**引文格式**:秦佳松,李为乐,周胜森,等.考虑InSAR地表形变的滑坡易发性评价负样本选取与结果优化[J].武汉大学学报(信息科学版),2026,51(6):1228-1241.DOI:10.13203/j.whugis20240494

**Citation**: QIN Jiasong, LI Weile, ZHOU Shengsen, et al. Negative Sample Selection and Result Optimization for Landslide Susceptibility Assessment Considering InSAR Surface Deformation[J]. Geomatics and Information Science of Wuhan University, 2026, 51(6):1228-1241.DOI:10.13203/j.whugis20240494

# 考虑InSAR地表形变的滑坡易发性评价负样本选取与结果优化

秦佳松[1] 李为乐[1,2] 周胜森[1] 单云锋[1] 何国庆[1] 范禄震[1]

1 成都理工大学地质灾害防治与地质环境保护全国重点实验室,四川 成都,610059

2 应急管理部滑坡灾害风险预警与防控实验室,四川 成都,610059

**摘 要**:滑坡易发性结果是滑坡灾害防控工作的重要参考,准确建模对灾害预警和风险管控至关重要。然而,在滑坡易发性建模过程中,非滑坡样本的选取存在较强的随机性,且未充分考虑滑坡负样本可能来源于潜在滑坡区域,导致评估结果的准确性受限。因此,提出了一种基于合成孔径雷达干涉测量(interferometric synthetic aperture radar, InSAR)的地表形变信息滑坡负样本采样策略,以四川省汉源县为例,采用小基线集InSAR技术生成年平均地表形变速率,并从形变速率极低的区域中生成滑坡负样本。结果表明,在极端梯度提升和随机森林(random forest,RF)模型中,基于InSAR的采样方法均显著提升了模型的预测性能。其中,InSAR-RF模型的预测精度最高,受试者工作特征曲线下面积为0.847,准确率为89.8%,相比于传统的缓冲区采样方法,分别提高了5.5%和5.6%。此外,通过将InSAR数据引入模型优化滑坡易发性制图结果,并利用沙普利加性解释法解释模型内部的决策机制,进一步增强了滑坡易发性评价的科学性和可靠性。所提方法为InSAR技术在机器学习预测滑坡易发性中的应用提供了新思路。

**关键词**:滑坡易发性;负样本;InSAR;极端梯度提升;随机森林;沙普利加性解释法

**中图分类号**:P237 **文献标识码**:A **收稿日期**:2025-05-19

**DOI**:10.13203/j.whugis20240494 **文章编号**:1671-8860(2026)06-1228-14

# Negative Sample Selection and Result Optimization for Landslide Susceptibility Assessment Considering InSAR Surface Deformation

*QIN Jiasong*[1] *LI Weile*[1,2] *ZHOU Shengsen*[1] *SHAN Yunfeng*[1] *HE Guoqing*[1] *FAN Luzhen*[1]

1 State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China

2 Laboratory of Landslide Risk Early-Warning and Control, Ministry of Emergency Management, Chengdu 610059, China

**Abstract**: **Objectives**: Landslide is one of the most frequent and destructive geological hazards in the world, posing significant threats to human lives, property, and natural environments. Accurate landslide susceptibility modeling is critical for disaster early warning and risk management. However, a major challenge in this process is the selection of non-landslide samples, which is often arbitrary and overlooks the fact that these samples might originate from potentially unstable areas, leading to reduced model accuracy. **Methods**: To address this, a novel negative sample selection strategy based on interferometric synthetic aperture radar (InSAR)-derived ground deformation data is proposed. Taking Hanyuan County, Sichuan Province as the study area, annual average ground deformation rates are generated using the small baseline subset InSAR technique. Negative samples are extracted from the regions which exhibit extremely low deformation rates, representing areas unlikely to experience landslides. **Results**: Results demonstrate that the InSAR-based

**基金项目**:四川省重点研发项目(2023YFS0435);自然资源部-四川省合作项目(SCDZRS-2023);地质灾害防治与地质环境保护国家重点实验室自主研究课题(SKLGP2022Z007)。

**第一作者**:秦佳松,硕士生,主要从事地质灾害预测评价研究。2663288045@qq.com

**通信作者**:李为乐,博士,教授。liweile08@mail.cdut.edu.cn

sampling method significantly improves predictive performance in both extreme gradient boosting (XGBoost) and random forest (RF) models. The InSAR-RF model achieves the highest accuracy, with the area under the curve of 0.847 and the overall accuracy of 89.8%, representing 5.5% and 5.6% improvement, respectively, over traditional buffer-based sampling approaches. Additionally, to overcome the lack of dynamic features in conventional landslide susceptibility assessment, InSAR data are integrated with machine learning models, and an evaluation matrix is constructed to refine susceptibility mapping. The decision mechanisms within the models are further analyzed using shapley additive explanations to enhance interpretability. **Conclusions**: The InSAR-based negative sampling strategy not only improves model accuracy, but also strengthens the scientific reliability of landslide susceptibility assessments by integrating real-world deformation data. It offers a new approach for landslide hazard early warning and risk management.

**Key words**: landslide susceptibility; negative sample; InSAR; extreme gradient boosting; random forest; Shapley additive explanations algorithm

滑坡是全球范围内最常见且破坏性极大的地质灾害之一，对人类生命和财产安全构成严重威胁[1-2]。滑坡易发性评价为准确预测潜在滑坡的空间分布提供了有效途径，是滑坡灾害风险定量评估和监测预警的基础，在防灾减灾、土地规划和资源利用等领域发挥着重要作用[3-4]。在滑坡易发性评价中，许多研究都侧重于数据驱动模型，尤其是基于机器学习的模型，这些模型擅长处理复杂的非线性问题[5-6]。然而，机器学习模型对数据的真实性要求较高，输入模型的滑坡样本和负样本数据必须具有较高的准确性[7]。滑坡样本即正样本，通常来源于实地调查和遥感影像解译；非滑坡样本即负样本，无法直接获取，通常需要依赖后期策略，从未发生滑坡区域进行采样获得[8-9]。早期的研究通常采用随机抽取的方式来确定非滑坡样本，但这种方法未考虑到所选的非滑坡样本可能属于潜在滑坡，难以确保负样本的可靠性[10]。此外，也有研究尝试根据滑坡区与非滑坡区的地理环境特征差异来选取滑坡负样本，但由于滑坡成因的复杂性和不同地区主导因素的差异，单纯依赖地理环境特征难以充分解释滑坡发生的概率差异[11]。

最新研究表明，利用物理模型能够显著提高滑坡负样本选择的科学性和可靠性。例如，瞬态降雨浅层滑坡稳定性模型、三维边坡稳定性分析模型等物理模型在边坡稳定性分析中的应用，可以从安全系数较高的区域选取滑坡负样本，这些方法生成的滑坡负样本比随机采样更为合理[12-13]。然而，物理驱动模型也存在明显的局限性，一方面，准确刻画边坡稳定性的物理机制需要大量的先验知识和参数标定工作；另一方面，获取水文和岩土数据的成本高昂、过程复杂，这些限制迫使研究人员依赖近似数据，从而带来了不确定性，在复杂或区域性环境中更为突出[14]。因此，寻找一种更直接、更易于解释的方法，在更大尺度上生成滑坡负样本显得尤为重要[15]。

近年来，合成孔径雷达干涉测量(interferometric synthetic aperture radar, InSAR)技术被广泛应用于城市沉降、地下水监测、地震探测等领域，因其高精度毫米级的地表形变监测能力，成为滑坡灾害预警和防治的重要工具[16-17]。在滑坡研究中，InSAR不仅能够监测现有滑坡的形变特征，还可通过长期监测数据识别潜在滑坡，并结合环境触发因素深入解析滑坡的致灾机理和时空演化规律。同时，多源遥感数据的融合以及天-空-地一体化监测网络的协同工作，进一步提升了监测数据的时空分辨率和准确性[18]。目前，将InSAR与机器学习模型有效结合起来用于滑坡易发性评价的研究尚处于探索阶段。传统滑坡易发性评价主要依赖地质和水文等静态数据，缺乏对滑坡动态形变特征的考虑，降低了评价结果的可靠性[19]。已有研究表明，集成InSAR技术可通过更新滑坡清单、交叉验证、与环境因子联合训练等方式提高滑坡易发性模型的准确性，也可利用列联合矩阵或经验公式对结果进行加权叠加[20-22]。然而，关于如何利用InSAR形变数据进行滑坡负样本选择的研究仍较为稀缺。

在滑坡易发性评价领域，尽管机器学习模型展现出了卓越的预测性能，但其黑箱性问题仍限制了其在实际应用中的可解释性和决策透明度[23]。现有研究大多侧重于模型精度的提升，但对各评价因子与滑坡形成机制的内在关联缺乏系统揭示，这不仅影响了对模型决策过程的理解与信任，也在一定程度上限制了对滑坡发生条件的深入认识[24-25]。早期的研究尝试使用皮尔逊相关系数、地理探测器等方法来解释模型输出，但这些方法仅限于分析区域尺度上主要因素的重要性，难以解释滑坡的空间分布模式[26]。近年来，学者们针对机器学习可解释性问题提出了多

