关系会导致多重共线性的发生^[76],因子重要性会被影响,导致无法正确理解特征^[77]。因此,在构建易发性评价模型前,对选取的13个因子进行共线性检验。通过计算因子间的方差膨胀因子(variance inflation factor, VIF)及容忍度(tolerance, TOL)对多重共线性进行检验,计算公式为

$$VIF = \frac{1}{1 - R_i^2} \quad (5)$$

$$TOL = \frac{1}{VIF} \quad (6)$$

式中: R_i 为评价因子间的方差。

基于SPSS26计算方差膨胀因子和容忍度开展多重共线性检验,计算结果如表2所示。当VIF > 10或TOL < 0.1时,表示因子间存在共线性问题^[78-79]。检验结果显示,本文中选取的13个因子不存在共线性问题。

表2 评价因子共线性检验表

Table 2 Table for evaluating collinearity of factors

序号	因子	TOL	VIF
1	高程	0.387 3	2.581 8
2	坡度	0.319 2	3.132 5
3	地形起伏度	0.625 0	1.600 0
4	地貌	0.288 0	3.472 4
5	坡形	0.825 2	1.211 8
6	地层	0.808 9	1.236 2
7	距构造距离	0.899 8	1.111 3
8	距水系距离	0.771 6	1.296 0
9	等高线密度	0.395 5	2.528 4
10	植被覆盖度	0.657 6	1.520 7
11	土地利用类型	0.801 0	1.248 5
12	距道路距离	0.853 1	1.172 2
13	距建筑距离	0.661 3	1.512 2

3.2.2 易发性评价建模

为避免传统统计分析方法的非线性问题^[31]和机器学习模型数据过拟合、评价结果不确定性等缺点,基于巩义市康店镇黄土地质灾害发育特征及前人工作结果,选取CatBoost模型、XGBoost模型和LightGBM模型开展易发性评价,评价因子统一使用5 m×5 m栅格,选取85处崩塌,将其崩塌范围覆盖的2 100个栅格同2 100个非灾害点栅格构成数据集,按照7:3的比例将数据集拆分为训练集和测试集。将训练集输入模型,用测试集进行精度验证,训练完成后,将整个研究区数据集输入模型,完成巩义市康店镇易发性评价工作。

3.2.3 易发性评价结果

图3和评价因子分析表明,黄土丘陵地貌,上新统(Qp₃^d)地层,高程在120~180 m,坡度在35°~45°,地形起伏度在10~20 m,距构造距离在0~100 m,NDVI值在0.4~0.6,距道路距离50~100 m,距水系距离500~1 000 m,距建筑物距离0~50 m,等高线密度5~20,建设用地和凹型坡地面积大的黄土区最易发生地质灾害。

通过自然断点法将CatBoost、XGBoost和LightGBM 3种模型评价结果分为极低易发、低易发、中易发、高易发和极高易发共5个等级。如图4和表3所示,3种模型评价结果的高及极高易发区具有较高的一致性,主要集中在研究区东部及北部切坡建房等人类活动强烈的冲沟两侧,受人类工程活动影响尤为明显。研究区内的85处崩塌点几乎全

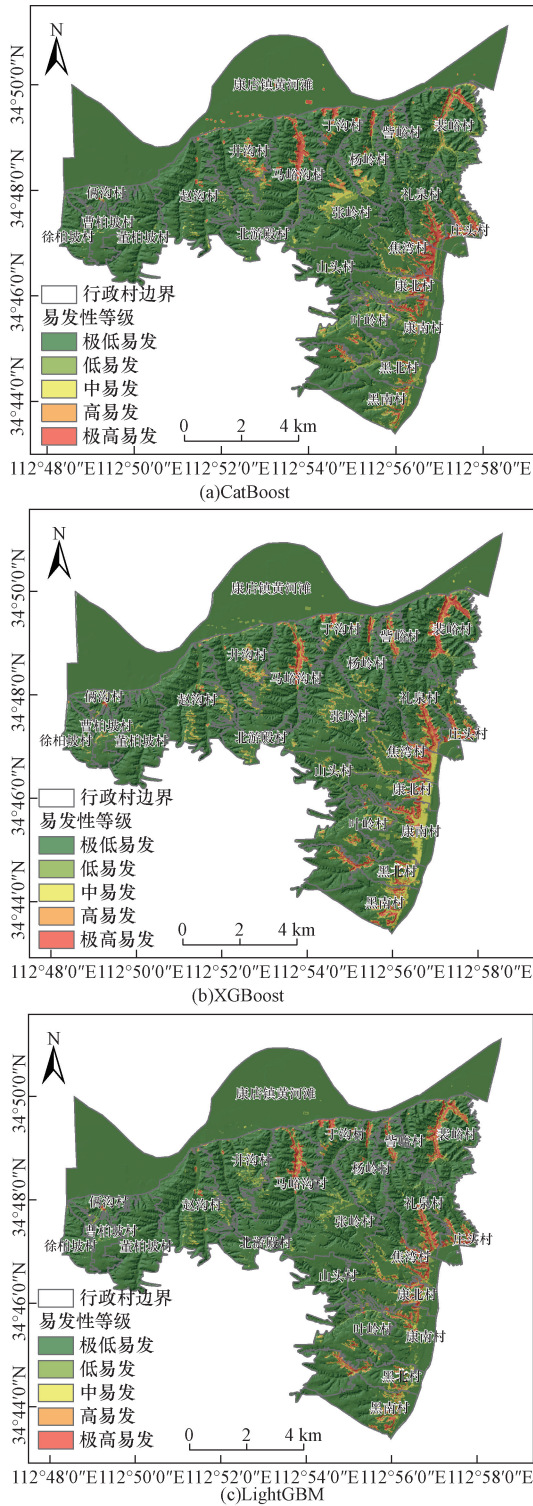


图4 康店镇黄土地质灾害易发性分区图
Fig. 4 Zoning map for susceptibility to loess geological hazards in Kangdian Town

部位于3种模型高和极高易发区域内, CatBoost模型和LightGBM模型的易发性分布图显示出高度一致性, XGBoost模型在东南部的冲洪积倾斜平原带预测出了中易发性, 在该区域对易发性的评估存在过高估计。

表3 地质灾害易发性分区统计表
Table 3 Statistical table for zoning geological hazard susceptibility

模型	评价等级	面积/km ²	面积占比/%	崩塌数量/处	崩塌数量占比/%
CatBoost	极低易发	95.44	87.44	2	2.35
	低易发	6.47	5.93	1	1.18
	中易发	2.22	2.04	1	1.18
	高易发	1.53	1.40	0	0
	极高易发	3.48	3.19	81	95.29
XGBoost	极低易发	92.93	85.14	2	2.35
	低易发	6.75	6.18	0	0
	中易发	3.66	3.35	4	4.71
	高易发	2.31	2.11	9	10.59
	极高易发	3.50	3.21	70	82.35
LightGBM	极低易发	97.57	89.39	1	1.18
	低易发	5.19	4.75	5	5.88
	中易发	1.95	1.79	1	1.18
	高易发	1.46	1.34	4	4.71
	极高易发	2.98	2.73	74	87.06

将评价结果和野外调查对比分析是重要的检验方法^[80]。通过野外实地调查发现, 黄土丘陵区地质环境复杂、地形起伏较大、水力侵蚀作用较强、人口分布集中, 当地居民多习惯筑窑居住, 居民房前屋后切坡斜坡坡度多在60°~90°, 斜坡岩性又以第四系粉土或粉质黏土最普遍, 开挖后斜坡稳定性降低, 多处于临界稳定状态, 切坡后又极少进行支护加固或削坡卸荷等防治措施, 叠加强降雨等外部扰动, 极易诱发崩塌灾害, 严重威胁居民生命及财产安全。康店镇85处崩塌地质灾害点中, 存在切坡建房活动的有80处, 占有地质灾害点的94.12%。因此, 切坡建房成为研究区崩塌地质灾害的重要诱发因素, 与易发性分区分布规律一致。

