https://doi.org/10.3799/dqkx.2025.306

基于频率比与K-means聚类的无需非滑坡样本的滑坡易发性评价新方法

魏亚兰[1],姜元俊[2*],阿比尔的[3,4]

1. 重庆交通大学河海学院,重庆 400074
2. 中国科学院、水利部成都山地灾害与环境研究所,四川成都 610213
3. 重庆交通大学国家内河航道整治工程技术研究中心,重庆 400074
4. 重庆交通大学山区公路水运交通地质减灾重庆市教委重点实验室,重庆 400074

摘　要: 为缓解传统滑坡易发性评价对非滑坡样本依赖导致的空间代表性不足与精度不确定问题,提出一种无需非滑坡样本的FR-K-means易发性评价方法,并以随机森林(RF)作为对比.该方法基于频率比(FR)量化关键因子与滑坡发生的条件概率,构建综合因子数据集;进一步利用K-means对致灾环境进行相似性聚类,并结合自然断点法实现易发性分区.以四川平武县为研究区、通江县为验证区的结果表明,FR-K-means在极高易发区滑坡捕捉能力优于RF,滑坡密度分别为0.124和0.098个/km^2(RF为0.112和0.092个/km^2).同时,该方法可清晰刻画流域两岸高风险带,有效减少流域内部极低易发区出现高易发值,减少非孕灾区误判.研究表明FR-K-means具有较好的空间聚焦能力与区域迁移性,可为复杂地质环境区滑坡风险快速制图与精细化管控提供补充思路.

关键词: 滑坡易发性;非滑坡样本采样;频率比;K-means聚类;滑坡密度;工程地质.

中图分类号: P642.22　　**文章编号:** 1000-2383(2026)10-3997-16　　**收稿日期:** 2025-10-09

Integrated Frequency Ratio and K-Means Clustering for Landslide Susceptibility Mapping without Non-Landslide Samples

Wei Yalan[1], Jiang Yuanjun[2*], ABI Erdi[3,4]

1. *School of River and Ocean, Chongqing Jiaotong University, Chongqing* 400074, *China*
2. *Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu* 610213, *China*
3. *National Engineering Research Center for Inland Way Regulation, Chongqing Jiaotong University, Chongqing* 400074, *China*
4. *Key Laboratory of Geological Hazards Mitigation of Mountainous Highway and Waterway, Chongqing Municipal Education Commission, Chongqing Jiaotong University, Chongqing* 400074, *China*

Abstract: To alleviate the insufficient spatial representativeness and accuracy uncertainty caused by the reliance on non-landslide samples in traditional landslide susceptibility assessment, this study proposes a non-landslide-sample-free FR-K-means

基金项目: 重庆市自然科学基金院士专项(Nos.CSTB2023YSZX-JCX0004,CSTC2021YSZX-JCYJX0002);重庆市研究生联合培养基地建设项目(No.JDLHPYJD2022004).

作者简介: 魏亚兰(2000-),女,硕士研究生,研究方向:滑坡易发性研究.ORCID:0009-0000-0863-5658. E-mail:weiyl_008@163.com

* **通讯作者:** 姜元俊(1982-),男,博士,研究员,研究方向:滑坡启动机理及预警技术.E-mail:yjjiang@imde.ac.cn

引用格式: 魏亚兰,姜元俊,阿比尔的,2026.基于频率比与K-means聚类的无需非滑坡样本的滑坡易发性评价新方法.地球科学,51(10):3997-4012.

Citation: Wei Yalan, Jiang Yuanjun, ABI Erdi, 2026. Integrated Frequency Ratio and K-Means Clustering for Landslide Susceptibility Mapping without Non-Landslide Samples. *Earth Science*, 51(10): 3997-4012.

susceptibility evaluation method, with the random forest (RF) model adopted for comparison. The frequency ratio (FR) model is used to quantify the conditional probabilities between key environmental factors and landslide occurrence, thereby constructing an integrated factor dataset. The K-means algorithm is then applied to cluster similar hazard-inducing environmental conditions, and the natural breaks method is employed to delineate landslide susceptibility zones. Using Pingwu County as the study area and Tongjiang County as the validation area in Sichuan Province, the results indicate that the FR-K-means method outperforms the RF model in capturing landslides within very-high-susceptibility zones, with landslide densities of 0.124 and 0.098 events/km^2, respectively, compared with 0.112 and 0.092 events/km^2 for RF. In addition, the proposed method clearly delineates high-risk belts along both sides of river basins and effectively reduces the misclassification of extremely low-susceptibility areas within basin interiors as high-susceptibility zones, thereby decreasing false identification in non-hazardous regions. The results demonstrate that the FR-K-means method exhibits favorable spatial focusing capability and regional transferability, providing a complementary approach for rapid landslide risk mapping and refined management in complex geological environments.

Key words: landslide susceptibility; non-landslide sample; frequency ratio; K-means; landslide density; engineering geology.

0 引言

滑坡是世界上发生频率最高、危害最为严重的地质灾害之一,对人类生命安全及财产造成严重威胁.由于我国疆域广阔,地形地貌复杂多样,极端天气频发,使我国成为受滑坡地质灾害影响最严重的国家之一(徐胜华等, 2020).据自然资源部统计结果显示,截至2022年,在全国登记在册的地质灾害隐患数据中,滑坡地质灾害发生频次最多,占比高达40%,对自然环境和人类社会造成严重危害.易发性评价是对滑坡进行风险评估以及潜在滑坡定位的基础(刘雅婷等, 2025),可综合分析研究区的各种地质环境因素、历史滑坡数据、滑坡物理规律等,以预测研究区未来发生滑坡的概率(Lu *et al.*, 2024),其预测与分析结果可为滑坡灾害风险管理、国土空间规划布局和滑坡监测提供重要的决策依据.

目前,滑坡易发性的评价模型主要分为3类:知识驱动模型、物理模型以及数据驱动模型(Li *et al.*, 2017).知识驱动模型主要依据专家的知识给滑坡相关的不同因素进行主观的评分与加权,进而得到滑坡易发性评价结果,是一种定性的评价方法.然而不同专家知识存在差异,评价结果往往存在较强主观性与不确定性,且缺乏可重复性(Mirus *et al.*, 2024);物理模型方法则通过分析滑坡诱发条件对边坡稳定性的影响机理,使用物理模型来定量评估滑坡发生的可能性.但是,此方法对区域地质力学参数的精度和完整性要求较高,而这些参数在大多数区域难以获取,导致其在大尺度应用中存在较大限制(古腾飞等, 2024).相比之下,数据驱动的方法则通过历史滑坡和非滑坡区域的地质环境来定量评价滑坡的易发性.人工智能技术的快速发展,使机器学习算法在滑坡易发性评价中得到广泛应用(Huang *et al.*, 2018),例如支持向量机(徐胜华等, 2020)、随机森林(Corominas *et al.*, 2014)、以及深度学习模型(Nguyen *et al.*, 2025)等通过处理大量地质、气候、地形等多源数据,能够有效提高滑坡易发性评价的精度与可靠性.数据驱动的方法因与机器学习结合使其可应用于大区域滑坡灾害易发性评价,精度较高,因此成为区域尺度滑坡易发性评价的标准方法.然而,数据驱动模型的关键在于滑坡样本和非滑坡样本的选取,其质量直接影响模型的学习效果和最终的预测精度(Zhu *et al.*, 2024).在当前的研究中,滑坡样本数据通常来源于历史滑坡、遥感影像与现场验证,这些数据相对准确(Zhao *et al.*, 2022),而非滑坡样本无法直接获取,存在较大的不确定性(苏志萍等, 2025),若选择的非滑坡样本位于滑坡频率较高的地区,会削弱模型的学习能力,进而影响其预测性能.

为减小非滑坡样本选择带来的偏差,近年来已有大量工作围绕"如何更合理地采样非滑坡点"展开.Zhou *et al.* (2018)采用随机采样来进行非滑坡样本的选取,该方法实现简便,但获取非滑坡点的数据随机性较强,使选取的数据质量可靠性低,可能会将潜在的滑坡归类为非滑坡,从而阻碍模型学习有效特征的能力.林荣福等 (2020)通过缓冲区采样法进行非滑坡点的选取,在距离已有历史滑坡点一定缓冲区外随机选择,但缓冲区距离的确定具有较强的主观性(缪亚敏等, 2016),导致滑坡易发性预测结果的不确定性增加.为进一步提高非滑坡点采样的质量,部分研究采用稳定区域采样法,如

Choi *et al*.(2012)、周晓亭等(2022)等人提出在河道、沟谷等低坡度区域选择非滑坡样本,从而在一定程度上提高非滑坡样本的稳定性.但此类方法可能在特定地貌类型过度采样,使样本存在整体代表性不足,从而削弱模型的泛化能力(Lu *et al*., 2024).近年来,研究者提出了多种优化采样方法,通过信息量(IV)(黄发明等, 2024)、逻辑回归模型(王悦等, 2024)、半监督学习(吴宏阳等, 2024)以及集成采样(Fu *et al*., 2023)等技术方法来改进非滑坡样本的不确定性,这些方法在提高模型精度和稳健性方面取得了积极进展.

