

泥石流易发性间谍技术随机森林模型研究

陈宇涛, 李 宁, 常 鸣, 幸夫诚, 向 晗, 李晋生, 杨 瑞, 陈子龙

Study on debris flow susceptibility based on SPY-RF model: A case study of the upper Minjiang River Basin

CHEN Yutao, LI Ning, CHANG Ming, XING Fucheng, XIANG Han, LI Jinsheng, YANG Rui, and CHEN Zilong

在线阅读 View online: <https://doi.org/10.16031/j.cnki.issn.1003-8035.202410044>

您可能感兴趣的其他文章

Articles you may be interested in

岷江上游汶川地震前后泥石流易发性评价

Susceptibility assessment of debris flow in the upper reaches of the Minjiang River before and after the Wenchuan earthquake

赵佳忆, 田述军, 李凯, 侯鹏鹏 中国地质灾害与防治学报. 2024, 35(1): 51-59

基于自组织特征映射网络-随机森林模型的滑坡易发性评价

Evaluation on landslide susceptibility based on self-organizing feature map network and random forest model: A case study of Dayu County of Jiangxi Province

何书, 鲜木斯艳·阿布迪克依木, 胡萌, 陈康 中国地质灾害与防治学报. 2022, 33(1): 132-140

基于Logistic回归和随机森林的清江流域长阳库岸段堆积层滑坡易发性评价

Susceptibility assessment of colluvium landslides along the Changyang section of Qingjiang River using Logistic regression and random forest methods

曾斌, 吕权儒, 寇磊, 艾东, 许汇源, 袁晶晶 中国地质灾害与防治学报. 2023, 34(4): 105-113

基于斜坡单元与随机森林模型的元阳县崩滑地质灾害易发性评价

Evaluation of geological hazard susceptibility of collapse and landslide in Yuanyang County using slope units and random forest modeling

刘帅, 朱杰勇, 杨得虎, 马博 中国地质灾害与防治学报. 2023, 34(4): 144-150

基于机器学习的伊犁河谷黄土区泥石流易发性评估

Susceptibility assessment of debris flow disaster based on machine learning models in the loess area along Yili Valley

李志, 陈宁生, 侯儒宁, 吴铭洋, 张瀛玉龙, 杜鹏 中国地质灾害与防治学报. 2024, 35(3): 129-140

D-InSAR与随机森林模型耦合的活动性滑坡识别方法探究

Active landslide identification with a combined method of D-InSAR and random forest model

高星月, 王世杰, 高鹏程 中国地质灾害与防治学报. 2022, 33(5): 102-108



关注微信公众号, 获得更多资讯信息

DOI: 10.16031/j.cnki.issn.1003-8035.202410044

陈宇涛, 李宁, 常鸣, 等. 泥石流易发性间谍技术随机森林模型研究——以岷江上游为例[J]. 中国地质灾害与防治学报, 2025, 36(4): 137-149.

CHEN Yutao, LI Ning, CHANG Ming, et al. Study on debris flow susceptibility based on SPY-RF model: A case study of the upper Minjiang River Basin[J]. The Chinese Journal of Geological Hazard and Control, 2025, 36(4): 137-149.

泥石流易发性间谍技术随机森林模型研究 ——以岷江上游为例

陈宇涛¹, 李 宁¹, 常 鸣², 幸夫诚¹, 向 吟¹, 李晋生¹, 杨 瑞¹, 陈子龙³

(1. 西华大学应急管理学院, 四川 成都 610039; 2. 地质灾害防治与环境保护国家重点实验室(成都理工大学), 四川 成都 610059; 3. 四川省华地建设工程有限公司, 四川 成都 610036)

摘要: 泥石流作为一种由强降雨或冰雪融化引发的高浓度非均质流体, 具有复杂的形成和运动过程。评估泥石流的易发性对于灾害监测与应对具有重要的实际意义。传统方法难以准确预测泥石流的发生, 因此近年来机器学习算法在该领域的应用逐渐增多。文章以岷江上游为例, 提出一种基于间谍技术(SPY)的随机森林模型 SPY-RF, 用于构建泥石流易发性评价系统。SPY 方法通过对未标记数据进行伪负样本生成, 克服了不平衡数据集在负样本获取上的局限性, 提高了模型的区分能力。在研究中选取 14 个评价因子, 如沟壑密度、岩性、流域面积等, 结合遥感影像和地质灾害数据构建泥石流数据集。通过 SPY 技术优化负样本的获取, 结合随机森林模型对泥石流易发性进行建模。结果显示: SPY-RF 模型和 RF 模型的 AUC 值分别为 0.98, 0.93, 且 SPY-RF 模型性能指标整体优于 RF 模型, SPY-RF 模型在预测泥石流易发性方面表现出较高的精确度和稳定性, 极高易发区与现有泥石流点的分布吻合, 在极低和低风险区域也能识别泥石流点。在负样本获取和筛选策略上, 采用 SPY 技术显著提高了负样本的质量, 从而提升了模型的预测精度和可靠性。为岷江上游地区泥石流风险管理提供了参考依据。

关键词: 泥石流; 易发性; 随机森林; 间谍技术; 岷江上游流域

中图分类号: P642.23 **文献标志码:** A **文章编号:** 1003-8035(2025)04-0137-13

Study on debris flow susceptibility based on SPY-RF model: A case study of the upper Minjiang River Basin

CHEN Yutao¹, LI Ning¹, CHANG Ming², XING Fucheng¹, XIANG Han¹, LI Jinsheng¹,
YANG Rui¹, CHEN Zilong³

(1. School of Emergency Management, Xihua University, Chengdu, Sichuan 610039, China; 2. State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, Sichuan 610059, China; 3. Sichuan Huadi Construction Engineering Co. Ltd., Chengdu, Sichuan 610036, China)

Abstract: Debris flow is a high-concentration, heterogeneous, multiphase flow typically triggered by intense rainfall or snowmelt. Its complex formation and movement processes make accurate susceptibility assessment vital for disaster monitoring and mitigation. Traditional methods often fall short in predictive accuracy, leading to a growing adoption of machine learning

收稿日期: 2024-10-31; 修订日期: 2024-11-29

投稿网址: <https://www.zgdzzyfzxb.com/>

基金项目: 地质灾害防治与地质环境保护国家重点实验室开放基金项目(SKLGP2023K010); 四川省自然科学基金项目(2023NSFSC0809)

第一作者: 陈宇涛(2000—), 男, 四川成都人, 硕士研究生, 主要从事地质灾害评价等方面的研究。E-mail: 776636130@qq.com

通讯作者: 李 宁(1990—), 男, 河南商丘人, 工学博士, 硕士生导师, 主要从事地震地质灾害调查与评价、地质灾害致灾机制与长期演化效应、地质灾害风险评价与管控等方面的研究工作。E-mail: 790626903@qq.com

algorithms in this field in recent years. This study proposes a debris flow susceptibility assessment model, SPY-RF, which integrates the random forest (RF) algorithm with the spy technique (SPY), using the upper Minjiang River Basin as a case study. The SPY method addresses the common issue of class imbalance by generating high-quality pseudo-negative samples from unlabeled data, thereby enhancing the model's classification performance. A total of fourteen assessment factors, including gully density, lithology, area, and others, were selected based on geological disaster data and remote sensing imagery to construct a comprehensive debris flow dataset. The SPY technique was utilized to optimize the negative sample selection process, which was then combined with the RF model to evaluate susceptibility. The findings indicate that the SPY-RF model outperforms the traditional RF model, achieving an *AUC* of 0.98 compared to 0.93. The predicted distribution of extremely high susceptibility areas aligns closely with the current debris flow points, indicating that the SPY-RF model predicts debris flow susceptibility with greater accuracy and stability. Additionally, the model also successfully identifies debris flow occurrences in low-risk and extremely low-risk susceptibility areas. The quality of negative samples was greatly increased by using SPY technology in terms of negative sample acquisition and filtering techniques, which raised the prediction accuracy and dependability of the model. The proposed SPY-RF model serves as a useful guidance for managing the risk of debris flows in the upper Minjiang River Basin.

