

# 基于随机森林模型的藏东南地区滑坡易发性评价及主控因素分析

杜鹏<sup>1,2</sup>, 陈宁生<sup>1,3</sup>, 伍康林<sup>1,4</sup>, 李志<sup>1,2</sup>, 张瀛玉龙<sup>1,2</sup>

1. 中国科学院、水利部成都山地灾害与环境研究所, 成都 610041; 2. 西藏大学工学院, 拉萨 850000;  
3. 高原科学与可持续发展研究院, 西宁 810016; 4. 中国科学院大学工程科学学院, 北京 100049

**[摘要]** 为了获取藏东南地区的滑坡易发性分区, 深入探究研究区滑坡发育的主控因素。本文采用随机森林模型对藏东南地区的滑坡易发性进行研究。首先通过现场调查、遥感解译和文献查询等手段, 共发现研究区滑坡 306 处, 其次使用频率比模型(FR)对研究区易发性做快速评判, 在其低易发区进行非滑坡样本的选取, 使得随机森林模型得到较好的原始数据集, 并通过计算权重来制作研究区易发性图, 分析滑坡发生的主控因素, 最后绘制 ROC 曲线和计算线下曲线的面积(AUC)对模型的准确性进行验证。模型预测结果显示, 研究区高易发区主要集中在易贡藏布、帕隆藏布交汇处和雅鲁藏布江大拐弯的地方。在对特征重要性因素排序中, 前三分别是断层密度、高程和坡度, 模型的 AUC=0.940, 因此有较好的准确性。根据结果分析, 研究区因长期受到构造作用影响, 地质构造复杂, 东西两侧均是巨型断层, 南侧为喜马拉雅前缘断裂带, 断裂带发育, 尤其在高易发区断层纵横交错, 多重挤压造成岩石破碎, 节理发育。因此, 易贡藏布、帕隆藏布交汇处和雅鲁藏布江大拐弯区域为滑坡灾害高易发区具有合理性。该模型对藏东南地区滑坡的发生和灾害的治理提供了理论依据, 对该地区防灾减灾提供了理论支撑。

**[关键词]** 藏东南地区; 随机森林模型; 滑坡易发性; 断层密度

**[分类号]** P642.22

**[文献标志码]** A

## Evaluation of landslide susceptibility in southeast Tibet based on a random forest model

DU Peng<sup>1,2</sup>, CHEN Ningsheng<sup>1,3</sup>, WU Kanglin<sup>1,4</sup>, LI Zhi<sup>1,2</sup>, ZHANG Yingyulong<sup>1,2</sup>

1. Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Ministry of Water Resources, Chengdu 610041, China;

2. College of Engineering, Tibet University, Lhasa 850000, China;

3. Academy of Plateau Science and Sustainability, Xining 810016, China;

4. College of Engineering Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

**[收稿日期]** 2023-05-11。

**[基金项目]** 国家自然科学基金—国际(地区)合作与交流项目(42361144880); 国家重点研发计划(2023YFC3008301); 中国电建核心技术攻关“揭榜挂帅”项目(DJ-HXGG-2022-02); 中国科学院加德满都科教中心资助项目(131C11KY5B20200033); 西藏自治区重大科技专项(XZ202101ZD0013G)。

**[第一作者]** 杜鹏(1997—), 男, 硕士研究生, 研究方向: 岩土工程, E-mail: 1837841266@qq.com。

**[通信作者]** 陈宁生(1965—), 男, 研究员, 研究方向: 山地灾害形成机理与防治, E-mail: chennsh@imde.ac.cn。

**[引用格式]** 杜鹏, 陈宁生, 伍康林, 等, 2024. 基于随机森林模型的藏东南地区滑坡易发性评价及主控因素分析[J]. 成都理工大学学报(自然科学版), 51(2): 328-344.

Du P, Chen N S, Wu K L, et al., 2024. Evaluation of landslide susceptibility in southeast Tibet based on a random forest model[J]. Journal of Chengdu University of Technology (Science & Technology Edition), 51(2): 328-344.

**Abstract:** In this study, we investigated the main factors influencing the development of landslides in southeast Tibet by using the random forest model to determine the susceptibility of the area to them. We identified 306 landslides in the area through field investigations, the interpretation of remote sensing data, and the literature. Moreover, we used the frequency ratio model to quickly determine the susceptibility of the area to landslides, and chose non-landslide samples for regions with a low susceptibility to them to provide a suitable dataset to train the random forest model. The weights of the parameters were calculated to generate a map of the susceptibility of the area to landslides to analyze the main factors controlling their occurrence. The accuracy of the proposed model was verified by plotting its ROC curve and calculating the AUC of the offline curve. The results of predictions by the model showed that highly susceptible areas were mainly concentrated at the confluence of Yigong Zangbo and Palong Zangbo, and in the large bend of Yarlung Zangbo River. The density of faults, elevation, and slope were, respectively, the three most important features with regard to the susceptibility to landslides, and the AUC of 0.940 of the model showed that it was accurate. Our analysis showed that the study area had a complex geological structure owing to the long-term influence of tectonic action, with large faults on the east and west sides, and the Himalayan frontal fault zone on the south side. With the development of the fault zones, especially in the region that was highly susceptible to landslides, the faults intersected, and multiple extrusions led to the fragmentation of rocks and the development of joints. Therefore, the intersection of Yigong Zangbo and Palong Zangbo as well as the region containing the large bend of Yarlung Zangbo River were highly susceptible to landslides. The proposed model provided a theoretical basis for accurately predicting landslides and appropriately managing disasters in southeast Tibet.

**Key words:** southeast Tibet; random forest model; landslide susceptibility; fault density

滑坡、泥石流等灾害作为世界范围内广泛分布且极具破坏力的地质灾害类型,不但对山区地貌演化过程具有重大影响,而且经常对当地人造成重大的生命财产损失(Aleotti and Chowdhury, 1999),同时也对社会的稳定性与可持续发展带来了严重的损害。中国是世界上受自然灾害影响最为严重的国家之一,尤其在西南地区,地质灾害发生率极高。根据铁永波等(2022)的统计数据,西南地区地质灾害隐患点的比例达到中国灾害总数的 30%,其中滑坡约 5.0 万处,占 57.5%,滑坡成为西南地区的主要地质灾害。

青藏高原东南部由于其特殊的地质构造、地形地貌等原因,受到构造性隆升、地震活动、冰川、气候和人类活动等因素的影响,使得该地区滑坡发育频繁(汪发武等, 2022)。该地区曾发生过著名的易贡大滑坡(李俊等, 2018; Li et al., 2020; 袁浩等, 2023)、帕隆藏布 102 滑坡群(祝介旺等, 2010)、白格滑坡等(王盈等, 2019; 钟政等, 2021),严重影响着当地人民生活生命和生命财产安全。藏东南地区是中国公路、铁路、水电等国家重大工程的重点工作区域。该地区迫切的开发需求和脆弱的生态环境之间的矛盾十分尖锐,因此对于地

质灾害的调查和科考需求不仅是该地区,也是全国人民的迫切期盼,深入研究该地区致灾因素和致灾机理,细化灾害发生条件,成为学者专家的重要任务(张冬梅和张莉, 2019),灾害预警和灾害分区成了当下主要的研究目标。

传统的灾害预警方法,主要依赖专家学者的经验和对环境地质的认识程度。近年来,随着科技的发展,机器学习算法作为一种基于数据驱动的科学方法,逐渐成为了灾害预测的新兴技术(Bzdok et al., 2018)。通过对大量的灾害数据进行分析 and 挖掘,机器学习能够提取出影响灾害发生的关键因素,并建立预测模型。机器学习方法相对于传统的灾害预测,有着速度快、准确度高、可靠性强、有效性强和智能性高的特点,并可运用于灾害的各个领域,如滑坡(李阳春等, 2021; 周萍等, 2022; 刘海知等, 2022)、崩塌(林报嘉等, 2020)、泥石流(高泽民等, 2021; 王向前等, 2023)、冰川等(宋轩宇等, 2023)。滑坡易发性评价方法包括定性评价、半定量评价和定量评价 3 种类型。其中,定量评价在滑坡易发性区划中是最主要的方法。随着 GIS 技术的快速发展,将其应用于滑坡灾害易发性评价领域,有助于进一步提高定量

评价的准确性。王文俊等(2003)基于 GIS 平台,采用模糊综合评判法开展四川省珙县崩塌、滑坡易发性评价,其研究结果与实际情况较为吻合。杨硕等(2021)利用随机森林模型对乌江高陡岸坡滑坡进行了地质灾害易发性评价,模型结果与现场调查一致。邓念东等(2020)在陕西省洋县地区,通过使用频率比—随机森林模型对洋县进行滑坡易发性评价,认为利用该模型划分易发区合理,区内滑坡呈聚集性发育。

常用的易发性区划评价方法包括层次分析法、信息量模型、频率比、逻辑回归、随机森林、支持向量机和人工神经网络等。在机器学习算法中,不同方法之间也存在着不同的差异。周萍等(2022)在对四川省理县的滑坡易发性评价研究中,发现随机森林模型相较于径向基函数神经网络模型有着更为可靠的预测效果,在区域地质灾害易发性评价中更具有优势。在以鲁甸地震为例的地震滑坡易发性评价中,其结果显示,随机森林模型相较于逻辑回归、K 邻近和朴素贝叶斯在灵敏度、精确度、准确度、ROC 曲线和 AUC 值上均表现较好(吉日伍呷等,2022)。黄发明等(2022)在对陕西省延长县滑坡预测分析中发现,相较于其他机器学习模型,随机森林模型可靠性更高。综上所述,相较于其他易发性区划评价方法,随机森林模型有

着更好的预测效果,因此,本文对藏东南地区滑坡易发性评价选取随机森林模型进行预测,并分析其主控因素,为该地区的滑坡灾害发生预测和滑坡防治提供有效的理论指导(Pourghasemi and Kerle, 2016; Merghadi et al., 2020)。