在上述各类方法中,非滑坡样本的构建都需要依赖"当前未被观测到的滑坡",潜在的滑坡单元被纳入非滑坡样本的风险在理论上难以避免.因此,相较于依赖正负样本构建的监督学习模型,无监督的聚类方法为缓解上述问题提供了另一种思路:在不预先设定非滑坡样本的情况下,仅依赖环境因子对研究区进行分区,从而为滑坡易发性评价构建一个更为灵活的框架(Gu *et al*., 2024;刘雅婷等, 2025).在这一类方法中,聚类过程侧重于在多维环境因子空间中识别"条件相近"的区域,再通过引入历史滑坡信息,对不同环境类型的相对易发性进行评估,从而避免在建模阶段对非滑坡样本做出假设.在众多聚类方法中,K-means聚类因其实现简单、可扩展性强且易于与多源栅格因子耦合等优点,被广泛应用于滑坡易发性及相关空间分区研究中.通过对环境因子进行标准化处理,K-means能够将环境条件相近的栅格单元归为同一类,再结合各类中滑坡点的分布情况,刻画不同环境类型的相对危险程度.

综上所述,本文在充分肯定现有非滑坡样本优化方法与监督模型的基础上,尝试提出一种基于频率比(FR)与K-means聚类的滑坡易发性评价方法(FR-K-means方法):不显式采样非滑坡样本,而是在多维环境因子空间中引入由滑坡清单驱动的FR先验,通过K-means对研究区进行环境类型分区,并利用各分区内的滑坡密度构建归一化易发性指数,实现对滑坡易发性的定量分级.该方法通过致灾因子的空间相似性划分,将"滑坡先验"嵌入无监督分区过程,一定程度上减弱了非滑坡样本构建带来的不确定性,可作为传统监督式易发性评价方法的一种补充.

1　研究方法

滑坡易发性评价的核心在于识别与滑坡发生具有强相关性的因子条件,并在空间上划定此类条件的分布区域.Guzzetti *et al*.(2005)强调了滑坡易发性评估过程中因子与滑坡发生的紧密关系,并指出通过识别与滑坡发生相关的环境因子,可以有效地评估滑坡风险.现有研究表明,滑坡多发区域往往具有一定的地质、地貌和气候一致性,因此在环境条件相似单元中开展滑坡概率建模,已成为一种有效的评价策略(Wang *et al*., 2017, 2025).基于此,本文采用K-means聚类算法,以滑坡因子构建环境相似性分区,实质上是对空间中滑坡因子的自动划分.在划分出的每一子区中,滑坡发生的频次或密度可作为该区滑坡易发程度的直接量化指标,这种基于频率的赋值方法在滑坡区划中得到广泛应用(Youssef *et al*., 2015).为进一步增强聚类结果与滑坡发生概率之间的相关性,本文将基于频率比(FR)模型计算的初步概率值作为引导因子加入聚类分析,相当于将FR模型的先验概率信息融合入环境分区过程,有助于提高聚类单元内部的致灾一致性,并为后续基于滑坡密度的易发性等级划分提供更合理的分区基础.

1.1　研究思路

为减弱传统滑坡易发性评价中非滑坡样本采样不确定性对结果的影响,本文提出一种无需显式构建非滑坡样本集的易发性评价方法.总体思路为:首先基于历史滑坡点与环境因子计算各因子的频率比值,得到表征滑坡先验信息的初步概率值,并与标准化后的多源因子共同构建多维因子集;其次,采用K-means聚类算法在因子空间中对研究区栅格进行环境相似性分区;随后,在各聚类单元内统计滑坡密度并进行归一化处理以构建相对滑坡易发性指数;最后,利用自然断点法对易发性指数进行分级,并结合独立滑坡数据和传统随机森林方法对结果进行验证与对比分析.整体技术流程如图1所示.

1.2　特征因子预处理

1.2.1　地层岩性数据的量化处理:蒙特卡洛模拟方法　在滑坡易发性评价中,地层岩性作为重要的控制性因子,多以定性或分类属性存在,难以直接参与基于数值计算的模型训练或聚类分析.为了有效地将该因子量化为具有物理意义的数值,本文引

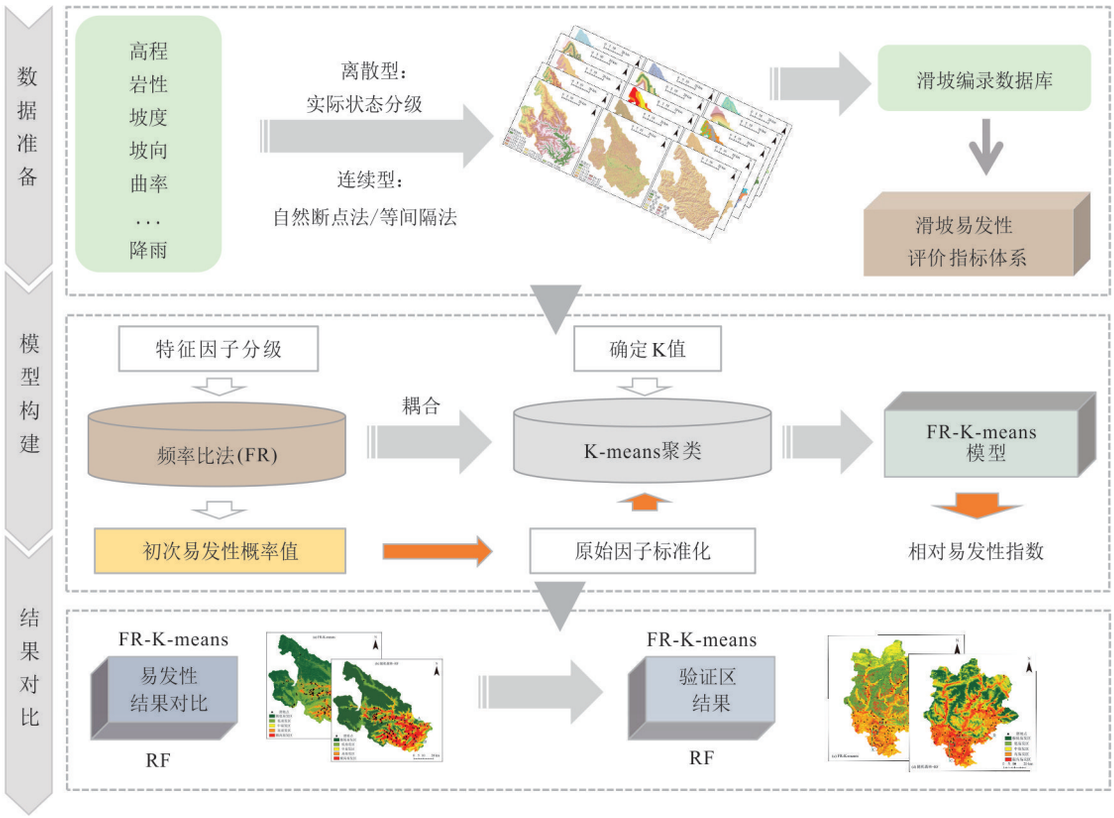

图 1 滑坡易发性技术流程图

Fig.1 Technical flowchart for landslide-susceptibility assessment

入基于地质力学参数的蒙特卡洛模拟方法.

蒙特卡洛方法是一种基于概率统计的数值模拟技术,适用于处理自然系统中存在不确定性或参数变异性的场景.其基本原理是在预设的参数取值区间或概率分布下,通过大量随机抽样,模拟变量的可能值序列,从而近似估计其统计行为特征.结合该理论,本文依据不同岩性类型在区域工程地质资料中所具有的典型抗剪强度参数黏聚力(C)与内摩擦角(φ)的统计范围,构建每类岩性的物理参数分布模型.在均匀分布假设下,对各类岩性参数执行 500 次蒙特卡洛随机抽样,生成具有地质力学意义的数值序列,并以其均值作为该岩性类别的数值表达.

1.2.2 频率比法 频率比法是计算每个滑坡特征因子在不同分级区间内滑坡发生的概率,对滑坡的分布与各个影响因子分级之间的空间关系进行分析(Akinci *et al.*, 2021).频率比(FR)表示的是影响因子在某一分级区间内滑坡发生的面积和研究区总滑坡面积的比值以及与该分级下的面积和研究区总面积的比值之比.频率比公式及计算过程如公式(1)所示:

$$FR=\frac{N_{ij}/N_{\mathrm{r}}}{A_{ij}/A_{\mathrm{r}}}\ , \tag{1}$$

式中:FR为频率比值;N_{ij}为第i个影响因子中第j类发生滑坡的面积;N_{r}为研究区滑坡的总面积;A_{ij}表示第i个影响因子中第j类所占的面积;A_{r}表示研究区内总的面积.

1.3 K-means聚类

K-means聚类是一种经典的无监督学习算法,广泛应用于空间分区、图像处理、模式识别和地学建模等领域.该算法的核心思想是通过最小化样本点与其所属聚类中心之间的误差平方和(Sum of Squared Errors, SSE),以实现最优的样本划分(Huang *et al.*, 2024).其计算如公式(2)所示:

$$SSE=\min\sum_{i=1}^{k}\sum_{x_j\in C_i}\|x_j-\mu_i\|^2\ , \tag{2}$$

式中:k表示聚类数;C_i表示第i个聚类;x_j为属于该聚类的样本点;μ_i为聚类C_i的中心;$\|x_j-\mu_i\|^2$表示样本与聚类中心的欧几里得距离平方.K-means聚类的执行流程主要包括以下 4 个步骤:

(1)在样本空间中随机选择k个样本点作为初始聚类中心;(2)计算所有样本与当前聚类中心之

间的距离,并将其划分到最近的聚类中心所代表的类别中;(3)对每个聚类内的所有样本点计算均值,更新聚类中心的位置;(4)重复步骤(2)和(3),直到聚类中心位置趋于稳定或达到预设的最大迭代次数,算法终止.