Keywords: debris flow; susceptibility; random forest; spy technique; upper Minjiang River Basin

0 引言

泥石流主要是由强降雨或冰雪消融形成的高浓度、宽级配的多相非均质流体,其物质组成级配跨 6 个数量级,形成及运动过程极为复杂,常以快速流动性、强冲击性以及突发性不断侵蚀山区的环境,破坏作用强烈,治理难度大^[1]。因此,通过对泥石流易发性进行评估以预测其发生的空间概率,对于泥石流灾害的监测与应对具有重要作用,也是该研究领域的难点和热点问题。随着近年来极端天气的不断加剧,使得我国本就严峻的泥石流灾害形势也变得愈发严峻。泥石流易发性也更加难以用传统方法准确判断和评价,因此,如何准确描述泥石流的易发性在此背景下则显得尤为重要。

传统的泥石流易发性评价常用的方法有专家打分法^[2]、层次分析法^[3-4]、模糊综合评判法^[5-7]、加权信息熵法和多元回归方法等^[8-10]。从因子权重的确定角度,主要可以分为两类:一类是依赖于专家的经验知识来分配权重,另一类则是将专家的经验与数学统计分析方法结合,以计算因子的权重。两种方法的核心都是通过确定因子权重来评估泥石流易发性,然而由于专家经验的区域性,使得这类方法的适用性有限,难以在不同地理环境中广泛应用^[11]。这种方法在相当长的时间内发挥了重要作用,尤其在数据有限和技术手段相对落后的情况下。然而,随着极端天气事件的频繁发生,泥石流的突发性和快速性使得这些传统方法在应对实时灾害时常显得滞后,无法提供足够的预测精确度和时效性。但

近年来,机器学习算法也在泥石流的易发性评估方面获得了应用,包括人工神经网络、支持向量机、决策树、随机森林和卷积神经网络等都已经成功应用到了地质灾害的分析中^[12-17]。以机器学习为基础的统计分析方法能够客观地揭示泥石流发生条件与易发性之间的非线性映射关系,有效排除了人为主观因素的干扰,从而提升了评估结果的精确度。通常机器学习方法的训练集往往包含数量相等的泥石流流点和非泥石流流点。在泥石流易发性评估的研究中,研究者往往只重视已发生或可能发生的正样本,而对不会发生灾害的负样本则很少给予关注^[18]。在前人的研究中,建模所需的负样本通过随机抽取未标记样本来获取^[19-20]。这种方法将优质的正样本与潜在的噪声负样本等同对待,从而使得不合理的负样本采集方式成为数据集中噪声的主要来源之一。这类噪声数据通常会导致模型出现异常学习等问题^[21-22]。因此,如何优化样本库以提高模型性能成为机器学习领域的关键研究方向之一。2024 年,于海坤等人为提高滑坡样本集的质量,采用了 4 种采样策略:随机采样、滑坡缓冲区采样、信息量约束采样和合成少数类过采样,旨在为滑坡易发性评估中的样本挑选带来新的方法^[23]。2024 年洪浩源等人通过比较某点与灾害点正样本的地理环境相似度和不相似度,分别衡量正样本和负样本的可信度,并基于互斥法设定可信度阈值以收集训练样本,取得有效成果^[24]。通过改进样本的选择策略,可以显著降低噪音数据对模型的影响。这为泥石流灾害的预测提供了更广泛的前景,并有助于提高模型

的实际应用能力。

岷江上游地区的地质结构非常复杂,常常受到活断层的影响。此外,山地生态环境的脆弱性使得该区域泥石流灾害时有发生。2019年8月20日,岷江上游暴发了危害性、破坏性极强的群发泥石流,都汶高速一带、213国道及30省道等多条干线公路受到严重破坏^[25]。2023年6月26日,位于岷江上游流域的绵虬镇板子沟、威州镇新桥沟两处突发山洪泥石流灾害,造成2人遇难、3人失联,紧急避险转移群众900余人^[26]。因此,本文充分考虑研究区的复杂自然地理条件,结合实际情况,选取面积、沟壑密度、圆度、Melton比率、岩性、距道路距离、距断层距离、坡向、归一化植被指数、水流强度指数、地形湿度指数、降雨、土地利用类型、曲率14个评价因子作为岷江上游流域区域泥石流易发性评价指标,提出间谍技术(SPY)构建可靠非泥石流样本,采用随机森林模型,基于流域单元作为评价单元对岷江上游流域泥石流进行易发性评价,评价结果可为该区域未来的泥石流灾害预警提供指导参照。

1 研究区概况

岷江上游流域纵跨四川省阿坝藏族羌族自治州的汶川县、茂县和松潘县。岷江上游流域位于四川盆地与青藏高原的边缘区域,在40~50 km水平范围内海拔从750 m陡变至5200 m,地表起伏变化显著,属于典型的高山峡谷地带。同时,该地区位于川西地槽,历经多次构造变动,导致地质结构破碎和基础不稳定。此外,由于季风气候的影响,夏季经常出现强降雨,这种降水分布的不均匀性使得泥石流等自然灾害的发生风险显著增加。岷江是研究区内的主要河流,也是长江上游的一条重要支流。研究区岷江上游干流全长约295 km,流域面积7326 km²,根据干流河道纵剖面特征可将干流由上游向下游分为北段、中段和南段(图1)。

2 数据与研究方法

2.1 数据源

2.1.1 泥石流数据集

泥石流数据集是建立泥石流易发性评估模型的关键,能够帮助识别泥石流发生的地点,并阐明指标因子与泥石流发生之间的联系。在遥感影像中,泥石流具有显著的地貌特征,可以明确区分为形成区、流动区和堆积区。研究主要以堆积区特征进行解译。根据沟口的地貌特征,即发现沟口有明显的堆积扇,则判别其为泥石流沟^[27]。因此,本文根据岷江上游2021年高分二号

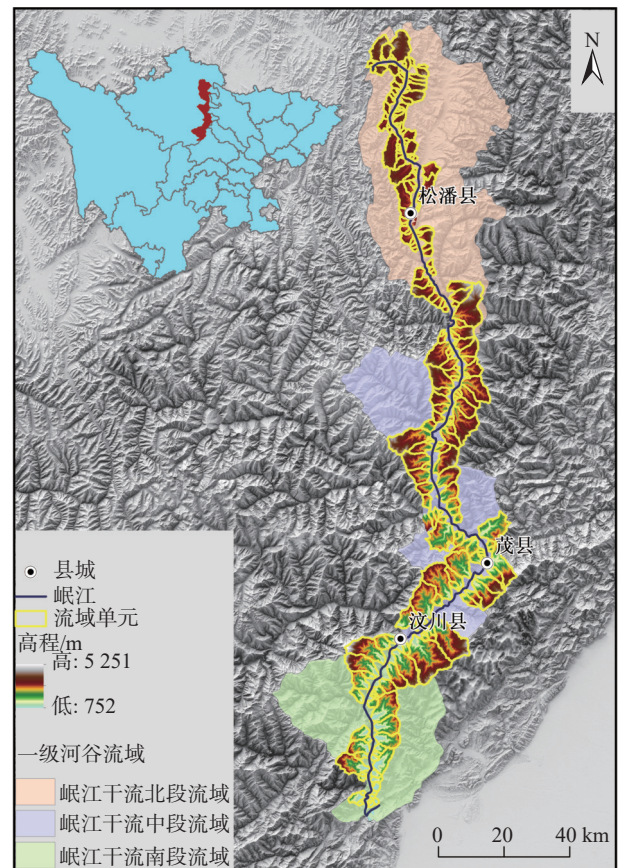


图1 岷江上游区域地理位置及泥石流流域分布图

Fig. 1 Geographic location of the upper Minjiang River Basin and distribution of debris flow watersheds

卫星影像、实地考察等方法,结合从中国科学院资源环境科学与数据中心(<https://www.resdc.cn/>)获取的分辨率为12.5 m×12.5 m的DEM数据及研究区地质灾害分布数据分别对泥石流沟进行解译,最终在研究区内共解译出113个泥石流沟。

2.1.2 评价因子

通过小流域形成泥石流的孕灾条件,从地形地貌、降雨、地质、植被覆盖和人类活动5个方面选取了14个评价因子。(1)地形地貌数据,以研究区空间分辨率为12.5 m的DEM数据,通过ArcGIS10.8提取得到流域面积、沟壑密度、圆度、Melton比率、坡向、曲率,并通过水文处理得到区域的水流强度指数(SPI)和地形湿度指数(TWI);(2)降雨数据,通过ArcGIS10.8裁剪四川省年均降雨量(国家气象科学数据中心获取)得到研究区年均降雨量矢量数据;(3)地质:从中国科学院资源环境科学与数据中心(<http://www.resdc.cn/>)获取研究区断层分布与地层岩体数据,经坐标转换得到距断层距离、地层岩体数据;(4)植被覆盖数据,选取研究区时间为2024年8月分辨率为30 m的Landsat8近红外和远

红外波段计算得到植被覆盖指数(NDVI);通过裁剪全球 30 m 地表覆盖数据,得到研究区土地利用类型矢量数据 (<http://www.globeland30.org/home/background.aspx>) (5)人类活动数据,从 2021 年高分二号卫星影像提取研究区道路信息,得到距道路距离。表 1 说明了评价因子的选取依据。