### 1 区域背景

研究区位于西藏东南部,包括林芝市波密县全域、墨脱县、米林县和巴宜区部分地区。研究区内山势陡峻、群山重叠、沟谷纵横、构造发育,按地貌形态属高山河谷地貌,帕隆藏布、易贡藏布和雅鲁藏布江贯穿整个地区。海拔在 599~7 782 m 之间,岭谷高差一般在 2 000 m 以上,最大岭谷高差达 5 409 m;最高峰是南迦巴瓦峰,海拔 7 782 m。区内海拔 5 000 m 以上的高峰有数座,而在海拔较低的墨脱境内,一些谷地海拔却不足 600 m。

研究区总面积约 27 039 km<sup>2</sup>,在构造上包括念青唐古拉块体、雅鲁藏布江缝合带和喜马拉雅造山带。其构造变形以挤压逆冲和走滑为特征,穿过该区域的断裂带主要包括嘉黎深切断裂、迫龙—旁辛断裂、墨脱断裂、雅鲁藏布江断裂、易贡—鲁朗断裂、达德—阿尼扎断裂和日卡—索通断裂。由于特殊的地质构造,导致该地区地质灾害发育。研究区地理位置及概况如图 1 所示。

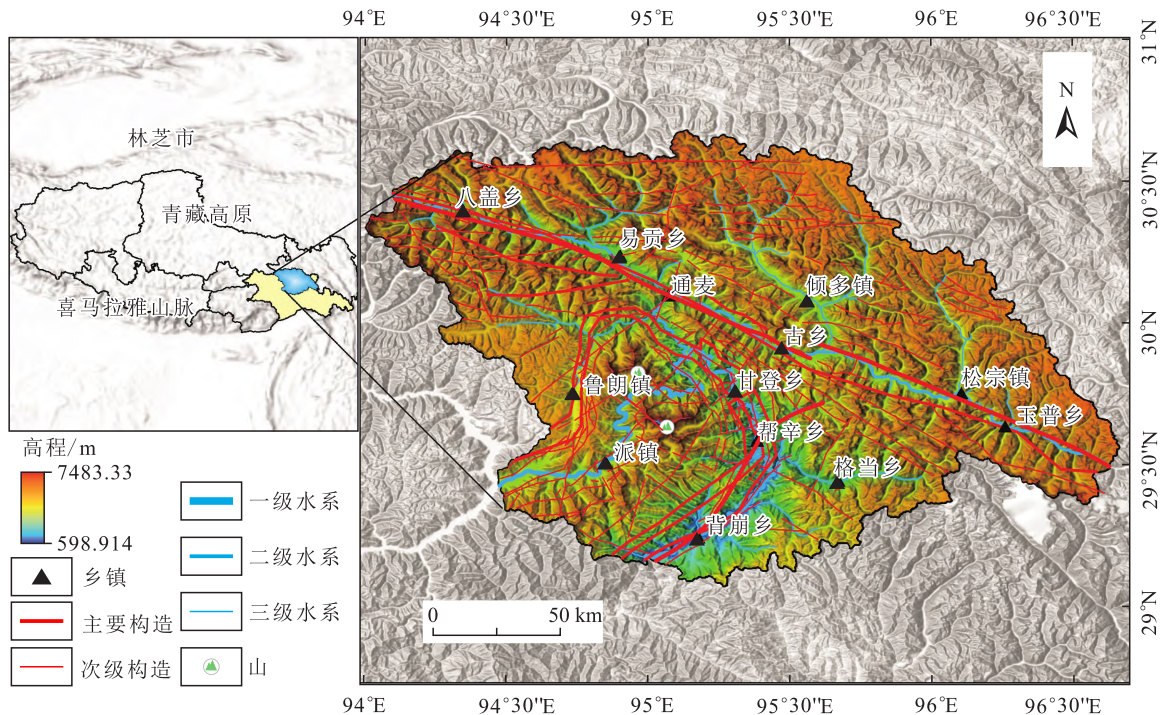


图 1 研究区概况  
Fig. 1 Overview of the study area

## 2 研究方法

### 2.1 频率比模型

频率比模型(FR)是一种用于统计推断的概率模型。它基于观察到的事件频率,通过计算事件之间的比率来推断概率。其值(FR 值)表示各级评价因子对滑坡发生的重要程度,其计算公式为

$$FR_i = \frac{\frac{n_i}{N}}{\frac{s_i}{S}} \quad (1)$$

式中: $n_i$  为分类内滑坡面积; $N$  为研究区滑坡总面积; $s_i$  为分类面积; $S$  为研究区总面积。若  $FR \leq 1$ ,表明该分区对滑坡灾害发生的作用较弱; $FR > 1$ ,表明该分区对与滑坡灾害发生的作用较强(马啸等,2020)。

### 2.2 随机森林模型

随机森林是一种集成(Ensemble)学习的方法,其通过以决策树为基学习器,利用随机特征选择和自助采样技术进行训练,在保持决策树的低偏差的同时,控制了过拟合问题。随机森林是决策树的集成,每次从收集的训练样本中随机且有放回地选择一部分样本用于生成决策树训练子集。并最终将决策树训练子集组合起来进行投票,最终分类结果由得票数所决定(图 2),其主要分为以下 5 步(武晓岩和李康,2006;李亭等,2014;郑迎凯等,2020)。

(1)从原始数据中使用自助采样法(即采用有放回方式随机抽取训练数据集中的样本)多次构造出  $K$  个大小相等的训练数据子集。

(2)对于每个训练数据子集,使用 CART 算法构建一棵决策树。在构建每一个节点时,随机从所有特征中选择(预先设定) $M$  个特征进行划分(其中  $M$  一般远小于总特征数),从而使得每个节点的形成独立性得到提高。

(3)对于每个节点,从  $M$  个特征中选取最优特征进行分裂。最优特征通常按照信息增益或者基尼指数等准则进行计算,并选取能够获得最大准则值的特征来进行分裂。

(4)重复执行 2~3 步,直到得到  $K$  棵决策树。

(5)预测时,将待预测样本输入到生成的所有决策树中,得到  $K$  个预测结果,则最终的分类或回归结果由这  $K$  个结果投票决定。

相对于传统的滑坡易发性评价机器学习方

法,比如决策树、支持向量机和人工神经网络模型,随机森林模型采用了对样本和特征进行随机抽样的技术。这 2 个过程的随机抽样对降低模型数据噪声和异常值敏感性有着显著效果,在提高预测准确度方面大有帮助,也成功解决了过拟合问题。

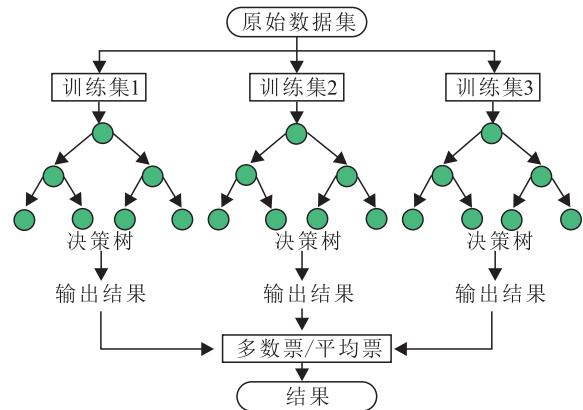


图 2 随机森林模型流程图

Fig. 2 Flowchart of the random forest model

### 2.3 自然间断点法

自然间断点法,也称为 Jenks 自然断点法,是一种用于数据分类和分级的方法。其原理基于最小化类别内部差异且最大化类别之间差异的思想,该方法的主要步骤如下。

(1)将数据按照某个特征值进行排序。

(2)根据排序后的数据,尝试所有可能的分割点,并计算每个分割点对应的类别内部差异的度量(通常使用方差、平方误差和离差平方和等)。

(3)找到使得类别内部差异最小化且类别之间差异最大化的最佳分割点。

(4)将数据集根据最佳分割点划分成多个类别。

(5)重复上述步骤,直至达到预定的分割数目。

通过自然间断点法,数据可以被划分为若干个不同的类别,使得每个类别内部的差异最小,并且不同类别之间的差异最大化。这样做的目的是为了能够更好地展示数据的分布特征,有助于理解和解释数据的含义。

## 3 原始数据选取

### 3.1 滑坡样本选取

藏东南地区由于其独特的构造背景条件、复

杂的地形地貌、多样的气候分区,地质灾害非常发育,具有明显的区域性和地形特征。从空间和时间分布规律来看,藏东南地区的滑坡明显受到该区地质地貌背景和气候条件的控制。藏东南地区易发生滑坡的区段都是河谷深切、山势陡峻、相对高差大的地区,尤其是高山和极高山地区。

研究区内共分布了 306 处滑坡,根据《滑坡防治工程勘查规范》(中华人民共和国国家质量监督检验检疫总局和中国国家标准化管理委员会, 2016),按滑坡体积可分为小型、中型、大型、特大型和巨型滑坡 5 种类型。其分布图如图 3 所示。

306 处滑坡主要为中小型滑坡,部分为大型滑坡,零星分布有巨型和特大型滑坡。其中小型滑坡 84 处,占比 27.5%;中型滑坡 115 处,占比 37.6%;大型滑坡 97 处,占比 31.7%;特大型滑坡 9 处,占比 2.9%;巨型滑坡 1 处,占比 0.3%。研究区内已查明物质组成的滑坡基本为土质滑坡,岩质滑坡仅分布有 4 处,分别为通易公路 8#(94°57'2.98"E,30°9'44.41"N)、拉月村委会背后滑坡(94°51'39.81"E,29°58'51.84"N)、菊汤蒙崩滑群(95°18'19.36"E,29°43'26.73"N)、达木乡珠村滑坡(95°25'40.21"E,29°31'1.29"N)。