种方法,如部分依赖图、局部可解释模型无关解释和沙普利加性解释(Shapley additive explanations, SHAP)等[27]。其中,SHAP方法因其强大的可视化能力和从全局与局部角度提供模型解释的能力得到了广泛应用,并在多个领域取得了良好效果[28]。

四川省汉源县因脆弱的地质条件、季节性降雨和周期性水库填埋等因素,滑坡灾害频发。因此,本文以汉源县为研究区,提出了一种基于InSAR的滑坡负样本采样策略,并结合极端梯度提升(extreme gradient boosting, XGBoost)和随机森林(random forest, RF)两种常用的机器学习方法构建滑坡易发性模型。为验证采样方法的有效性,本文将基于InSAR的滑坡负样本采样方法与广泛使用的缓冲区采样方法进行对比研究。此外,为解决滑坡易发性评价中缺乏动态特征的问题,本文将InSAR数据与机器学习模型相结合,通过构建评估矩阵优化滑坡易发性制图结果,并利用SHAP方法解释模型内部的决策机制。本文的创新之处在于,扩展了对InSAR形变数据与机器学习方法集成的理解,并进一步阐明了InSAR数据在滑坡负样本采样中的作用。

## 1 研究区与数据来源

### 1.1 研究区概况

汉源县隶属于四川省雅安市,位于四川省西南部,地理坐标为102°16′E~103°01′E,29°05′N~29°43′N,地理位置如图1所示。

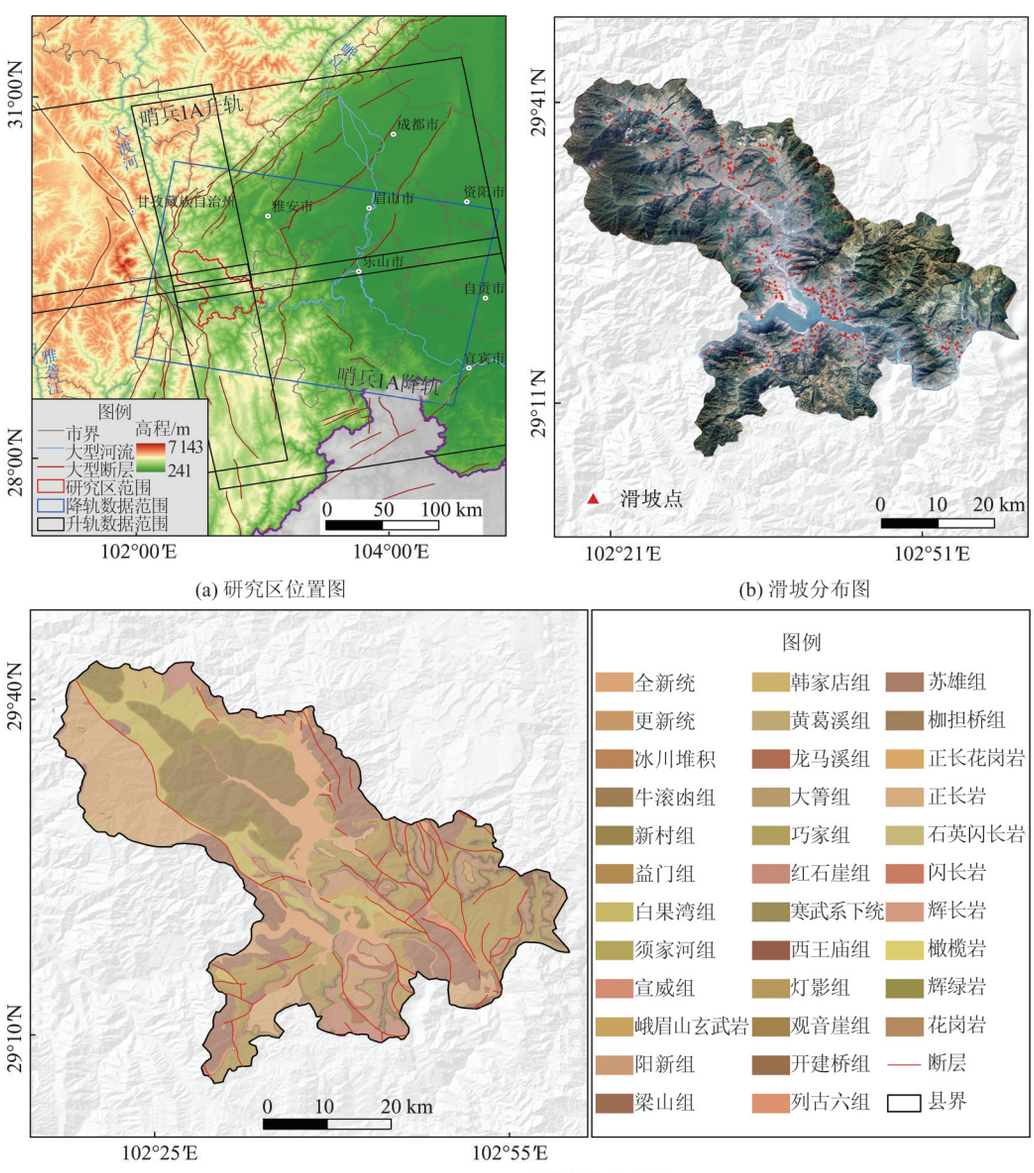



图1 研究区位置、滑坡分布及地质图

Fig. 1 Location, Landslide Distribution, and Geological Map of the Study Area

汉源县属于亚热带季风性湿润气候，多年平均气温为18 ℃，降水集中在5月—10月[29]。汉源县地处横断山脉北段东缘，大渡河横穿东西，流沙河纵贯南北，形成四周群山环绕、中部地势低平的河谷区。以汉源县县城为中心，海拔逐渐向周边递增，研究区内出露地层齐全，从第四系至震旦系地层均有不同程度的出露；区内地质构造较复杂，其主要活动性断裂有金坪断裂、汉源-甘洛断裂、流沙河隐伏断裂和宜坪-万坪断裂。

### 1.2　数据来源

通过高分辨率遥感影像、滑坡报告和野外调查，在汉源县共识别出319处滑坡，覆盖面积12.03 $km^2$，占整个研究区域的0.54%，如图1(b)所示，主要以中小型滑坡为主，季节性降雨、水位周期性波动以及人为活动是滑坡的主要诱发因素。

在滑坡易发性评价中，选取合适的滑坡影响因子是前提条件[30]。然而，目前尚未形成统一的因子选取标准[31-32]。本文依据汉源县的自然环境和滑坡分布特征以及相关文献参考，初步选取高程、坡度、坡向、地形起伏度、工程地质岩组、距道路距离、距断层距离、距河流距离、地形湿度指数、土地利用类型、归一化植被指数、年均降雨量12个因子作为评价指标(图2)。其中，工程地质岩组依据《工程岩体分级标准》(GB/T 50218—2014)[33]划分为坚硬岩(花岗岩、玄武岩、闪长岩、辉绿岩)、较坚硬岩(白云岩、灰岩、板岩)、较软弱岩(粉砂岩、泥灰岩)、软岩(泥岩、泥质页岩)、极软岩(砂土层)。滑坡影响因子信息如表1所示。

本文采用欧洲航天局发射的C波段哨兵1号雷达卫星影像进行地表形变探测，收集了2022-02-03—2024-10-02共57景升轨数据和2022-01-12—2025-02-25共91景降轨数据，数据覆盖范围如图1(a)所示。利用小基线集InSAR(small baseline subsets InSAR, SBAS-InSAR)技术估算研究区视线方向(line of sight, LOS)的形变速率，同时收集了30 m分辨率的ALOS数字高程模型数据，用于校正地形相位，并利用精密轨道数据消除轨道误差。

## 2　滑坡易发性建模方法

本文技术路线包括4个关键步骤：(1)数据收集与处理，包括更新滑坡数据库、筛选环境因子、获取地表形变信息；(2)样本集构建，集成滑坡样本和不同采样方式下的非滑坡样本，保持正样本与负样本比例为1∶1，并以7∶3的比例构建训练集和测试集；(3)易发性制图与优化，应用XGBoost和RF两种预测模型进行滑坡易发性预测，使用多种评价指标验证和比较预测模型的性能，选取最佳模型并结合InSAR数据构建评估矩阵以优化滑坡易发性评价结果；(4)模型解释，利用SHAP方法分析模型决策的内在机制。

