



引文格式: 李泽芝, 王新刚. 镇域尺度下秦巴山区堆积层滑坡易发性不同单元评价性能对比研究[J]. 西北地质, 2024, 57(1): 1-11. DOI: 10.12401/j.nwg.2023159

Citation: LI Zezhi, WANG Xingang. Comparative Study on Evaluation Performance of Different Units of Susceptibility of Accumulation Layer Landslide in Qinba Mountain Area at Town Scale[J]. Northwestern Geology, 2024, 57(1): 1-11. DOI: 10.12401/j.nwg.2023159

镇域尺度下秦巴山区堆积层滑坡易发性 不同单元评价性能对比研究

李泽芝, 王新刚*

(大陆动力学国家重点实验室, 西北大学地质学系, 陕西 西安 710069)

摘要: 秦巴山区堆积层滑坡数量多、分布广、密度大、频次高, 所造成的危害十分严重, 且具有孕灾条件复杂多样和部分灾害评价数据获取难度大等特征。笔者选取秦巴山区小岭镇作为研究区, 在地质灾害野外调查基础上, 结合堆积层滑坡区域特点, 采取栅格、斜坡两种单元类型, 因地制宜的提取了滑坡孕灾因子, 分析其相关性, 提选出坡度、坡高、坡面形态、斜坡结构类型、堆积层厚度、距道路、矿区、断裂的距离等 8 个因子作为堆积层滑坡特征因子, 运用随机森林模型方法对该镇域进行了滑坡易发性评价; 并通过评价结果频率比、ROC 曲线、易发性概率均值与标准差, 对栅格单元、斜坡单元两种单元类型的精度与准确性进行了验证, 结果表明: 两种评价单元的预测结果都有良好的表现, 但斜坡单元作为评价单元总体预测性能高于栅格单元, 栅格单元在灾害防治具体空间部署上有着更精细的参考。研究成果对秦巴山区镇域地质灾害风险评价工作有一定的理论和实践意义。

关键词: 易发性; 堆积层滑坡; 随机森林; 单元评价; 秦巴山区

中图分类号: P694

文献标志码: A

文章编号: 1009-6248(2024)01-0001-11

Comparative Study on Evaluation Performance of Different Units of Susceptibility of Accumulation Layer Landslide in Qinba Mountain Area at Town Scale

LI Zezhi, WANG Xingang*

(State Key Laboratory of Continental Dynamics, Department of Geology, Northwest University, Xi'an 710069, Shaanxi, China)

Abstract: The accumulation layer landslides in Qinba Mountain area are abundant, widely distributed and frequently, and the harm caused by them is very serious. Moreover, it is characterized by complex and diverse disaster pregnancy conditions and difficult to obtain some disaster evaluation data. Xiaoling Town, Qinba Mountain, was selected as the research area. The geological hazard field survey was taken as the basis. Combined with the regional characteristics of accumulation landslide, two element types, grid element and slope element, are

收稿日期: 2023-02-01; 修回日期: 2023-04-19; 责任编辑: 贾晓丹

基金项目: 国家自然科学基金项目“灌溉区黄土干湿循环劣化后蠕变特性及其促滑机理研究”(41902268)资助。

作者简介: 李泽芝(1998-), 男, 硕士研究生, 主要从事滑坡调查与机理研究。E-mail: 1642372052@qq.com。

* 通讯作者: 王新刚(1984-), 男, 教授, 博士生导师, 主要从事地质灾害机理与防控研究。E-mail: xgwang@nwu.edu.cn。

adopted. The landslide hazard factors were selected according to local conditions, and their correlation was analyzed. Eight factors, including slope, slope height, slope morphology, slope structure type, accumulation layer thickness, distance from road, mining area and fault, are selected as the characteristic factors of accumulation layer landslide. The random forest model method was used to evaluate the landslide susceptibility of the town area. In addition, the accuracy and accuracy of grid element and slope element were verified by frequency ratio, ROC curve, mean value and standard deviation of susceptibility probability of evaluation results. The results show that both evaluation elements have good performance in the re-prediction results, but the overall prediction performance of slope element as evaluation element is higher than that of grid element. In the specific spatial deployment of disaster prevention and control, more detailed reference comes from grid element. The research results have certain theoretical and practical significance for the risk assessment of geological hazards in towns in Qinba Mountains.