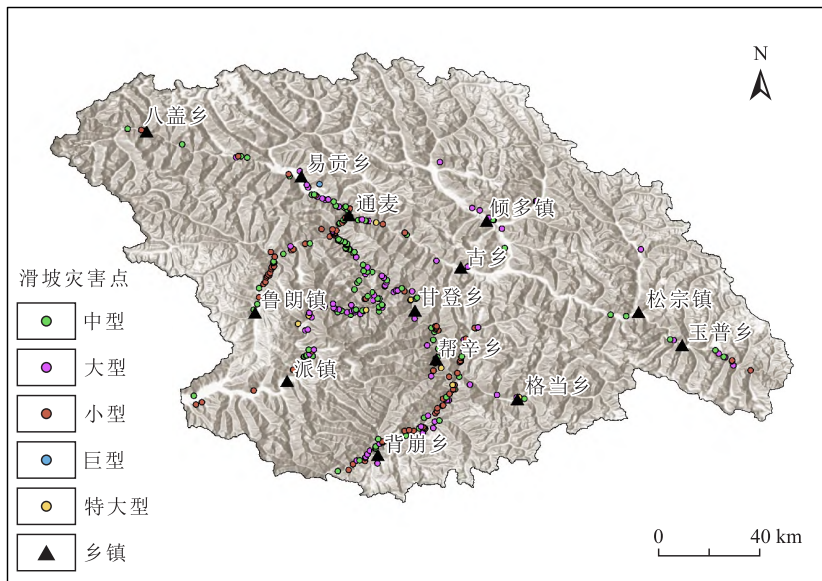


图 3 滑坡样本分布图  
Fig. 3 Distribution of landslides

### 3.2 评价因子选取及因子分析

滑坡灾害的发育是一个非常复杂的非线性过程,受到内外动力因素的控制。内部因素决定了滑坡的发展和发育状况,包括地质构造、坡体结构、地层岩性等,是滑坡发育的基础;外部因素则会诱发滑坡的发生,包括人类活动、大气降水和地震等(Ling et al., 2014)。本次选取滑坡评价因子时,考虑了藏东南地区特殊的地质构造条件和气候因素(张炳华等, 2023)。由于研究区多为无人区,人类活动影响较少,无法体现其差异性,因此,不选用人类活动作为其评价因子。而关于藏东南特别发育的地质构造方面,常见的方法是寻找灾害点与断层间的距离,但这不太科学,因为当遇到灾害隐患点周边有很多断层的时候,该指标难以确定。尤其是当多条断裂带穿过坡体和一条

断裂带穿过坡体在传统指标上反映不出差异。所以,较为合理地衡量地质构造的重要指标之一,应是断层密度。

首先,高程为滑坡发生的启动条件,为滑坡坡体滑动提供势能;其次,高程也反映了人类活动的范围,人们总是在海拔高度合适的范围活动。

坡向可以将坡分为阴坡和阳坡,或迎风坡和背风坡。不同的坡向有着不同的差异,包括降雨和植被等的不同。研究区滑坡主要发生在日照时间长的南坡。

坡度控制滑坡发育的原因主要是土体的剪应力增加,抗剪强度降低。当坡度处于一定的范围,会使得剪应力出现临界点,在其他因素的耦合作用下,滑坡就会发生滑动。研究区滑坡主要发生在 33°~42°之间。

地形湿度指数(TWI)是描述地表水分状况的方法。它基于地形特征,考虑了地表降水量、蒸发量以及土壤水分等因素,并通过计算高程梯度来反映地形的影响。其计算公式为:

$$TWI = \ln\left(\frac{\alpha}{\tan \beta}\right) \quad (2)$$

式中: $\alpha$ 是上游汇流面积; $\beta$ 是坡度角,可通过倾斜角度进行估算。TWI越高,说明地形越湿润。

地表粗糙度是指地面或其他物体表面的不平整程度或凹凸不平的程度。

地形起伏度( $R_{Fi}$ )代表该区域的地表起伏程度,其大小可以反映出该区域的地貌特征和地形复杂程度。其计算公式为:

$$R_{Fi} = H_{max} - H_{min} \quad (3)$$

式中: $H_{max}$ 指在范围内的最大高程值; $H_{min}$ 指在范围内的最小高程值。

曲率是用来衡量地质形态的变化和地表变形的性质。曲率的大小代表边坡的类型,当曲率 $<0$ 时为凸型坡,曲率 $>1$ 时为凹型坡,曲率 $=1$ 时为直线坡。

地震加速度是滑坡启动的重要条件之一,其

会导致土壤中颗粒的相互作用变弱,引起滑动平面上的应力变化,并且还可能会引起土体液化,剪切失稳等现象。

岩性是指岩石的种类和组成特征。多数学者认为,软岩因其内部结构不够紧密,缺乏足够的支撑力,当外界条件(例如降雨)给这些物质带来压力时,就很容易发生变形和滑动(左群超等,2018)。

断层是由于地震发生时,在岩石之间积累的压力被迫释放而形成的。其可以增加地质构造的复杂性,导致岩石、土壤力学特性的变化,同时可以改变区域的水文结构,对研究区的水循环和排水系统产生影响。断层密度表示单位面积内断层的长度,反映出断层在单位范围内的发育情况(Hermanns and Strecker,2011)。

年降水量是指研究区在一年内收集到的降水总量,以毫米为单位。它是评估该地区气候和水资源状况的重要指标之一。年降水量可根据降水测量仪器(如雨量计)记录的数据进行测量并计算得出。降雨量增加会导致土壤变得饱和,减少土壤的稳定性,从而增加滑坡的风险(表 1)。

表 1 原始数据获取来源  
Table 1 Sources of raw data

数据	评价因子编号	获取方式
断层		通过 1:25 万地质图(下载网站 <a href="http://dcc.ngac.org.cn/">http://dcc.ngac.org.cn/</a> )和文献提取
高程	A	通过 ASF Data Search( <a href="https://search.asf.alaska.edu/">https://search.asf.alaska.edu/</a> )下载
坡向	B	通过数字地形高程在 ArcGIS 软件中生成
坡度	C	
地形湿度指数	D	
地表粗糙度	E	
地形起伏度	F	
曲率	G	
地震加速度	H	来源于中国地震动参数区划图 GB18306—2015(网站 <a href="https://www.gb18306.net/">https://www.gb18306.net/</a> )
岩性	I	通过 1:25 万地质图(下载网站 <a href="http://dcc.ngac.org.cn/">http://dcc.ngac.org.cn/</a> )提取
断层密度	J	通过 ArcGIS 软件生成渔网网格法计算断层密度
年均降水量	K	通过国家气候中心( <a href="http://www.ncc-cma.net/cn/">http://www.ncc-cma.net/cn/</a> )下载

### 3.3 相关性分析

滑坡易发性评价模型中,需要保证各因子之间的相关性较小,因此需要进行相关性分析。计算相关性系数能反映各评价指标之间的相关性,本研究使用斯皮尔曼法计算各因子之间的相关性系数( $R$ )。根据系数大小,将相关性分为线性相关( $|R|=1$ )、高度相关( $0.8 \leq |R| < 1$ )、

显著相关( $0.5 \leq |R| < 0.8$ )、低度相关( $0.3 \leq |R| < 0.5$ )、微弱相关( $|R| < 0.3$ )。具体计算公式如下:

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (4)$$

式中:  $x, y$  为 2 个变量中的对应元素;  $N$  为样本总量。

根据 SPSS 软件中的相关性分析, 导入各评价因子的数据并计算出相关系数, 将显著相关及

以上的评价因子剔除掉, 根据结果可知地形起伏度、年均降水量和高程、坡度、地表粗糙度相关性较高。因此, 将这 2 项评价指标剔除, 最终得到 9 项评价因子进行易发性评价(图 4)。

	A	B	C	D	E	F	G	H	I	J	K
A	1	-0.048	0.19	-0.091	0.182	0.178	0.121	-0.257	-0.036	-0.222	-0.505
B	-0.048	1	-0.11	-0.032	-0.099	-0.095	-0.032	0.117	0.17	-0.027	-0.033
C	0.19	-0.11	1	-0.251	0.292	0.977	-0.017	-0.156	-0.068	-0.015	-0.17
D	-0.091	-0.032	-0.251	1	-0.253	-0.268	-0.126	-0.145	-0.015	0.042	-0.009
E	0.182	-0.099	0.292	-0.253	1	0.968	-0.02	-0.156	-0.066	-0.02	-0.155
F	0.178	-0.095	0.977	-0.268	0.968	1	-0.035	-0.126	-0.064	-0.039	-0.159
G	0.121	-0.032	-0.017	-0.126	-0.02	-0.035	1	-0.049	0.031	-0.034	0.018
H	-0.257	0.117	-0.156	-0.145	-0.156	-0.126	-0.049	1	0.161	0.012	0.317
I	-0.036	0.17	-0.068	-0.015	-0.066	-0.064	0.031	0.161	1	0.019	0.105
J	-0.222	-0.027	-0.015	0.042	-0.02	-0.039	-0.034	0.012	0.019	1	0.22
K	-0.505	-0.033	-0.17	-0.009	-0.155	-0.159	0.018	0.317	0.105	0.22	1

图 4 评价因子相关性热力图

Fig. 4 Heat map of correlations among the evaluative factors

通过上述相关性分析, 剔除后的评价因子剩余高程(A)、坡向(B)、坡度(C)、地形湿度指数(D)、地表粗糙度(E)、曲率(F)、地震加速度(G)、岩性(H)、断层密度(I), 作为原始数据输入(图 5)。

### 3.4 非滑坡样本选取

为了让模型能够输出更精确的预测结果, 在建模前需要选取一定量的正负样本, 正样本为本次研究获取的原始数据, 负样本为非滑坡点。负样本的选取非常关键, 能够帮助模型更好的泛化训练数据, 且能确保模型的可靠性和鲁棒性。但是由于在预测前不能准确地确定滑坡易发区的位置, 常常在随机选取的时候会存在部

分非滑坡点落在易发区, 导致模型预测精度降低。

因此, 本次研究为了满足模型的准确性, 通过计算各个因子的  $FR$  值(表 2), 并进行整合得到基于频率比模型的易发性指数, 快速识别研究区易发区的模糊范围(图 6)(朱宇航等, 2022)。根据自然断点法, 将该区域的滑坡分成了 5 个等级(图 7), 包括低、较低、中、较高和高易发区。尽管无法完全准确地刻画研究区内的滑坡易发区域, 但相较于传统的随机选点有着更高的准确性, 通过该模型的鉴定, 在这些低易发区和较低易发区的范围内进行非滑坡点的随机抽样, 从而最大程度地提高所选点的准确性。