3.2.4 模型性能分析

(1) 精度评价。模型准确性评估至关重要, 由混淆矩阵得到的ROC曲线、准确率、精确率、AUC、 F_1 分数、召回率等被作为重要量化评估指标^[81]。选择受试者工作特征曲线(receiver operating characteristic curve, ROC)^[82]、准确率、精确率、召回率和 F_1 分数对研究区易发性评价结果进行模型精度检验, 其中, ROC曲线下面积(area under curve, AUC)体现模型预测精度, AUC值范围为[0.5, 1], AUC值越高表明模型预测精度越高, AUC值在0.90以上时表明有很高的准确性和非常好的预测效果^[83-84], 准确率为预测正确的占比, 精确率为实际正例占预测正例的比例, 召回率是所有实际正例被准确预测的占比, F_1 分数是精确率和召回率的加权调和平均值, 用于评估模型性能。如图5和表4所示, 3种模型的AUC值、准确率、精确率、召回率和 F_1 分

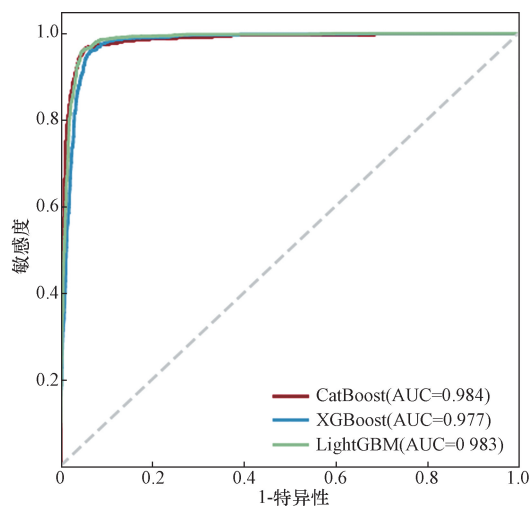


图5 ROC 曲线
Fig. 5 ROC curve

表4 性能评价结果

Table 4 Performance assessment results

模型	AUC 值	准确率/%	精确率/%	召回率/%	F_1 分数/%
CatBoost	0.984	95.28	95.69	94.76	95.22
XGBoost	0.977	93.58	94.54	92.39	93.45
LightGBM	0.983	94.27	96.04	92.25	94.11

数均大于0.9,表明模型评价方法和结果合理性较高,可信度较好。CatBoost 模型和 LightGBM 模型在5种精度评价指标上的表现均优于 XGBoost 模型,LightGBM 模型在精确率上略高于 CatBoost 模型,但在其他指标表现上均逊于 CatBoost。总体来说,CatBoost 模型表现出最优性能,其 AUC、准确率、精确率、召回率和 F_1 分数分别为 0.984、0.953、0.957、0.948、0.952。

(2)野外验证分析。根据实地调查,康店镇新增三处崩塌灾害点,分别为庄头村崩塌点、焦湾村崩塌点和马峪沟村崩塌点。新增崩塌点形成的主要原因基本一致:人工切坡形成的高陡边坡,破坏了原有应力平衡,在坡肩、坡面等应力集中处出现拉张裂缝、变形破裂,在暴雨或久雨等因素的扰动下,雨水入渗增加潜在危岩土体自重,近饱和土体抗剪强度降低,后缘拉张裂缝和垂直节理裂隙上下贯通,最终诱发崩塌。新增崩塌点未参与模型训练,可以很好地评估3种模型预测能力,进一步验证评价结果的准确性。

三处新增的崩塌点位置和现场照片如图6所示。庄头村崩塌点位于山脚,原始坡高20 m,切坡高

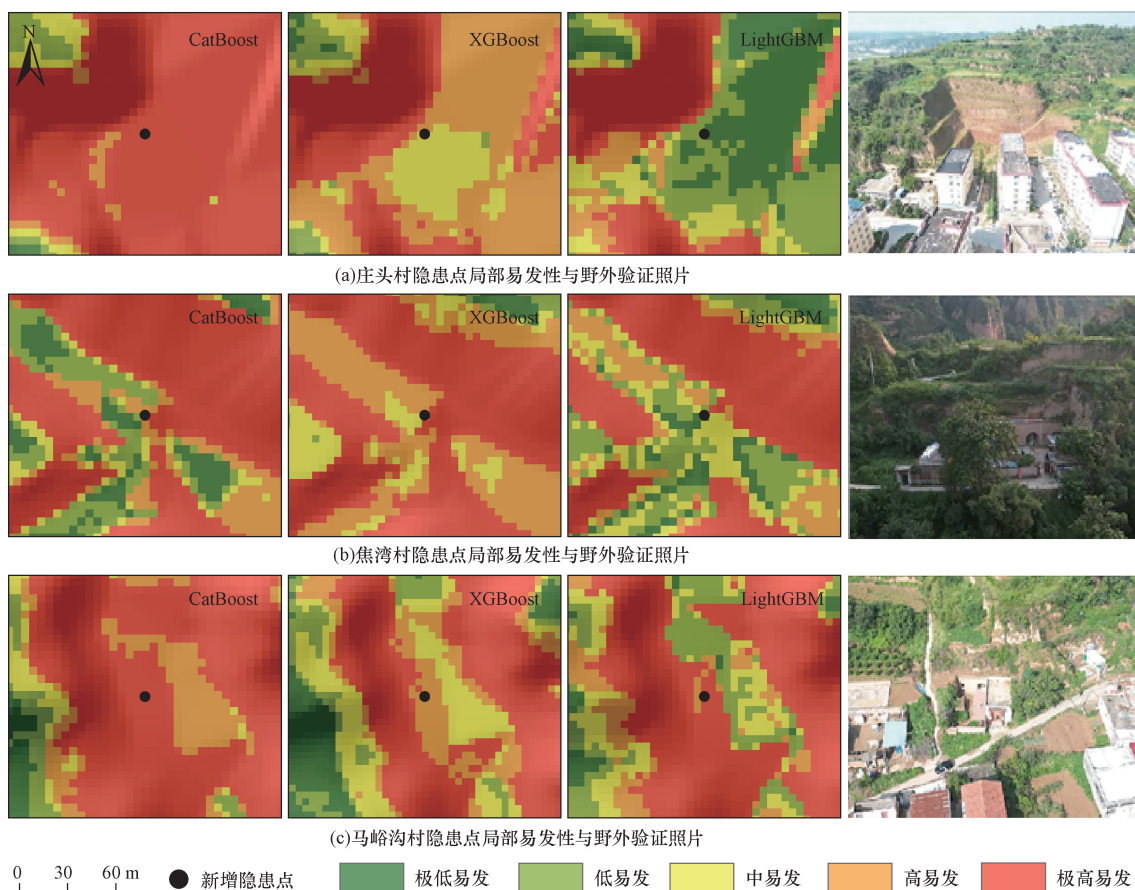


图6 易发性评价结果图及野外照片

Fig. 6 Results and field photos of the susceptibility assessment

度 10 m,坡度 75°,威胁坡脚房屋 30 余间,该点发生崩塌灾害后及时进行了工程治理。该新增崩塌点在 CatBoost 和 XGBoost 模型预测结果的极高易发区,在 LightGBM 模型预测结果的高易发区;焦湾村新增崩塌点位于山腰,切坡高度 10 m,坡度 85°,威胁坡脚房屋 27 间,该新增崩塌点在 CatBoost 和 XGBoost 模型预测结果的极高易发区,在 LightGBM 模型预测结果的中易发区;马峪沟村新增崩塌点切坡高度 10 m,坡度 85°,坡体强风化,存在再次剥坠落可能,威胁坡脚房屋 5 间,该新增崩塌点在 CatBoost 和 LightGBM 模型预测结果的极高易发区,在 XGBoost 模型预测结果的高易发区。从 3 种模型评价结果等级划分情况来看,CatBoost 模型对三处新增的崩塌点空间位置进行了很好的预测,效果最优,易发性评价结果与实际情况更为一致。

综上所述,与 XGBoost 和 LightGBM 模型相比,CatBoost 模型在分级结果的合理性、预测的准确性及与野外实际的吻合度等方面均有优明显优势,更适合对研究区开展易发性评价。

3.3 康店镇崩塌易发性评价

3.3.1 特征全局解释

基于 SHAP 算法对 CatBoost 模型易发性评价结果进行解释,得到每个因子的 Shapley 值,如图 7 所示。横轴为 Shapley 值,正值/负值表示因子对预测结果具有正向/负向影响,值越大表示影响越大,每个因子值的大小用颜色表示,红色对应高值,蓝色对应低值。距建筑距离、高程和距道路距离较小时,Shapley 值较大,表示距建筑距离、高程和距道路距离的减小有利于地质灾害的发生。等高线密度、地形起伏度、坡度、植被覆盖度值越大,Shapley 值越大,表明其增大易诱发地质灾害。