1.4　易发性评估方法

由于传统的模型性能评估指标,如受试者工作特征曲线(ROC)及其对应的AUC值(Xiao *et al.*, 2020; 付智勇等, 2023),通常依赖明确划分的正负样本集,因此在不采样非滑坡样本的条件下,其适用性受到限制(Woodard *et al.*, 2023).相较之下,基于滑坡密度的对比分析方法更适合评价不同易发性等级对滑坡的空间识别效果,也能更直观反映模型对滑坡敏感区域的实际区划能力(Mirus *et al.*, 2024).因此为评价FR-K-means模型的易发性区划效果,同时与随机森林(RF)方法进行比较,本文采用仅以滑坡清单为基础的验证方法,综合使用分级滑坡密度与基于面积排序的捕获效率指标对模型进行评估.

1.4.1　分级滑坡密度与单调性检验　将历史滑坡点叠加至易发性分区图中,对每一易发性等级k(极低、低、中、高、极高)统计其对应的面积N_k以及包含的滑坡数量L_k,计算该等级的滑坡密度D_k:

$$D_k=\frac{L_k}{N_k},k=1,2,3\cdots \quad (3)$$

滑坡密度反映了不同易发性等级对滑坡事件的富集能力,是检验分级合理性与空间聚焦性的基础指标.在合理的易发性分级下,滑坡密度应随等级由低到高近似单调递增,高、极高易发区的单位面积滑坡密度显著高于中、低等级;该思路已在多项滑坡易发性研究中作为常用的验证手段(Mirus *et al.*, 2024).

本文分别对FR-K-means与RF两种方法的易发性结果进行上述统计,比较各等级滑坡密度的梯度与单调性,并进一步分析高、极高易发区的滑坡富集程度,以评价不同方法对高风险区域的空间识别能力.

1.4.2　面积排序与滑坡捕获效率　在分级滑坡密度检验的基础上,为更直观地衡量模型在有限面积内对滑坡的“捕获效率”,本文采用基于栅格易发性指数排序的面积－滑坡累积分布分析方法.具体步骤如下:

(1)将研究区内所有栅格单元按易发性值从高到低排序;(2)在排序序列上依次累加栅格面积,得到任意面积占比a(如前10%、20%等)所对应的栅格集合,并统计该集合中包含的滑坡数量及其占全部滑坡的比例;(3)以面积占比为横轴、累计滑坡占比为纵轴,绘制面积－滑坡累积曲线,该曲线在形态上等价于常用的success-rate曲线;(4)计算面积－滑坡累积曲线下面积,记为AUC_{SR},并在若干代表性面积阈值(如前10%、20%)处记录对应的滑坡捕获率,用于比较不同模型在相同面积约束下的滑坡识别能力.

这类基于易发性指数排序的success-rate曲线,以及相应的曲线下面积(AUC),已在滑坡易发性研究中被广泛采用(Chung and Fabbri, 2003; Frattini *et al.*, 2010),作为不依赖分类阈值的精度评价指标,可在仅依赖滑坡清单的前提下,定量刻画模型对高易发区域的空间聚焦能力.

本文在FR-K-means与RF两种方法的易发性结果上均构建面积－滑坡累积曲线,通过比较二者在相同面积占比下的滑坡捕获率及AUC_{SR}值,评价不同模型在高风险区域空间聚焦性和面积利用效率方面的差异.

2　研究区概况与数据

2.1　研究区概况

平武县隶属于四川省绵阳市,地处绵阳市西北部,地理坐标为东经103°50′～104°58′,北纬31°59′～33°02′,总面积约5 974 km².平武县地处盆周山区,具有典型的山地地貌景观.全县地貌类型以山地为主,地质构造位置处于扬子准地台的西缘龙门山褶皱断裂带、昆仑秦岭的西秦岭－摩天岭东西构造带,以及松潘－甘孜地槽褶皱带的结合部位,地形起伏较大,整体地势为西北高、东南低,西北最高处近岷山山脉雪宝顶约为5 588 m,东南部河谷分布较广最低处海拔约600 m,相对高差近5 000 m.平武县属于北亚热带季风气候,气候温和,降水丰沛,年均降水量约为792 mm,6～9月是平武的汛期,月平均降雨量超过100 mm.复杂的地形条件和丰沛的降雨,加之边坡开挖和修建道路等人类活动,使平武县滑坡灾害频发.由中国科学院资源环境科学与数据中心提供的详细滑坡历史数据资料可知,截至2022年,平武县共登记在册滑坡事件808起,其中由降雨引起的滑坡共639起,大多为中小型滑坡.本文选取降雨型滑坡事件进行研究,研究区的地形特征和滑坡灾害的空间分布如图2所示.

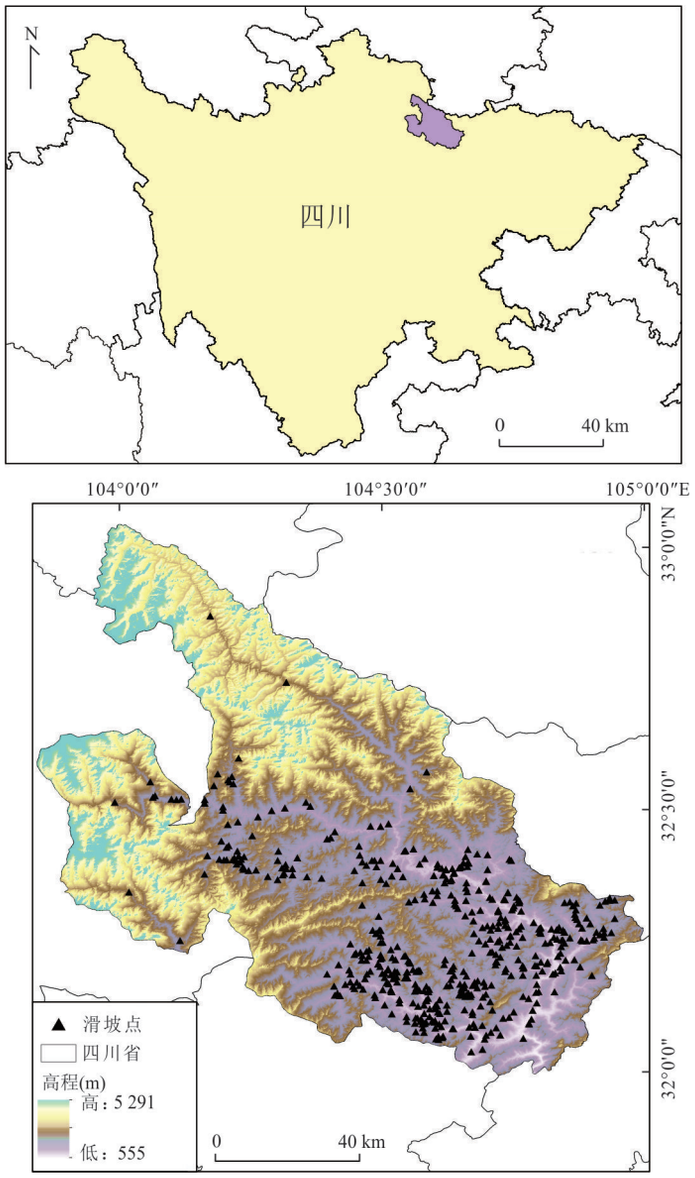

图 2 研究区概况及滑坡编录

Fig.2 Overview of the study area and landslides inventory

2.2 数据及来源

选择合适的滑坡影响因子是滑坡易发性建模的关键步骤.根据平武县滑坡致灾机理及相关的研究成果(Bao *et al.*, 2023; Huang *et al.*, 2025),本文从地形地貌、地质、水文环境条件、人类工程活动等方面选取了15个滑坡评价因子,包括高程、坡度、坡向、平面曲率、剖面曲率、归一化植被指数(NDVI)、地形湿度指数(TWI)、水流强度指数(SPI)、年均降雨量、土地利用、地层岩性、距道路距离、距断层距离、距水系距离、高程变异系数,其数据来源见表1.利用方差膨胀系数(VIF),对所选特征因子进行多重共线性检验,结果显示,所选特征因子的VIF值均小于10,说明各特征因子间具有较好的独立性,可作为滑坡易发性评价的试验数据.由于不同数据集的分辨率不同,收集到的滑坡灾害因子具有不同的尺度.为了简化后续的特征提取,所有的因子都被均匀地重采样至30 m的分辨率.研究区共6 773 035个栅格单元,将研究区栅格转点,通过多值提取至点,为每个栅格点分配各因子值.地层岩性和土地利用为两个分类变量,其他因子为连续变量.地层岩性考虑用黏聚力C和内摩擦角φ来量化表征各个岩性类型,土地利用则进行独热编码将各土地类型以0/1提取至各栅格点.