### 2.1　数据处理平台

本文利用GAMMA软件进行合成孔径雷达(synthetic aperture radar, SAR)影像配准、差分干涉、相位滤波、相位解缠及地理编码等SAR干涉步骤的处理。环境因子预处理在ArcGIS软件中完成，相关性分析则通过SPSS软件实现。利用Python 3.8.2中的数据处理与分析、数值运算、科学计算等库对数据进行读取、计算和预处理，滑坡易发性预测模型的构建则依托机器学习算法库进行。

### 2.2　环境因子筛选

在滑坡易发性评价中，评估因子的选择至关重要，所选因子不仅应能够反映对滑坡的影响，还应避免它们之间存在强烈的相关性[34]。为了评估所选因子之间的相关性，本文使用皮尔逊相关系数进行分析[35]。皮尔逊相关系数是一种测量多个连续变量之间线性关系的方法，用$r$表示，取值范围为$[-1,1]$，符号表示相关的方向，绝对值表示强度[36]。通常，$|r|<0.3$表示无相关性，$0.3\leqslant|r|<0.5$表示相关性较弱，$0.5\leqslant|r|\leqslant0.8$表示中等相关性，$|r|>0.8$表示相关性较强，皮尔逊相关系数矩阵如图3所示。由图3可知，研究区内大多数评估因子的相关系数绝对值均低于0.50，仅高程和年均降雨量的相关系数绝对值为0.87，表现出强烈的正相关性。考虑到高程对滑坡发生的物理意义较弱，因此，本文决定剔除高程，并选择剩余的11个评价因子用于汉源县滑坡易发性评价的建模。

### 2.3　滑坡负样本采样策略

1)缓冲区采样方法：以滑坡边界为中心，设置200 m缓冲区，在缓冲区外随机生成非滑坡样本[37-38]。

2)基于InSAR定义的稳定区采样方法：本文通过SBAS-InSAR技术计算研究区的地表形变速率。使用ArcGIS 10.8中的克里金插值工具对InSAR形变点进行了插值，填补了由于陡峭地形和高密度植被覆盖导致的SAR影像失相关的无数据区域。在LOS向的形变速率结果中，负值表示目标远离卫星方向运动，正值表示目标朝向卫星方向，该正负值并不能直接代表滑坡真实的运动方向。无论是滑坡沿地表向下的负向变形，还

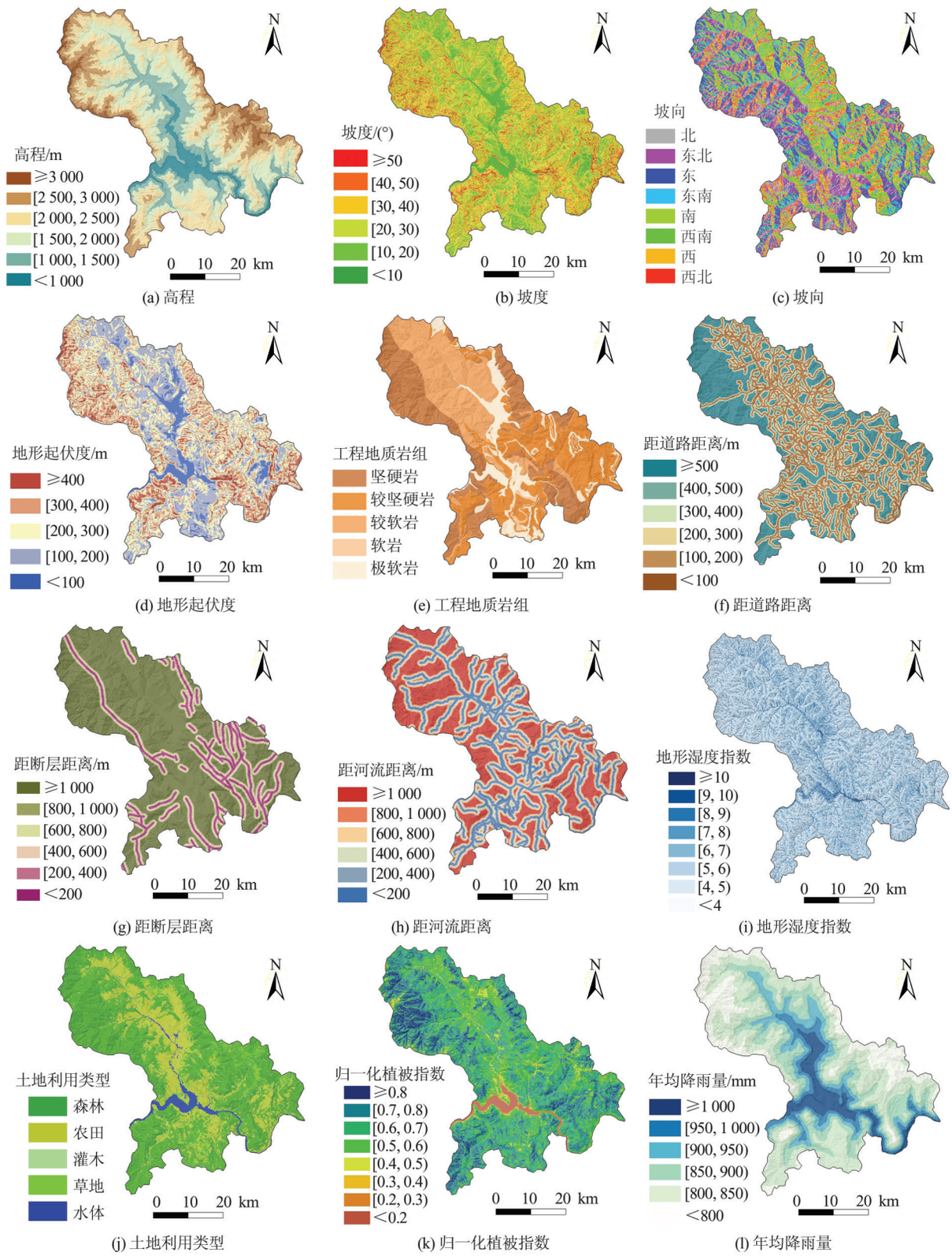



图2 评价指标分级

Fig. 2 Classification of Evaluation Indicators

是其前缘隆起的正向变形,都表征了滑坡或其他地貌的活动强度,因此,本文直接采用形变速率的绝对值表征地表的不稳定程度。为了更好地对形变速率进行分类,采用自然断点法将形变速率的绝对值划分为$V1$(极小)、$V2$(小)、$V3$(中)、$V4$(大)、$V5$(极大)5个等级。此外,考虑到不同卫星轨道对不同方向形变的敏感性不同,单一轨道的形变信息可能不完整,本文构建了一个组合矩

**表 1　影响因子数据源**

Table 1　Data Sources of Impact Factors

| 数据名称 | 空间分辨率 | 数据源 |
|---|---|---|
| 高程 | 30 m | http://earthdata.nasa.gov |
| 坡度<br>坡向<br>地形起伏度<br>地形湿度指数 | 30 m | 基于高程数据提取 |
| 工程地质岩组<br>距断层距离 | 1:250 000 | https://www.cgs.gov.cn/ |
| 距道路距离<br>距河流距离 | 30 m | https://www.openstreetmap.org/ |
| 土地利用类型 | 30 m | https://www.resdc.cn/ |
| 归一化植被指数 | 30 m | http://www.gscloud.cn |
| 年均降雨量 | 30 m | http://chrsdata.eng.uci.edu/ |

阵[21-22],通过整合升轨与降轨观测数据,实现地表形变等级的联合评估。升轨、降轨形变速率组合矩阵如表2所示。最终,形变量非常小的区域被定义为稳定区,并在该区域内随机生成非滑坡样本。

**表 2　升轨、降轨形变速率组合矩阵**

Table 2　Combination Matrix of Ascending and Descending Orbit Deformation Rates

| 等级 | 等级 | | | | |
|---|---|---|---|---|---|
| | *V*5 | *V*4 | *V*3 | *V*2 | *V*1 |
| *V*5 | 5 | 5 | 5 | 5 | 5 |
| *V*4 | 5 | 4 | 4 | 4 | 4 |
| *V*3 | 5 | 4 | 3 | 3 | 3 |
| *V*2 | 5 | 4 | 3 | 2 | 2 |
| *V*1 | 5 | 4 | 3 | 2 | 1 |