Keywords: susceptibility; landslide of accumulation layer; random forest; unit evaluation; Qinba Mountains

秦巴山区地质构造强烈,地形起伏大,地貌形态与岩土体结构类型复杂多样,岩体节理裂隙发育,风化严重,降雨丰富,人类工程活动活跃,导致区内灾害频发,尤其是以第四纪堆积层滑坡为主的地质灾害(隐患)具有数量多、分布广、密度大、频次高等特点,所造成的危害严重,是中国地质灾害的重灾区之一(强菲等, 2015; 赵力行等, 2020; 刘伟等, 2021)。因此,秦巴山区堆积层滑坡具有孕灾条件复杂多样、部分灾害评价数据获取难度大等特征(张世林, 2020),如斜坡堆积层厚度、斜坡结构类型。

滑坡易发性评价模型建立流程为:区域滑坡数据及孕灾条件编录→评价因子提取与选择→评价模型选取→模型参数推定→评价结果精度验证与分析等步骤,可见评价模型是十分重要的一环,关乎评价结果的准确性与可信度(黄发明等, 2022)。目前滑坡易发性评价模型主要分为知识驱动模型、物理驱动模型、数据驱动模型3类,其中,数据驱动模型以区域基础环境因子为数据输入,通过机器学习模型、数理统计模型等得到评价结果(Saro Lee et al., 2015; 张林梵等, 2022)。据研究表明,机器学习模型的预测性能明显高于数理统计模型与传统的知识驱动模型,其中随机森林模型在预测精度、提高泛化能力等方面较其他机器学习模型有着更好的表现,如Merghadi等(2018)以北非米拉盆地为例,比较了基于随机森林、梯度提升机、逻辑回归、神经网络和支持向量机5种评价模型下滑坡易发性的预测结果,得出随机森林模型预测性能更高的结论。

目前,滑坡易发性评价时所使用区划单元类型主

要为栅格单元与斜坡单元(田述军等, 2019; 常志璐等, 2023)。栅格单元区划在空间上破坏了评价区域斜坡的自然属性,但划分时操作简单,在给单元赋值时有更强的唯一性,信息损失较少;斜坡单元划分时相较于栅格更复杂,给单位赋值时可能造成信息损失,但能够很好的反应区域滑坡易发性的地形分布情况,对某些基于斜坡的特征因子有更好的表征,如斜坡结构、堆积层厚度等。

近年来,随着1:10 000地质灾害风险评价的开展,评价区域的尺度也转向镇域级,意味着对地质灾害风险区的划分有了相对县域级评价更高的精度要求,主要表现在对全域内每个斜坡孕灾地质条件的掌握(陕西省自然资源厅, 2021)。不同地质环境背景下滑坡的主要孕灾因子不尽相同(黄润秋等, 2007),如黄土地区由于土体的水敏性、灌溉等因素所致滑坡较为普遍(段钊等, 2018),故土地利用类型为黄土地区滑坡易发性评价的重要因子;而秦巴山区多以堆积层碎石土-基岩的二元结构斜坡为主(范立民等, 2004; 唐睿旋, 2017)层间的渗透性差异使该地区以堆积层滑坡为主,故堆积层厚度为堆积层滑坡易发性评价的重要因子。

基于上述分析,笔者选取秦巴山区柞水县小岭镇作为典型研究区,结合堆积层滑坡特点,采取栅格单元、斜坡单元两种单元类型,运用随机森林模型方法对该区域进行了滑坡易发性评价与精度分析,得到镇域滑坡空间概率,并探讨易发性评价在两种划分方法下的差异,研究成果对当地政府部门减灾防灾措施部署等具有重要意义。