表 1 评价因子的选取与依据
Table 1 Selection and description of evaluation factors

孕灾条件	评价因子	选取依据
	流域面积	在泥石流发育和形成的影响因素中沟道流域面积的大小,不仅影响沟道比降、主沟长度等其它地貌形态指标,还对强降雨的汇集、松散物质的分布等地质、水文条件产生影响 ^[28]
	沟壑密度	沟壑密度反映了地形起伏和切割程度以及岩土体的松散程度,是水流汇集能力的体现。沟壑密度越大,地表越破碎,地表物质稳定性降低更易于发生泥石流灾害 ^[29] ;其公式为: $D_s = \frac{\sum L}{A}$, 式中: $\sum L$ 表示每个流域单元的沟谷长度之和, A 表示流域面积
	圆度	圆度为泥石流流域的重要形态特征之一,其指的是流域面积与其相同周长的圆的面积之比 其计算公式: $R = \frac{4\pi A}{P^2}$, 其中 A 为流域面积; P 代表的则为流域周长
	Melton比率	Melton比率常用来描述流域单元内地势陡峻程度, Melton比率的值越大泥石流启动时具有的势能越强,其公式: $M = \frac{dH}{\sqrt{S}}$, 其中 dH 为流域内的高差; S 表示流域面积大小 ^[30]
地形地貌	坡向	不同坡向的坡面在太阳辐射、植被覆盖和水文条件上差异明显。阳坡因接受更多辐射, 岩土风化较严重, 水土流失加剧, 土壤稳定性差, 容易引发泥石流 ^[30-31]
	曲率	曲率作为影响坡面稳定性和水流运动的关键因子之一, 能直接影响泥石流的发生概率和分布模式
	水流强度指数	水流强度指数(stream power index, SPI)是一个描述流水侵蚀能力的指标, 在泥石流研究中, SPI不仅可以反映水流对滑坡体的冲击和侵蚀作用, 还可以帮助揭示水流在泥石流发生过程中的重要作用 ^[32]
	地形湿度指数	地形湿度指数(topographic wetness index, TWI)是评估地区水分条件的有效工具, 它考虑了地表径流模式、地下水埋藏深度和地形起伏等因素; 较高的SPI通常意味着该区域水分积聚较多, 增加了土壤饱和的风险, 从而易引发泥石流 ^[33]
降雨	年均降雨	降雨是公认的泥石流的主要诱发因素之一, 大量的不稳定体会在降雨条件下进一步失稳而形成泥石流
地质	断层层距离	断层层距离是岩土体受构造运动影响程度的一个量化反映; 通常情况下靠近断层的岩土体会更加破碎, 也更有可能成为泥石流的潜在物源
	岩体	岩体构成了泥石流形成发育的物源基础, 岩土体的岩性及其结构特征决定流域内斜坡岩土体的强度、变形破坏特征等条件
植被覆盖数据	归一化植被指数	本文选用归一化植被指数(normalized difference vegetation index, NDVI)来反映研究区的植被状况; 植被稀疏的地区, 基岩暴露、水土流失现象严重, 岩土体更容易失稳, 从而引发泥石流
人类活动数据	土地利用类型	土地利用类型通过影响水文条件、土壤稳定性和水流路径, 进而对泥石流发生的风险产生直接或间接的作用
	距道路距离	道路距离可以反映人类工程活动对岩土体的影响; 过度开挖不仅会破坏边坡稳定, 还可能导致废弃物堆积形成新的不稳定区域, 增加泥石流发生的风险

2.2 研究方法

2.2.1 评价单元

在泥石流易发性评价模型中,合理的评价单元选择至关重要,研究涵盖栅格单元、斜坡单元、特殊条件单元及流域单元^[33-37]。其中,流域单元能够充分包含泥石流的几何特性,并有效揭示其与相关指标之间的关系^[38]。因此,本文以选择的流域单元为主要研究依据,并利用 ArcGIS10.8 的水文处理方法,对研究单元的 12.5 m DEM 数据加以划分。在水文分析中,阈值的选择对于流域单元的划分至关重要。阈值在水文处理方法中通常用于确定流域内水流的积水区,即在特定的高程差或面积范围内,水流是否能够汇集成流。通过设置一个合适的阈值,可以确保水流积聚的区域得到正确划分,从而避免过度细分或过度简化流域单元^[39]。通过不同阈值选择,研究结果显示 5 000 是最佳的阈值。并根据高分二号卫星图像,将最终的研究范围界定在 226 个

小流域单元中(图 1)。

利用 ArcGIS10.8 软件的区域统计功能依据确定的流域单元对共 14 个流域因子进行离散化处理,其中面积、沟壑密度、圆度、Melton 比率、距道路距离、距断层距离、归一化植被指数、水流强度指数、地形湿度指数、降雨、曲率取平均值作为流域单元中相对应因子的值,在流域单元内,地层岩体和土地利用类型的取值采用频数最多的值进行确定。所有栅格图层通过自然间断法重分类,最终实现了空间分辨率统一为 12.5 m×12.5 m,如图 2 所示。

2.2.2 随机森林模型(RF)

随机森林(random forest, RF)是通过构建多个分类器并综合它们的预测结果来提高整体预测准确性的算法。从原始数据集随机抽取多个子集作为训练集,对于每个训练集,使用随机特征子集构建一棵决策树,在构建每棵决策树时,从可用特征中随机选择一个子集,这

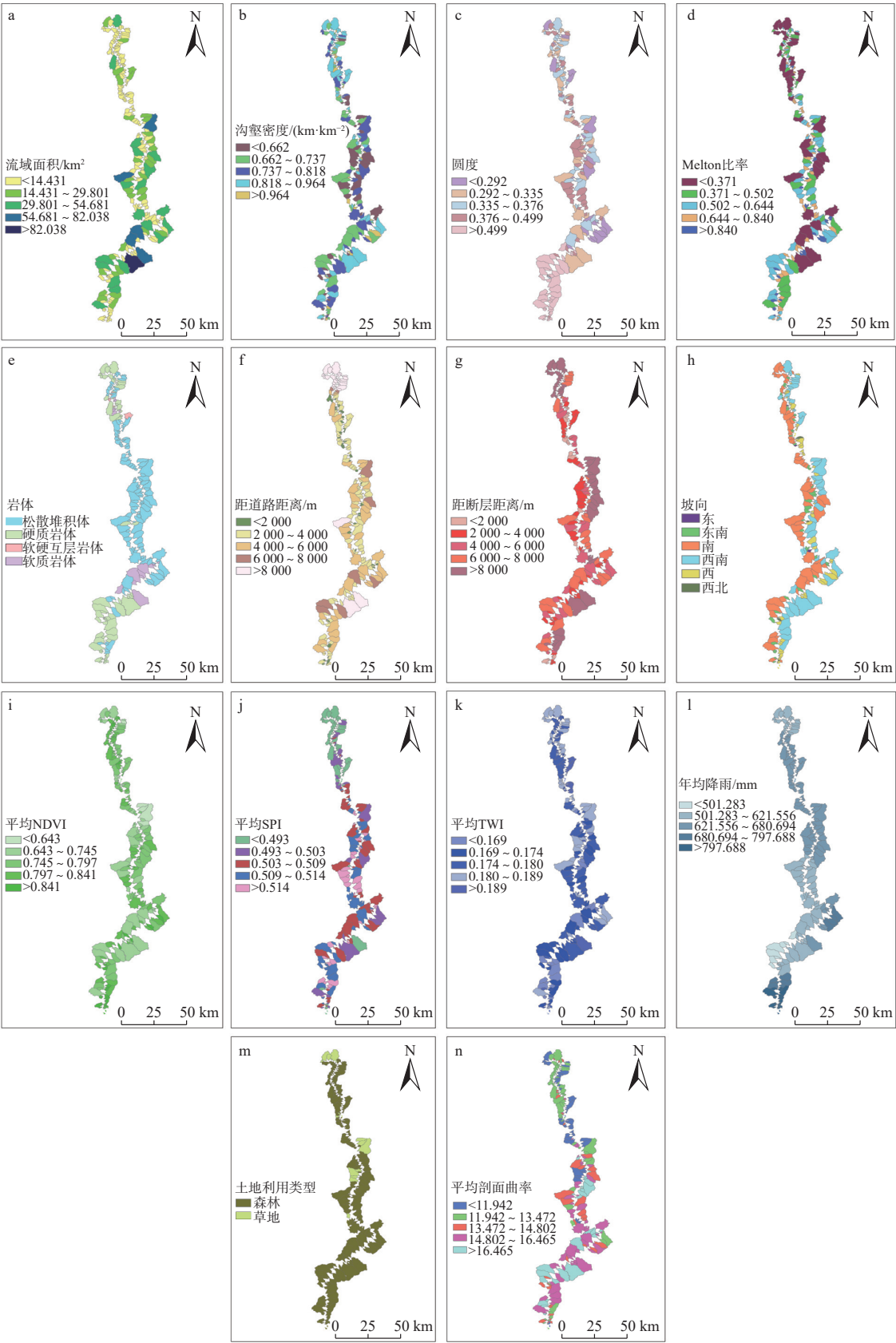


图 2 泥石流流域因子分级图

Fig. 2 Classification map of debris flow watershed factors

有助于减少过拟合^[40]。通过使用训练集训练每棵决策树,直到达到停止条件。当对新数据进行预测时,采用 Bagging 集成方法,将各棵决策树的训练结果整合在一起,最终通过投票或均值计算的方式来决定预测值的分类或回归结果^[41]。