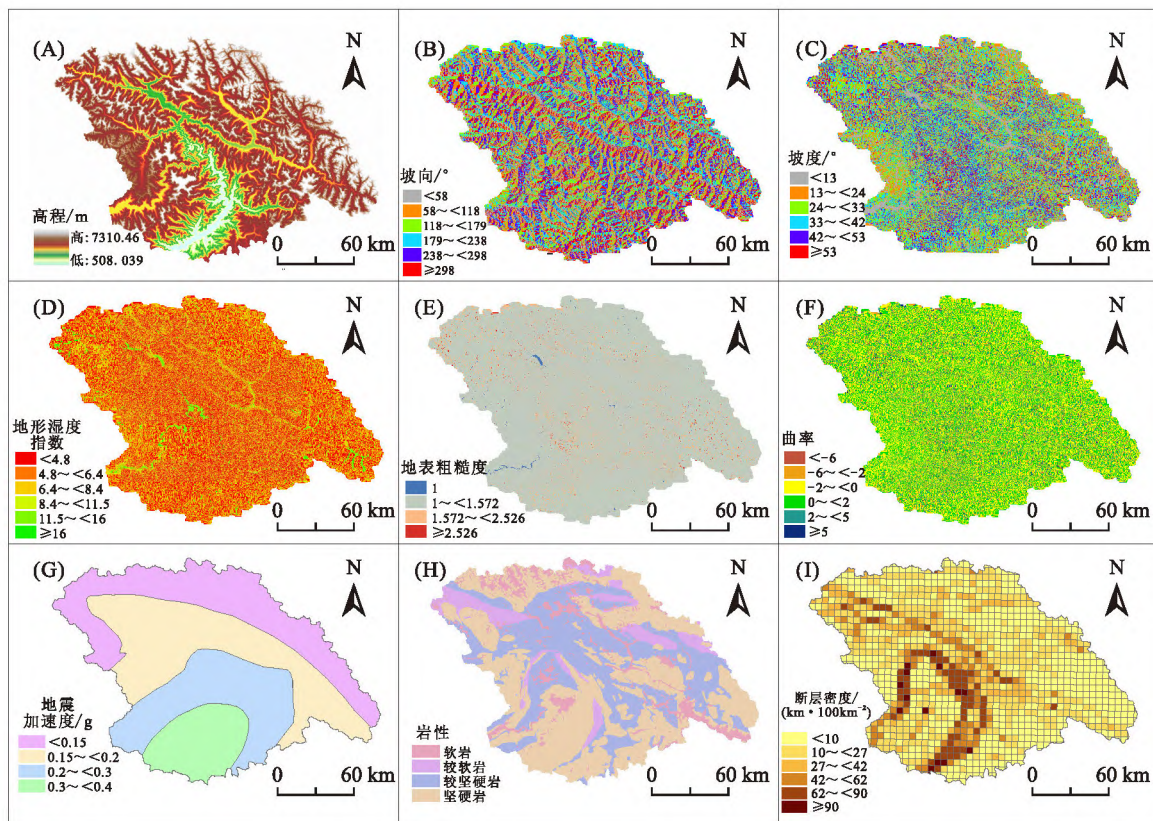


图 5 滑坡评价因子图

Fig. 5 Map of factors influencing landslides

(A) 高程; (B) 坡向; (C) 坡度; (D) 地形湿度指数; (E) 地表粗糙度; (F) 曲率; (G) 地震加速度; (H) 岩性; (I) 断层密度

表 2 评价因子频率比值表

Table 2 Frequency ratios of the evaluative factors

因子	分类	FR	因子	分类	FR	因子	分类	FR
A	<2214	7.51	D	<4.8	1.09	G	<0.15	0.07
	2 214~<3 120	3.52		4.8~<6.4	0.75		0.15~<0.2	1.69
	3 120~<3 786	0.48		6.4~<8.4	0.90		0.2~<0.3	1.12
	≥3286	0.04		8.4~<11.5	0.76		≥0.3	1.50
B	<58	0.56	H	软岩	0.31	I	<19	0.33
	58~<118	0.90		较软岩	1.71		19~<38	0.96
	118~<179	1.35	较坚硬岩	1.69	38~<57		3.26	
	179~<238	2.02	坚硬岩	0.60	57~<76		4.01	
	238~<298	0.76	F	<-6	1.25		76~<96	1.54
	≥298	0.30		-6~<-2	0.82		≥96	0.16
C	<13	0.22		-2~<0	1.16	I	19~<38	0.96
	13~<24	0.45		0~<2	0.82		38~<57	3.26
	24~<33	1.04	2~<5	1.24	57~<76		4.01	
	33~<42	1.80	≥5	0.75	76~<96		1.54	
	42~<53	0.96	E	1	1.54		≥96	0.16
	≥53	1.10		1~<1.6	1.01		I	19~<38
C	<13	0.22		1.6~<2.5	0.80	38~<57		3.26
	13~<24	0.45		≥2.5	2.39	57~<76		4.01
	24~<33	1.04	F	<-6	1.25	76~<96		1.54
	33~<42	1.80		-6~<-2	0.82	≥96	0.16	
	42~<53	0.96		-2~<0	1.16	I	19~<38	0.96
	≥53	1.10		0~<2	0.82		38~<57	3.26
C	<13	0.22	2~<5	1.24	57~<76		4.01	
	13~<24	0.45	≥5	0.75	76~<96		1.54	
	24~<33	1.04	E	1	1.54		≥96	0.16
	33~<42	1.80		1~<1.6	1.01		I	19~<38
	42~<53	0.96		1.6~<2.5	0.80	38~<57		3.26
	≥53	1.10		≥2.5	2.39	57~<76		4.01

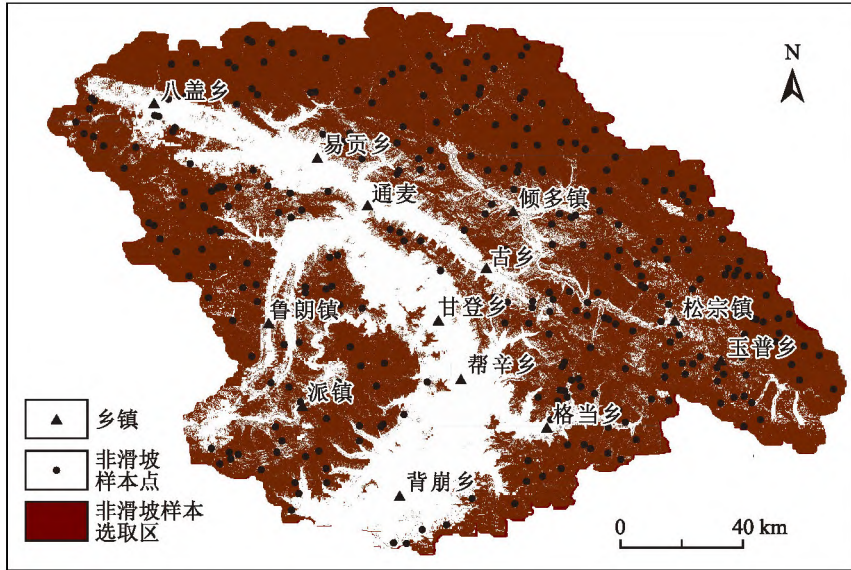


图 6 非滑坡样本分布图  
 Fig. 6 Distribution of non-landslide samples

## 4 模型运算结果

### 4.1 滑坡易发性结果

通过随机森林模型的预测,并运用 ArcGIS 叠加整合 (Pawluszek and Borkowski, 2017), 利用自然间断点法将其分为低易发性、较低易发性、中易发性、较高易发性和高易发性 5 个部分 (Chen et al., 2016), 从而将其生成为研究区滑坡易发性区划图 (图 7)。根据计算,其滑坡低易发区占研究区总面积的 22.71%, 约 6 139.3 km<sup>2</sup>、较低易发区占研究区总面积的 37.83%, 约

10 229.5 km<sup>2</sup>; 中易发区占研究区总面积的 23.43%, 约 6 335.6 km<sup>2</sup>; 较高易发区占研究区总面积的 10.47%, 约 2 830.9 km<sup>2</sup>; 高易发区占研究区总面积的 5.56%, 约 1 503.4 km<sup>2</sup> (图 7)。研究区中 93% 的滑坡发生在中易发区以上地区, 各个分区的滑坡密度在易发性由低到高的区域内, 滑坡密度呈逐渐上升的趋势。研究区高易发区和较高易发区主要集中在易贡藏布、帕隆藏布交汇处和雅鲁藏布江大拐弯区域。该地区地质条件复杂, 断层密度大, 侵蚀作用强烈, 气候交替变化频繁, 导致该地区地质灾害发育严重 (表 3)。

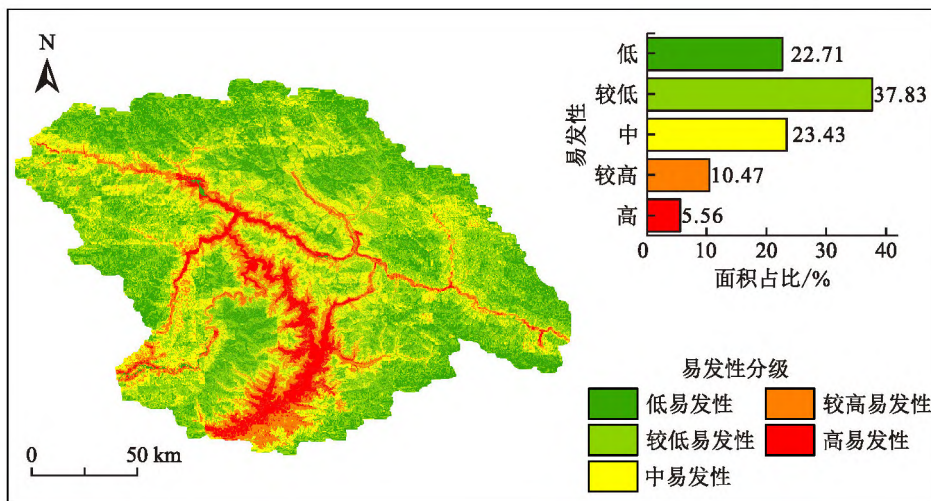


图 7 研究区滑坡易发性分级图  
 Fig. 7 Map of classification of the susceptibility of the study area to landslides