表 1　滑坡因子数据来源

Table 1　Data sources for landslide factors

分类	特征因子	分辨率	数据源
地形地貌	高程、坡度、坡向、平面曲率、剖面曲率、高程变异系数	30 m	https://www.gscloud.cn/ 地理空间数据云
地质因子	地层岩性	1:50 000	http://www.activefault-datacenter.cn/
	距断层距离	1:50 000	
	土地利用		
水文环境	SPI	30 m	https://earthexplorer.usgs.gov/
	TWI	30 m	https://www.resdc.cn/
	NDVI	1:50 000	https://www.openstreetmap.org/
	年均降雨量	30 m	http://www.globallandcover.com/
	距水系距离		
人类工程	距道路距离	1:50 000	https://www.openstreetmap.org/

第1类因子为地形地貌因子,包括高程、坡度、坡向、曲率及高程变异系数等.高程在滑坡发生中起着重要作用,不同的高程区域存在显著的气候垂直分布差异,尤其是气温、降水量和植被覆盖的变化,这些都对滑坡的发生风险产生影响.坡度是影响斜坡稳定性的关键因素之一.坡度越大,坡体的稳定性越差,滑坡发生的风险越高.较大的坡度往往伴随着水土流失的现象,更加剧了滑坡的可能性.坡向通过影响斜坡的日照、温度和水流分布来间接影响土壤的稳定性.阳坡通常较为干燥,土壤容易失水,可能导致土壤的强度下降,容易发生滑坡.而阴坡则积水较多,湿润的土壤也可能增加滑坡发生的风险.平面曲率与剖面曲率反映地表形态的凹凸性,进而控制坡面水流汇聚或分散状况,影响滑坡发生的概率.高程变异系数描述地形起伏复杂度,地形越破碎,越可能发生滑坡.

第2类因子为水文因子,包括NDVI、TWI、SPI以及年均降雨量.NDVI反映植被覆盖状况,高的NDVI(归一化植被指数)表示较好的植被覆盖,能够有效增强土壤的稳定性,降低滑坡发生的可能性.反之,低NDVI值代表植被稀疏,土壤易于受侵蚀,增加了滑坡的风险.TWI与SPI反映地表水分积聚与水流强度,影响土体含水率和抗剪强度,是降雨型滑坡的重要控制因子.本文主要对平武县由降雨诱发的滑坡进行研究,故降水量对滑坡发生具有极其重要影响.较高的降水量会导致土壤的含水量增加,进而降低土壤的稳定性,增加滑坡发生的风险.

第3类因子为地质与人类活动相关因子,包括地层岩性、土地利用、距道路距离、距水系距离和距断层距离.地层岩性直接影响岩土体的强度与变形特性,不同岩性区域的斜坡稳定性差异较大,软岩层更易发生滑坡.土地利用类型通过改变植被与土体扰动状况影响斜坡稳定性.道路、水系与断层的邻近性反映了人类扰动、卸荷作用或结构弱带,对滑坡发展具有显著诱发作用.各滑坡评价因子如图3所示.

2.3　特征因子分级

开展频率比(FR)分析,需要将部分连续因子进行离散化处理,以构建各分级下滑坡发生频度的统计关系.依据已有的滑坡易发性研究(李文彬等,2021;贾卓等, 2025),自然断点法(Jenks)与等间隔法两种分级法的应用较为广泛.自然断点法能够依据数据本身的分布特性自动识别最佳分段点,被广泛用于高程、坡度、曲率等连续因子的分级;而等间隔法则常用于距道路、距水系等距离类因子的分级,具有实现简便和可比性较强的特点.因此,本研究对高程、坡度、曲率、NDVI等因子采用自然断点法进行分级,对距道路、距断层及距水系距离采用等间隔方式进行分级,并据此计算各分级区间的频率比值.需要说明的是,本文的FR-K-means流程中,因子分级仅用于FR统计阶段,而非直接作为建模输入.在K-means聚类过程中,因子及FR值均以连续形式参与聚类,并在聚类前进行统一标准化处理.因此,因子分级可能带来的边界偏差在聚类阶段被进一步减弱,最终易发性分区主要取决于因子在高维特征空间中的聚类结构,而非分级的精细程度.这一点在一定程度上减弱了因子分级处理对最终结果的影响.同时,笔者未来在工作中也将尝试引入更具机理约束的分级方案,以进一步提高因

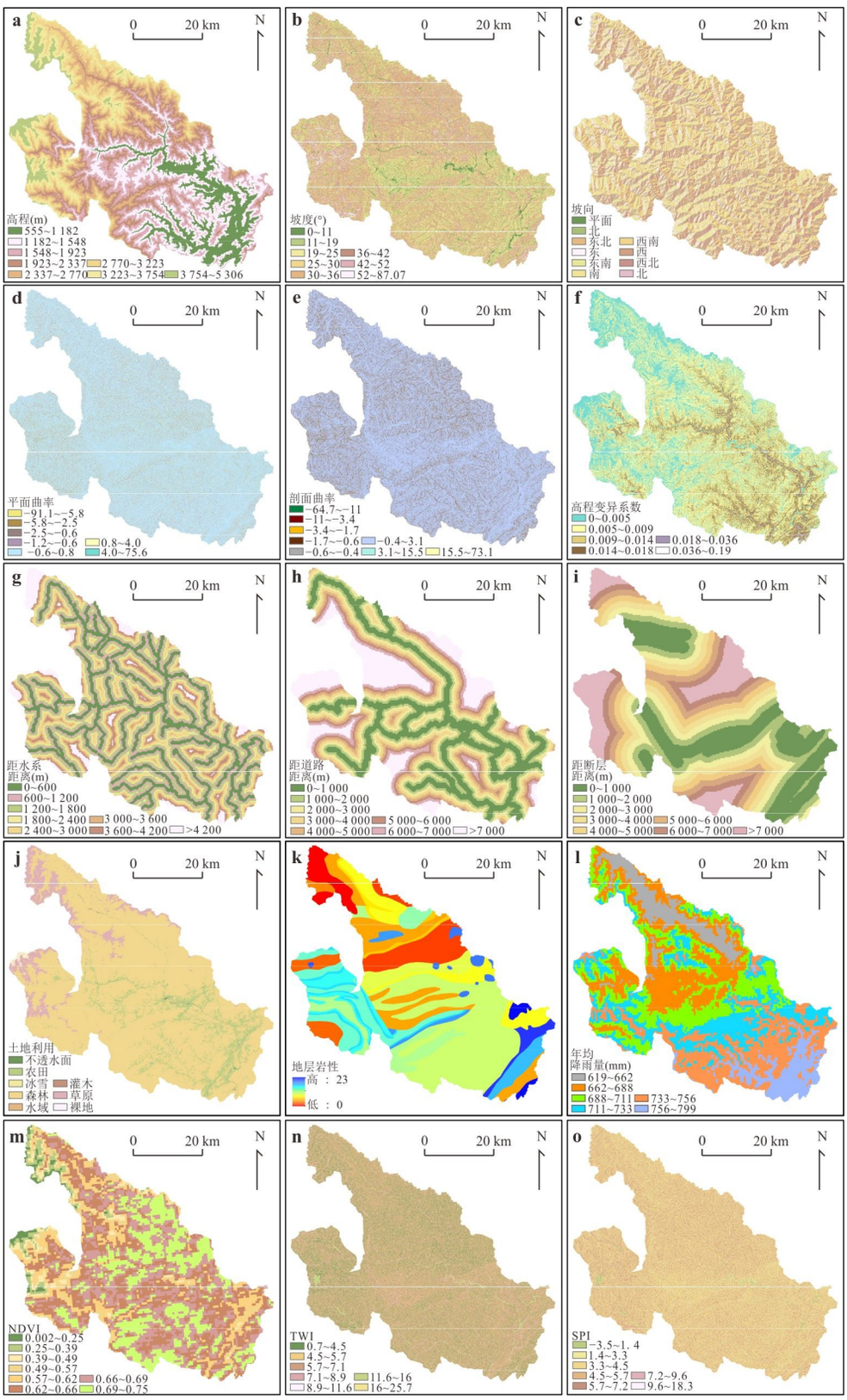

图 3 滑坡因子分级图

Fig.3 Landslide factor classification diagrams

a.高程;b.坡度;c.坡向;d.平面曲率;e.剖面曲率;f.高程变异系数;g.距水系距离;h.距道路距离;i.距断层距离;j.土地利用;k.地层岩性;l.年均降雨量;m.NDVI;n.TWI;o.SPI

子分级的物理合理性和解释力.

3　分析与讨论

为验证本文提出的无需显式采样非滑坡数据的易发性评价方法的可行性,本文依托平武县地质数据,依次对频率比建模结果、FR-K-means分区结果及最终的FR-K-means易发性区划结果进行可视化与统计分析,并与传统易发性模型随机森林(RF)进行效果对比.

3.1　FR结果分析

为了方便FR的计算,通过自然断点法将连续因子划分为5~9类.基于上述所选择的因子及研究区滑坡数据计算FR值,进而求取滑坡易发指数,并将其归一化后得到各栅格单元的初步滑坡易发性值.该值可视为一种反映历史滑坡响应特征的经验性先验因子:一方面,它综合了已有滑坡在多环境因子维度上的联合敏感性;另一方面,相当于在多维因子空间中引入了一个以滑坡发生频度为核心的概率维度,使不同因子组合在"是否有利于滑坡发生"这一意义上得以区分.因此,该因子本身具有较强的灾害空间分异能力.