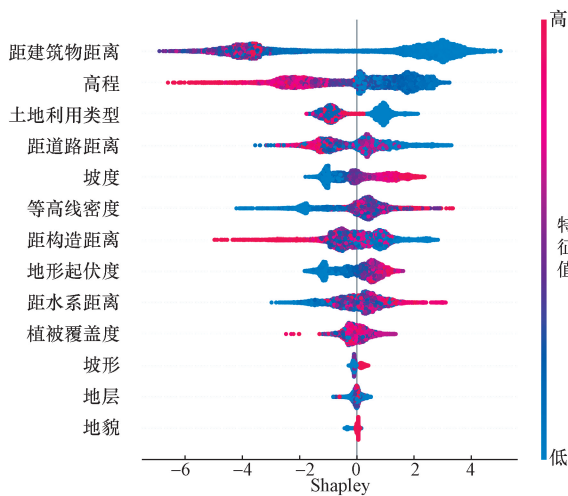


图 7 SHAP 特征分析

Fig. 7 SHAP feature analysis

取每个样本的 Shapley 绝对值的平均值,可得到因子特征重要性排序结果,如图 8 所示。排名前六的因子分别为距建筑物距离、高程、距道路距离、土地利用类型、等高线密度和距构造距离,其余因子对研究区易发性的影响较小。距建筑物距离因子特征重要性排名第一位,进一步表明切坡建房、建设窑洞等人类工程活动是康店镇地质灾害的重要诱发因素。

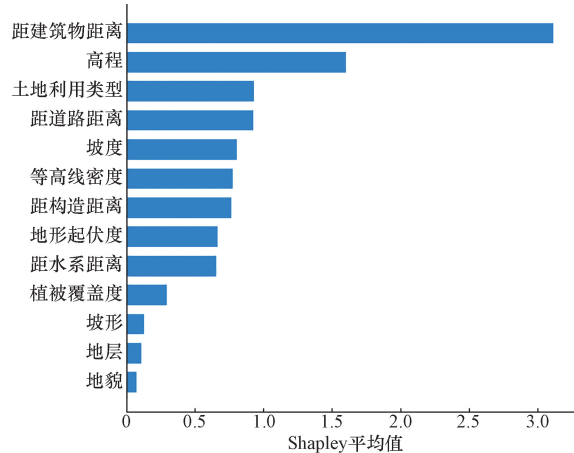


图 8 基于 SHAP 的特征重要性

Fig. 8 Feature importance based on SHAP

3.3.2 特征依赖性分析

特征单依赖性分析描述单个特征对易发性评价结果的影响,对 CatBoost 模型中最重要的 6 个因子进行特征单依赖性分析,如图 9 所示,距建筑物距离 0 ~ 50 m、高程在 104 ~ 154 m、距道路距离 0 ~ 1 000 m、土地利用类型为建设用地(值为 1)、等高线密度 9 ~ 74、距构造距离在 0 ~ 1 333 m 时,Shapley > 0,发生地质灾害的可能性较大。

双依赖图可描述两个因子在交互作用下对评价结果的影响,本文中基于特征重要性排名第一的距建筑物距离因子进行双依赖性分析,如图 10 所示,x 轴为距建筑物距离,y 轴为距建筑物距离的 Shapley 值,垂直着色条带为另一个因子的值,用颜色区分值的大小。距建筑物距离与距道路距离、坡度、等高线密度、距构造距离、地形起伏度、地层、地貌存在明显的交互关系。对于相同的距建筑物距离,距道路距离、距构造距离越小,Shapley 值越大,两个因子的交互作用对地质灾害的贡献越大。当距建筑物距离一定时,地貌为冲洪积倾斜平原,地层为第四系时,Shapley 值越大,更易发生地质灾害。对于相同的距建筑物距离,等高线密度越大、坡度越大、地形起伏度越大,地形越陡峭,地质灾害发生的可能性越高。