表 3 滑坡易发性结果分区统计表  
Table 3 Statistics of zoning for susceptibility to landslides

易发性分区	分区面积/km <sup>2</sup>	面积占比/%	分区内滑坡数/处	分区内滑坡占比/%	分区内滑坡密度/(处·100 km <sup>-2</sup> )
低易发区	6139.3	22.71	1	0.33	0.016
较低易发区	10229.5	37.83	4	1.31	0.039
中易发区	6335.6	23.43	14	4.58	0.220
较高易发区	2830.9	10.47	69	22.55	2.437
高易发区	1503.4	5.56	218	71.24	14.5

#### 4.2 评价因子重要性排序

根据决策树各个节点的平均基尼值的减少量,来计算评价因子的重要性,并对所计算的因子权重进行归一化处理。模型预测结果显示,断层密度、高程和坡度对滑坡发育起着主要控制作用,坡向和曲率对该地区滑坡发育影响较小,通过模型计算结果对本次选用的 9 个评价因子进行重要性排序,因子重要性排序从高到低依次为断层密度(35.63%)、高程(21.96%)、坡度(13.68%)、地形湿度指数(6.24%)、地震加速度(5.53%)、岩性(4.95%)、地形粗糙度(4.39%)、曲率(3.87%)、坡向(3.75%)。而对于影响滑坡发育普遍认为较为重要的岩性而言,通常认为,滑坡主要发生在软弱岩层,但是在构造发育和地震频发的区域,硬岩区域也同样发育滑坡,例如汶川地震区,大型滑坡主要发生在硬岩分布区,这是强震等内部因素诱发地质灾害与天然或暴雨诱发地质灾害的最大区别。虽然较软岩区域滑坡较发育,但相差不大,因此,在各因子排序中,岩性在重要性排序中位置较后是合理的。

#### 4.3 ROC 曲线验证

在本次随机森林滑坡模型的预测中,AUC 值为 0.940,表明该模型预测效果较好且具有较高的可靠性。ROC 曲线是一种常用的检测方法,其横坐标表示 1-特异性,纵坐标表示灵敏度。ROC 曲线数值范围为 0~1,曲线越陡峭代表模型预测精度越高(武雪玲等,2013)。模型的优劣可根据 ROC 曲线线下面积(AUC)取值来判断。若 AUC 值越高则该训练模型的预测能力越好。例如 AUC>0.9,代表预测效果非常好;0.8<AUC≤0.9,代表预测效果较好;0.5<AUC≤0.8,代表该模型预测效果较差;当 AUC 等于 0.5,说明其模型结果不具有参考价值;而当 AUC 小于 0.5,则说明模型不符合真实情况(扈秀宇等,2019)。

#### 4.4 结果验证与分析

##### 4.4.1 结果验证

为验证模型的准确性,选择了河谷深切、山势陡峻且相对高差较大的河段作为重点研究区域。在这些地理条件特殊的区域中,精选了多处滑坡进行详细调查和分析。有助于加强对滑坡灾害形成机制的理解,并提供有效案例来验证模型的适用性。河谷深切以及山势陡峻的特征使得这些地区更容易发生滑坡灾害,因此通过针对这些地方的研究,可以更好地了解滑坡灾害的发展规律以及其与地理条件之间的关联。基于此,分别选取了研究区典型的易贡藏布流域的易贡滑坡、帕隆藏布流域的 102 滑坡和野外调查发现的雅鲁藏布江流域的巴日村滑坡、易贡藏布流域的 K47 滑坡、H03 滑坡(图 8)。

##### (1) 典型案例——易贡滑坡

易贡滑坡发生在易贡藏布左岸(图 9),平均坡度为 50°,其投影面积为 0.57 km<sup>2</sup>,平均厚度为 180 m,体积达 3×10<sup>8</sup> m<sup>3</sup>。滑坡附近有 5 条断层,其中 3 条断层穿过扎木弄沟流域内。断层密度为 67.165 km/100 km<sup>2</sup>,是区域平均密度的 3.42 倍。由于扎木弄沟处于嘉黎深切断裂带和易贡—鲁朗断裂交汇点附近,在其沟源有喜马拉雅早期的花岗岩侵入,又受到了持续的构造作用的影响,特别是嘉黎深切断裂带、易贡—鲁朗断裂和达嫩—则普断裂复合控制作用,导致花岗岩非常破碎,有利于滑坡泥石流灾害的发生。2000 年易贡滑坡发生前,区域内存在地貌凸出体 155.87×10<sup>6</sup> m<sup>3</sup>。这些地貌凸出体为滑坡的形成提供了物源条件,并且陡峻的边坡为滑坡的发生提供了势能条件。

##### (2) 典型案例——102 滑坡

102 滑坡发生在帕隆藏布右岸(图 10)。滑坡前缘最低海拔 2 120 m,后缘最高海拔 2 525 m,相对高差达 405 m。滑坡体长 550 m,宽 380 m,平均坡度为 35°,滑坡面积为 3.81 km<sup>2</sup>,体积为

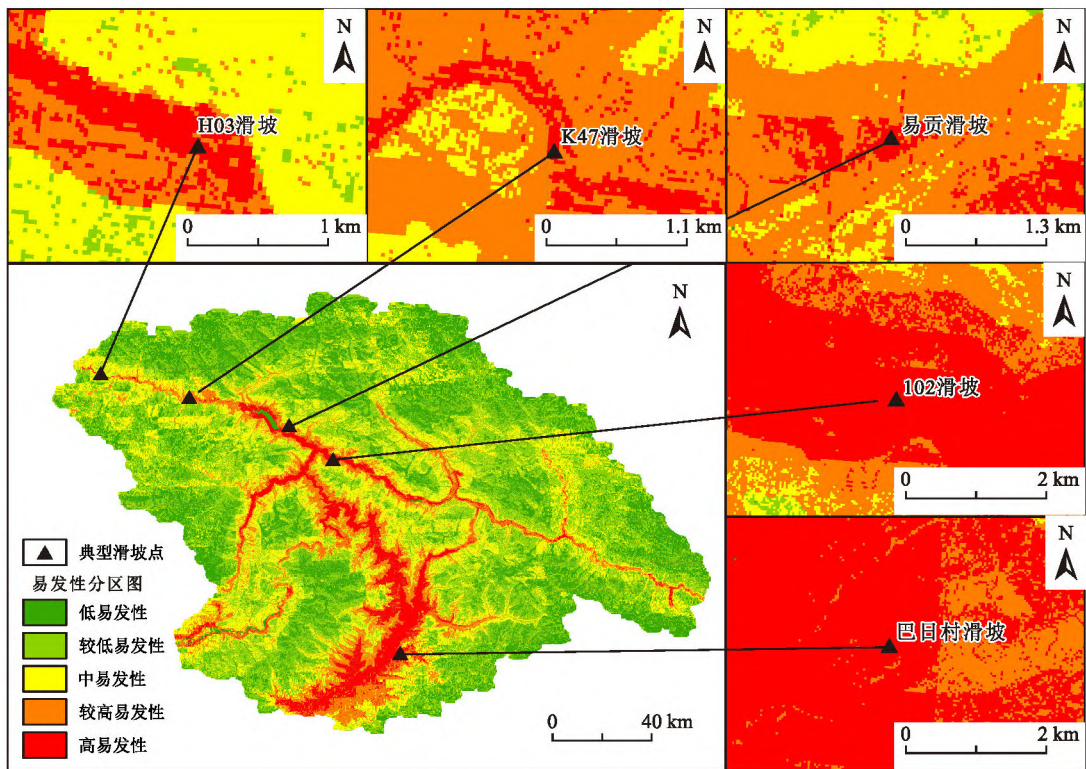


图 8 典型滑坡易发性局部图  
Fig. 8 Map of typical locations that were susceptible to landslides

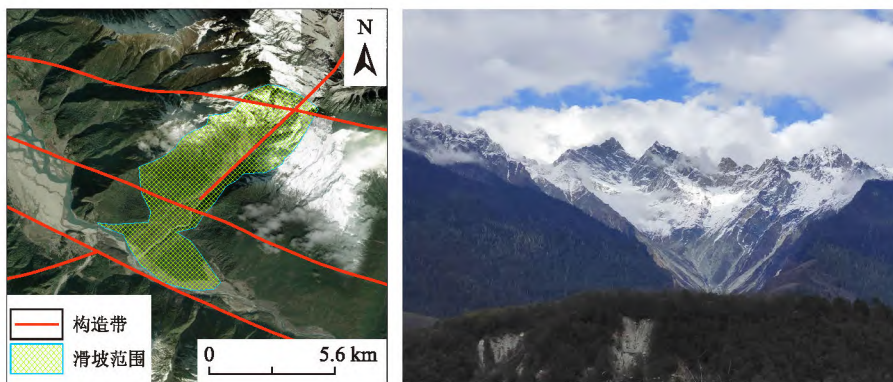


图 9 易贡滑坡图  
Fig. 9 Map of the landslide in Egon

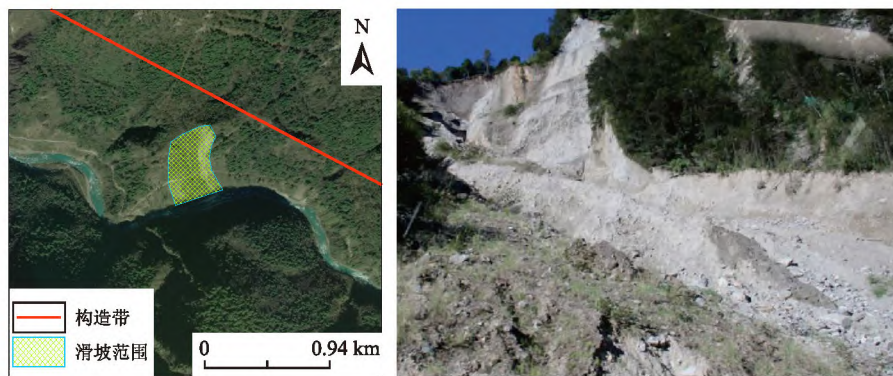


图 10 102 滑坡图  
Fig. 10 Map of 102 landslides

$510 \times 10^4 \text{ m}^3$ 。在区域地质构造上,恰好处于板块构造的结合部位,周围有 4 条规模较大的断裂带,并且断层密度为  $50.472 \text{ km}/100 \text{ km}^2$ ,是平均断层密度的 2.57 倍,在海拔 2 450 m 高程处平台至河床间的陡坡上,由老冰碛组成的平台与现代帕隆藏布河面有 400 m 的高差,使斜坡地带的物质具有很高的势能。斜坡地带坡体基本均大于  $30^\circ$ ,即 102 滑坡在发生灾害前存在大量地貌凸出体,总量约为  $1.8 \times 10^7 \text{ m}^3$ 。这些地貌凸出体为滑坡的形成提供了物源条件,并且陡峻的边坡为滑坡的发生提供了势能条件。