FR概率值的空间分布如图4所示.由图可知,高概率区域与历史滑坡点具有较好重合性,尤其在河谷附近表现突出,说明单纯FR模型已能够在宏观上合理刻画研究区滑坡易发性格局,也验证了FR结果作为基线易发性图的有效性.在此基础上,本文进一步将FR概率作为先验因子引入K-means聚类,有助于在保持整体易发性分布合理的前提下,将"滑坡频发的环境类型"在多因子空间中凸显出来,引导后续FR-K-means模型与滑坡先验产生直接关联.

3.2　FR-K-means分区结果分析

仅依赖FR模型,滑坡易发性评价主要基于各单因子分级与滑坡发生之间的一维相关性,难以充分反映多因子组合下的环境差异;而若只使用15项环境因子进行K-means聚类,则聚类结果本质上是一种"环境类型分区",聚类过程并未与滑坡直接联系起来.为同时兼顾多因子环境差异与滑坡发生的先验信息,本文在FR模型的基础上构建FR-K-means模型.

具体而言,在对所有环境因子进行标准化处理后,将FR概率值与15项地形、地貌、地质及人类活动因子共同构建多维因子集,作为K-means聚类的

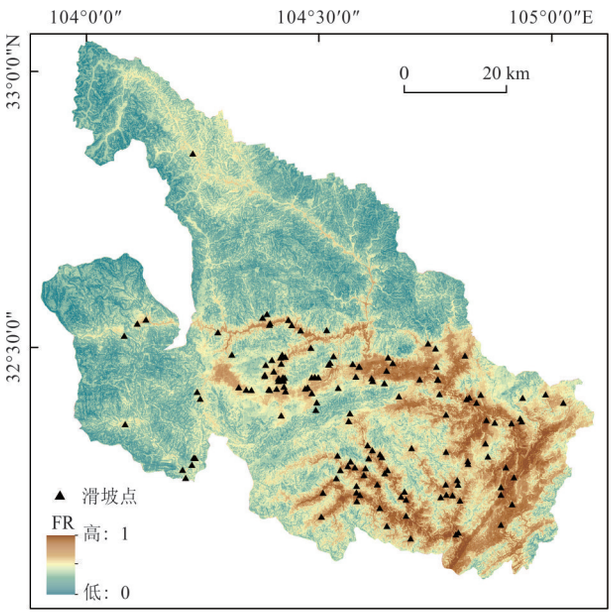

图4　FR概率图

Fig.4　FR probability

输入特征空间.FR值的引入相当于在环境因子空间中增加了一个由滑坡数据主导的因子,促使K-means在进行环境相似性划分时,优先将"环境条件相近且历史滑坡较为集中的栅格单元"归为同一聚类分区,从而使聚类结果不仅区分环境类型,而且对滑坡易发性的空间聚集特征更加敏感.

基于上述设置进行聚类分区,得到的分区结果如图5所示.在上述设置下对研究区进行K-means聚类,得到的环境分区结果如图5所示.结果表明,K-means能够有效识别出一系列在因子空间中相对一致的地貌单元,聚类分区在空间上具有良好的完整性和连贯性.河流、河谷侧坡以及高海拔积雪区等典型地貌单元在聚类图中呈现出清晰边界(图5中A、B、C所示),滑坡点明显聚集在特定聚类单元内,尤其是在东南部河谷和西部山区,聚类结果与历史滑坡点空间分布高度重合,表明该分区在孕灾环境表达方面具有较好的合理性.基于此类环境相似性分区计算各类滑坡密度,能够为后续相对滑坡易发性指数的构建提供稳定的空间单元.需要指出的是,尽管传统聚类分析中多采用肘部法或轮廓系数等方法来选择最优聚类数,但在滑坡易发性评价中,模型效果更依赖于能否实现对空间差异的充分识别.鉴于本研究不采样非滑坡样本且以高风险带的空间聚焦为优化目标,笔者以滑坡密度响应与面积效率作为选择K的主准则:在K=10、15、20、25、30、35等多种设置下,比较各等级区(尤其极高/

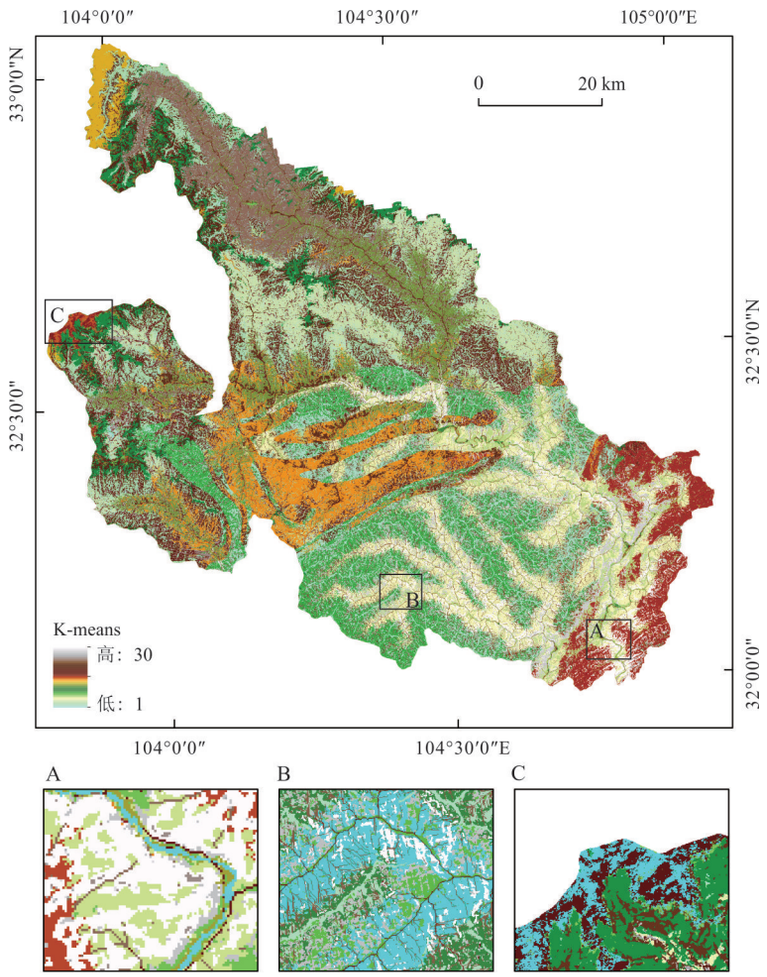

图 5 FR-K-means 聚类分区结果

Fig.5 FR-K-means clustering results

高区）对应滑坡密度与面积占比，优先选择“更高密度、较小面积”的分区方案．结果表明，K=30时，高、极高易发区在面积保持适中的同时具有更高的单位面积滑坡密度，且空间结构最为合理，因此选取K=30作为后续 FR-K-means 模型的聚类数．

3.3 FR-K-means模型易发性结果

本文将494个历史滑坡点空间叠加至聚类结果图中，统计每一聚类分区内实际发生的滑坡数量L_i与该分区包含的栅格总数N_i，计算其滑坡密度D_i，作为该区域滑坡风险的初步度量指标．为使不同分区的滑坡密度具有可比性，并提升其在空间建模中的解释能力，本文对所有分区的滑坡密度值进行归一化处理，得到统一量纲下的相对滑坡易发性指数．易发性指数值范围为[0，1]，其数值越大表示该分区滑坡发生的相对频度越高，地貌环境越敏感．基于归一化后的相对滑坡易发性指数值，本文进一步采用自然断点法（Natural Breaks）（Singh *et al.*，2021）对指数进行分级划分，将研究区划分为5个滑坡易发性等级：极高、高、中、低、极低．区划结果如图 6a，其各分区面积占比如饼图 6e$_1$所示．

3.4 RF模型易发性结果

为全面评估本文提出的FR-K-means模型的预测性能与适用价值，本节选用监督式学习模型——随机森林（Random Forest，RF）作为对比方法．随机森林是一种典型的机器学习方法，具有较强的非线性建模能力和高泛化性能，在滑坡易发性研究中被广泛应用（邓明东等，2024）．其基本原理是：通过Bootstrap重采样技术从原始样本中构建多个训练子集，分别训练若干决策树，并对所有树的预测结果进行投票融合，得到最终分类结果．模型训练过程中，每棵树的分裂节点均从所有特征中随机抽取部分变量进行最优划分，从而有效缓解过拟合问题并提升模型鲁棒性（Fu *et al.*，2025）．为确保两种方法的可比性，RF模型与FR-K-means模型方法使用