1 研究区及模型特征因子分析

1.1 研究区概况

小岭镇地处秦巴山区腹地(图 1), 区内地形陡峭、沟壑纵横, 其地质构造受东西展布的山阳—凤镇等深大断裂影响较大, 其滑坡孕灾条件也较为复杂: 区域内斜坡以二元结构的陡坡为主, 丰富的矿产资源与稀少的平缓地带, 不可避免的形成了区内削坡建房频繁、切坡修路密集、毁林造地普遍、矿山开发持续的现象, 这些现象导致了斜坡开挖形成临空面致使开挖处剪应力集中, 加之堆积层大都为渗透性强的松散碎石土, 矿区附近更有矿渣堆积使得堆积层自重增大, 孔隙水压力增强, 导致堆积层滑坡在研究区内极易发育, 经统计发现(表 1), 研究区地质灾害以堆积层滑坡为主, 占区内滑坡的 92.86%, 主要分布在道路两侧的陡坡处(图 2)。

1.2 随机森林模型简介

随机森林是一种集成 Bagging 算法, 可用于分类问题与回归问题(方匡南等, 2011)。在分类问题中, 旨在生成众多决策树, 并通过对建模数据集的样本观测和特征变量分别进行随机抽样, 每次抽样结果均为一棵树, 且每棵树都会生成符合自身属性的规则和分类结果, 而森林最终集成所有决策树的规则和分类结果, 实现随机森林算法的分类。其主要特点在于数据采集随机、特征选取随机。

2 模型数据制备

2.1 数据来源及用途

将镇域孕灾因子数据赋予评价单元作为随机森林分类模型的样本数据, 其中孕灾因子为模型特征因子, 区域滑坡环境特征因子数据源主要包括遥感影像数据、表面数字高程模型、地质图、野外调查数据等, 其分类及用途见表 2。