(3)野外滑坡案例——巴日村滑坡

巴日村滑坡位于西莫日河的右岸(图 11),靠近墨脱县巴日村附近。滑坡的海拔范围从最低点 991 m 到最高点 1 323 m,相对高差为 332 m。滑坡的平均坡度为  $38^\circ$ ,面积约为  $21.02 \times 10^4 \text{ m}^2$ ,体积估计为  $445.6 \times 10^4 \text{ m}^3$ 。在区域地质构造方面,滑坡被墨脱次级断裂带穿过,并且周围还存在着 4 条断裂带,断层密度达到了  $43.89 \text{ km}/100 \text{ km}^2$ ,是平均断层密度的 2.24 倍。滑坡中部存在一些地貌突出的形态,给滑坡发育提供了势能条件。目前,滑坡仅表层部分发生了滑动,并且滑坡后缘出现了明显的裂缝,整个滑坡滑动可能会导致西莫日河的堵塞。

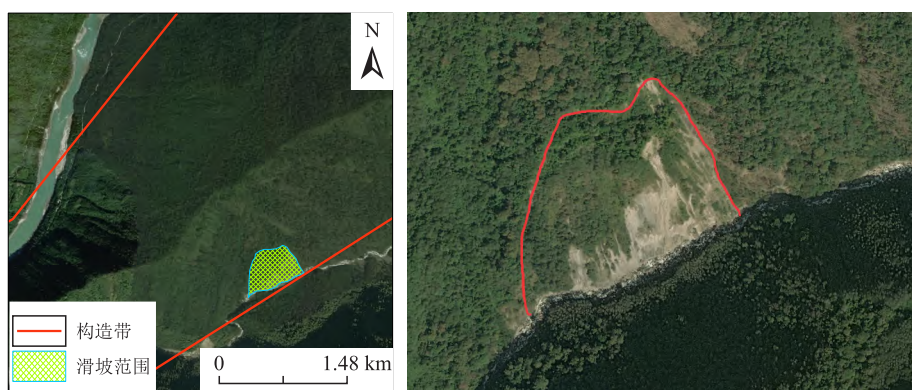


图 11 巴日村滑坡图  
Fig. 11 Map of the landslide in Bari village

(4)野外滑坡案例——K47 滑坡

K47 滑坡位于八盖乡嘛村附近(图 12),其前缘海拔最低约为 2 603 m,后缘海拔最高达到 2 830 m,高差约为 227 m。滑坡面积为  $7.04 \times 10^4 \text{ m}^2$ ,平均厚度为 15 m,平均坡度为  $51^\circ$ 。据估计,滑坡的体积约为  $105.7 \times 10^4 \text{ m}^3$ 。根据初步判定,滑坡属于浅层牵引式滑坡类型。在区域构造方面,滑坡周

围存在 3 条断层,其中嘉黎断裂带贯穿滑坡区域,断层密度为  $65.13 \text{ km}/100 \text{ km}^2$ ,是平均断层密度的 3.32 倍。由于断裂带的影响,滑坡区域中的基岩主要是糜棱岩。现场调查显示,在滑坡发生之前,滑坡上存在大量突出的地貌形态。这些突出形态为滑坡的形成提供了物源,并且陡峭的边坡为滑坡的发生提供了势能条件。

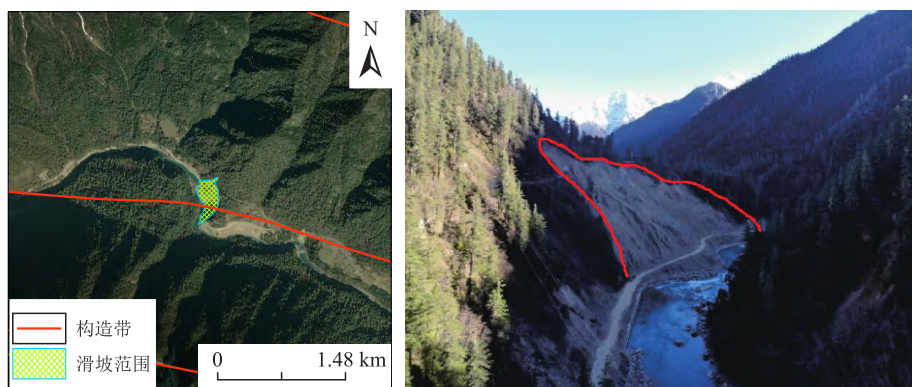


图 12 K47 滑坡图  
Fig. 12 Map of the K47 landslide

(5)野外滑坡案例——H03 滑坡

H03 滑坡位于八盖乡日卡村附近(图 13), 303 省道 K989 处。滑坡前缘最低海拔为 3 006 m,后缘最高海拔为 3 203 m,相对高差为 197 m。滑坡面积约为  $8.03 \times 10^4 \text{ m}^2$ ,长度为 336 m,平均坡度为  $39^\circ$ ,滑坡体积估计为  $120.4 \times 10^4 \text{ m}^3$ ,属于大型滑坡。在区域构造

方面,嘉黎断裂带穿过了滑坡区域,相对断层密度为  $40.336 \text{ km}/100 \text{ km}^2$ ,是平均断层密度的 2.06 倍。受到断裂带的影响,滑坡后缘垮塌,出现地貌堆积体,滑坡体上裂隙发育,稳定性较差,滑坡中部有道路穿过,前缘受到易贡藏布河冲刷,为滑坡的发育提供了势能和动能条件。

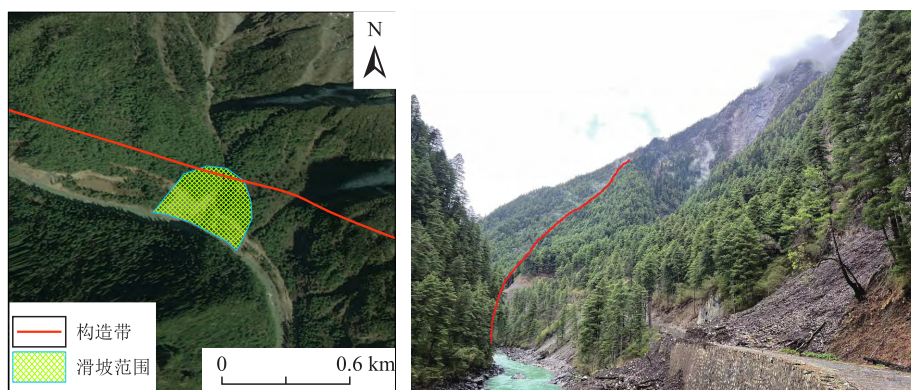


图 13 H03 滑坡图

Fig. 13 Map of H03 landslide

4.4.2 滑坡发育与主控因素分析

为了深入探究研究区滑坡灾害的发育与主控因素,本研究选择将研究区河道细分为 13 段,并通过对灾害发展密度的统计,对研究区滑坡灾害的主要影响因素进行系统分析(表 4),更全面地理解滑坡灾害的形成机制和空间分布规律。首先,将研究区按照河道的特征和地貌条件进行划

分。通过分段的方式,有助于捕捉到河道沿线不同区域的滑坡活动情况,从而为进一步探讨滑坡灾害的发展提供有力的数据基础。同时,通过统计每个河道段的滑坡灾害发育密度,可以准确评估不同区域的滑坡灾害程度及其对整个研究区的影响程度;通过分析各段之间的差异,可以揭示出可能的主控因素并加以验证。

表 4 研究区滑坡样本分布表  
Table 4 Distribution of landslides in the study area

分段	河道长度/km	数量/处	占比/%	密度/(处·km <sup>-1</sup> )
通麦至拉月曲汇口	12.4	17	5.56	1.37
培龙沟至结东	23.1	25	8.17	1.09
隆巴曲至通麦	21.3	20	6.54	0.94
则隆弄至错嘎列卜	36.1	26	8.50	0.72
帮辛乡至背崩乡	78.1	51	16.67	0.65
错嘎列卜至巴玉村	60.1	35	11.44	0.58
巴玉村至帮辛乡	52.8	27	8.82	0.51
101 道班至通麦	19.1	9	2.94	0.47
鲁朗河上游至拉月曲汇口	90.7	28	9.15	0.31
金珠藏布	88.4	24	7.84	0.27
八盖乡至隆巴曲	90.9	11	3.59	0.12
羌纳乡至则隆弄	59.6	6	1.96	0.10
米堆至 101 道班	292.1	27	8.82	0.09