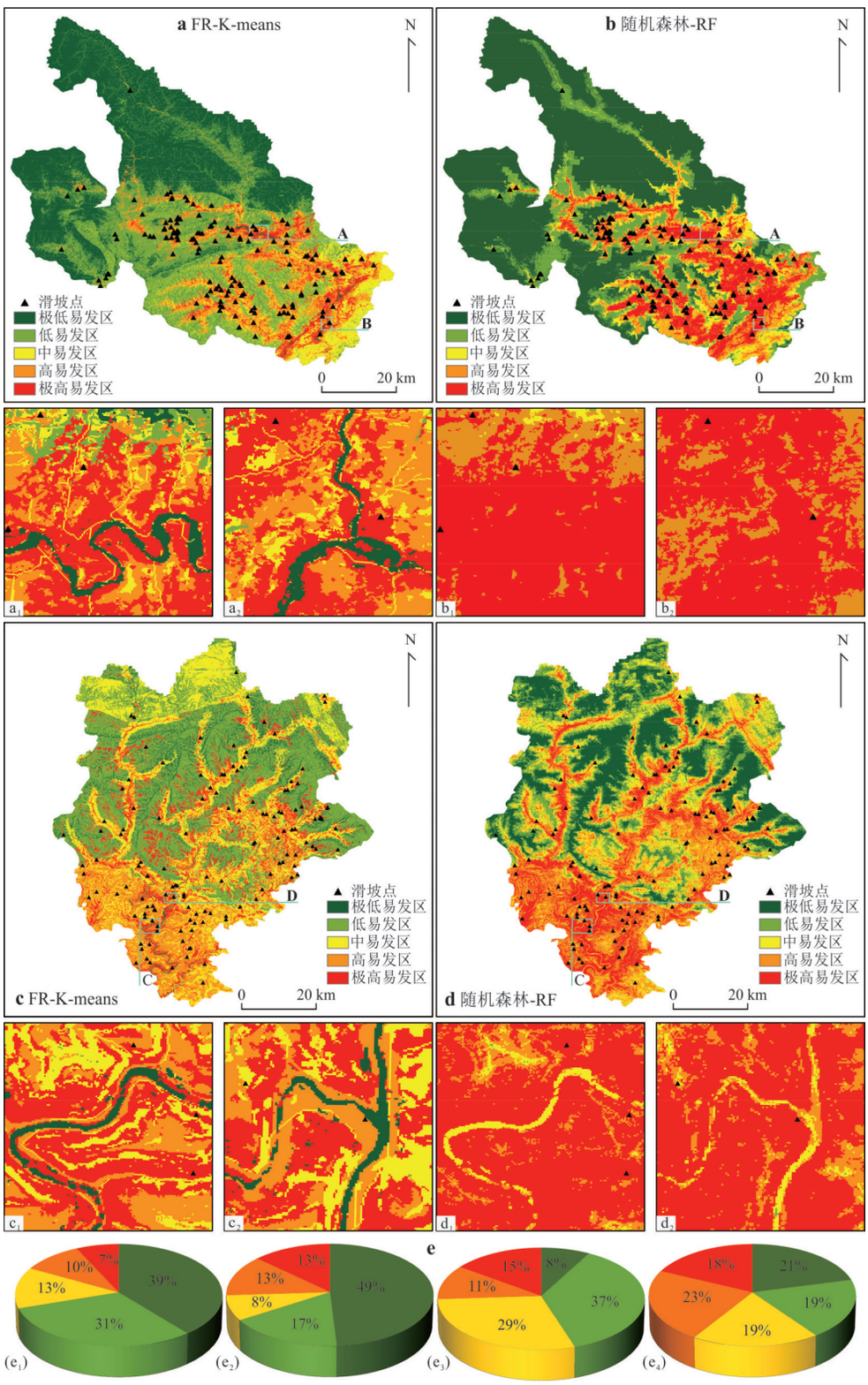

图6　滑坡易发性图

Fig.6　Landslide susceptibility mapping results

a.平武县FR-K-means易发性结果；a_1.A典型区域细节图；a_2.B典型区域细节图；b. 平武县RF易发性结果；b_1.A典型区域细节图；b_2.B典型区域细节图；c. 通江县FR-K-means易发性结果；c_1.C典型区域细节图；c_2. D典型区域细节图；d. 通江县RF易发性结果；d_1. C典型区域细节图；d_2.D典型区域细节图；e. 各易发性结果分区面积占比；e_1. 平武县-FR-K-means；e_2. 平武县-RF；e_3. 通江县-FR-K-means；e_4. 通江县-RF

完全相同的滑坡因子集及历史滑坡点.非滑坡样本则采用缓冲区排除策略(Ke *et al.*, 2025),在距滑坡点 1 000 m 以外区域随机采样,数量与滑坡点保持 1∶1 配比.利用训练好的 RF 模型对研究区进行预测,模型输出结果同样经自然断点法划分为 5 个滑坡易发性等级,与 FR-K-means 方法区划标准保持一致,便于后续各易发性分区结果的对比分析,其结果如图 6b,各分区面积占比如饼图 6e$_2$所示.

3.5 结果对比

本文将 144 个独立验证滑坡点分别叠加至两种方法的易发性区划结果上,开展对比分析.结果显示,两种方法的评价结果具有一定的相似性:极高区都主要位于平武县的东南部和中部.该区域地质条件复杂,河流密布,加上降雨充沛以及日益频繁的人类工程活动,给滑坡灾害提供有利的孕灾环境.

在此基础上,FR-K-means 方法在高风险区的滑坡聚集能力表现出一定的优势.以极高易发区为例,FR-K-means 方法的滑坡点密度为 0.124 个/km^2,高于随机森林(RF)模型的 0.112 个/km^2.该结果表明,FR-K-means 方法能够在较小的面积内捕捉更多的滑坡事件,呈现出较强的风险聚焦能力(图 7a,图 7b).此外,如上图 6e 所示,饼状图分别展示了 FR-K-means 方法和随机森林方法经自然断点法得到的各易发性分级的面积占比.从面积利用效率角度看,FR-K-means 方法能够以较小的面积划定高风险区域,避免过度扩展高易发区面积.如图 6e 所示,FR-K-means 的高、极高易发区面积分别约为 10% 和 7%(图 6e$_1$),均小于 RF 模型对应的 13% 和 13%(图 6e$_2$).这表明,FR-K-means 更加精准地聚焦于滑坡发生概率较高的区域.

另一方面,基于易发性指数排序构建的面积—滑坡累积分布曲线表明(图 7c),在相同的面积占比条件下,FR-K-means 曲线整体略高于 RF 曲线,其 AUC$_{SR}$指标也相对更大,说明在滑坡易发性高的前若干面积(如前 10%~20%)中,FR-K-means 能够覆盖更多的滑坡事件,并具有更强的空间聚焦性.这一结果与基于面积占比和滑坡密度的分析相互印证,进一步为 FR-K-means 方法的有效性提供了支撑.

从空间分布结构来看,FR-K-means 在流域两岸及河谷边缘的表现更为突出(图 6a$_1$, 图 6a$_2$).该方法在长期侵蚀作用显著的河谷边缘区域划定了连续的高、极高易发带,这些区域往往具有较大的坡度、较为破碎的岩性及较弱的植被覆盖,是典型的高风险地貌单元;而流域中心及其他整体稳定区域通常被划分为低至极低易发区,避免了大面积零散高易发性值的出现.相比之下,RF 结果在部分区域存在一定程度的空间“过预测”现象(图 6b$_1$, 图 6b$_2$),即局部环境条件相对稳定的区域被划入中高易发等级,使得风险区划的空间边界略显宽泛.

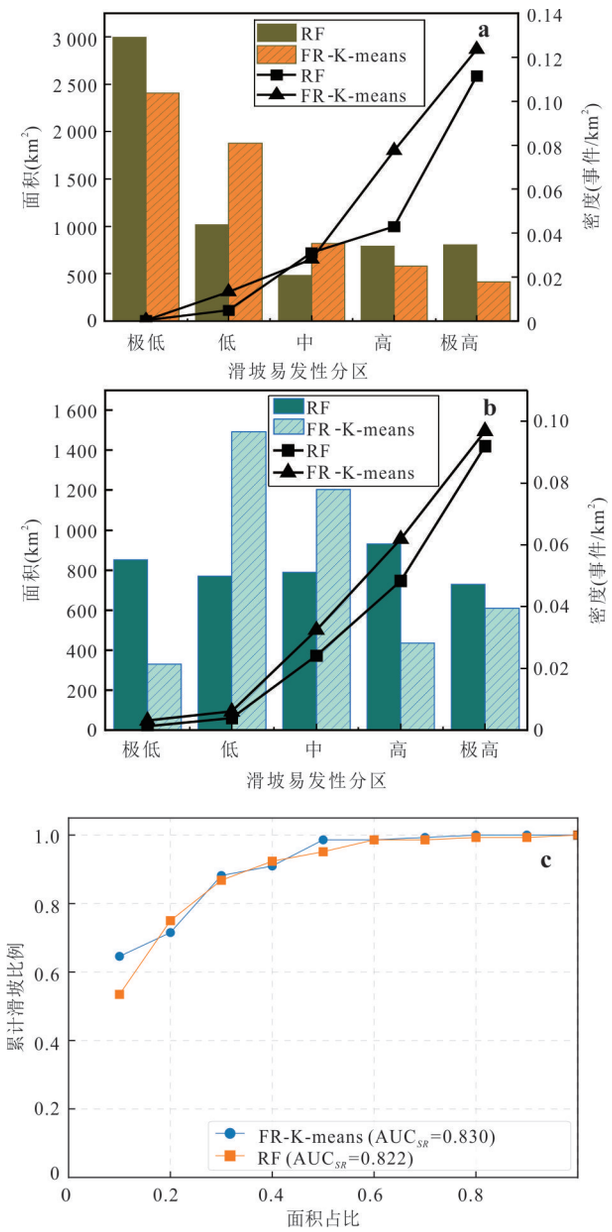

图 7 研究区结果对比(a),验证区结果对比(b)和面积—滑坡累积分布曲线(c)

Fig.7 Comparison of landslide susceptibility results in the study area (a), comparison of landslide susceptibility results in the validation (b) and area-cumulative mapping results in the landslide distribution curve (c)

综上所述,FR-K-means方法在滑坡易发性评估中表现出较强的空间聚焦能力和较好的面积利用效率,尤其是在高易发区的滑坡捕捉能力方面具有一定的优势.虽然RF模型在传统监督式框架下依然是一个强有力的模型,但在非滑坡样本难以高质量采样的情况下,FR-K-means提供了一种具有实践潜力的互补方法,尤其在高风险区域的划定上展现了其精准性和有效性.