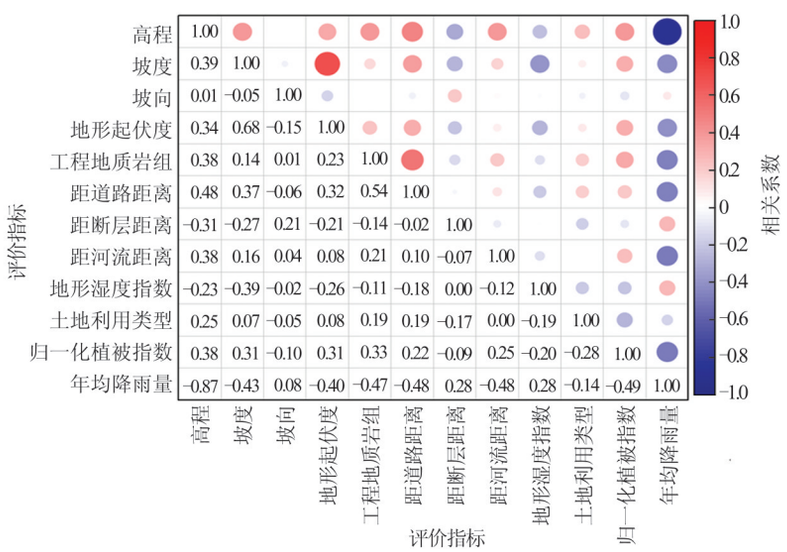



图 3　皮尔逊相关系数矩阵

Fig. 3　Pearson Correlation Coefficient Matrix

## 2.4　滑坡易发性预测模型

### 2.4.1　XGBoost模型

XGBoost是一种基于梯度提升树的高效算法,通过构建多棵决策树并利用梯度下降法逐步优化损失函数,从而实现强大的预测能力[39]。XGBoost的优化目标是通过最小化损失函数来学习最佳模型,损失函数的计算公式为:

$$L(\theta)=\sum_{i=1}^{n} l(y_i,\hat{y}_i)+\sum_{t=1}^{T}\Omega(f_t) \tag{1}$$

式中,$y_i$是第$i$个样本的真实标签;$\hat{y}_i$是模型预测的标签;$l(y_i,\hat{y}_i)$是预测值与实际值之间的损失;$T$是决策树的数量;$f_t$是第$t$棵树;$\Omega(f_t)$是模型复杂度的正则化项,防止过拟合。

### 2.4.2　RF模型

RF是一种流行的集成学习算法,用于解决分类和回归问题,已成功应用于许多滑坡易发性评估研究[40]。RF的基本原理是通过从原始数据中随机抽取多个子集,训练多个决策树,并通过多数投票法将各树的预测结果合并,最终得出分类或预测结果,计算公式为:

$$P=\max\left(p_i=\frac{1}{k}\sum_{j=1}^{k}p_{ij}\right) \tag{2}$$

式中,$P$表示最终的分类概率;$p_i$是类别$i$的平均概率;$k$是决策树的数量;$p_{ij}$是第$j$棵决策树对类别$i$的预测概率;$I$是所有类别的集合,$i\in I$。

## 2.5　模型性能评估

### 2.5.1　受试者工作特征(receiver operating characteristic, ROC)曲线

ROC曲线是评价模型精度最常用的方法之一,表示特异性与敏感性的综合指标。通过ROC

曲线下面积(area under curve, AUC)判断模型的总体指标,当AUC值小于等于0.5时,模型预测结果与现实相反;当AUC值大于0.5并接近1时,模型预测结果接近现实,即说明模型精度较高[11]。

#### 2.5.2 混淆矩阵

混淆矩阵是衡量易发性模型准确度较为直观的方法,它将模型的预测结果与观察到的结果进行比较,并将结果分为4个不同的类别:真阳性(true positive, TP)、真阴性(true negative, TN)、假阳性(false positive, FP)和假阴性(false negative, FN)。$N_{TP}$、$N_{TN}$、$N_{FP}$、$N_{FN}$分别为结果判断为TP、TN、FP、FN的数量。准确率(accuracy, ACC)是分类模式最常用的性能指标[12],它表示模型的精度,即模型正确识别的数量与样本总数比值。一般来说,模型的准确率越高,其性能就越好,ACC的计算公式为:

$$E_{ACC}=\frac{N_{TP}+N_{TN}}{N_{TP}+N_{FP}+N_{TN}+N_{FN}} \tag{3}$$

### 2.6 SHAP方法

SHAP是一种可解释机器学习方法,基于合作博弈理论中的Shapley值,通过计算各特征对预测结果的贡献度,为每个特征分配贡献值,从而揭示模型的决策机制[41]。因此,应用SHAP方法有助于探究影响滑坡发生的关键因素及其影响程度,计算公式为:

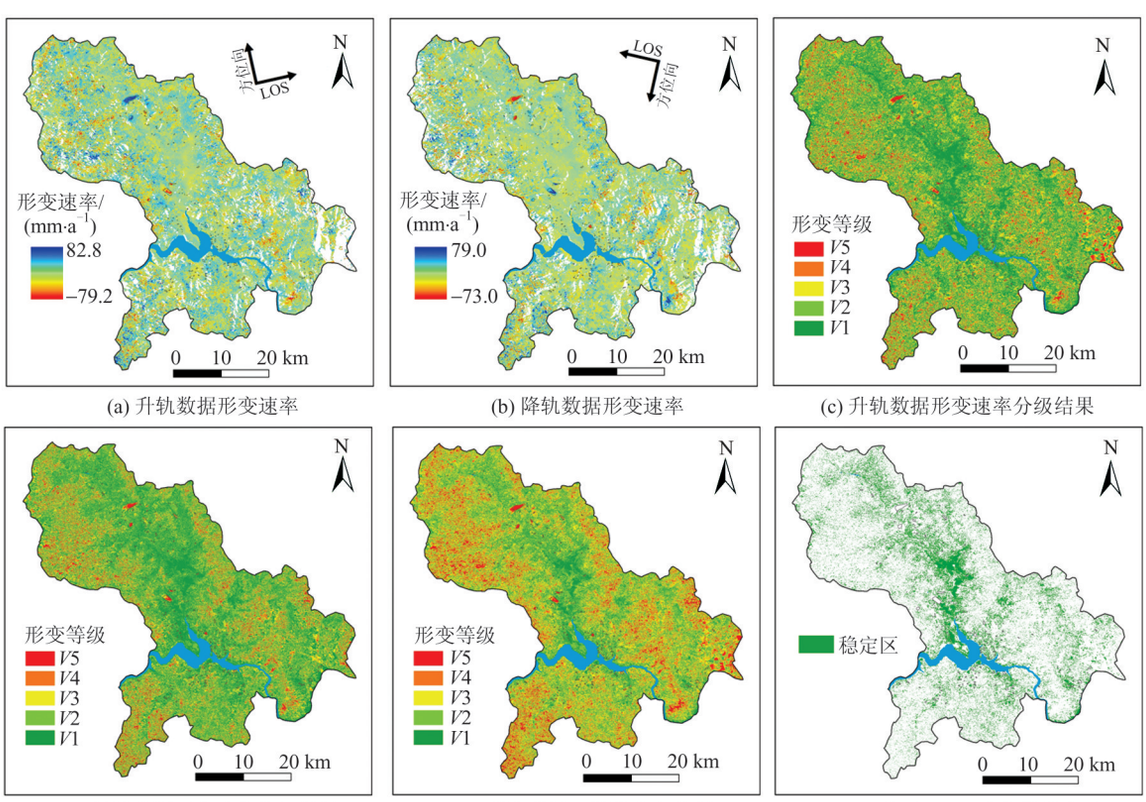
$$\phi_i(f)=\sum_{S\subseteq C\setminus\{j\}}\frac{|S|!(|C|-|S|-1)!}{|C|!}\cdot \left[f(S\cup\{k\})-f(S)\right] \tag{4}$$

式中,$\phi_i(f)$是特征$i$的SHAP值,表示特征$i$对预测结果$f$的贡献;$S$是特征集合,表示特征$j$在其他特征上的所有子集;$C$是所有特征的集合;$f(S\cup\{k\})$表示在$S$的基础上增加特征$k$后的预测结果;$f(S)$是模型在特征子集$S$上的预测结果。

## 3 滑坡易发性评价结果

### 3.1 InSAR定义的稳定区

通过SBAS-InSAR技术获取了研究区在观测期内的形变水平,结果如图4所示。图4中,升轨数据的最大年形变速率分别为−79.2 mm/a和82.8 mm/a(图4(a)),降轨数据的最大年形变速率分别为−73.0 mm/a和79.0 mm/a(图4(b))。依据§2.3所述方法,分别将升轨、降轨数据的形变速率划分为5级(图4(c)、4(d)),并通过联合矩阵




(a) 升轨数据形变速率 (b) 降轨数据形变速率 (c) 升轨数据形变速率分级结果

(d) 降轨数据形变速率分级结果 (e) 联合升轨、降轨分级结果 (f) 基于InSAR定义的稳定区

图4 SBAS-InSAR形变结果

Fig. 4 SBAS-InSAR Deformation Results

进行整合(图 4(e))。最终,变形程度从大到小的比例依次为 4.24%、14.68%、28.86%、35.29%、16.93%。本文将变形等级最小的区域定义为稳定区,在该区域内生成滑坡负样本。然而,在滑坡区域内,仍存在变形速率最慢的单元,为避免滑坡和非滑坡样本之间的重叠,将滑坡区域中变形最慢的单元排除在稳定区之外(图 4(f))。