研究区内发育了306处滑坡,通过统计研究区滑坡样本分布及密度可知,该区域内的通麦至拉月曲汇口、培龙沟至结东、隆巴曲至通麦区域滑坡密度大。该地区属于雅鲁藏布江大拐弯地区、易贡藏布和帕隆藏布江的交汇处,是高断层密度区,且大型以上规模的滑坡主要分布于色东普至加热萨乡段(雅鲁藏布江大拐弯),这一区域位于喜马拉雅山脉、横断山脉和念青唐古拉山脉三大山脉的交汇处,处在东喜马拉雅构造结北缘,是构造应力最集中的地段。受构造应力的作用,大拐弯顶端为隆升的核心区,隆升造成了新的不稳定和不平衡,是灾害发育的主控因素。强烈的多期次构造变形变质作用,使岩体相互挤压破碎或位移,在这一区域形成了大量的断层、节理等结构面。纵横交错的断层产生了地垒效应,将岩体剪切使其破裂,缝合带内的片麻岩、板岩、蛇绿岩受构造挤压破碎,稳定性较差,容易发生滑坡崩塌。地震波倾向于沿着断裂带传播,诱发滑坡崩塌,如汶川地震后,巨大的能量迅速沿龙门山断裂带向东北方向扩展,沿断裂带的公路、山体、村庄遭到了严重破坏,并造成了严重的次级效应。地貌上,这一区域为深切峡谷段,容易受河流侧蚀形成凌空面,造成重力失稳,诱发滑坡发生。

## 5 讨论

在本次研究中,选取藏东南部分区域进行滑坡灾害的易发性评价,通过应用随机森林模型,并结合研究区特征和频率比值模型等,选取了较好的原始数据进行输入,生成滑坡易发性分区图,并按照模型贡献率对特征进行排序。在该区域中,高易发区主要集中在易贡藏布、帕隆藏布交汇处和雅鲁藏布江大拐弯区域,根据9个因子的特征重要性排序,发现断层密度是该地区滑坡发育的主控因素。在该地区,由于长期受到构造作用影响,地质构造极其复杂。东西两侧均是巨型断层,南侧为喜马拉雅前缘断裂带,北侧为冈底斯造山带。该地区整体构成为晚新生代三叠纪,其中四川盆地南部和喜马拉雅北麓沉积为早期始新世时期的古老地层,而南北两侧的晚白垩世—第四纪大量岩浆活动异常发达,形成了一系列岩浆岩体,造成了复杂的地理构造和多样的地貌景观,产状极其丰富。在这种特殊的地质情况下,使得滑坡发育,地质灾害频发成为常态。

机器学习在地质灾害评价方面有着较长的历史,并且目前已经比较成熟。相对于传统模型的主观性,机器学习模型具有许多突出的优点,它能够处理大量数据,高效地对数据进行统计和组合,从而提高了预测精度。机器学习很少遗漏某些重要因素且不会受到主观因素的影响,具有广泛的适用性和可靠性。另外,机器学习还具有良好的自适应性,可以通过有监督的学习方法逐步优化,以更好地适应环境和预测变化。

滑坡灾害的发生是一个极其复杂的过程,涉及水文学、力学、地质学等多学科、多领域的知识,因此,仅通过部分因子和模型对滑坡灾害的预测是片面的。此外,原始数据的精度也会对模型预测结果造成影响。在广泛获取原始滑坡数据时,没有完整的获取流程,滑坡数据的获取仍然存在主观性,因此,在滑坡易发性评价中,机器学习算法仍存在较大的误差。未来,应该系统化完善机器学习算法在滑坡易发性预测中应用,包括全面获取原始滑坡样本、准确选择非滑坡样本数据等,以尽量减少人为客观因素的影响,提高滑坡易发性在时间和空间方面的准确性。

## 6 结论

本研究以藏东南地区为研究区,采用随机森林模型评估了研究区滑坡灾害的易发性,并通过使用频率比值法,对研究区滑坡易发性做快速评判,在其低易发区选取非滑坡样本,确保模型能够获得良好的原始数据,得出以下结论:

a. 根据模型预测结果显示,研究区93%的滑坡位于中易发区以上区域。在高易发区,滑坡密度达14.5处/100 km<sup>2</sup>,而低易发区滑坡密度仅为0.016处/100 km<sup>2</sup>。由此可以看出,易发区和滑坡密度之间存在正比关系,即易发区等级越高,相应的滑坡密度也越高。在这个研究区中,高易发区主要集中在易贡藏布、帕隆藏布交汇处和雅鲁藏布江大拐弯区域,滑坡事件在这些地方更加频繁和密集,通过对这些地区的重点观察和监测,能够更好地预测和控制滑坡事件的发生,以保证人民生命财产的安全。

b. 通过模型贡献率排序的结果显示,在研究区中,影响滑坡发生的因素前三依次为断层密度、高程和坡度。研究区的断裂带发育程度较高,特别是在高易发区,断层纵横交错,形成了多重挤压

的情况,导致研究区岩石破碎、节理发育。此外,由于构造活动活跃,研究区地表隆升现象严重,河流冲刷力度强,地形坡度陡峭,高差很大,这些因素为滑坡的发育提供了能量。

c. 通过对随机森林模型进行 ROC 曲线评估,结果显示:在本次研究中,随机森林模型 ROC 曲线线下面积 AUC 值为 0.94,说明随机森林模型的评价方法在本次研究中精确度较高。因此,对于滑坡地质灾害易发性的研究而言,随机森林模型具有较高的精确度和可靠度。

d. 为了验证和分析随机森林模型的性能,本研究通过选取 2 处典型滑坡点和 3 个野外调查滑坡点,对模型进行进一步的实地验证。通过对这 5 处滑坡点进行断层密度的测量,发现它们的断层密度均大于  $40 \text{ km}/100 \text{ km}^2$ ,远高于研究区断层密度的平均值  $19.61 \text{ km}/100 \text{ km}^2$ 。表明这些滑坡所处的地质构造环境更加复杂和活跃,同时也说明随机森林模型的结果是合理和准确的。

### [ 参 考 文 献 ]

- 邓念东,崔阳阳,郭有金,2020. 基于频率比—随机森林模型的滑坡易发性评价[J]. 科学技术与工程,20(34):13990—13996.
- Deng N D, Cui Y Y, Guo Y J, 2020. Frequency ratio-random forest-model-based landslide susceptibility assessment[J]. Science Technology and Engineering, 20(34):13990—13996. (in Chinese)
- 高泽民,丁明涛,杨国辉,等,2021. 川藏铁路孜热—波密段泥石流灾害危险性评价[J]. 工程地质学报,29(2):478—485.
- Gao Z M, Ding M T, Yang G H, *et al.*, 2021. Hazard assessment of debris flow along Zire-Bomi section of Sichuan-Tibet railway[J]. Journal of Engineering Geology, 29(2):478—485. (in Chinese)
- 扈秀宇,秦胜伍,窦强,等,2019. 基于 GIS 和随机森林模型的泥石流敏感性分析——以吉林省洮南市北部山区为例[J]. 水土保持通报,39(5):204—210.
- Hu X Y, Qin S W, Dou Q, *et al.*, 2019. Susceptibility analysis of debris flow based on GIS and random forest—A case study of a mountainous area in northern Taonan city, Jilin Province[J]. Bulletin of Soil and Water Conservation, 39(5):204—210. (in Chinese)
- 黄发明,胡松雁,闫学涯,等,2022. 基于机器学习的滑坡易发性预测建模及其主控因子识别[J]. 地质科技通报,41(2):79—90.
- Huang F M, Hu S Y, Yan X Y, *et al.*, 2022. Landslide susceptibility prediction and identification of its main environmental factors based on machine learning models[J]. Bulletin of Geological Science and Technology, 41(2):79—90. (in Chinese)
- 吉日伍呷,田宏岭,韩继冲,2022. 基于不同机器学习算法的地震滑坡易发性评价——以鲁甸地震为例[J]. 昆明理工大学学报(自然科学版),47(2):47—56.
- Jiri W G, Tian H L, Han J C, 2022. Evaluation of the susceptibility of earthquake landslides based on different machine learning algorithms—Taking Ludian earthquake as an example[J]. Journal of Kunming University of Science and Technology (Natural Science), 47(2):47—56. (in Chinese)
- 李俊,陈宁生,刘美,等,2018. 2000 年易贡乡扎木弄沟滑坡型泥石流主控因素分析[J]. 南水北调与水利科技,16(6):187—193.
- Li J, Chen N S, Liu M, *et al.*, 2018. Analysis of main factors for landslide-triggered debris flow in Zhamunong gully on April 9th 2000 [J]. South-to-North Water Transfers and Water Science & Technology, 16(6):187—193. (in Chinese)
- 李亭,田原,邬伦,等,2014. 基于随机森林方法的滑坡灾害危险性区划[J]. 地理与地理信息科学,30(6):25—30.
- Li T, Tian Y, Wu L, *et al.*, 2014. Landslide susceptibility mapping using random forest [J]. Geography and Geo-Information Science, 30(6):25—30. (in Chinese)
- 李阳春,刘黔云,李潇,等,2021. 基于机器学习的滑坡崩塌地质灾害气象风险预警研究[J]. 中国地质灾害与防治学报,32(3):118—123.
- Li Y C, Liu Q Y, Li X, *et al.*, 2021. Exploring early warning and forecasting of meteorological risk of landslide and rockfall induced by meteorological factors by the approach of machine learning[J]. The Chinese Journal of Geological Hazard and Control, 32(3):118—123. (in Chinese)
- 林报嘉,刘晓东,杨川,等,2020. XGBoost 机器学习模型与 GIS 技术结合的公路崩塌灾害易发性研究[J]. 公路,65(7):20—26.
- Lin B J, Liu X D, Yang C, *et al.*, 2020. Avalanche susceptibility assessment of highway based on XGBoost machine learning model and GIS method[J]. Highway, 65(7):20—26. (in Chinese)
- 刘海知,徐辉,包红军,等,2022. 机器学习分类算法在降雨型滑坡预报中的应用[J]. 应用气象学报,33(3):282—292.