最后,为进一步验证FR-K-means的适用性和稳健性,本文将通江县作为独立验证区,沿用平武县相同的因子选择、聚类方法与样本处理流程开展实验.结果如图所示(图6c, 图6d),在通江县的极高易发区,FR-K-means方法的滑坡密度达0.098个/km^2,依然优于RF方法的0.092个/km^2(图8),并能够准确区分流域两岸与内部区域的滑坡风险差异(图$6c_1$, 图$6c_2$),表现出良好的滑坡敏感性区划能力,验证了该方法在不同地质区域的迁移性和实用性.

3.6　讨论

在不显式采样非滑坡样本集的前提下,FR-K-means方法通过引入FR值并在多因子空间内进行环境相似性划分,实现了滑坡孕灾环境表达.结果表明,该方法在研究区及独立验证区均取得了较好的滑坡易发性结果,高、极高易发等级对滑坡事件具有显著的空间聚焦能力,在非滑坡样本难以精确获取的情形下,提供了一种具有实际可行性的滑坡易发性评价思路.

与此同时,该方法仍需要进一步完善.首先,K-means本质上假设各簇在特征空间中近似球形分布,并采用欧氏距离度量相似性,在某些高度非线性或非凸形的环境类型中可能难以充分刻画其内部结构;其次,聚类数K对结果具有较强敏感性,受地形地貌、因子维度及目标尺度等多重因素影响,尚难给出统一的选取准则,本研究选取高和极高易发区对应更高滑坡密度的分区数作为最优K值,虽然提升了高风险的区域划分精度,但此会影响模型的自动化与迁移效率.再者,FR先验依赖于现有滑坡清单,若滑坡清单存在漏报或时效性问题,可能导致部分潜在高风险环境类型在建模阶段未被充分表达,从而影响局部结果的精度.故未来可在获取更完备的滑坡数据库条件下,来对本方法进行进一步对比和检验也是十分必要的.

从方法拓展的角度看,FR-K-means也为后续融合非滑坡样本的研究提供了一个相对稳健的起点.一方面,历史滑坡点与FR-K-means所划定的极低或低易发区的重合度较低,可被视为候选“相对稳定区域”.在此基础上,未来可以尝试在这些区域内构建高置信度的非滑坡样本集,并与RF、XGBoost等监督学习模型联合使用,在统一的评价框架下系统量化比较“无需非滑坡样本”的FR-K-means思路与传统监督学习方法之间的性能差异,从而更全面地刻画不同非滑坡构建方法对易发性预测结果的影响.另一方面,FR-K-means输出的聚类单元本身也可作为一种分区基础:在每一分区内部引入半监督学习、主动学习或迭代式样本修正等策略,根据模型在不同聚类环境中的分类置信度,逐步优化非滑坡样本的空间分布,有望在保持方法可解释性的同时,进一步提升整体预测精度和稳健性.

综上,FR-K-means是建立在传统FR与监督学习方法基础上的一种尝试:一方面,FR、RF等依赖显式正负样本构建的监督式方法仍然是滑坡易发性评价的重要技术路线;另一方面,在非滑坡样本构建存在不确定性的背景下,FR-K-means结果表明,在仅依赖滑坡清单和环境因子信息的条件下,有获得具有较准确易发性区划结果的可能.未来若能在本方法划定的稳定区基础上进一步构建高置信度非滑坡样本,并与多种监督学习及时序建模方法相结合,将有望在“无非滑坡样本”和“高质量非滑坡样本”两类思路之间搭建起更紧密的联系,推动滑坡易发性评价方法体系的整体完善与发展.

4　结论

本文针对传统滑坡易发性评价方法中非滑坡样本选取不确定性对结果的影响,提出了一种创新的滑坡易发性建模方法,旨在实现无非滑坡样本的滑坡易发性评估.通过结合因子概率响应与空间聚类特征,构建了一种基于频率比(FR)模型与K-means聚类算法的滑坡易发性评估方法.主要结论如下:

(1)核心创新与方法优势:本文提出的方法以基于已有滑坡及其因子进行频率比(FR)得到的初次概率值为先验,结合K-means聚类实现环境因子的空间分区,并通过分区滑坡密度构建相对易发性指数.在无需显式采样非滑坡点的前提下,该方法

实现了滑坡风险的定量化与精细区划,一定程度上避免了传统模型中非滑坡样本选取不确定性的问题.与依赖非滑坡样本的随机森林(RF)方法相比,本方法在高风险区划定上更为聚焦,以更小面积锁定更高密度的滑坡点,表现出科学性与实用性.

(2)流域尺度的区划价值:该方法能够有效凸显滑坡在流域空间上的差异性,精准识别并锁定河流两岸的高风险带,同时合理剔除流域内部的稳定区域,避免了高易发性区域的过度扩展.其结果与实际孕灾环境高度一致,展现出良好的空间解释力与工程应用价值.

(3)区域适应性与稳定性:通过在通江县的验证,本文方法在不同地貌背景下仍保持较高的适应性和稳定性.尽管地形、气候等因子存在差异,FR-K-means方法依然能够有效适应这些变化,并对不同区域的滑坡易发性进行准确划分.这表明该方法具有较强的区域尺度适用性,为滑坡易发性评估提供了可靠的技术支持.

References

Akinci, H., Yavuz Ozalp, A., 2021. Landslide Susceptibility Mapping and Hazard Assessment in Artvin (Turkey) Using Frequency Ratio and Modified Information Value Model. *Acta Geophysica*, 69(3): 725−745. https://doi.org/10.1007/s11600-021-00577-7

Choi, J., Oh, H. J., Lee, H. J., et al., 2012. Combining Landslide Susceptibility Maps Obtained from Frequency Ratio, Logistic Regression, and Artificial Neural Network Models Using ASTER Images and GIS. *Engineering Geology*, 124: 12−23. https://doi.org/10.1016/j.enggeo.2011.09.011

Chung, C. F., Fabbri, A. G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. *Natural Hazards*, 30(3): 451−472. https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b

Corominas, J., van Westen, C., Frattini, P., et al., 2014. Recommendations for the Quantitative Analysis of Landslide Risk. *Bulletin of Engineering Geology and the Environment*, 73(2): 209−263. https://doi.org/10.1007/s10064-013-0538-8

Deng, M. D., Ju, N. P., Wu, T. W., et al., 2024. Vulnerability Evaluation under Different Landslide Sample Points and Polygon Expression Modes. *Earth Science*, 49(5): 1565−1583 (in Chinese with English abstract).

Frattini, P., Crosta, G., Carrara, A., 2010. Techniques for Evaluating the Performance of Landslide Susceptibility Models. *Engineering Geology*, 111(1−4): 62−72. https://doi.org/10.1016/j.enggeo.2009.12.004

Fu, Y., Fan, Z. H., Li, X. Z., et al., 2025. The Influence of Non-Landslide Sample Selection Methods on Landslide Susceptibility Prediction. *Land*, 14(4): 722. https://doi.org/10.3390/land14040722

Fu, Z. J., Wang, F. W., Dou, J., et al., 2023. Enhanced Absence Sampling Technique for Data-Driven Landslide Susceptibility Mapping: A Case Study in Songyang County, China. *Remote Sensing*, 15(13): 3345. https://doi.org/10.3390/rs15133345

Fu, Z. Y., Li, D. Q., Wang, S., et al., 2023. Landslide Susceptibility Assessment Based on Multitemporal Landslide Inventories and TrAdaBoost Transfer Learning. *Earth Science*, 48(5): 1935−1947 (in Chinese with English abstract).

Gu, T. F., Duan, P., Wang, M. G., et al., 2024. Effects of Non-Landslide Sampling Strategies on Machine Learning Models in Landslide Susceptibility Mapping. *Scientific Reports*, 14: 7201. https://doi.org/10.1038/s41598-024-57964-5

Gu, T. F., Duan, P., Wang, W. P., et al., 2024. Review of Methods for Landslide Susceptibility Assessment. *Journal of Natural Disasters*, 33(6): 1−16 (in Chinese with English abstract).

Guzzetti, F., Reichenbach, P., Cardinali, M., et al., 2005. Probabilistic Landslide Hazard Assessment at the Basin Scale. *Geomorphology*, 72(1−4): 272−299. https://doi.org/10.1016/j.geomorph.2005.06.002

Huang, F. M., Zeng, S. Y., Yao, C., et al., 2024. Uncertainties of Landslide Susceptibility Prediction Modeling: Influence of Different Selection Methods of "Non-Landslide Samples". *Advanced Engineering Sciences*, 56(1): 169−182 (in Chinese with English abstract).