基于 SHAP 算法研究特征重要性排序与特征依赖关系,能够从全局解释各个特征对易发性结果的影响机理,有利于制定地质灾害防灾减灾措施。

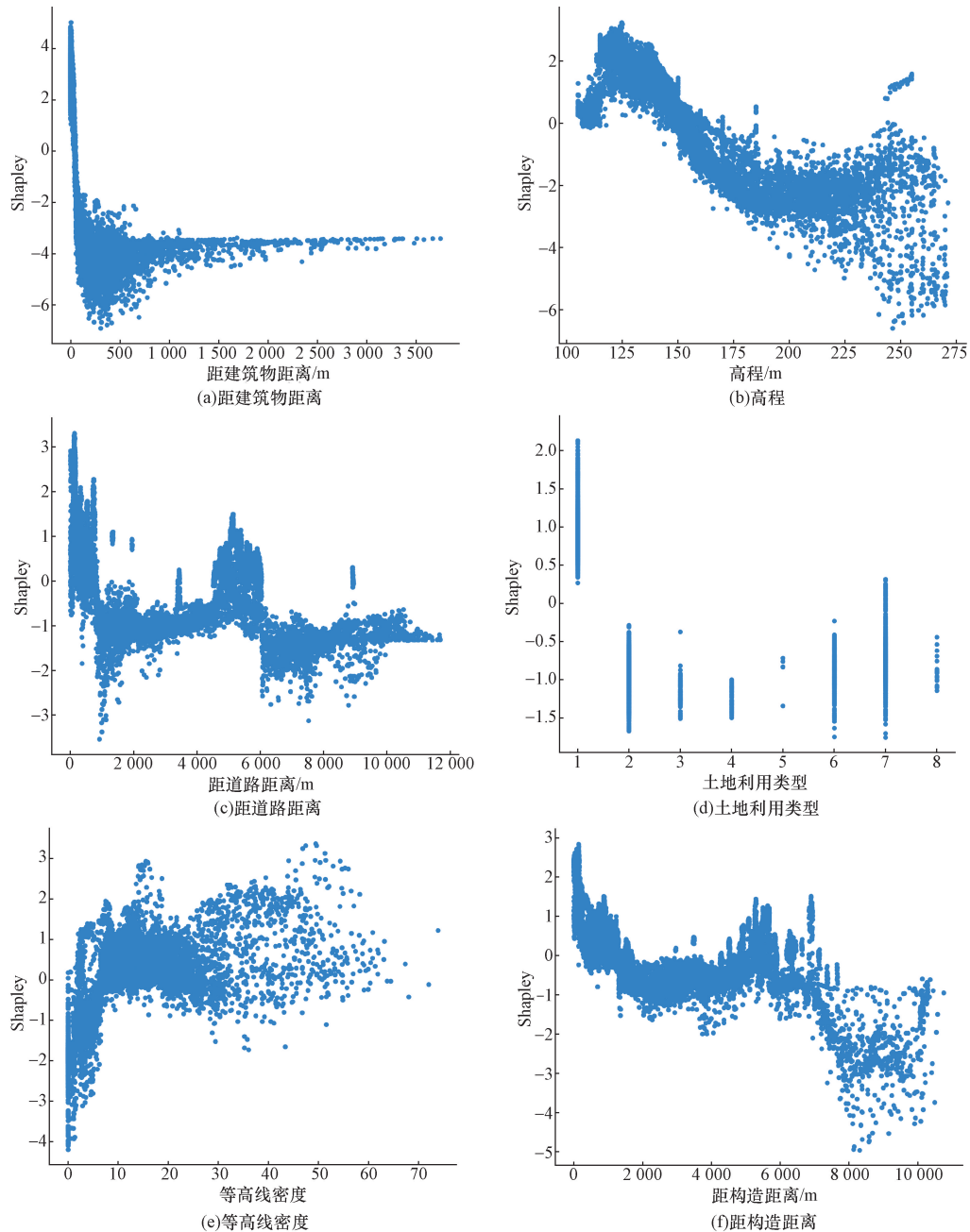


图9 主要影响因子的单依赖性结果

Fig. 9 Single-dependent results of main influencing factors

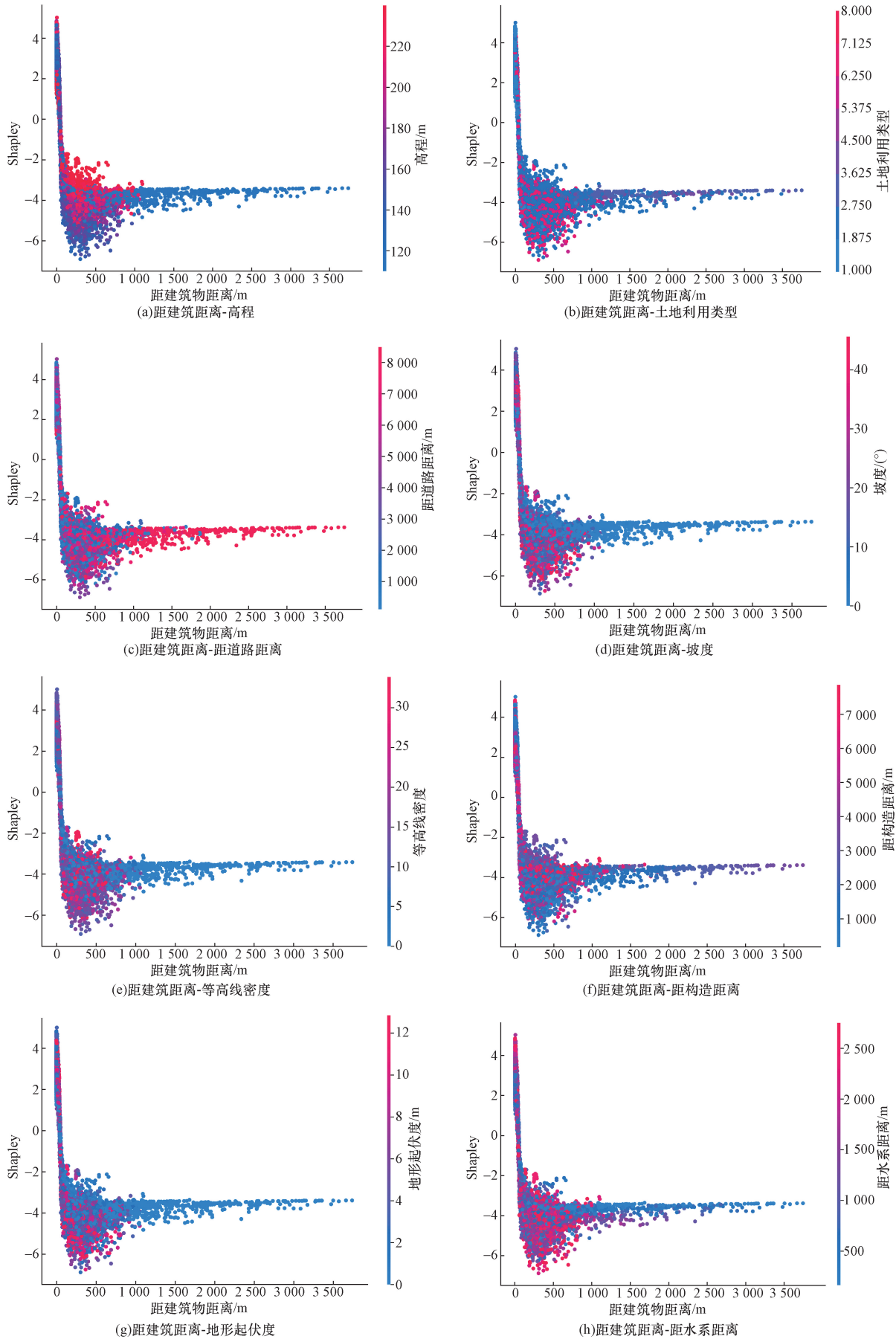
4 结论

(1) 3种算法模型地质灾害易发性评价结果的高和极高易发区具有较高的一致性,主要集中在研究区东部及北部切坡建房等人类活动强烈的冲沟两侧。易发性评价结果与野外调查情况一致,94.12%的地质灾害点存在切坡建房人类工程活动,切坡建房是研究区地质灾害发生的重要诱因。

(2) 与XGBoost和LightGBM模型相比,CatBoost模型在分级结果的合理性、预测的准确性及与野外实际的吻合度等方面均表现最优,其AUC值、准确

率、精确率、召回率和 F_1 分数分别为0.984、0.953、0.957、0.948和0.952,其极高易发区域和高易发区域面积占比分别为3.19%和1.40%,更适合对研究区开展地质灾害易发性评价。

(3) 综合因子分析结果和特征依赖性分析结果可知,黄土丘陵地貌,上更新统(Qp_3^{al})地层,高程在104~154 m,坡度在 $35^\circ \sim 45^\circ$,地形起伏度在10~20 m,距构造距离在0~100 m,NDVI值在0.4~0.6,距道路距离50~100 m,距水系距离500~1000 m,距建筑物距离0~50 m,等高线密度9~20,建设用地和凹型坡地面积大的黄土区最易发生地质灾害。



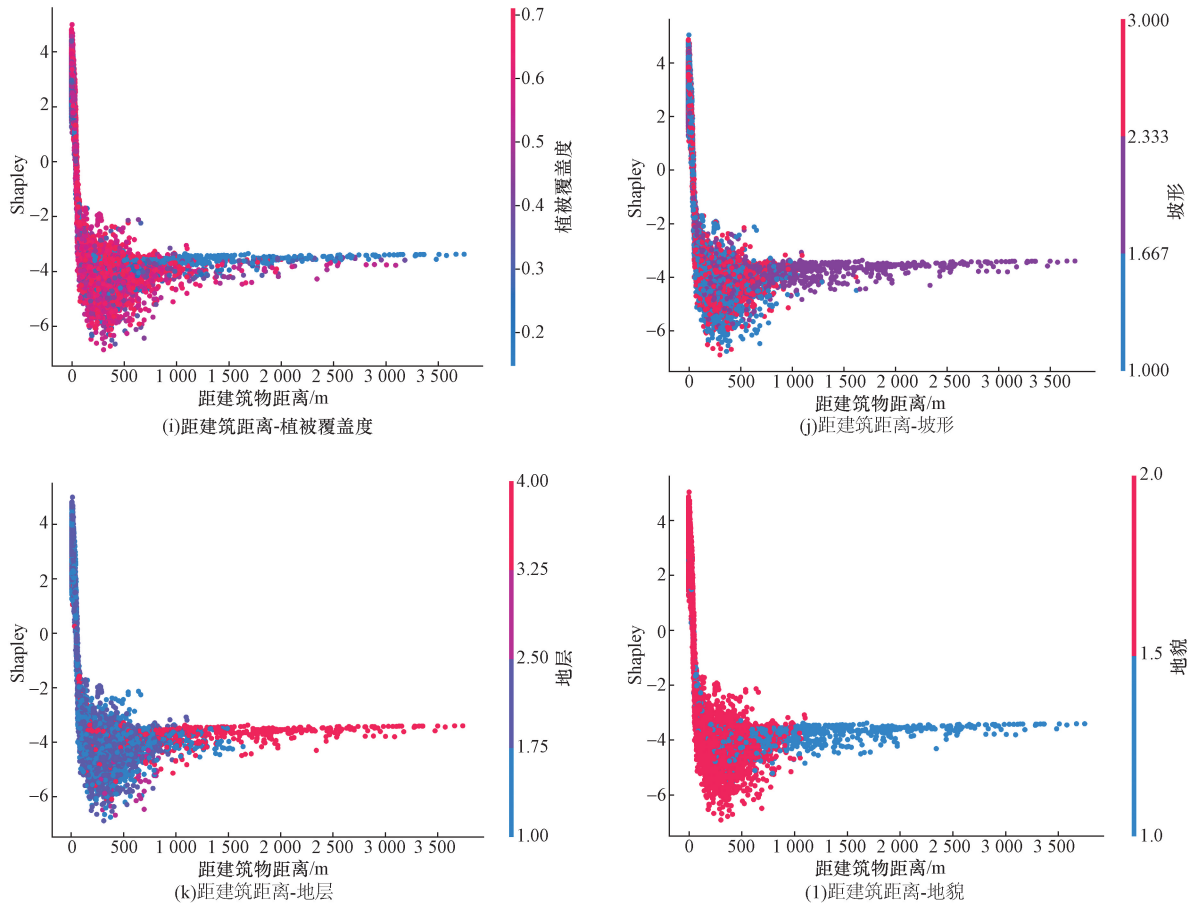


图 10 距建筑物距离与其他影响因子的双依赖性结果

Fig. 10 Double-dependence results of distance from buildings and other main influencing factors