2.2.3 间谍技术(SPY)

基于 SPY 技术的负样本获取策略如下:首先,从泥石流样本中抽取一定比例(本文参考前人研究设定为 15%)作为未标记样本,这部分样本被称为间谍样本^[42]。这些间谍样本与其他未标记样本一起被视为非泥石流样本。在此基础上,采用随机森林(RF)算法训练分类器,以对所有样本进行预测。以间谍样本的最低预测概率为界,低于该阈值的样本将被视为可靠的负样本^[43]。通过间谍样本和未标记样本中的潜在泥石流样本在分类器的预测结果上会表现出相似性的假定,评估间谍样本的概率阈值,能够有效识别出可靠的负样本。

2.2.4 模型精度评价

(1)准确率、精确率、召回率和 F1 分数

泥石流易发性评估过程中的一个重要部分是模型的验证和性能评估。通常,二分类模型的性能通过混淆矩阵来评估^[44]。混淆矩阵由 4 个参数组成:真阳性(true positive, TP)是模型预测为泥石流且实际为泥石流的数量;假阴性(false negative, FN)是模型预测为非泥石流但实际为泥石流的数量;假阳性(false positive, FP)是模型预测为泥石流但实际为非泥石流的数量;真阴性(true negative, TN)是模型预测为非泥石流且实际为泥石流的数量。基于此,通过四个统计指标对每个模型的性能进行评估:准确率、精确率、召回率和 F1 分数。表 2 显示了每个指标的描述。

表 2 模型精度评价指标
Table 2 Model accuracy evaluation indicators

指标	公式	描述
准确率	$\frac{TP+TN}{TP+FP+TN+FN}$	计算准确预测的样本的百分比
精确率	$\frac{TP}{TP+FP}$	计算每个预测的阳性样本中的 TP 样本百分比
召回率	$\frac{TP}{TP+FN}$	计算每个真阳性样本中 TP 样本的百分比
F1 分数	$\frac{2 \times \text{精确率} \times \text{召回率}}{\text{精确率} + \text{召回率}}$	表示准确率和召回率调和平均值,值范围从 -1 到 1

(2)ROC 值和 AUC 曲线

在 ROC 曲线中,纵轴为真阳性率(true positive rate, TPR),横轴为假阳性率(false positive rate, FPR),它反映了模型在各种分类阈值设置下的效果。假阳性率表示被错误预测为阳性样本的比例,真阳性率表示被准确

预测为阳性样本的比例。ROC 曲线越接近左上角,表明模型性能越好^[45]。ROC 曲线下的面积称为曲线下面积(area under the curve, AUC)值,它对模型性能提供了全面的评估。当 AUC 值为 0 时,表示模型的预测性能与随机猜测相当;当 AUC 值为 1 时,表示模型可以完美区分阳性和阴性样本。AUC 值越接近 1,模型的预测能力越强。

2.2.5 研究流程

根据随机森林模型的分类要求,将经历泥石流事件的流域标记为“1”,而未经历泥石流事件的流域则标记为“0”,从而将泥石流易发性评价转化为二分类问题^[46]。在研究区的 226 个小流域单元中,113 个流域曾发生泥石流,另 113 个流域未发生泥石流,形成了正样本与负样本数量相等的样本数据集。在将样本数据集构建为训练集和测试集的过程中,合理的划分比例分配至关重要,因为不合理的划分比例可能会显著影响模型的精度和准确度。因此,参考相关文献的研究,基础数据集按照 70% 用于训练集和 30% 用于测试集的比例进行划分。这一划分策略不仅确保了样本的均衡性,还为随机森林模型的泥石流易发性评估提供了可靠的训练与测试基础,从而增强了模型的泛化能力和预测精度^[47]。

基于前文提出的 SPY 泥石流负样本选取方法,获取基础数据集中的可靠负样本的获取过程可分为以下两个步骤:(1)选取 15% 的泥石流样本作为间谍样本与非泥石流样本组成负样本集,本文负样本集数量为 129,与剩下的 85% 作为泥石流正样本(数量为 97)一起构建数据集,通过 RF 模型对研究区域中的每个样本赋予相应概率值;(2)以间谍样本的最低概率作为界点,概率值低于界点的样本即为可靠负样本。本文对间谍样本预测的结果确定的概率界点为 0.294。在研究区域中所有样本低于间谍样本的概率界点 0.294 的样本数量为 78,因此,生成了 78 个可靠的负样本,并在正样本中随机抽取了 78 个正样本,从而以 1:1 的比例构建了 SPY-RF 模型的基础数据集。

随机森林模型采用了 Python 编程语言,在 Scikit-learn 架构下建立,并通过超参数调优,定义了随机森林模型的最优配置:最大决策树数量为 1 273,最大特征数设为 0.5,最大深度为 6,最小节点样本分割数为 5。基于这些最优参数,分别构建了传统随机森林(RF)和改进的随机森林(SPY-RF)泥石流易发性评价模型。计算每个流域单元的泥石流易发性指数,并绘制了泥石流易发性分级图,以便进行系统的易发性评估。技术路线图如图 3 所示。

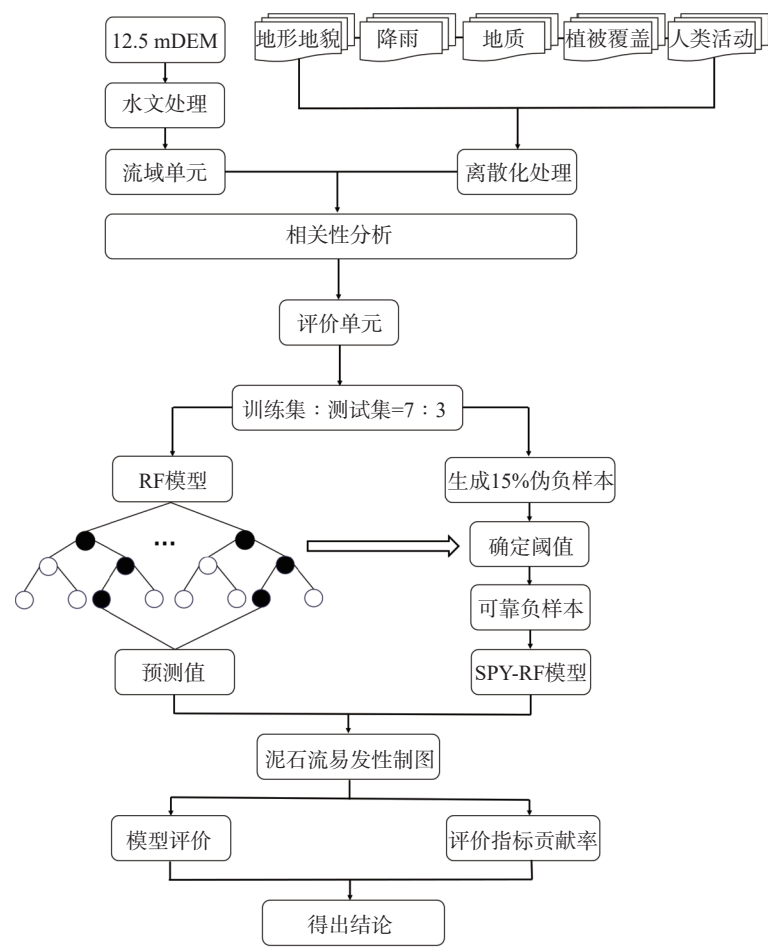


图3 技术路线图
Fig. 3 Technology roadmap

3 泥石流易发性分析

3.1 泥石流易发性评价指标相关性分析

为选择最具预测能力的评价因子,提升模型预测的准确性,对初步所选择的14个因子进行相关性分析。使用Origin绘图软件的Correlation Plot插件,可以得到14个评价指标之间的相关系数矩阵,并可视化得到图4。正相关在图中以红色标识,负相关以蓝色标识。相关系数的大小与图像的大小直接相关。由图可知,所有评价因子的相关系数均小于0.48,相关系数较弱,表明因子交互作用较小,选择的评价因子较为合理。

3.2 易发性模型制图分析

在RF和SPY-RF模型的泥石流易发性模型建立后,利用模型计算了研究区各流域单元的泥石流易发性指数。基于ArcGIS10.8软件中的自然断点法,将泥石流易发性指数划分为五个等级,依次为极低易发区、低易发区、中易发区、高易发区和极高易发区,如图5所示。通过研究区泥石流易发性分级图显示,基于SPY-