### 3.2　滑坡易发性制图

本文采用栅格单元作为评价单元,将所有模型预测的滑坡发生概率值采用自然断点法分为 *S*1(极低)、*S*2(低)、*S*3(中)、*S*4(高)、*S*5(极高)5个易发性级别[42-43]。滑坡易发性分级结果如图 5 所示,总体而言,滑坡灾害点聚集的区域绝大多数位于所划分的中易发区及其以上区域。各模型下的研究区滑坡易发性分级规律类似,低易发区主要分布在人类活动影响较小、海拔较高的山区,中易发区大多分布在距水系和道路较远的地区或支流附近,而等级为 *S*4 的易发区主要分布在主河流、道路两岸的河谷地区,呈条带状分布。不同模型在细节上存在差异,基于缓冲区采样的方法,分区效果不够明确,中易发区的占比较大。

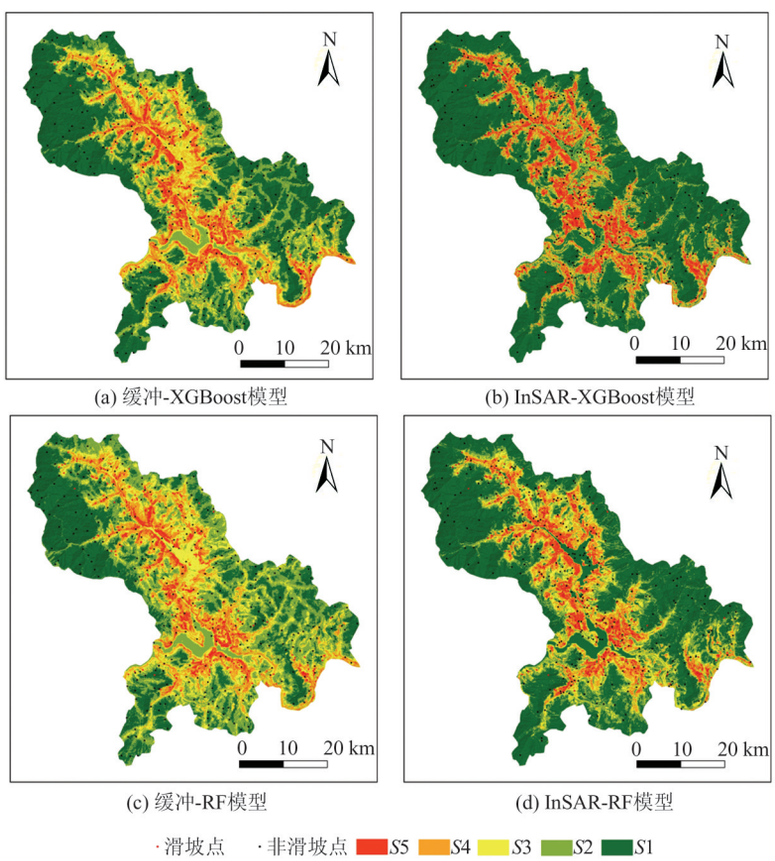



图 5　滑坡易发性分级结果

Fig. 5　Classification Results of Landslide Susceptibility

### 3.3　精度验证结果

ROC 曲线如图 6 所示,由图 6 可知,缓冲区采样法和基于 InSAR 定义的稳定区采样法的 AUC 值均高于 0.75,对于不同机器学习模型,基于 InSAR 的负样本采样法的 AUC 值显著高于缓冲区采样法。其中,InSAR-RF 模型的 AUC 值最高,达到 0.847,比缓冲区取样法高出 5.5%。

混淆矩阵结果如图 7 所示,由图 7 可知,缓冲区采样法和基于 InSAR 定义的稳定区采样法的 ACC 值均高于 80%。其中,基于 InSAR 的负样本选取方法在 RF 模型中的 ACC 值最高,达到 89.8%,比缓冲区采样法高出 5.6%。这一结果表明,基于 InSAR 的负样本选取方法显著提升了模型的精度。

各易发性等级分区情况的统计结果如表 3 所示,其中,频率比为各分区内滑坡数量占比与分区面积占比的比值。由表 3 可知,随着易发性等级的提高,频率比逐渐增大。对于相同采样方法,不同模型的表现有所不同,在易发性等级为 *S*4、*S*5 的区域,RF 模型的频率比均高于 XGBoost 模型。在基于 InSAR 的非滑坡采样方法中,RF 模型在易发性等级为 *S*5 的易发区其频率比最大,达到 8.98。

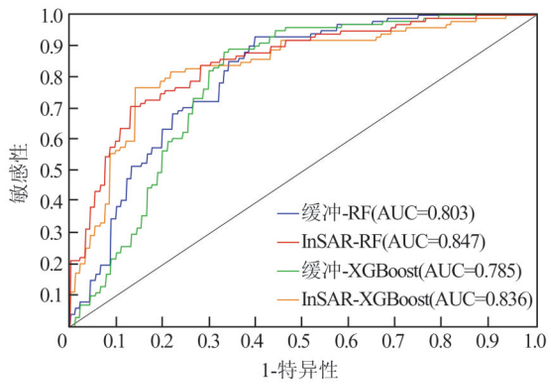
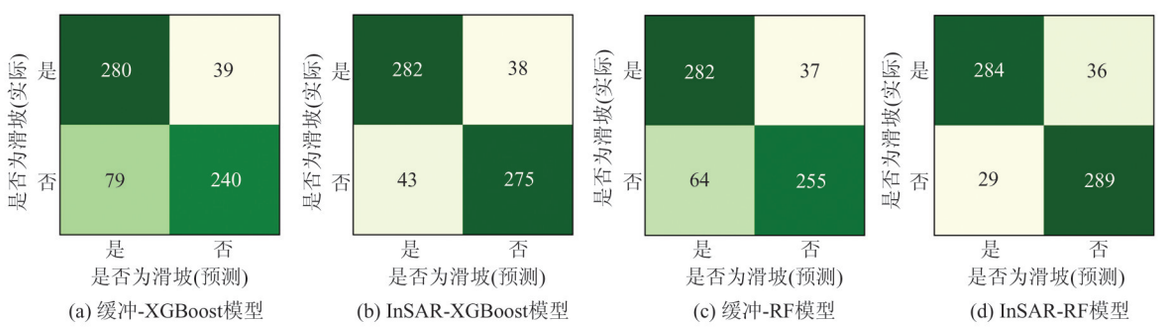



图6 ROC曲线

Fig. 6 ROC Curve

### 3.4 滑坡易发性结果优化

通过对比形变速率与滑坡易发性评价结果可以看出,LOS向形变速率能准确标注研究区内的潜在滑坡位置,与实际地表形变吻合度较高。依据表2中的矩阵,将升、降轨数据获得的形变速率相结合,并划分为$V1$、$V2$、$V3$、$V4$、$V5$ 5个等级,同样将InSAR-RF模型的滑坡易发性预测结果也分为$S1$、$S2$、$S3$、$S4$、$S5$ 5个等级。将上述两者结合起来,构建了滑坡易发性联合评估矩阵[20-22],见表4。图8展示了滑坡易发性优化结果。结果表明,基于InSAR形变速率与滑坡易发性评价结果相结合的联合评估方法不仅能精确定位高风险区域,进一步降低评估结果中出现误报的可能性,还能增加滑坡易发性评价的实时性,使得易发性分区图能够持续更新,从而实现对汉源县滑坡易发性的动态、准确评估。




(a) 缓冲-XGBoost模型 (b) InSAR-XGBoost模型 (c) 缓冲-RF模型 (d) InSAR-RF模型

图7 混淆矩阵

Fig. 7 Confusion Matrix

### 3.5 结果验证

为了验证易发性结果的可信度,本文选取了两个典型滑坡进行对比分析,选取的滑坡编号分别为HP01、HP02,验证结果如图9所示。其中,HP01为活动性滑坡,整体变形显著,局部发生了垮塌失稳(图9(a)),2022—2025年间该滑坡的最大形变速率达−51.9 mm/a(图9(b)),处于快速蠕变阶段,危险性较大;HP02为古老滑坡(图9(e)),2022—2025年间未见明显形变(图9(f))。从优化前后的对比验证结果可以看出:(1)未考虑动态因子的评估结果对HP01的易发性等级预测较低(图9(c)),未能充分反映其活动性特征;对HP02的易发性等级预测较高(图9(g)),可能存在一定的过度预测。(2)考虑了动态因子的评估结果显著提高了HP01的易发性等级,更准确地反映了滑坡的活动性和风险(图9(d));而对HP02的易发性等级有所降低(图9(h)),更符合其非活动性的实际情况。由此可见,动态因子的引入显著降低了评估结果中出现TP和TN的情况,提高了模型的精确性。