(4) SHAP 特征重要性排名前六的因子分别为距建筑物距离、高程、距道路距离、土地利用类型、等高线密度和距构造距离,是研究区黄土地质灾害的主控因子。

(5) 由特征全局解释可知,当距建筑距离、高程和距道路距离减小,等高线密度、地形起伏度、坡度、植被覆盖度值增大时,Shapley 值越大,越易诱发地质灾害。

(6) 由特征依赖性分析可知,特征重要性排名第一的因子距建筑物距离与距道路距离、坡度、等高线密度、距构造距离、地形起伏度、地层、地貌存在明显的交互关系。当距建筑物距离一定时,存在以下规律:距道路和距构造距离越小,两因子交互作用对地质灾害的贡献越大;冲洪积倾斜平原地貌分布区中的第四系地层更易发生地质灾害;等高线密度越大、坡度越大、地形起伏度越大,地形越陡峭的区域,地质灾害发生的可能性越高。

参考文献

[1] 张茂省, 李同录. 黄土滑坡诱发因素及其形成机理研究[J]. 工程地质学报, 2011, 19(4): 530-540.

Zhang Maosheng, Li Tonglu. Study on the inducing factors and formation mechanism of loess landslides [J]. Journal of Engineering Geology, 2011, 19(4): 530-540.

[2] Wang J, Gu T, Zhang M, et al. Experimental study of loess disintegration characteristics [J]. Earth Surface Processes and Landforms, 2019, 44(6): 1317-1329.

[3] Leng Y, Peng J, Wang S, et al. Development of water sensitivity index of loess from its mechanical properties [J]. Engineering Geology, 2021, 280: 105918.

[4] Ma P, Peng J, Zhuang J, et al. Initiation mechanism of loess mudflows by flume experiments [J]. Journal of Earth Science, 2022, 33(5): 1166-1178.

[5] 马鹏辉, 彭建兵. 论黄土地质灾害链(一)[J]. 自然灾害学报, 2022, 31(2): 1-11.

Ma Penghui, Peng Jianbing. On the loess geological hazard chain (part 1) [J]. Journal of Natural Disasters, 2022, 31(2): 1-11.

[6] 薛强, 张茂省. 延安淹土安滑坡监测预警及变形特征[J]. 西北地质, 2018, 51(2): 220-226.

Xue Qiang, Zhang Maosheng. Monitoring, early warning, and deformation characteristics of Yan'an Yantuan landslide [J]. Northwest Geology, 2018, 51(2): 220-226.

[7] 黄晓维. 中国致命性滑坡和黄土高原滑坡的灾害与风险评价[D]. 兰州: 兰州大学, 2019.

Huang Xiaowei. Disaster and risk assessment of fatal landslides in

- China and loess plateau landslides [D]. Lanzhou: Lanzhou University, 2019.
- [8] Peng J, Wang S, Wang Q, et al. Distribution and genetic types of loess landslides in China [J]. *Journal of Asian Earth Sciences*, 2019, 170: 329-350.
- [9] Malamud B D, Turcotte D L, Guzzetti F, et al. Landslide inventories and their statistical properties [J]. *Earth Surface Processes and Landforms*, 2004, 29(6): 687-711.
- [10] Guzzetti F, Mondini A C, Cardinali M, et al. Landslide inventory maps: new tools for an old problem [J]. *Earth-Science Reviews*, 2012, 112(1): 42-66.
- [11] 黄发明, 殷坤龙, 蒋水华, 等. 基于聚类分析和支持向量机的滑坡易发性评价 [J]. *岩石力学与工程学报*, 2018, 37(1): 156-167.
Huang Faming, Yin Kunlong, Jiang Shuihua, et al. Landslide susceptibility evaluation based on cluster analysis and support vector machine [J]. *Journal of Rock Mechanics and Engineering*, 2018, 37(1): 156-167.
- [12] 谭新平, 徐张建, 王小秋, 等. 厚黄土与煤炭采空区地质灾害易发性评价 [J]. *煤田地质与勘探*, 2017, 45(5): 112-120.
Tan Xinping, Xu Zhangjian, Wang Xiaoqiu, et al. Evaluation of the susceptibility of geological hazards in thick loess and coal goaf [J]. *Coalfield Geology and Exploration*, 2017, 45(5): 112-120.
- [13] 蔡怀恩, 张继文, 秦广平, 等. 延安黄土丘陵沟壑区地质灾害易发性浅析 [J]. *地震工程学报*, 2015, 37(S1): 68-73.
Cai Huaen, Zhang Jiwen, Qin Guangping, et al. A preliminary analysis of geological hazard susceptibility in the loess hilly and gully region of Yan'an [J]. *Journal of Earthquake Engineering*, 2015, 37(S1): 68-73.
- [14] 李会. 黄土高原区地质灾害易发性研究——以临夏地区为例 [D]. 呼和浩特: 内蒙古农业大学, 2016.
Li Hui. Study on the susceptibility of geological hazards in the Loess Plateau region; a case study of Linxia area [D]. Hohhot: Inner Mongolia Agricultural University, 2016.
- [15] 齐娜, 胡良柏, 赵喜娟. 基于 CF 指数的天水盆地黄土滑坡灾害易发性评价 [J]. *冶金管理*, 2022(13): 85-87.
Qi Na, Hu Liangbai, Zhao Xijuan. Vulnerability assessment of loess landslide disasters in Tianshui Basin based on CF index [J]. *Metallurgical Management*, 2022(13): 85-87.
- [16] 孟晓捷, 张新社, 曾庆铭, 等. 基于加权信息量法的黄土滑坡易发性评价——以 1: 5 万天水市麦积区为例 [J]. *西北地质*, 2022, 55(2): 249-259.
Meng Xiaojie, Zhang Xinshe, Zeng Qingming, et al. Evaluation of susceptibility to loess landslides based on weighted information method; a case study of 1:50 000 Maijipu in Tianshui City [J]. *Northwest Geology*, 2022, 55(2): 249-259.
- [17] 徐敏, 崔玉龙, 袁文华, 等. 基于 GIS 和信息量法的伊犁黄土滑坡易发性评价 [J]. *河南城建学院学报*, 2021, 30(4): 25-31.
Xu Min, Cui Yulong, Yuan Wenhua, et al. Vulnerability assessment of Yili loess landslide based on GIS and information quantity method [J]. *Journal of Henan Urban Construction University*, 2021, 30(4): 25-31.
- [18] 王正刚. 基于 GIS 的黄土滑坡易发性定量评价模型 [J]. *科技与企业*, 2013(21): 187, 190.
Wang Zhenggang. Quantitative evaluation model for susceptibility of loess landslides based on GIS [J]. *Science and Enterprise*, 2013(21): 187, 190.
- [19] 陈思名, 霍艾迪, 张佳, 等. 基于 InSAR 的宁夏黄土丘陵区(西吉县)滑坡隐患早期识别 [J]. *科学技术与工程*, 2022, 22(12): 4721-4728.
Chen Siming, Huo Aidi, Zhang Jia, et al. Early identification of landslide hazards in the loess hilly area of Ningxia (Xiji County) based on InSAR [J]. *Science Technology and Engineering*, 2022, 22(12): 4721-4728.
- [20] 尹玉玲, 徐素宁, 王军, 等. 典型黄土丘陵区地质灾害隐患识别与时序监测 [J]. *水文地质工程地质*, 2023, 50(2): 141-149.
Yin Yuling, Xu Suning, Wang Jun, et al. Identification and temporal monitoring of geological hazards in typical loess hilly areas [J]. *Hydrogeology and Engineering Geology*, 2023, 50(2): 141-149.
- [21] 张威, 胡舫瑞, 綦巍, 等. 基于 XGBoost 和云模型的地质灾害易发性评价 [J]. *中国地质灾害与防治学报*, 2023, 34(6): 136-145.
Zhang Wei, Hu Fangrui, Qi Wei, et al. Geological hazard susceptibility evaluation based on XGBoost and cloud model [J]. *Chinese Journal of Geological Hazards and Prevention*, 2023, 34(6): 136-145.
- [22] Huang F, Zhang J, Zhou C, et al. A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction [J]. *Landslides*, 2020, 17(1): 217-229.
- [23] Segoni S, Pappafico G, Luti T, et al. Landslide susceptibility assessment in complex geological settings: sensitivity to geological information and insights on its parameterization [J]. *Landslides*, 2020, 17(10): 2443-2453.
- [24] 贺鹏, 童立强, 郭兆成, 等. GIS 支持下基于层次分析法的西藏札达地区滑坡灾害易发性评价研究 [J]. *科学技术与工程*, 2016, 16(25): 193-200.
He Peng, Tong Liqiang, Guo Zhaocheng, et al. GIS based analytic hierarchy process (AHP) based landslide hazard vulnerability assessment in Zhada, Xizang [J]. *Science Technology and Engineering*, 2016, 16(25): 193-200.
- [25] 李文彦, 王喜乐. 频率比与信息量模型在黄土沟壑区滑坡易发性评价中的应用与比较 [J]. *自然灾害学报*, 2020, 29(4): 213-220.
Li Wenyan, Wang Xile. Application and comparison of frequency ratio and information model in landslide susceptibility evaluation in loess gully region [J]. *Journal of Natural Disasters*, 2020, 29(4): 213-220.
- [26] 王佳佳, 殷坤龙, 肖莉丽. 基于 GIS 和信息量的滑坡灾害易发性评价——以三峡库区万州区为例 [J]. *岩石力学与工程学报*, 2014, 33(4): 797-808.