RF模型得到的泥石流易发性区域与实际发生的极高易发区范围大致一致。这一结果进一步验证了模型的有效性,并为泥石流风险管理提供了重要依据。

泥石流易发性评价结果亦可通过统计方法进行分析。表3分析了各模型泥石流易发性等级的流域单元数量、各等级所占的面积以及对应的泥石流点,并统计每个易发性等级的泥石流点占总数的比例与相应面积占比之间的比率,得出各易发性等级的泥石流密度。统计图如图6所示。从表3和图6可以观察到各易发性等级内泥石流的分布特点与密度差异。RF模型中高易发和极高易发区域分别占比29.15%和24.97%,这两个等级的泥石流个数合计为91个,占泥石流总数的80%以上。然而,RF模型在低易发和极低易发等级的预测结果较为保守,例如极低易发区域没有泥石流发生。相比之下,SPY-RF模型在极高易发等级覆盖了40.86%的面积,流域个数达到74个,并识别了72个泥石流点,表明它对高风险区域更敏感。此外,SPY-RF模型在极低易发和低易发区域的泥石流密度分别为 0.46×10^{-2} 个/ km^2

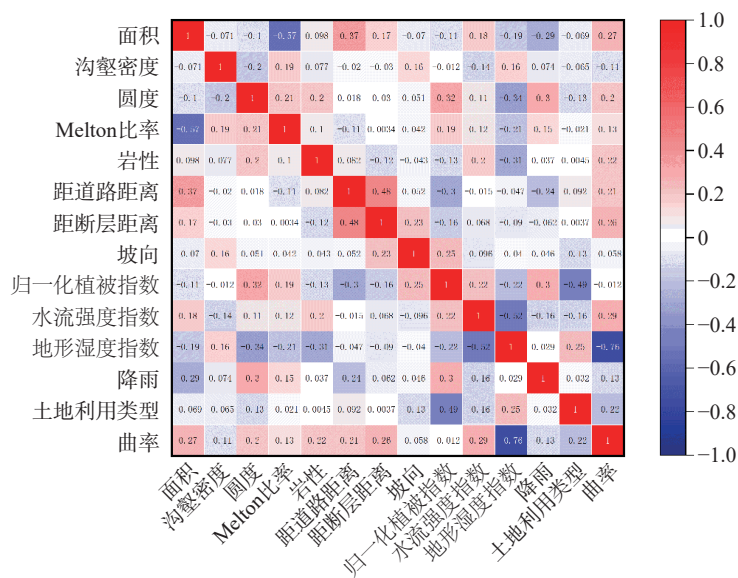


图 4 泥石流流域因子相关性分析

Fig. 4 Correlation analysis of debris flow catchment factors

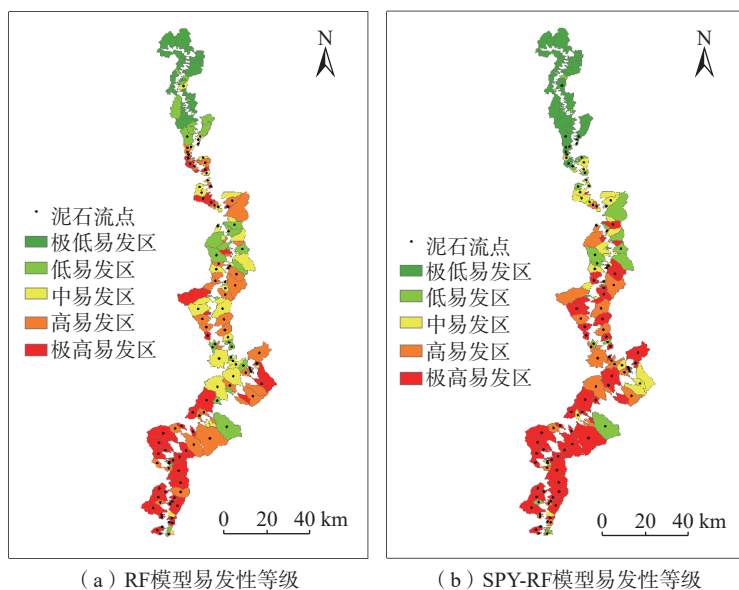


图 5 泥石流易发性分级图

Fig. 5 Debris flow susceptibility grading map

和 0.85×10^{-2} 个/ km^2 , 表明它在低风险区域也能识别泥石流灾害。综合来看, SPY-RF 模型比 RF 模型在不同风险等级上表现得更好, 该模型能够更全面地反映不同易发性等级的泥石流, 适合更广泛的应用场景。不仅适合聚焦于高风险区域的精确预测, 也能在低风险区域显示出更高的敏感性和预测能力。

3.3 模型精度验证

根据前文研究方法, 本研究将样本数据集按照 7:3 的比例, 构建 70% 的训练数据集以及 30% 的测试数据集。分别统计 TP、FN、FP、TN 值, 并计算模型性能

指标准确率、精确率、召回率和 F1 分数, 如图 7 所示。在泥石流易发性评价中, SPY-RF 的整体表现优于 RF。尽管 RF 在召回率上表现最佳, SPY-RF 的高准确率、高精确率和高 F1 分数使其在实际应用中更具优势。虽然 RF 能识别所有易发区域, 但 SPY-RF 的综合优势表明其在实际应用中能够提供更平衡、更可靠的评估结果。

本文基于 RF 和 SPY-RF 的泥石流易发性评价模型的 ROC 曲线如图 8 所示。本文 SPY-RF 模型的 AUC 值为 0.98 略高于 RF 模型的 AUC 值 0.93。在泥石流易发性评估中, SPY-RF 模型表现最为优越。因此, 接下来

表 3 不同模型泥石流易发性等级及泥石流分布特点

Table 3 Debris flow distribution across susceptibility classes and characteristics predicted by different models

模型	易发性等级	流域个数	面积/km ²	面积占比/%	泥石流个数	泥石流占比/%	泥石流密度/(10 ⁻² 个/km ²)
RF	极低易发	51	747.28	10.20	0	0.00	0.00
	低易发	40	1 236.98	16.88	6	5.31	0.49
	中易发	41	1 376.40	18.79	16	14.16	1.16
	高易发	50	2 135.76	29.15	47	41.59	2.20
	极高易发	44	1 829.57	24.97	44	38.94	2.40
SPY-RF	极低易发	70	1 305.84	16.46	6	5.31	0.46
	低易发	22	943.89	12.88	8	7.08	0.85
	中易发	28	808.26	11.03	8	7.08	0.99
	高易发	30	1 374.51	18.76	19	16.81	1.38
	极高易发	74	2 893.5	40.86	72	63.72	2.49

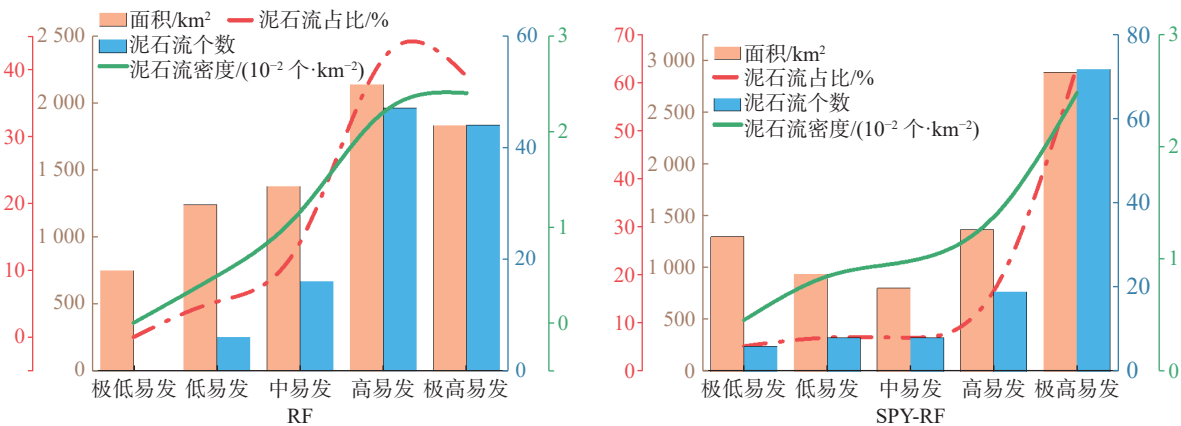


图 6 不同模型泥石流易发性等级统计图

Fig. 6 Statistical comparison of debris flow susceptibility levels across different models