### 3.6 SHAP可解释性结果

本文通过Shapley值,从全局角度量化每个因素对模型预测的贡献和重要性,结果如图10所示。图10中的每个点代表一个样本,并显示特征值的大小及其对应的SHAP值,点的颜色根据因子值的大小变化,较大的值显示为红色。正的SHAP值表示该特征对滑坡发生有正向影响,SHAP绝对值越大,表示该特征的贡献越大。结果表明,所选因子对滑坡的发生有显著贡献,并呈现一定规律。例如,年均降雨量与SHAP值呈正相关;随着距道路距离和距河流距离的减小,SHAP值呈增大趋势,表明这些因子对滑坡发生的影响较大。

对所有样本的SHAP值进行绝对值平均处理,得到因子重要性排序图,结果如图11所示,可以直观地比较各因子的贡献度。由图11可知,年均降雨量是滑坡的主要影响因素,其SHAP值的绝对值均值为0.160,年均降雨量对滑坡影响机制具有双重性:一方面,降雨通过地表径流冲刷导致坡面物质流失,另一方面,降雨通过渗透作用改变岩土体内部水文条件,从而影响边坡稳定性。地形起伏度是第二大影响因素,其平均SHAP值的绝对值均值为0.062,反映了地形在垂

**表3　易发性分区统计结果**

Table 3　Statistical Results of Susceptibility Zoning

| 模型 | 易发性等级 | 发生滑坡数量/个 | 占总滑坡比例/% | 分区面积/km² | 占研究区总面积的比例/% | 频率比 |
|---|---|---|---|---|---|---|
| 缓冲-XGBoost | *S*1 | 8 | 2.51 | 916.70 | 41.17 | 0.06 |
| | *S*2 | 20 | 6.27 | 519.82 | 23.34 | 0.27 |
| | *S*3 | 61 | 19.12 | 348.19 | 15.63 | 1.22 |
| | *S*4 | 140 | 43.89 | 329.25 | 14.79 | 2.97 |
| | *S*5 | 90 | 28.21 | 112.84 | 5.07 | 5.56 |
| InSAR-XGBoost | *S*1 | 13 | 4.08 | 1 267.42 | 56.92 | 0.07 |
| | *S*2 | 16 | 5.02 | 330.12 | 14.82 | 0.34 |
| | *S*3 | 44 | 13.79 | 232.71 | 10.45 | 1.32 |
| | *S*4 | 87 | 27.27 | 224.13 | 10.07 | 2.71 |
| | *S*5 | 159 | 49.84 | 172.42 | 7.74 | 6.44 |
| 缓冲-RF | *S*1 | 3 | 0.94 | 657.10 | 29.51 | 0.03 |
| | *S*2 | 11 | 3.45 | 699.86 | 31.43 | 0.11 |
| | *S*3 | 55 | 17.24 | 456.04 | 20.48 | 0.84 |
| | *S*4 | 148 | 46.40 | 325.33 | 14.61 | 3.18 |
| | *S*5 | 102 | 31.97 | 88.48 | 3.97 | 8.05 |
| InSAR-RF | *S*1 | 6 | 1.88 | 1 158.31 | 52.02 | 0.04 |
| | *S*2 | 12 | 3.76 | 396.91 | 17.82 | 0.21 |
| | *S*3 | 44 | 13.79 | 317.23 | 14.24 | 0.97 |
| | *S*4 | 119 | 37.31 | 247.12 | 11.10 | 3.36 |
| | *S*5 | 138 | 43.26 | 107.23 | 4.82 | 8.98 |

**表4　滑坡易发性联合评估矩阵**

Table 4　Combined Evaluation Matrix of Landslide Susceptibility

| 等级 | 等级 | | | | |
|---|---|---|---|---|---|
| | *V*5 | *V*4 | *V*3 | *V*2 | *V*1 |
| *S*5 | 5 | 5 | 4 | 3 | 2 |
| *S*4 | 5 | 4 | 4 | 3 | 1 |
| *S*3 | 4 | 4 | 3 | 2 | 1 |
| *S*2 | 3 | 3 | 2 | 1 | 1 |
| *S*1 | 2 | 1 | 1 | 1 | 1 |

直方向上的变化幅度,当地形起伏较大时,土体的重力和水力作用会引起不均衡,进而诱发土体的滑动和变形。第三大影响因素是距道路距离,其平均SHAP值的绝对值均值为0.046,道路基础设施建设通过边坡开挖和切坡作业改变滑坡体的应力状态,增大临空面并削弱结构稳定性,进而加剧边坡失稳的风险。此外,归一化植被指数、土地利用类型、和距河流距离因子也与滑坡发生密切相关,其SHAP值的绝对值均值均大于0.03,而坡度、坡向、地形湿度指数、距断层距离、工程地质岩组等因子与滑坡发生的相关性较低,对滑坡发生的影响较小。

## 4　结　语

本文提出了一种基于InSAR的采样策略,用于在四川省汉源县的滑坡易发性制图中生成滑坡负样本。通过SBAS-InSAR技术计算年平均地表形变速率,并从变形非常缓慢的区域选取滑坡负样本。研究表明,基于InSAR的采样方法显著提高了XGBoost和RF模型的预测效果,来自InSAR定义的稳定区的样本能更准确地代表非滑坡区域的特征。此外,本文通过引入InSAR数据优化模型结果,进一步提高了滑坡易发性评估的准确性,并利用SHAP方法对模型决策过程进行了可解释性分析。主要结论如下:

1)汉源县滑坡灾害高易发区主要集中在人类活动频繁且海拔较低的地区,为滑坡灾害的防控和治理提供了重要的空间信息支持。

2)相比于缓冲区采样法,采用SBAS-InSAR技术选取滑坡负样本能更准确地代表非滑坡区域的特征,从而提高了滑坡易发性模型的预测精度,RF模型在滑坡易发性评价中的预测精度明显高于XGBoost模型。

3)本文提出的基于InSAR的负样本采样方法不仅提高了模型的准确性,而且通过将InSAR

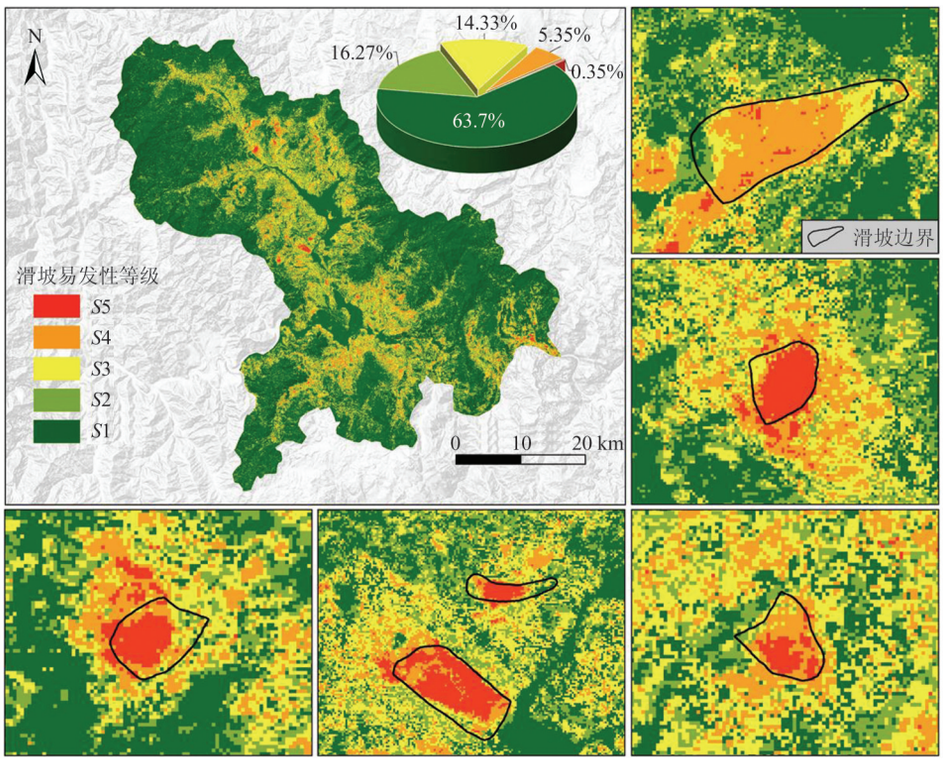
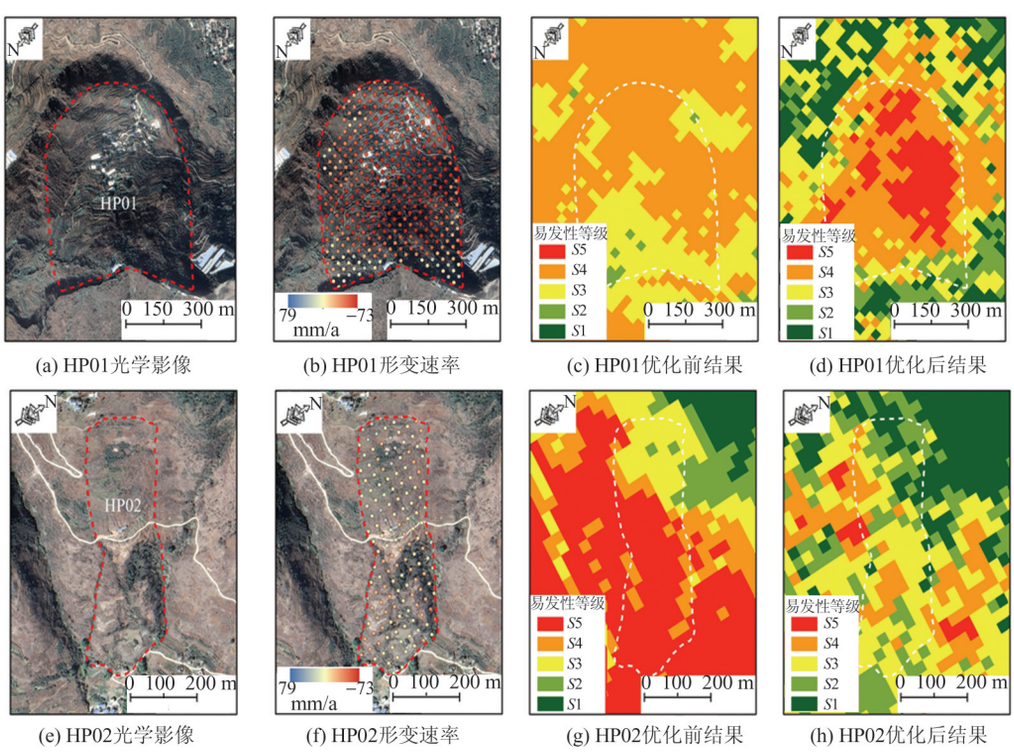