- Wang Jiajia, Yin Kunlong, Xiao Lili. Evaluation of landslide susceptibility based on GIS and information quantity: a case study of Wanzhou District in the Three Gorges Reservoir Area[J]. *Journal of Rock Mechanics and Engineering*, 2014, 33(4): 797-808.
- [27] 许冲, 戴福初, 姚鑫, 等. 基于GIS与确定性系数分析方法的汶川地震滑坡易发性评价[J]. *工程地质学报*, 2010, 18(1): 15-26.
- Xu Chong, Dai Fuchu, Yao Xin, et al. Evaluation of landslide susceptibility in Wenchuan earthquake based on GIS and deterministic coefficient analysis method [J]. *Journal of Engineering Geology*, 2010, 18(1): 15-26.
- [28] 庄建琦, 崔鹏, 葛永刚, 等. “5·12”汶川地震崩塌滑坡危险性评价——以都汶公路沿线为例[J]. *岩石力学与工程学报*, 2010, 29(S2): 3735-3742.
- Zhuang Jianqi, Cui Peng, Ge Yonggang, et al. Risk assessment of landslides and collapses during the Wenchuan earthquake on May 12th: a case study along the Duwen Highway[J]. *Journal of Rock Mechanics and Engineering*, 2010, 29(S2): 3735-3742.
- [29] 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.
- [30] 王世宝, 庄建琦, 郑佳, 等. 基于深度学习的CZ铁路康定—理塘段滑坡易发性评价[J]. *工程地质学报*, 2022, 30(3): 908-919.
- Wang Shibao, Zhuang Jianqi, Zheng Jia, et al. Landslide susceptibility evaluation of the Kangding Litang section of the CZ railway based on deep learning [J]. *Journal of Engineering Geology*, 2022, 30(3): 908-919.
- [31] 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. Modeling and interpretability of landslide susceptibility prediction based on self screening deep learning [J]. *Earth Science*, 2023, 48(5): 1696-1710.
- [32] 田春山, 刘希林, 汪佳. 基于CF和Logistic回归模型的广东省地质灾害易发性评价[J]. *水文地质工程地质*, 2016, 43(6): 154-161, 170.
- Tian Chunshan, Liu Xilin, Wang Jia. Evaluation of geological hazard susceptibility in Guangdong Province based on CF and logistic regression models [J]. *Hydrogeology and Engineering Geology*, 2016, 43(6): 154-161, 170.
- [33] 胡涛, 樊鑫, 王硕, 等. 基于径向基神经网络的思南县崩塌易发性评价[J]. *科学技术与工程*, 2019, 19(35): 61-69.
- Hu Tao, Fan Xin, Wang Shuo, et al. Evaluation of landslide susceptibility in Sinan County based on radial basis function neural network [J]. *Science Technology and Engineering*, 2019, 19(35): 61-69.
- [34] 邓念东, 石辉, 文强, 等. 信息量支持下的随机森林模型的崩塌易发性评价[J]. *科学技术与工程*, 2021, 21(6): 2210-2217.
- Deng Niandong, Shi Hui, Wen Qiang, et al. Evaluation of landslide susceptibility in a random forest model supported by information [J]. *Science Technology and Engineering*, 2021, 21(6): 2210-2217.
- [35] 陈芯宇, 师芸, 温永啸, 等. 基于确定性系数与支持向量机的滑坡易发性评价[J]. *科学技术与工程*, 2023, 23(2): 518-527.
- Chen Xinyu, Shi Yun, Wen Yongxiao, et al. Landslide susceptibility evaluation based on deterministic coefficients and support vector machines [J]. *Science Technology and Engineering*, 2023, 23(2): 518-527.
- [36] 杨永刚, 殷坤龙, 赵海燕, 等. 基于C5.0决策树-快速聚类模型的万州区库岸段乡镇滑坡易发性区划[J]. *地质科技情报*, 2019, 38(6): 189-197.
- Yang Yonggang, Yin Kunlong, Zhao Haiyan, et al. Landslide susceptibility zoning of towns and villages in the Ku'an section of Wanzhou District based on C5.0 decision tree fast clustering model [J]. *Geological Science and Technology Information*, 2019, 38(6): 189-197.
- [37] 周超, 殷坤龙, 曹颖, 等. 基于集成学习与径向基神经网络耦合模型的三峡库区滑坡易发性评价[J]. *地球科学*, 2020, 45(6): 1865-1876.
- Zhou Chao, Yin Kunlong, Cao Ying, et al. Evaluation of landslide susceptibility in the Three Gorges Reservoir area based on ensemble learning and radial basis function neural network coupling model [J]. *Earth Science*, 2020, 45(6): 1865-1876.
- [38] Pradhan B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS[J]. *Computers & Geosciences*, 2013, 51(2): 350-365.
- [39] 黄发明, 李金凤, 王俊宇, 等. 考虑线状环境因子适宜性和不同机器学习模型的滑坡易发性预测建模规律[J]. *地质科技通报*, 2022, 41(2): 44-59.
- Huang Faming, Li Jinfeng, Wang Junyu, et al. Modeling law for landslide susceptibility prediction considering the suitability of linear environmental factors and different machine learning models [J]. *Geological Science and Technology Bulletin*, 2022, 41(2): 44-59.
- [40] 林琴, 郭永刚, 吴升杰, 等. 基于梯度提升的优化集成机器学习算法对滑坡易发性评价: 以雅鲁藏布江与尼洋河两岸为例[J]. *西北地质*, 2024, 57(1): 12-22.
- Lin Qin, Guo Yonggang, Wu Shengjie, et al. Evaluation of landslide susceptibility based on gradient lifting optimized integrated machine learning algorithm: taking both banks of the Yarlung Zangbo River and Niyang River as examples [J]. *Geology of Northwest China*, 2024, 57(1): 12-22.
- [41] 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(2): 259-272.
- Wang Gaofeng, Guo Ning, Deng Bing, et al. Analysis of landslide susceptibility and accuracy in different combination model regions [J]. *Northwest Geology*, 2021, 54(2): 259-272.
- [42] 田述军, 张珊珊, 唐青松, 等. 基于不同评价单元的滑坡易发性评价对比研究[J]. *自然灾害学报*, 2019, 28(6): 137-145.