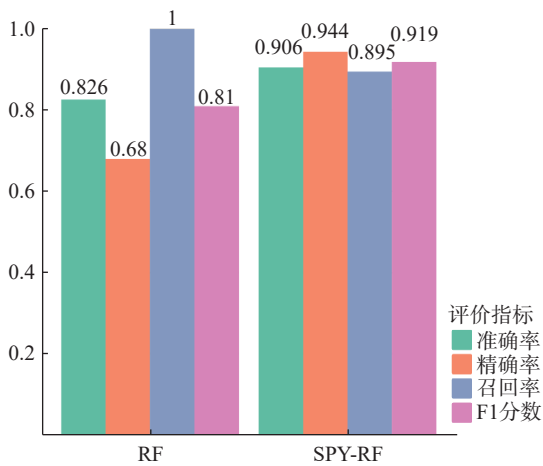


图 7 模型性能指标

Fig. 7 Model performance metrics

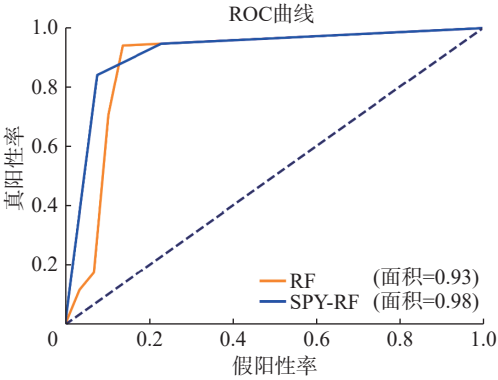


图 8 模型 AUC 值与 ROC 曲线

Fig. 8 Model AUC values vs. ROC curves

的分析将围绕 SPY-RF 模型的评估结果, 探讨泥石流易发性及各评价因子的贡献率。

3.4 泥石流评价指标贡献率分析

基于随机森林模型, 根据 Python 软件中 feature_

importances 函数计算泥石流易发性各评价指标的贡献率, 计算式见式(1):

$$IP = \frac{IP_i}{\sum_{i=1}^N IP_i}$$

(1)

式中: IP_i ——第*i*个因子的贡献率。

对研究区泥石流评价指标贡献率的分析结果表明,如图 9 所示,水流强度指数 (SPI)、降雨量、曲率和流域面积是影响泥石流易发性的重要因素,其中 SPI 的贡献率最大。

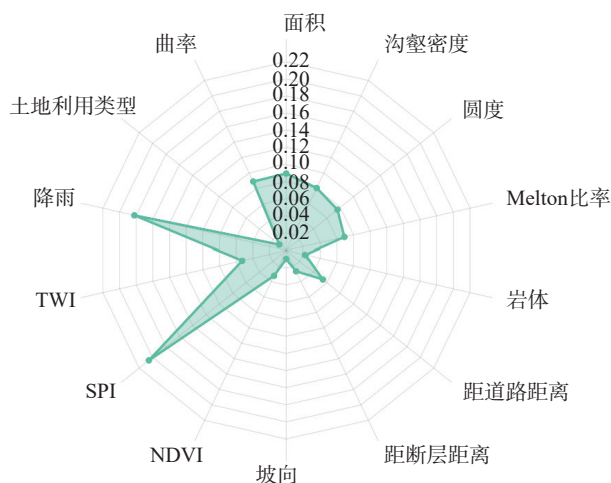


图 9 泥石流流域因子贡献率

Fig. 9 Contribution rates of debris flow susceptibility factors [47]

4 讨论

在泥石流易发性评价的研究中,研究者们通常将重点放在分析已经发生或有可能发生泥石流灾害的正样本上,而对未发生灾害的负样本关注较少。然而,忽视负样本的研究可能导致模型在灾害预测中的过度拟合,从而降低预测的准确性。负样本不仅能帮助识别灾害发生的潜在风险区,还能为评估灾害未发生的因素提供重要的参考,进而提升泥石流易发性评价的全面性和可靠性。本文采用间谍技术选择可靠负样本,构建泥石流易发性评价的基础数据集。如表 3 和图 6 所示,SPY-RF 模型相比于 RF 模型能更准确地识别极高易发性区域的泥石流事件,显著降低非易发区域被误判为易发区域的概率,增强了风险评估的可靠性。此外,基于间谍技术的模型在准确率、精确率、召回率和 AUC 值等指标上均高于原模型。这一性能提升归因于间谍技术在负样本选择和特征提取中的优势,使得模型更精准地识别易发与非易发区域。这些结果进一步验证了采用间谍技术进行泥石流易发性评估的有效性,为未来的研究提供了重要的理论支持。

在本文泥石流评价指标贡献率分析中,水流强度指数的贡献率最大,表明水流强度在泥石流发生中起到了关键作用,因为强水流可以加剧土壤侵蚀和泥石流的形成。从图 2 和图 5 可以看出,松潘等地区水流强度较弱,水流对地表的侵蚀作用有限;该区域属于高原气候,

年降水量相较四川盆地其他地区较少;地形上,许多区域呈现较为平缓的曲率,如山脊和平原地带;同时,松潘地区流域面积较小,尤其在一些分散的山谷区域,流域面积相对有限。这些因素共同作用,有效抑制了泥石流灾害的发生。然而,在汶川等下游深切河谷区域,地势陡峭,部分地区呈现较为明显的负曲率,流域面积相对较大,水流能够在较大范围内汇聚并增加冲刷强度;此外,汶川地区夏季降水量集中,进一步增强了水流的侵蚀作用。这些综合因素导致汶川地区泥石流灾害的发生风险显著增加。贡献率较小的因素包括土地利用类型、坡向、距断层距离和岩体,其中土地利用类型和坡向的贡献较低,可能与研究区域的土地利用类型和坡向相对均匀或者影响泥石流发生的其他因素更加突出有关;距断层距离和岩体的贡献率最低,说明在特定研究区,断层和岩体的空间分布可能没有形成显著的泥石流易发性模式。若研究区的泥石流主要受降雨、地形等因素影响,而断层和岩体的分布未能形成明显的影响特征,则其贡献率自然较低。这一分析结果为未来研究中选择和应用影响因子提供了重要依据。

5 结果

(1)本文通过建立 RF 和 SPY-RF 模型对岷江上游流域泥石流进行易发性评价,通过对模型进行参数调优以及负样本优化,SPY-RF 模型和 RF 模型的 AUC 值分别为 0.98、0.93,且 SPY-RF 模型性能指标整体优于 RF 模型。表明在负样本获取和筛选策略上,采用 SPY 技术显著提高了负样本的质量,从而提升了模型的精确度和稳定性。

(2)分析模型的预测结果发现,RF 模型主要集中在高易发和极高易发区域,分别占比 29.15% 和 24.97%,这两个等级的泥石流个数合计为 91 个,占泥石流总数的 80% 以上。相比之下,SPY-RF 模型在极高易发等级覆盖了 40.86% 的面积,流域个数达到 74 个,并识别了 72 个泥石流点,表明它对极高易发区域更敏感。SPY-RF 模型通过优化负样本的选择,使得模型在各个易发性等级上表现更加全面,尤其是在高风险区域的预测准确性显著提高。

(3)对研究区泥石流评价指标贡献率进行分析发现,贡献率最大的 4 个因子分别是水流强度指数、降雨、曲率和面积,贡献率最小的 4 个因子分别是土地利用类型、坡向、距断层距离和岩体。在制定岷江上游流域泥石流地质灾害风险管控措施时,应着重考虑区内受降雨和地形等因素。这些主要因素对泥石流的爆发具有重要的作用,而其他因子的作用则相对较小,因此针

对性地优化降雨管理和地形改造措施将是提升流域泥石流防治效果的关键。

参考文献(References):