图8 滑坡易发性优化结果

Fig. 8 Optimization Results of Landslide Susceptibility




(a) HP01光学影像 (b) HP01形变速率 (c) HP01优化前结果 (d) HP01优化后结果

(e) HP02光学影像 (f) HP02形变速率 (g) HP02优化前结果 (h) HP02优化后结果

图9 易发性分区结果验证

Fig. 9 Validation of Susceptibility Zoning Results

数据与易发性结果相结合,增强了滑坡易发性评价的科学性和可靠性,为滑坡灾害预警和管理提供了新的思路和方法。

4)通过应用SHAP可解释性模型,对滑坡易发性预测结果进行了深入分析,量化了各个因子对滑坡发生的贡献,揭示了年均降雨量、地形起伏度、距道路距离等关键因素对滑坡易发性的影响,为进一步优化滑坡易发性评估提供了有力支持。

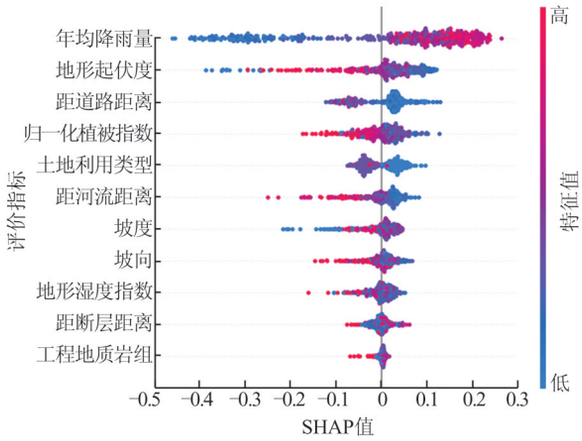



图 10　SHAP 汇总图

Fig. 10　SHAP Summary Plot

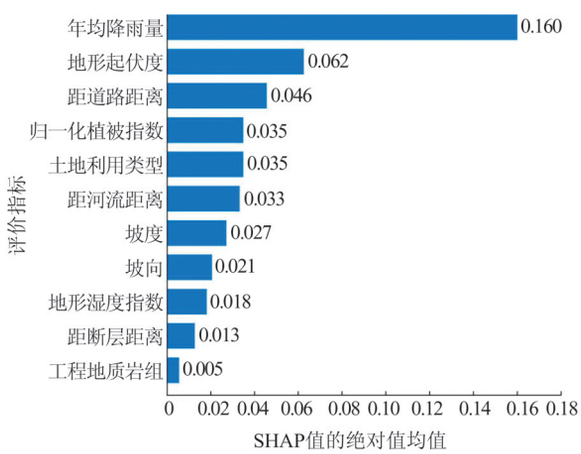



图 11　影响因子重要性排序

Fig. 11　Importance Ranking of Impact Factors

## 参考文献


[1] 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. The Xinmocun Landslide on June 24，2017 in Maoxian，Sichuan：Characteristics and Failure Mechanism[J]. *Chinese Journal of Rock Mechanics and Engineering*，2017，36(11)：2612-2628.

[2] 许强．对滑坡监测预警相关问题的认识与思考[J]．工程地质学报，2020，28(2)：360-374.
XU Qiang. Understanding the Landslide Monitoring and Early Warning：Consideration to Practical Issues[J]. *Journal of Engineering Geology*，2020，28(2)：360-374.

[3] 瞿伟，边子策，李久元，等．融合 Transformer 与 CNN 的正样本缓冲滑坡易发性评价模型：以宝鸡市为例[J]．武汉大学学报(信息科学版)，2026，51(4)：726-740.
QU Wei，BIAN Zice，LI Jiuyuan，et al. Landslide Susceptibility Assessment Model Based on Positive Sample Buffering with Transformer-CNN Integration：A Case Study of Baoji City[J]. *Geomatics and Information Science of Wuhan University*，2026，51(4)：726-740.

[4] 郭飞，吴迪，葛民荣，等．连续变量因子分级和机器学习模型对滑坡易发性评价精度的影响[J]．武汉大学学报(信息科学版)，2026，51(2)：236-248.
GUO Fei，WU Di，GE Minrong，et al. Influence of Continuous Variable Factor Classification and Machine Learning Model on the Accuracy of Landslide Susceptibility Evaluation[J]. *Geomatics and Information Science of Wuhan University*，2026，51(2)：236-248.

[5] 徐胜华，马钰，刘纪平，等．多尺度特征学习的轻量化滑坡易发性评价方法[J]．武汉大学学报(信息科学版)，2025，50(6)：1175-1190.
XU Shenghua，MA Yu，LIU Jiping，et al. A Lightweight Network Based on Multi-scale Feature Learning for Landslide Susceptibility Assessment[J]. *Geomatics and Information Science of Wuhan University*，2025，50(6)：1175-1190.

[6] HUANG F M，LI R H，CATANI F，et al. Uncertainties in Landslide Susceptibility Prediction：Influence Rule of Different Levels of Errors in Landslide Spatial Position[J]. *Journal of Rock Mechanics and Geotechnical Engineering*，2024，16(10)：4177-4191.

[7] ALVIOLI M，LOCHE M，JACOBS L，et al. A Benchmark Dataset and Workflow for Landslide Susceptibility Zonation[J]. *Earth-Science Reviews*，2024，258：104927.

[8] LIU L L，ZHAO S L，YANG C，et al. Quantifying Uncertainty in Landslide Susceptibility Mapping Due to Sampling Randomness[J]. *International Journal of Disaster Risk Reduction*，2024，114：104966.

[9] LU Y X，XU H G，WANG C，et al. A Novel Strategy Coupling Optimised Sampling with Heterogeneous Ensemble Machine-Learning to Predict Landslide Susceptibility[J]. *Remote Sensing*，2024，16(19)：3663.

[10] GU T F，DUAN P，WANG M G，et al. Effects of Non-landslide Sampling Strategies on Machine Learning Models in Landslide Susceptibility Mapping[J]. *Scientific Reports*，2024，14：7201.

[11] ZHANG Q，HE Y，ZHANG L F，et al. A Landslide Susceptibility Assessment Method Considering the Similarity of Geographic Environments Based on Graph Neural Network[J]. *Gondwana Research*，

2024, 132: 323-342.

[12] YANG L, CUI Y L, XU C, et al. Application of Coupling Physics-Based Model TRIGRS with Random Forest in Rainfall-Induced Landslide-Susceptibility Assessment[J]. *Landslides*, 2024, 21(9): 2179-2193.

[13] LIU S L, WANG L Q, ZHANG W G, et al. A Physics-Informed Data-Driven Model for Landslide Susceptibility Assessment in the Three Gorges Reservoir Area[J]. *Geoscience Frontiers*, 2023, 14(5): 101621.

[14] LIU S L, WANG L Q, ZHANG W G, et al. Physics-Informed Optimization for a Data-Driven Approach in Landslide Susceptibility Evaluation[J]. *Journal of Rock Mechanics and Geotechnical Engineering*, 2024, 16(8): 3192-3205.

[15] GUO Z Z, TIAN B X, ZHU Y H, et al. How Do the Landslide and Non-landslide Sampling Strategies Impact Landslide Susceptibility Assessment? —A Catchment-Scale Case Study from China[J]. *Journal of Rock Mechanics and Geotechnical Engineering*, 2024, 16(3): 877-894.