- Liu H Z, Xu H, Bao H J, *et al.*, 2022. Application of machine learning classification algorithm to precipitation-induced landslides forecasting[J]. *Journal of Applied Meteorology Science*, 33(3):282-292. (in Chinese)
- 马啸, 王念秦, 李晓抗, 等, 2022. 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*, 55(3):335-344.
- Ma X, Wang N Q, Li X K, *et al.*, 2022. Assessment of landslide susceptibility based on RF-FR Model: Taking Lueyang County as an example[J]. *Northwestern Geology*, 55(3):335-344. (in Chinese)
- 宋轩宇, 许民, 康世昌, 等, 2023. 基于机器学习的冰冻圈典型流域水文过程模拟研究[J]. *地学前缘*, 30(4):451-469.
- Song X Y, Xu M, Kang S C, *et al.*, 2023. Modeling of hydrological processes in cryospheric watersheds based on machine learning[J]. *Earth Science Frontiers*, 30(4):451-469. (in Chinese)
- 铁永波, 徐伟, 向炳霖, 等, 2022. 西南地区地质灾害风险“点面双控”体系构建与思考[J]. *中国地质灾害与防治学报*, 33(3):106-113.
- Tie Y B, Xu W, Xiang B L, *et al.*, 2022. The thoughts on construction of “double-control of point and zone” system of geological hazard risk in southwest China [J]. *The Chinese Journal of Geological Hazard and Control*, 33(3):106-113. (in Chinese)
- 汪发武, 陈也, 刘伟超, 等, 2022. 藏东南高位远程滑坡动力学特征及研究难点[J]. *工程地质学报*, 30(6):1831-1841.
- Wang F W, Chen Y, Liu W C, *et al.*, 2022. Characteristics and challenges to dynamics of long-runout landslides with high-altitude in southeast Tibet[J]. *Journal of Engineering Geology*, 30(6):1831-1841. (in Chinese)
- 王文俊, 向喜琼, 黄润秋, 等, 2003. 区域崩塌滑坡的易发性评价——以四川省珙县为例[J]. *中国地质灾害与防治学报*, 14(2):33-36.
- Wang W J, Xiang X Q, Huang R Q, *et al.*, 2003. Areal landslides susceptibility analysis—A case study in Hong County, Sichuan Province [J]. *The Chinese Journal of Geological Hazard and Control*, 14(2):33-36. (in Chinese)
- 王向前, 丁明涛, 何雨枫, 2023. 基于 RF 模型的甘孜藏族自治州乡城县泥石流易发性评价[J]. *灾害学*, 38(3):222-226+234.
- Wang X Q, Ding M T, He Y F, 2023. Assessment of debris flow vulnerability in Xiangcheng County, Ganzi Tibetan Autonomous Prefecture based on RF model[J]. *Journal of Catastrophology*, 38(3):222-226+234. (in Chinese)
- 王盈, 金家梁, 袁仁茂, 2019. 藏东南地区地质灾害空间分布及影响因素分析[J]. *地震研究*, 42(3):428-437+456.
- Wang Y, Jin J L, Yuan R M, 2019. Analysis on spatial distribution and influencing factors of geological disasters in southeast Tibet[J]. *Journal of Seismological Research*, 42(3):428-437+456. (in Chinese)
- 武雪玲, 任福, 牛瑞卿, 等, 2013. 斜坡单元支持下的滑坡易发性评价支持向量机模型[J]. *武汉大学学报: 信息科学版*, 38(12):1499-1503.
- Wu X L, Ren F, Niu R Q, *et al.*, 2013. Landslide spatial prediction based on slope units and support vector machines[J]. *Geomatics and Information Science of Wuhan University*, 38(12):1499-1503. (in Chinese)
- 武晓岩, 李康, 2006. 基因表达数据判别分析的随机森林方法[J]. *中国卫生统计*, 23(6):491-494.
- Wu X Y, Li K, 2006. The application of random forests for the classification of gene expression data[J]. *Chinese Journal of Health Statistics*, 23(6):491-494. (in Chinese)
- 杨硕, 李德营, 严亮轩, 等, 2021. 基于随机森林模型的乌江高陡岸坡滑坡地质灾害易发性评价[J]. *安全与环境工程*, 28(4):131-138.
- Yang S, Li D Y, Yan L X, *et al.*, 2021. Landslide susceptibility assessment in high and steep bank slopes along Wujiang River based on random forest model[J]. *Safety and Environmental Engineering*, 28(4):131-138. (in Chinese)
- 袁浩, 郭长宝, 吴瑞安, 等, 2023. 西藏易贡高位远程滑坡研究进展与展望[J]. *地质通报*, 42(10):1757-1773.
- Yuan H, Guo C B, Wu R A, *et al.*, 2023. Research progress and prospects of the giant Yigong long run-out landslide, Tibetan Plateau, China[J]. *Geological Bulletin of China*, 42(10):1757-1773. (in Chinese)
- 张炳华, 张懿铨, 谷昌军, 等, 2023. 基于随机森林与特征选择的藏东南土地覆被分类方法及精度评价[J]. *地理科学*, 43(3):388-397.
- Zhang B H, Zhang Y L, Gu C J, *et al.*, 2023. Land cover classification based on random forest and feature optimism in the Southeast Qinghai-Tibet Plateau[J]. *Scientia Geographica Sinica*, 43(3):388-397. (in Chinese)
- 张冬梅, 张莉, 2019. 对话姚檀栋: 走近第二次青藏高原综合科学考察[J]. *科学通报*, 64(27):2765-2769.
- Zhang D M, Zhang L, 2019. Tandong Yao: The second Ti-

- betan Plateau scientific expedition and research[J]. Chinese Science Bulletin, 64(27): 2765–2769. (in Chinese)
- 郑迎凯, 陈建国, 王成彬, 等, 2020. 确定性系数与随机森林模型在云南芒市滑坡易发性评价中的应用[J]. 地质科技通报, 39(6): 131–144.
- Zheng Y K, Chen J G, Wang C B, *et al.*, 2020. Application of certainty factor and random forest model in landslide susceptibility evaluation in Mangshi, Yunnan Province[J]. Bulletin of Geological Science and Technology, 39(6): 131–144. (in Chinese)
- 钟政, 陈宁生, 倪华勇, 等, 2021. 金沙江白格滑坡工程地质与水文地质特征[J]. 科学技术与工程, 21(13): 5262–5269.
- Zhong Z, Chen N S, Ni H Y, *et al.*, 2021. Engineering geology and hydrogeological characteristics of Jinsha River in Baige Landslide[J]. Science Technology and Engineering, 21(13): 5262–5269. (in Chinese)
- 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会, 2016. 滑坡防治工程勘查规范: GB/T 32864—2016 [S]. 北京: 中国标准出版社.
- General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China, 2016. Code for Geological Investigation of Landslide Prevention: GB/T 32864—2016 [S]. Beijing: Standards Press of China. (in Chinese)
- 周萍, 邓辉, 张文江, 等, 2022. 基于信息量模型和机器学习方法的滑坡易发性评价研究——以四川理县为例[J]. 地理科学, 42(9): 1665–1675.
- Zhou P, Deng H, Zhang W J, *et al.*, 2022. Landslide susceptibility evaluation based on Information Value model and Machine Learning method: A case study of Lixian County, Sichuan Province [J]. Scientia Geographica Sinica, 42(9): 1665–1675. (in Chinese)
- 祝介旺, 苏天明, 张路青, 等, 2010. 川藏公路 102 滑坡失稳因素与治理方案研究[J]. 水文地质工程地质, 37(3): 43–47.
- Zhu J W, Su T M, Zhang L Q, *et al.*, 2010. Studies of factors influencing the stability of No. 102 landslide and controlling scheme at the Sichuan-Tebet highway[J]. Hydrogeology & Engineering Geology, 37(3): 43–47. (in Chinese)
- 朱宇航, 殷坤龙, 郭子正, 等, 2022. 基于滑坡易发性评价因子特征选取非滑坡样本的方法 [P]. 湖北省: CN115481211A, 2022–12–16.
- 左群超, 叶天竺, 冯艳芳, 等, 2018. 中国陆域 1: 25 万分幅建造构造图空间数据库[J]. 中国地质, 45(增刊 1): 1–26.
- Zuo Q C, Ye T Z, Feng Y F, *et al.*, 2018. Spatial database of serial suite-tectonic map-sheets of mainland China (1: 250 000)[J]. Geology in China, 45(S1): 1–26. (in Chinese)
- Aleotti P, Chowdhury R, 1999. Landslide hazard assessment: Summary review and new perspectives[J]. Bulletin of Engineering Geology & the Environment, 58(1): 21–44.
- Bzdok D, Altman N, Krzywinski M I, 2018. Statistics versus machine learning[J]. Nature Methods, 15(4): 233–234.
- Chen W, Chai H, Zhao Z, *et al.*, 2016. Landslide susceptibility mapping based on GIS and support vector machine models for the Qianyang County, China[J]. Environmental Earth Sciences, 75(6): 1–13.
- Hermanns R L, Strecker M R, 2011. Structural and lithological controls on large Quaternary rock avalanches (sturzstroms) in arid northwestern Argentina[J]. Geological Society of America Bulletin, 111(6): 934–948.
- Li J, Chen N S, Zhao Y, *et al.*, 2020. A catastrophic landslide triggered debris flow in China's Yigong: Factors, dynamic processes and tendency[J]. Earth Sciences Research Journal, 24(1): 71–82.
- Ling P, Niu R, Bo H, *et al.*, 2014. Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China [J]. Geomorphology, 204(1): 287–301.
- Merghadi A, Yunus A P, Dou J, *et al.*, 2020. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance [J]. Earth-Science Reviews, 207: 103225.
- Pawluszek K, Borkowski A, 2017. Impact of DEM-derived factors and analytical hierarchy process on landslide susceptibility mapping in the region of Roznow Lake, Poland[J]. Natural Hazards, 86(2): 919–952.
- Pourghasemi H R, Kerle N, 2016. Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran [J]. Environmental Earth Sciences, 75: 185.

(责任编辑: 黄维婷)