Huang, Y., Zhao, L., 2018. Review on Landslide Susceptibility Mapping Using Support Vector Machines. *CATENA*, 165: 520−529. https://doi.org/10.1016/j.catena.2018.03.003

Jia, Z., Cheng, Z. J., Chang, Z. L., et al., 2025. Modeling and Uncertainty in Landslide Susceptibility Prediction Considering Coupling Mode of Landslide Types. *Earth Science*, 50(6): 2311−2329 (in Chinese with English abstract).

Ke, C. Y., Sun, P., Zhang, S., et al., 2025. Influences of Non-Landslide Sampling Strategies on Landslide Susceptibility Mapping: A Case of Tianshui City, Northwest of China. *Bulletin of Engineering Geology and the Environment*, 84(3): 123. https://doi.org/10.1007/

s10064-025-04147-9

Li, L. P., Lan, H. X., Guo, C. B., et al., 2017. A Modified Frequency Ratio Method for Landslide Susceptibility Assessment. *Landslides*, 14(2): 727－741. https://doi.org/10.1007/s10346-016-0771-x

Li, W. B., Fan, X. M., Huang, F. M., et al., 2021. Uncertainties of Landslide Susceptibility Modeling under Different Environmental Factor Connections and Prediction Models. *Earth Science*, 46(10): 3777－3795 (in Chinese with English abstract).

Lin, R. F., Liu, J. P., Xu, S. H., et al., 2020. Evaluation Method of Landslide Susceptibility Based on Random Forest Weighted Information. *Science of Surveying and Mapping*, 45(12): 131－138 (in Chinese with English abstract).

Liu, Y. T., Chen, C. F., He, Q. X., et al., 2025. Landslide Susceptibility Evaluation Considering Positive and Negative Sample Optimization. *Acta Geodaetica et Cartographica Sinica*, 54(2): 308－320 (in Chinese with English abstract).

Lu, J. G., He, Y., Zhang, L. F., et al., 2024. Ensemble Learning Landslide Susceptibility Assessment with Optimized Non-Landslide Samples Selection. *Geomatics, Natural Hazardsand Risk*, 15(1): 2378176. https://doi.org/10.1080/19475705.2024.2378176

Miao, Y. M., Zhu, A. X., Yang, L., et al., 2016. Sensitivity of BCS for Sampling Landslide Absence Data in Landslide Susceptibility Assessment. *Mountain Research*, 34(4): 432－441 (in Chinese with English abstract).

Mirus, B. B., Belair, G. M., Wood, N. J., et al., 2024. Parsimonious High-Resolution Landslide Susceptibility Modeling at Continental Scales. *AGU Advances*, 5(5): e2024AV001214. https://doi. org/10.1029/2024av001214

Nguyen, D. D., Nguyen, M. D., Nguyen, T. V., et al., 2025. Enhanced Landslide Spatial Prediction Using Hybrid Deep Learning Model and SHAP Analysis: A Case Study of the Tuyen Quang-Ha Giang Expressway, Vietnam.*Journal of the Indian Society of Remote Sensing*, 53(5): 1647－1666. https://doi.org/10.1007/s12524-024-02055-1

Singh, P., Sharma, A., Sur, U., et al., 2021. Comparative Landslide Susceptibility Assessment Using Statistical Information Value and Index of Entropy Model in Bhanupali-Beri Region, Himachal Pradesh, India. *Environment,Development and Sustainability*, 23(4): 5233－5250. https://doi.org/10.1007/s10668-020-00811-0

Su, Z. P., Yang, C. S., Wang, Z. Q., 2025. Landslide Susceptibility Assessment in Nujiang Canyon Using Negative Sample Optimization and Machine Learning Models. *Journal of Geo-Information Science*, 27(4): 979－993 (in Chinese with English abstract).

Wang, Q., Wang, Y., Niu, R. Q., et al., 2017. Integration of Information Theory, K-Means Cluster Analysis and the Logistic Regression Model for Landslide Susceptibility Mapping in the Three Gorges Area, China. *Remote Sensing*, 9(9): 938. https://doi.org/10.3390/rs9090938

Wang, R. T., Xi, W. F., Huang, G. C., et al., 2025. Landslide Susceptibility Evaluation Based on the Combination of Environmental Similarity and BP Neural Networks. *Land*, 14(4): 839. https://doi. org/10.3390/land14040839

Wang, Y., Cao, Y., Xu, F. D., et al., 2024. Reservoir Landslide Susceptibility Prediction Considering Non-Landslide Sampling and Ensemble Machine Learning Methods. *Earth Science*, 49(5): 1619－1635 (in Chinese with English abstract).

Woodard, J. B., Mirus, B. B., Crawford, M. M., et al., 2023. Mapping Landslide Susceptibility over Large Regions with Limited Data. *Journal of Geophysical Research: Earth Surface*, 128(5): e2022JF006810. https://doi.org/10.1029/2022jf006810

Wu, H. Y., Zhou, C., Liang, X., et al., 2024. Evaluation of Landslide Susceptibility Based on Sample Optimization Strategy. *Geomatics and Information Science of Wuhan University*, 49(8): 1492－1502 (in Chinese with English abstract).

Xiao, T., Segoni, S., Chen, L. X., et al., 2020. A Step beyond Landslide Susceptibility Maps: A Simple Method to Investigate and Explain the Different Outcomes Obtained by Different Approaches. *Landslides*, 17(3): 627－640. https://doi. org/10.1007/s10346-019-01299-0

Xu, S. H., Liu, J. P., Wang, X. H., et al., 2020. Landslide Susceptibility Assessment Method Incorporating Index of Entropy Based on Support Vector Machine: A Case Study of Shaanxi Province. *Geomatics and Information Science of Wuhan University*, 45(8): 1214－1222 (in Chinese with English abstract).

Youssef, A. M., Pradhan, B., Jebur, M. N., et al., 2015. Landslide Susceptibility Mapping Using Ensemble Bivariate and Multivariate Statistical Models in Fayfa Area, Saudi Arabia. *Environmental Earth Sciences*, 73(7): 3745－3761. https://doi. org/10.1007/s12665-014-3661-3

Zhao, P. X., Masoumi, Z., Kalantari, M., et al., 2022. A GIS-Based Landslide Susceptibility Mapping and Variable Importance Analysis Using Artificial Intelligent Training-Based Methods. *Remote Sensing*, 14(1): 211. https://doi.org/10.3390/rs14010211

Zhou, C., Yin, K. L., Cao, Y., et al., 2018. Landslide Susceptibility Modeling Applying Machine Learning Methods: A Case Study from Longju in the Three Gorges Reservoir Area, China. *Computers & Geosciences*, 112: 23－37. https://doi.org/10.1016/j.cageo.2017.11.019

Zhou, X. T., Huang, F. M., Wu, W. C., et al., 2022. Regional Landslide Susceptibility Prediction Based on Negative Sample Selected by Coupling Information Value Method. *Advanced Engineering Sciences*, 54(3): 25－35 (in Chinese with English abstract).

Zhu, Y. C., Sun, D. L., Wen, H. J., et al., 2024. Considering the Effect of Non-Landslide Sample Selection on Landslide Susceptibility Assessment. *Geomatics, Natural Hazards and Risk*, 15(1): 2392778. https://doi.org/10.1080/19475705.2024.2392778

中文参考文献

邓明东，巨能攀，吴天伟，等，2024. 不同滑坡样本点和多边形表达模式下的易发性评价. 地球科学，49(5): 1565－1583.

付智勇，李典庆，王顺，等，2023. 基于多时空滑坡编录和TrAdaBoost迁移学习的滑坡易发性评价. 地球科学，48(5): 1935－1947.

古腾飞，段平，王伟鹏，等，2024. 滑坡易发性评价方法综述. 自然灾害学报，33(6): 1－16.

黄发明，曾诗怡，姚池，等，2024. 滑坡易发性预测建模的不确定性：不同“非滑坡样本”选择方式的影响. 工程科学与技术，56(1): 169－182.

贾卓，程志金，常志璐，等，2025. 考虑滑坡类型耦合方式的滑坡易发性预测建模与不确定性分析. 地球科学，50(6): 2311－2329.

李文彬，范宣梅，黄发明，等，2021. 不同环境因子联接和预测模型的滑坡易发性建模不确定性. 地球科学，46(10): 3777－3795.

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

刘雅婷，陈传法，何青鑫，等，2025. 顾及正负样本优化的滑坡易发性评价. 测绘学报，54(2): 308－320.

缪亚敏，朱阿兴，杨琳，等，2016. 滑坡危险度评价对BCS负样本采样的敏感性. 山地学报，34(4): 432－441.

苏志萍，杨成生，王子倩，2025. 结合负样本优化与机器学习模型的怒江洲峡谷段滑坡易发性评价. 地球信息科学学报，27(4): 979－993.

王悦，曹颖，许方党，等，2024. 考虑非滑坡样本选取和集成机器学习方法的水库滑坡易发性预测. 地球科学，49(5): 1619－1635.

吴宏阳，周超，梁鑫，等，2024. 基于样本优化策略的滑坡易发性评价. 武汉大学学报(信息科学版)，49(8): 1492－1502.

徐胜华，刘纪平，王想红，等，2020. 熵指数融入支持向量机的滑坡灾害易发性评价方法：以陕西省为例. 武汉大学学报(信息科学版)，45(8): 1214－1222.

周晓亭，黄发明，吴伟成，等，2022. 基于耦合信息量法选择负样本的区域滑坡易发性预测. 工程科学与技术，54(3): 25－35.