[16] 王磊, 吕广涵, 杜建涛, 等. 黄河上游玛沁段活动滑坡InSAR探测与时空特征分析[J]. 武汉大学学报(信息科学版), 2025, 50(11): 2226-2235.
WANG Lei, LÜ Guanghan, DU Jiantao, et al. InSAR Detection and Spatiotemporal Characteristics of Active Landslides in the Maqin Section of the Upper Yellow River[J]. *Geomatics and Information Science of Wuhan University*, 2025, 50(11): 2226-2235.

[17] 何国庆, 李为乐, 陆会燕, 等. 基于Spearman滤波与HNSW-DBSCAN聚类的滑坡隐患InSAR形变区自动识别[J]. 武汉大学学报(信息科学版), 2025, 50(8): 1682-1693.
HE Guoqing, LI Weile, LU Huiyan, et al. Automated Identification of InSAR Deformation Concentration Zones for Landslide Hazards Using Spearman Coefficient Filtering and HNSW-DBSCAN Clustering[J]. *Geomatics and Information Science of Wuhan University*, 2025, 50(8): 1682-1693.

[18] XU Q, ZHAO B, DAI K R, et al. Remote Sensing for Landslide Investigations: A Progress Report from China[J]. *Engineering Geology*, 2023, 321: 107156.

[19] GUO Z Z, WANG H J, HE J, et al. PSLSA V2.0: An Automatic Python Package Integrating Machine Learning Models for Regional Landslide Susceptibility Assessment[J]. *Environmental Modelling & Software*, 2025, 186: 106367.

[20] WU X L, QI X S, PENG B, et al. Optimized Landslide Susceptibility Mapping and Modelling Using the SBAS-InSAR Coupling Model[J]. *Remote Sensing*, 2024, 16(16): 2873.

[21] SHAN Y F, XU Z, ZHOU S S, et al. Landslide Hazard Assessment Combined with InSAR Deformation: A Case Study in the Zagunao River Basin, Sichuan Province, Southwestern China[J]. *Remote Sensing*, 2024, 16(1): 99.

[22] DAI C, LI W L, LU H Y, et al. Landslide Hazard Assessment Method Considering the Deformation Factor: A Case Study of Zhouqu, Gansu Province, Northwest China[J]. *Remote Sensing*, 2023, 15(3): 596.

[23] AN B S, ZHANG Z J, XIONG S Q, et al. Landslide Susceptibility Mapping Based on Ensemble Learning in the Jiuzhaigou Region, Sichuan, China[J]. *Remote Sensing*, 2024, 16(22): 4218.

[24] 刘瑞阳, 许强, 蒲川豪, 等. 湖南资兴市2024年7月台风"格美"诱发滑坡特征及其地质控制作用[J]. 武汉大学学报(信息科学版), 2025, DOI: 10.13203/j.whugis20250027.
LIU Ruiyang, XU Qiang, PU Chuanhao, et al. Characteristics of Landslides Induced by Typhoon 'Gaemi' in Zixing, Hunan, July 2024, and Their Geological Control[J]. *Geomatics and Information Science of Wuhan University*, 2025, DOI: 10.13203/j.whugis20250027.

[25] 丁明涛, 江辉, 李振洪, 等. 融合卷积模糊神经网络与SHAP特征优化的地震诱发滑坡易发性评价[J]. 武汉大学学报(信息科学版), 2025, DOI: 10.13203/j.whugis20250017.
DING Mingtao, JIANG Hui, LI Zhenhong, et al. Earthquake-Induced Landslide Susceptibility Assessment by Fusing Convolutional Fuzzy Neural Network and SHAP Feature Optimization Strategy[J]. *Geomatics and Information Science of Wuhan University*, 2025, DOI: 10.13203/j.whugis20250017.

[26] YANG S L, TAN J Y, LUO D Y, et al. Sample Size Effects on Landslide Susceptibility Models: A Comparative Study of Heuristic, Statistical, Machine Learning, Deep Learning and Ensemble Learning Models with SHAP Analysis[J]. *Computers & Geosciences*, 2024, 193: 105723.

[27] WEN H J, LIU B, DI M R, et al. A SHAP-Enhanced XGBoost Model for Interpretable Prediction of Coseismic Landslides[J]. *Advances in Space Research*, 2024, 74(8): 3826-3854.

[28] SUN D L, CHEN D L, ZHANG J L, et al. Landslide Susceptibility Mapping Based on Interpretable

Machine Learning from the Perspective of Geomorphological Differentiation[J]. *Land*, 2023, 12(5): 1018.

[29] YE F, FU W X, ZHOU H F, et al. The "8·21" Rainfall-Induced Zhonghaicun Landslide in Hanyuan County of China: Surface Features and Genetic Mechanisms[J]. *Landslides*, 2021, 18(10): 3421-3434.

[30] YANG C, LIU L L, HUANG F M, et al. Machine Learning-Based Landslide Susceptibility Assessment with Optimized Ratio of Landslide to Non-landslide Samples[J]. *Gondwana Research*, 2023, 123: 198-216.

[31] LIU Q, TANG A P, HUANG D L. Exploring the Uncertainty of Landslide Susceptibility Assessment Caused by the Number of Non-Landslides[J]. *Catena*, 2023, 227: 107109.

[32] CHANG Z L, HUANG J S, HUANG F M, et al. Uncertainty Analysis of Non-landslide Sample Selection in Landslide Susceptibility Prediction Using Slope Unit-Based Machine Learning Models [J]. *Gondwana Research*, 2023, 117: 307-320.

[33] 中华人民共和国住房和城乡建设部. 工程岩体分级标准: GB/T 50218—2014[S]. 北京: 中国计划出版社, 2015.
Ministry of Housing and Urban-Rural Development of the People's Republic of China. Standard for Engineering Classification of Rock Mass: GB/T 50218—2014[S]. Beijing: China Planning Press, 2015.

[34] ZHAO F C, MIAO F S, WU Y P, et al. Landslide Dynamic Susceptibility Mapping in Urban Expansion Area Considering Spatiotemporal Land Use and Land Cover Change[J]. *Science of the Total Environment*, 2024, 949: 175059.

[35] ULLAH M, TANG B Z, HUANGFU W C, et al. Machine Learning-Driven Landslide Susceptibility Mapping in the Himalayan China-Pakistan Economic Corridor Region[J]. *Land*, 2024, 13(7): 1011.

[36] ZHAO F C, MIAO F S, WU Y P, et al. Refined Landslide Susceptibility Mapping in Township Area Using Ensemble Machine Learning Method Under Dataset Replenishment Strategy[J]. *Gondwana Research*, 2024, 131: 20-37.

[37] 段中满, 罗伟奇, 陈雅娜, 等. 基于信息量和机器学习的滑坡灾害易发性分析: 以湖南省中西部地区为例[J]. 湖南师范大学自然科学学报, 2024, 47(5): 111-120.
DUAN Zhongman, LUO Weiqi, CHEN Ya'na, et al. Analysis of Landslide Disaster Susceptibility Based on Information Quantity and Machine Learning: Take the Central and Western Regions of Hunan Province as an Example[J]. *Journal of Natural Science of Hunan Normal University*, 2024, 47(5): 111-120.

[38] MA S Y, XU C, SHAO X Y. Spatial Prediction Strategy for Landslides Triggered by Large Earthquakes Oriented to Emergency Response, Mid-term Resettlement and Later Reconstruction[J]. *International Journal of Disaster Risk Reduction*, 2020, 43: 101362.

[39] LIU B, GUO H X, LI J L, et al. Application and Interpretability of Ensemble Learning for Landslide Susceptibility Mapping Along the Three Gorges Reservoir Area, China[J]. *Natural Hazards*, 2024, 120(5): 4601-4632.

[40] MA S Y, SHAO X Y, XU C. Estimating the Quality of the Most Popular Machine Learning Algorithms for Landslide Susceptibility Mapping in 2018 Mw 7.5 Palu Earthquake [J]. *Remote Sensing*, 2023, 15(19): 4733.

[41] ZHANG J Y, MA X L, ZHANG J L, et al. Insights into Geospatial Heterogeneity of Landslide Susceptibility Based on the SHAP-XGBoost Model [J]. *Journal of Environmental Management*, 2023, 332: 117357.

[42] KE C Y, HE S, QIN Y G. Comparison of Natural Breaks Method and Frequency Ratio Dividing Attribute Intervals for Landslide Susceptibility Mapping [J]. *Bulletin of Engineering Geology and the Environment*, 2023, 82(10): 384.

[43] WANG Q, WANG Y, NIU R Q, et al. Integration of Information Theory, K-Means Cluster Analysis and the Logistic Regression Model for Landslide Susceptibility Mapping in the Three Gorges Area, China [J]. *Remote Sensing*, 2017, 9(9): 938.