- Tian Shujun, Zhang Shanshan, Tang Qingsong, et al. Comparative study on landslide susceptibility evaluation based on different evaluation units [J]. *Journal of Natural Disasters*, 2019, 28(6): 137-145.
- [43] 岳韦霆, 任超, 梁月吉, 等. 基于可解释机器学习模型的南宁市野火灾害易发性研究[J]. *科学技术与工程*, 2024, 24(2): 858-870.
- Yue Weiting, Ren Chao, Liang Yueji, et al. Wildfire hazard susceptibility in nanning based on interpretable machine learning model[J]. *Science Technology and Engineering*, 2024, 24(2): 858-870.
- [44] 胡杨, 张紫昭, 林世河. 基于证据权与逻辑回归耦合的新疆伊犁河谷地区滑坡易发性评价[J]. *工程地质学报*, 2023, 31(4): 1350-1363.
- Hu Yang, Zhang Zizhao, Lin Shihe. Evaluation of landslide susceptibility in the Ili River Valley of Xinjiang based on evidence weight and logistic regression coupling [J]. *Journal of Engineering Geology*, 2023, 31(4): 1350-1363.
- [45] Wu Y, Ke Y, Chen Z, et al. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping[J]. *Catena*, 2020, 187: 104396.
- [46] Samia J, Temme A, Bregt A, et al. Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory[J]. *Landslides*, 2017, 14(2): 547-558.
- [47] 陈绪钰, 倪化勇, 李明辉, 等. 基于加权信息量和迭代自组织聚类的地质灾害易发性评价[J]. *灾害学*, 2021, 36(2): 71-78.
- Chen Xuyu, Ni Huayong, Li Minghui, et al. Geological hazard susceptibility evaluation based on weighted information and iterative self-organizing clustering [J]. *Journal of Hazards*, 2021, 36(2): 71-78.
- [48] 伍剑波, 孙强, 张泰丽, 等. 地形起伏度与滑坡发育的相关性——以丽水市滑坡为例[J]. *华东地质*, 2022, 43(2): 235-244.
- Wu Jianbo, Sun Qiang, Zhang Taili, et al. Research for the correlation between relief amplitude and landslides: a case study of Lishui City[J]. *East China Geology*, 2022(2): 235-244.
- [49] 范强, 巨能攀, 向喜琼, 等. 证据权法在滑坡易发性分区中的应用——以贵州桐梓河流域为例[J]. *灾害学*, 2015, 30(1): 124-129.
- Fan Qiang, Ju Nengpan, Xiang Xiqiong, et al. Application of evidence weight method in landslide susceptibility zoning: a case study of Tongzi River Basin in Guizhou Province [J]. *Journal of Disaster Science*, 2015, 30(1): 124-129.
- [50] 许强, 李为乐, 董秀军, 等. 四川茂县叠溪镇新磨村滑坡特征与成因机制初步研究[J]. *岩石力学与工程学报*, 2017, 36(11): 2612-2628.
- Xu Qiang, Li Weile, Dong Xiujun, et al. Preliminary study on the characteristics and genesis mechanism of the Xinmo Village landslide in Diexi Town, Mao County, Sichuan Province [J]. *Journal of Rock Mechanics and Engineering*, 2017, 36(11): 2612-2628.
- [51] 许强, 董秀军, 李为乐, 等. 基于天-空-地一体化的重大地质灾害隐患早期识别与监测预警[J]. *武汉大学学报(信息科学版)*, 2019, 44(7): 957-966.
- Xu Qiang, Dong Xiujun, Li Weile, et al. Early identification, monitoring, and early warning of major geological hazards based on the integration of sky, air, and earth[J]. *Journal of Wuhan University (Information Science Edition)*, 2019, 44(7): 957-966.
- [52] 张岩岩, 文海家, 麻超超, 等. 基于多源数据的蔡家坝特大型滑坡成因机制研究及稳定性评价[J]. *岩石力学与工程学报*, 2018, 37(9): 2048-2063.
- Zhang Yanyan, Wen Haijia, Ma Chaochao, et al. Study on the genesis mechanism and stability evaluation of the Caijiaba super large landslide based on multi-source data [J]. *Journal of Rock Mechanics and Engineering*, 2018, 37(9): 2048-2063.
- [53] 张泽方, 钱志宽, 魏勇, 等. 考虑最优影响因素组合的滑坡易发性评价——以水城区为例[J]. *科学技术与工程*, 2023, 23(10): 4091-4099.
- Zhang Zefang, Qian Zhikuan, Wei Yong, et al. Landslide susceptibility evaluation considering optimal combination of influencing factors: a case study of Shuicheng District[J]. *Science Technology and Engineering*, 2023, 23(10): 4091-4099.
- [54] 李辉, 翟星, 李琛曦, 等. 河北省泥石流灾害易发性云模型评价方法: 以邢台赵沟村泥石流为例[J]. *科学技术与工程*, 2024, 24(25): 10884-10891.
- Li Hui, Zhai Xing, Li Chenxi, et al. Development law of debris flow disaster and the evaluation method of its susceptibility: take mudslide in Zhaogou Village, Xingtai as an example[J]. *Science Technology and Engineering*, 2024, 24(25): 10884-10891.
- [55] 吴问楚, 王朝, 管后春, 等. 基于 GIS 与加权信息量模型的汤口断裂南段地区地质灾害易发性评价[J]. *科学技术与工程*, 2024, 24(26): 11121-11130.
- Wu Wenchu, Wang Chao, Guan Houchun, et al. Geological hazard vulnerability assessment in southern part of Tangkou fault based on GIS and weighted information value model [J]. *Science Technology and Engineering*, 2024, 24(26): 11121-11130.
- [56] 王向前, 丁明涛, 何雨枫, 等. 基于 RF 模型的甘孜藏族自治州乡城县泥石流易发性评价——以鲁东片区为例[J]. *灾害学*, 2023, 38(3): 222-234.
- Wang Xiangqian, Ding Mingtao, He Yufeng, et al. Assessment of debris flow susceptibility in Xiangcheng County, Garze Tibetan Autonomous Prefecture based on RF model: a case study of Ludong area [J]. *Journal of Disaster Studies*, 2023, 38(3): 222-234.
- [57] 杨迎冬, 晏祥省, 王宇, 等. 云南省地质灾害特征及形成规律研究[J]. *灾害学*, 2021, 36(3): 131-139.
- Yang Yingdong, Yan Xiangsheng, Wang Yu, et al. Research on the characteristics and formation laws of geological hazards in Yunnan Province [J]. *Journal of Disaster Science*, 2021, 36(3): 131-139.
- [58] Wu Y, Ke Y, Chen Z, et al. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping[J]. *Catena*, 2020, 187: 104396.
- [59] Samia J, Temme A, Bregt A, et al. Do landslides follow landslides? Insights in path dependency from a multi-temporal land-