- [1] CUI Peng, DANG Chao, ZHUANG Jianqi, et al. Landslide-dammed lake at Tangjiashan, Sichuan Province, China (triggered by the Wenchuan Earthquake, May 12, 2008): Risk assessment, mitigation strategy, and lessons learned [J]. *Environmental Earth Sciences*, 2012, 65(4): 1055–1065.
- [2] JIANG Weiguo, RAO Pingzeng, CAO Ran, et al. Comparative evaluation of geological disaster susceptibility using multi-regression methods and spatial accuracy validation [J]. *Journal of Geographical Sciences*, 2017, 27(4): 439–462.
- [3] 陈哲锋, 郭朝旭. 基于模糊层次分析法的泥石流易发性地形因子敏感性分析 [J]. 防灾科技学院学报, 2023, 25(1): 21–30. [CHEN Zhefeng, GUO Chaoxu. Sensitivity analysis of terrain factors of debris flow vulnerability based on fuzzy analytic hierarchy process [J]. *Journal of Institute of Disaster Prevention*, 2023, 25(1): 21–30. (in Chinese with English abstract)]
- [4] 崔志超, 王俊豪, 崔传峰, 等. 基于层次分析法和模糊数学相结合的甘肃东乡八丹沟泥石流易发性评价 [J]. 中国地质灾害与防治学报, 2020, 31(1): 44–50. [CUI Zhichao, WANG Junhao, CUI Chuanfeng, et al. Evaluation of the susceptibility of debris flow in Badan gully of Dongxiang County of Gansu based on AHP and fuzzy mathematics [J]. *The Chinese Journal of Geological Hazard and Control*, 2020, 31(1): 44–50. (in Chinese with English abstract)]
- [5] 李刚, 赵新生, 封丽华, 等. 基于模糊评价法优选泥石流防治方案——以青河县泥石流为例 [J]. 中国地质灾害与防治学报, 2014, 25(4): 18–23. [LI Gang, ZHAO Xinsheng, FENG Lihua, et al. Application of fuzzy mathematics in geological disaster prevention and treatment research for debris flow in Qinghe County [J]. *The Chinese Journal of Geological Hazard and Control*, 2014, 25(4): 18–23. (in Chinese with English abstract)]
- [6] 牛全福, 陆铭, 李月锋, 等. 基于灰色关联与粗糙依赖度的甘肃兰州市区泥石流危险性评价 [J]. 中国地质灾害与防治学报, 2019, 30(5): 48–56. [NIU Quanfu, LU Ming, LI Yuefeng, et al. Hazard assessment of debris flow in Lanzhou City of Gansu Province based on methods of grey relation and rough dependence [J]. *The Chinese Journal of Geological Hazard and Control*, 2019, 30(5): 48–56. (in Chinese with English abstract)]
- [7] 王念秦, 姚勇. 基于模糊数学和权的最小平方方法的泥石流易发性评价方法 [J]. 灾害学, 2008, 23(2): 5–9. [WANG Nianqin, YAO Yong. Method of debris-flow proneness evaluation based on fuzzy mathematics and least-square method [J]. *Journal of Catastrophology*, 2008, 23(2): 5–9. (in Chinese with English abstract)]
- [8] 于森, 邢会歌, 胡士瑜. 基于信息量-逻辑回归模型的泥石流易发性评价——以四川省石棉县为例 [J]. 人民长江, 2021, 52(12): 107–114. [YU Miao, XING Huige, HU Shiyu. Debris flow susceptibility assessment based on information value and logistic regression coupled model: Case of Shimian County, Sichuan Province [J]. *Yangtze River*, 2021, 52(12): 107–114. (in Chinese with English abstract)]
- [9] 张永宏, 葛涛涛, 田伟, 等. 基于地质大数据的泥石流灾害易发性评价 [J]. 计算机应用, 2018, 38(11): 3319–3325. [ZHANG Yonghong, GE Taotao, TIAN Wei, et al. Evaluation of susceptibility to debris flow hazards based on geological big data [J]. *Journal of Computer Applications*, 2018, 38(11): 3319–3325. (in Chinese with English abstract)]
- [10] 朱海丽. 海东市乐都区庙沟泥石流易发性评价 [J]. 中国水土保持, 2015(9): 55–58. [ZHU Haili. Evaluation on susceptibility of debris flow of Miaogou gully in Ledu District of Haidong City [J]. *Soil and Water Conservation in China*, 2015(9): 55–58. (in Chinese with English abstract)]
- [11] 徐艳琴, 白淑英, 徐永明. 基于两种方法的攀西泥石流易发性评价对比分析 [J]. 水土保持研究, 2018, 25(3): 285–291. [XU Yanqin, BAI Shuying, XU Yongming. Comparative analysis of debris flow susceptibility assessment based on two methods in Panxi District [J]. *Research of Soil and Water Conservation*, 2018, 25(3): 285–291. (in Chinese with English abstract)]
- [12] LI Yimin, JIANG Wenxue, FENG Xianjie, et al. Debris flow susceptibility mapping in Alpine canyon region: A case study of Nujiang Prefecture [J]. *Bulletin of Engineering Geology and the Environment*, 2024, 83(5): 169.
- [13] LI Yongwei, XU Linrong, SHANG Yonghui, et al. Debris flow susceptibility evaluation in meizoseismal region: A case study in Jiuzhaigou, China [J]. *Journal of Earth Science*, 2024, 35(1): 263–279.
- [14] XIONG Ke, ADHIKARI B R, STAMATOPOULOS C A, et al. Comparison of different machine learning methods for debris flow susceptibility mapping: A case study in the Sichuan Province, China [J]. *Remote Sensing*, 2020, 12(2): 295.
- [15] ZHANG Yonghong, GE Taotao, TIAN Wei, et al. Debris flow susceptibility mapping using machine-learning techniques in Shigatse area, China [J]. *Remote Sensing*, 2019, 11(23): 2801.
- [16] 郭鹏宁, 邢会歌, 李从江, 等. 基于OOD泛化性验证和深度全连接神经网络的泥石流易发性评价方法 [J]. 工程科学与技术, 2024, 56(4): 182–193. [GUO Pengning, XING Huige, LI Congjiang, et al. Methods for evaluating debris flow susceptibility based on OOD generalization verification and deep fully connected neural networks [J]. *China Industrial Economics*, 2024, 56(4): 182–193. (in Chinese with English abstract)]

- [17] 李志, 陈宁生, 侯儒宁, 等. 基于机器学习的伊犁河谷黄土区泥石流易发性评估 [J]. 中国地质灾害与防治学报, 2024, 35(3): 129 – 140. [LI Zhi, CHEN Ningsheng, HOU Runing, et al. Susceptibility assessment of debris flow disaster based on machine learning models in the loess area along Yili Valley [J]. The Chinese Journal of Geological Hazard and Control, 2024, 35(3): 129 – 140. (in Chinese with English abstract)]
- [18] 刘国栋, 秦胜伍, 孟凡奇, 等. 基于地理信息相似度的负样本采样策略在泥石流易发性评价中的应用 [J]. 工程地质学报, 2023, 31(2): 526 – 537. [LIU Guodong, QIN Shengwu, MENG Fanqi, et al. Application of geographic information similarity based absence sampling method to debris flow susceptibility mapping [J]. Journal of Engineering Geology, 2023, 31(2): 526 – 537. (in Chinese with English abstract)]
- [19] 李小龙, 宋国虎, 向灵芝, 等. 基于不同评价单元和灾害熵的泥石流危险性分析——以白龙江流域武都段为例 [J]. 中国地质灾害与防治学报, 2021, 32(6): 107 – 115. [LI Xiaolong, SONG Guohu, XIANG Lingzhi, et al. Hazard analysis of debris flows based on different evaluation units and disaster entropy: A case study in Wudu section of the Bailong River Basin [J]. The Chinese Journal of Geological Hazard and Control, 2021, 32(6): 107 – 115. (in Chinese with English abstract)]
- [20] 薛永安, 王玉洁, 朱婧聪, 等. 基于 CF 与 SVM 的小样本斜坡地质灾害敏感性评价研究 [J]. 太原理工大学学报, 2022, 53(4): 672 – 681. [XUE Yongan, WANG Yujie, ZHU Jingcong, et al. Study of slope geological hazard susceptibility valuation with small sample based on CF and SVM [J]. Journal of Taiyuan University of Technology, 2022, 53(4): 672 – 681. (in Chinese with English abstract)]
- [21] KE Ting, YANG Bing, ZHEN Ling, et al. Building high-performance classifiers using positive and unlabeled examples for text classification [C] // Advances in Neural Networks – ISNN 2012. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012: 187 – 195
- [22] LEE W S, LIU B. Learning with positive and unlabeled examples using weighted logistic regression [J]. DBLP, 2003.
- [23] 于海坤, 欧阳九发, 王丙千, 等. 不同采样策略下的区域滑坡易发性评价 [J]. 安全与环境工程, 2024, 31(5): 122 – 134. [YU Haikun, OUYANG Jiufa, WANG Bingqian, et al. Susceptibility assessment of regional landslides under different sampling strategies [J]. Safety and Environmental Engineering, 2024, 31(5): 122 – 134. (in Chinese with English abstract)]
- [24] 洪浩源, 王德生, 朱阿兴. 面向机器学习型区域滑坡易发性评价的训练样本采样方法 [J]. 地理学报, 2024, 79(7): 1718 – 1736. [HONG Haoyuan, WANG Desheng, ZHU Axing. A new training data sampling method for machine learning-based landslide susceptibility mapping [J]. Acta Geographica Sinica, 2024, 79(7): 1718 – 1736. (in Chinese with English abstract)]
- [25] 王东伟, 游勇, 李道凌, 等. 汶川县绵虬镇板子沟“8·20”大型泥石流堵河特征及危害性预测 [J]. 中国地质灾害与防治学报, 2022, 33(1): 58 – 66. [WANG Dongwei, YOU Yong, LI Daoling, et al. The river blockage characteristics of “8·20” large-scale debris flow and the hazard prediction in Banzi gully in Miansi Town, Wenchuan County [J]. The Chinese Journal of Geological Hazard and Control, 2022, 33(1): 58 – 66. (in Chinese with English abstract)]
- [26] 张宪政, 铁永波, 宁志杰, 等. 四川汶川县板子沟“6·26”特大型泥石流成因特征与活动性研究 [J]. 水文地质工程地质, 2023, 50(5): 134 – 145. [ZHANG Xianzheng, TIE Yongbo, NING Zhijie, et al. Characteristics and activity analysis of the catastrophic “6·26” debris flow in the Banzi catchment, Wenchuan County of Sichuan Province [J]. Hydrogeology & Engineering Geology, 2023, 50(5): 134 – 145. (in Chinese with English abstract)]
- [27] 曾超, 曹振宇, 苏凤环, 等. 四川及周边滑坡泥石流灾害高精度航空影像及解译数据集(2008—2020年) [J]. 中国科学数据, 2022, 7(2): 195 – 205. [ZENG Chao, CAO Zhenyu, SU Fenghuan, et al. A dataset of high-precision aerial imagery and interpretation of landslide and debris flow disaster in Sichuan and surrounding areas between 2008 and 2020 [J]. China Scientific Data, 2022, 7(2): 195 – 205. (in Chinese with English abstract)]
- [28] 姚振国, 刘建周, 牛贝贝, 等. 流域面积对沟道泥石流发育的影响分析 [J]. 资源环境与工程, 2019, 33(2): 217 – 219. [YAO Zhenguo, LIU Jianzhou, NIU Beibei, et al. Influence of drainage area on development of debris flow in gully [J]. Resources Environment & Engineering, 2019, 33(2): 217 – 219. (in Chinese with English abstract)]
- [29] 王向前, 丁明涛, 何雨枫. 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, 2023, 38(3): 222 – 226. [WANG Xiangqian, DING Mingtao, HE Yufeng. Assessment of debris flow vulnerability in Xiangcheng County, Ganzi Tibetan Autonomous Prefecture based on RF model [J]. Journal of Catastrophology, 2023, 38(3): 222 – 226. (in Chinese with English abstract)]
- [30] 王峰, 杨帆, 江忠荣, 等. 基于沟域单元的康定市泥石流易发性评价 [J]. 中国地质灾害与防治学报, 2023, 34(3): 145 – 156. [WANG Feng, YANG Fan, JIANG Zhongrong, et al. Susceptibility assessment of debris flow based on watershed units in Kangding City, Sichuan Province [J]. The Chinese Journal of Geological Hazard and Control, 2023, 34(3): 145 – 156. (in Chinese with English abstract)]
- [31] 毕鸿基, 聂磊, 曾超, 等. 基于三种多变量不安定指数分析模型的汶川县地质灾害易发性评价 [J]. 中国地质灾害与防治学报, 2022, 33(1): 123 – 131. [BI Hongji, NIE