- slide inventory[J]. *Landslides*, 2017, 14(2): 547-558.
- [60] 张岩岩, 文海家, 麻超超, 等. 基于多源数据的蔡家坝特大型滑坡成因机制研究及稳定性评价[J]. *岩石力学与工程学报*, 2018, 37(9): 2048-2063.
- Zhang Yanyan, Wen Haijia, Ma Chaochao, et al. Study on the genesis mechanism and stability evaluation of the Caijiaba super large landslide based on multi-source data [J]. *Journal of Rock Mechanics and Engineering*, 2018, 37(9): 2048-2063.
- [61] 吴树仁, 石菊松, 张春山, 等. 滑坡风险评估理论与技术[M]. 北京: 科学出版社, 2012.
- Wu Shuren, Shi Jusong, Zhang Chunshan, et al. *Landslide risk assessment theory and technology* [M]. Beijing: Science Press, 2012.
- [62] 金福喜, 罗滔, 李杰. 湘东南某市地形地貌对地质灾害发育规律的控制作用分析[J]. *中国地质灾害与防治学报*, 2014, 25(2): 115-118.
- Jin Fuxi, Luo Tao, Li Jie. Analysis of the control effect of topography and geomorphology on the development of geological hazards in a certain city in Southeastern Hunan [J]. *Chinese Journal of Geological Hazards and Prevention*, 2014, 25(2): 115-118.
- [63] 豆红强, 简文彬, 王浩, 等. 高植被覆盖区台风暴雨型滑坡成灾机制及预警模型研究综述[J]. *自然灾害学报*, 2023, 32(2): 1-15.
- Dou Hongqiang, Jian Wenbin, Wang Hao, et al. Summary of research on disaster mechanism and early warning model of typhoon rain landslide in high vegetation coverage area [J]. *Journal of Natural Disasters*, 2023, 32(2): 1-15.
- [64] 刘亚静, 刘红健. 基于信息量-随机森林模型的地震带地质灾害易发性评价: 以松潘-较场地震带为例[J]. *科学技术与工程*, 2024, 24(1): 143-154.
- Liu Yajing, Liu Hongjian. Evaluation of landslide susceptibility in Weixin County based on evidence weight and support vector machine model[J]. *Science Technology and Engineering*, 2024, 24(1): 143-154.
- [65] Feng L, Lin H, Zhang M, et al. Development and evolution of loess vertical joints on the Chinese Loess Plateau at different spatiotemporal scales[J]. *Engineering Geology*, 2020, 265: 105372.
- [66] 刘瑞, 罗书斌, 徐源, 等. 基于地理信息系统与斜坡单元的峨边县地质灾害易发性评价[J]. *科学技术与工程*, 2023, 23(18): 7678-7685.
- Liu Rui, Luo Shubin, Xu Yuan, et al. Geological hazard vulnerability assessment in Ebian County based on geographic information system and slope unit [J]. *Science Technology and Engineering*, 2023, 23(18): 7678-7685.
- [67] 李光辉, 铁永波. 基于信息量模型的综合地质灾害易发性建模方法对比研究[J]. *灾害学*, 2023, 38(3): 212-221.
- Li Guanghui, Tie Yongbo. Comparative study on comprehensive geological hazard susceptibility modeling methods based on information quantity model [J]. *Journal of Disaster Science*, 2023, 38(3): 212-221.
- [68] 孟凡奇, 高峰, 林波, 等. 基于 AHP 和信息量模型的地质灾害易发性评价——以鲁东片区为例[J]. *灾害学*, 2023, 38(3): 111-117.
- Meng Fanqi, Gao Feng, Lin Bo, et al. Evaluation of geological hazard susceptibility based on AHP and information model: a case study of Ludong Area [J]. *Disaster Science*, 2023, 38(3): 111-117.
- [69] 穆柯, 谢婉丽, 刘琦琦, 等. 基于 LR-RF 模型的滑坡易发性评价——以铜川市耀州区为例[J]. *灾害学*, 2022, 37(3): 212-218.
- Mu Ke, Xie Wanli, Liu Qiqi, et al. Landslide susceptibility evaluation based on LR-RF model: a case study of Yaozhou District, Tongchuan City [J]. *Journal of Disaster Science*, 2022, 37(3): 212-218.
- [70] 寸得欣, 令狐昌卫, 马一奇, 等. 基于 GIS 和加权信息量模型的富源县地质灾害易发性评价[J]. *科学技术与工程*, 2024, 24(18): 7563-7573.
- Cun Dexin, Linghu Changwei, Ma Yiqi, et al. Evaluation of susceptibility to geological hazard in Fuyuan County based on GIS and weighted information model [J]. *Science Technology and Engineering*, 2024, 24(18): 7563-7573.
- [71] 杜国梁, 张永双, 吕文明, 等. 基于加权信息量模型的藏东南地区滑坡易发性评价[J]. *灾害学*, 2016, 31(2): 226-234.
- Du Guoliang, Zhang Yongshuang, Lü Wenming, et al. Evaluation of landslide susceptibility in southeastern Tibet based on weighted information model [J]. *Journal of Disaster Science*, 2016, 31(2): 226-234.
- [72] 胡杨, 张紫昭, 林世河. 基于证据权与逻辑回归耦合的新疆伊犁河谷地区滑坡易发性评价[J]. *工程地质学报*, 2023, 31(4): 1350-1363.
- Hu Yang, Zhang Zizhao, Lin Shihe. Evaluation of landslide susceptibility in the Ili River Valley of Xinjiang based on evidence weight and logistic regression coupling [J]. *Journal of Engineering Geology*, 2023, 31(4): 1350-1363.
- [73] Wang Y, Duan H, Hong H. A comparative study of composite kernels for landslide susceptibility mapping: a case study in Yongxin County, China [J]. *Catena*, 2019, 183: 104217.
- [74] Wang X, Zhang L, Wang S, et al. Regional landslide susceptibility zoning with considering the aggregation of landslide points and the weights of factors [J]. *Landslides*, 2014, 11(3): 399-409.
- [75] 何万才, 赵俊三, 林伊琳, 等. 基于证据权和支持向量机模型的威信县滑坡易发性评价[J]. *科学技术与工程*, 2023, 23(15): 6350-6360.
- He Wancai, Zhao Junsan, Lin Yilin, et al. Evaluation of landslide susceptibility in Weixin County based on evidence weight and support vector machine model [J]. *Science Technology and Engineering*, 2023, 23(15): 6350-6360.
- [76] 朱钰, 郑屹然, 尹默. 统计学意义下的多重共线性检验方法[J]. *统计与决策*, 2020, 36(7): 34-36.
- Zhu Yu, Zheng Yiran, Yin Mo. Multi collinearity test method under statistical significance [J]. *Statistics and Decision*, 2020, 36(7): 34-36.
- [77] 贺攀, 郭荣昌, 张蕊, 等. 基于不同评价因子组合的铁路沿线滑坡危险性评价[J]. *兰州交通大学学报*, 2022, 41(5): 34-41.

- He Pan, Guo Rongchang, Zhang Rui, et al. Risk assessment of landslides along railway lines based on different combinations of evaluation factors [J]. *Journal of Lanzhou Jiaotong University*, 2022, 41(5): 34-41.
- [78] 周鑫. 金沙江上游茂顶河段滑坡成因机制及敏感性研究[D]. 长春: 吉林大学, 2019.
- Zhou Xin. Study on the formation mechanism and sensitivity of landslides in the Maoding River Section of the Upper Jinsha River [D]. Changchun: Jilin University, 2019.
- [79] 王文辉. 联合 SBAS-InSAR 与机器学习的滑坡隐患识别——以兰州市为例[D]. 兰州: 兰州交通大学, 2021.
- Wang Wenhui. Landslide hazard identification using SBAS InSAR and machine learning: a case study of Lanzhou City [D]. Lanzhou: Lanzhou Jiaotong University, 2021.
- [80] 薛强, 张茂省, 李林, 等. 基于斜坡单元与信息量法结合的宝塔区黄土滑坡易发性评价[J]. *地质通报*, 2015, 34(11): 2108-2115.
- Xue Qiang, Zhang Maosheng, Li Lin, et al. Evaluation of the susceptibility of loess landslides in Baota District based on the combination of slope units and information method [J]. *Geological Bulletin*, 2015, 34(11): 2108-2115.
- [81] 仇文岗, 何昱苇, 王鲁琦, 等. 基于水系分区的滑坡易发性机器学习分析方法: 以重庆市奉节县为例[J]. *地球科学*, 2023, 48(5): 2024-2038.
- Zhang Wengang, He Yuwei, Wang Luqi, et al. Machine learning analysis method for landslide susceptibility based on water system zoning: a case study of Fengjie County, Chongqing [J]. *Earth Science*, 2023, 48(5): 2024-2038.
- [82] 李成林, 刘严松, 赖思翰, 等. 基于 BP 神经网络模型的滑坡易发性评价方法[J]. *科学技术与工程*, 2023, 23(13): 5481-5492.
- Li Chenglin, Liu Yansong, Lai Sihan, et al. Landslide susceptibility evaluation methods based on BP neural network model[J]. *Science Technology and Engineering*, 2023, 23(13): 5481-5492.
- [83] 韩玲, 张庭瑜, 张恒. 基于 IOE 和 SVM 模型的府谷镇滑坡易发性分区[J]. *水土保持研究*, 2019, 26(3): 367-372.
- Han Ling, Zhang Tingyu, Zhang Heng. Landslide susceptibility zoning in Fugu Town based on IOE and SVM models [J]. *Soil and Water Conservation Research*, 2019, 26(3): 367-372.
- [84] 史进昌. 滑坡点采样模式对易发性评价结果的对比分析[J]. *灾害学*, 2023, 38(4): 50-55.
- Shi Jinchang. Comparative analysis of landslide sampling modes on susceptibility evaluation results [J]. *Journal of Disaster Science*, 2023, 38(4): 50-55.