- Lei, ZENG Chao, et al. Geological hazard susceptibility evaluation in Wenchuan area based on three models of multivariate instability index analysis [J]. The Chinese Journal of Geological Hazard and Control, 2022, 33(1): 123 – 131. (in Chinese with English abstract)]
- [32] XING Fucheng, LI Ning, ZHAO Boju, et al. An investigation into the susceptibility to landslides using integrated learning and Bayesian optimization: A case study of Xichang City [J]. *Sustainability*, 2024, 16(20): 9085.
- [33] CHEN Yutao, LI Ning, ZHAO Boju, et al. Comparison of informative modelling and machine learning methods in landslide vulnerability evaluation: A case study of Wenchuan County, China [J]. *Geocarto International*, 2024, 39(1): 2361714.
- [34] PALAMAKUMBURE D, FLENTJE P, STIRLING D. Consideration of optimal pixel resolution in deriving landslide susceptibility zoning within the Sydney Basin, New South Wales, Australia [J]. *Computers & Geosciences*, 2015, 82: 13 – 22.
- [35] ROSSI M, GUZZETTI F, REICHENBACH P, et al. Optimal landslide susceptibility zonation based on multiple forecasts [J]. *Geomorphology*, 2010, 114(3): 129 – 142.
- [36] 刘善军, 刘诗瑶, 魏恋欢, 等. 基于子流域单元的抚顺市泥石流易发性及危险性评价 [J]. 东北大学学报(自然科学版), 2024, 45(5): 713 – 720. [LIU Shanjun, LIU Shiyao, WEI Lianhuan, et al. Debris flow Susceptibility and hazard assessment in Fushun City based on hydrological response units [J]. *Journal of Northeastern University(Natural Science)*, 2024, 45(5): 713 – 720. (in Chinese with English abstract)]
- [37] 田尤, 黄海, 高波, 等. 考虑冻融侵蚀型物源的不同流域单元泥石流易发性评价——以藏东贡觉地区为例 [J]. 冰川冻土, 2024, 46(1): 40 – 51. [TIAN You, HUANG Hai, GAO Bo, et al. Susceptibility evaluation for debris flow in different watershed units considering freeze-thaw erosion type sources: Taking Gonjo area of eastern Tibet as an example [J]. *Journal of Glaciology and Geocryology*, 2024, 46(1): 40 – 51. (in Chinese with English abstract)]
- [38] 侯儒宁, 李志, 陈宁生, 等. 基于流域单元和堆叠集成模型的天山地区泥石流易发性评估建模 [J]. 地球科学, 2023, 48(5): 1892 – 1907. [HOU Runing, LI Zhi, CHEN Ningsheng, et al. Modeling of debris flow susceptibility assessment in Tianshan based on watershed unit and stacking ensemble algorithm [J]. *Earth Science*, 2023, 48(5): 1892 – 1907. (in Chinese with English abstract)]
- [39] 韩玲, 张恒, 张庭瑜, 等. 基于 DEM 的黄土沟道提取汇流量阈值确定方法 [J]. 长江科学院院报, 2020, 37(5): 174 – 179. [HAN Ling, ZHANG Heng, ZHANG Tingyu, et al. Determining runoff accumulation threshold for loess gully extraction based on DEM [J]. *Journal of Yangtze River Scientific Research Institute*, 2020, 37(5): 174 – 179. (in Chinese with English abstract)]
- [40] 饶姗姗, 冷小鹏. 基于 RSIV-RF 模型的凉山州泥石流易发性评价 [J]. 地质科技通报, 2024, 43(1): 275 – 287. [RAO Shanshan, LENG Xiaopeng. Debris flow susceptibility evaluation of Liangshan Prefecture based on the RSIV-RF model [J]. *Bulletin of Geological Science and Technology*, 2024, 43(1): 275 – 287. (in Chinese with English abstract)]
- [41] 赵佳忆, 田述军, 李凯, 等. 岷江上游汶川县地震前后泥石流易发性评价 [J]. 中国地质灾害与防治学报, 2024, 35(1): 51 – 59. [ZHAO Jiayi, TIAN Shujun, LI Kai, et al. Susceptibility assessment of debris flow in the upper reaches of the Minjiang River before and after the Wenchuan earthquake [J]. *The Chinese Journal of Geological Hazard and Control*, 2024, 35(1): 51 – 59. (in Chinese with English abstract)]
- [42] 高瑞源. 基于机器学习方法的区域泥石流易发性评价研究 [D]. 长春: 吉林大学, 2023. [GAO Ruiyuan. Study on the regional debris flow susceptibility evaluation based on machine learning methods [D]. Changchun: Jilin University, 2023. (in Chinese with English abstract)]
- [43] YAO Jingyu, QIN Shengwu, QIAO Shuangshuang, et al. Application of a two-step sampling strategy based on deep neural network for landslide susceptibility mapping [J]. *Bulletin of Engineering Geology and the Environment*, 2022, 81(4): 148.
- [44] 李坤, 赵俊三, 林伊琳, 等. 基于 RF 和 SVM 模型的东川泥石流易发性评价研究 [J]. 云南大学学报(自然科学版), 2022, 44(1): 107 – 115. [LI Kun, ZHAO Junsan, LIN Yiling, et al. Assessment of debris flow susceptibility in Dongchuan based on RF and SVM models [J]. *Journal of Yunnan University(Natural Sciences Edition)*, 2022, 44(1): 107 – 115. (in Chinese with English abstract)]
- [45] 饶品增, 曹冉, 蒋卫国. 基于地理加权回归模型的云南省地质灾害易发性评价 [J]. 自然灾害学报, 2017, 26(2): 134 – 143. [RAO Pinzeng, CAO Ran, JIANG Weiguo. Susceptibility evaluation of geological disasters in Yunnan Province based on geographically weighted regression model [J]. *Journal of Natural Disasters*, 2017, 26(2): 134 – 143. (in Chinese with English abstract)]
- [46] ERENER A, SIVAS A A, SELCUK-KESTEL A S, et al. Analysis of training sample selection strategies for regression-based quantitative landslide susceptibility mapping methods [J]. *Computers & Geosciences*, 2017, 104: 62 – 74.
- [47] 郭玲, 薛晔, 孙鹏翔. 基于 LA-GraphCAN 的甘肃省泥石流易发性评价 [J]. 地质科技通报, [2024-09-23] (2024-10-31). [GUO Ling, XUE Hua, SUN Pengxiang. Susceptibility evaluation of debris flow in Gansu Province based on LA-GraphCAN [J]. *Bulletin of Geological Science and Technology*, [2024-09-23] (2024-10-31). <https://doi.org/10.19509/j.cnki.dzkg.tb2024032410.19509/j.cnki.dzkg.tb20240324>. (in Chinese with English abstract)]