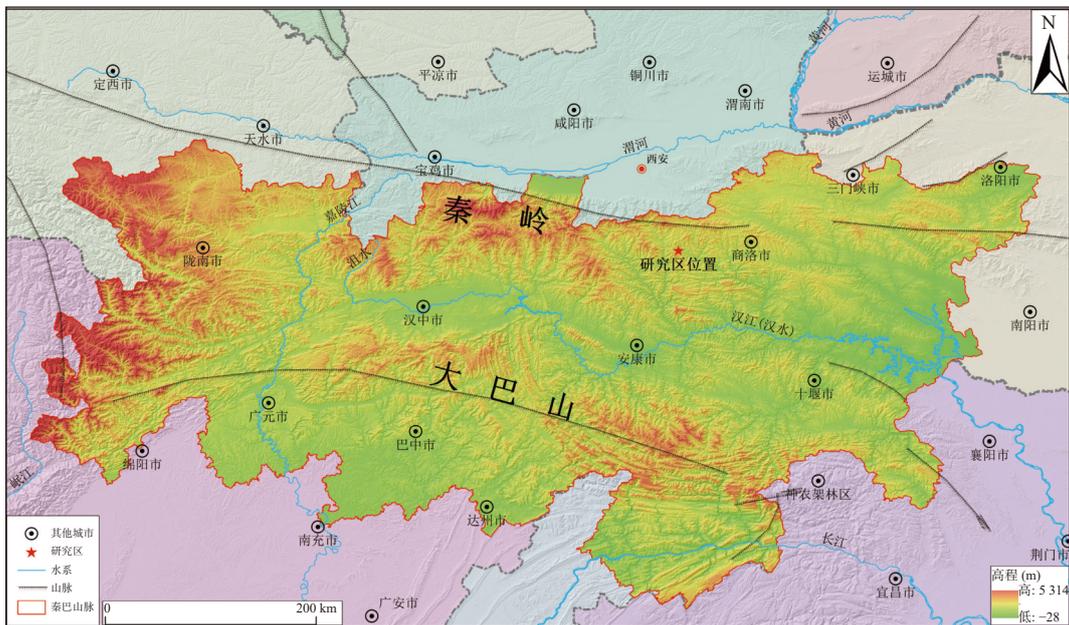


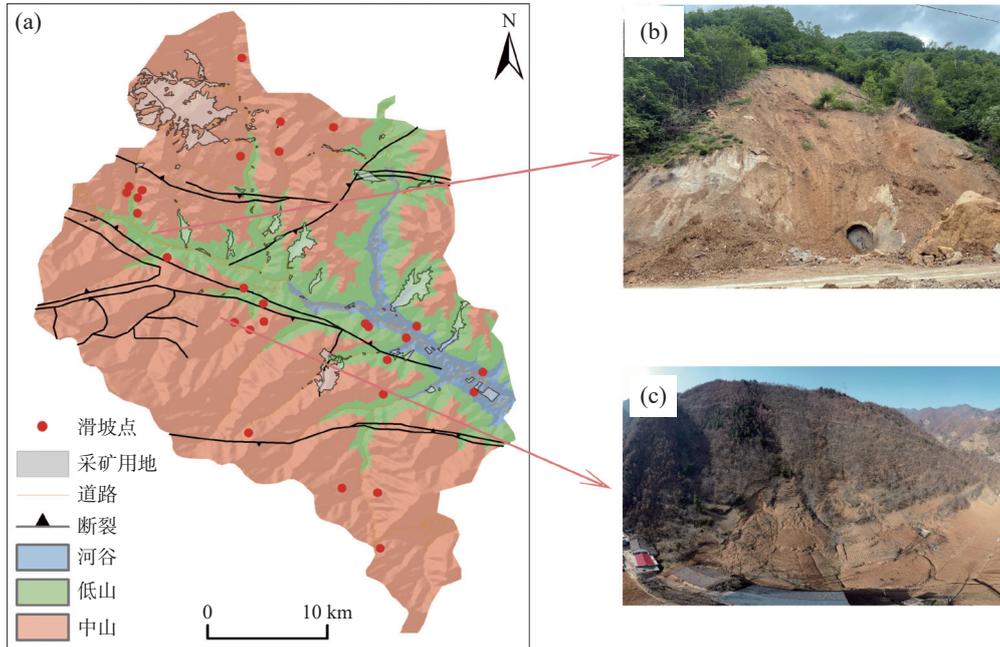
图1 研究区位置图

Fig. 1 Location of the study area

表 1 研究区滑坡规模分类

Tab. 1 Classification of landslide scale in study area

| | 个数 | 规模 | | | | | | 比例(%) |
|-------|----|-------|-------|-------|-------|-------|-------|--------|
| | | 大型(处) | | 中型(处) | | 小型(处) | | |
| | | 比例(%) | 比例(%) | 比例(%) | 比例(%) | 比例(%) | 比例(%) | |
| 堆积层滑坡 | 26 | 0 | 0 | 1 | 3.57 | 25 | 89.29 | 92.86 |
| 基岩滑坡 | 2 | 0 | 0 | 0 | 0 | 2 | 7.14 | 7.14 |
| 合计 | 28 | 0 | 0 | 1 | 3.57 | 27 | 96.43 | 100.00 |



a. 研究区现有堆积层滑坡分布; b. 岭丰村矿洞洞口堆积层滑坡; c. 罗庄三组堆积层滑坡

图2 研究区已有堆积层滑坡分布图

Fig. 2 The distribution of alluvial landslide in the study area

表2 数据类型及用途

Tab. 2 Data types and uses

| 数据类型 | 比例尺/分辨率 | 数据用途 |
|------|----------|-------------------------------|
| DEM | 5 m | 提取坡度、坡向、剖面曲率、河流水系等因子; 提取评价单元。 |
| 地质图 | 1:50 000 | 提取断裂等因子 |

2.2 评价单元

①栅格单元: 研究区栅格共计 4 576 590 个栅格; 提取出滑坡周界内所占栅格, 共计 6 192 个栅格单元(图 3)。

②斜坡单元: 对研究区 DEM 用水文分析模型进行斜坡单元初步划分, 再结合研究区卫星影像图与 10 米等高线进行人工校正(谷天峰等, 2013; 刘彬等, 2021), 最终将研究区划分为 729 个斜坡单元(图 4), 单个面积概况见表 3, 对其中滑坡所在斜坡单元进行标记, 共计 28 个斜坡单元。

2.3 特征因子

滑坡的发生是一个由多种因素作用下产生的非常复杂的非线性过程, 要通过机器学习模型对全区发生滑坡的概率进行分级, 特征因子的选取至关重要(Ahmed et al., 2016; 吴润泽等, 2021), 当根据研究区滑坡的孕灾特点, 有针对性的选取。在前人研究基础

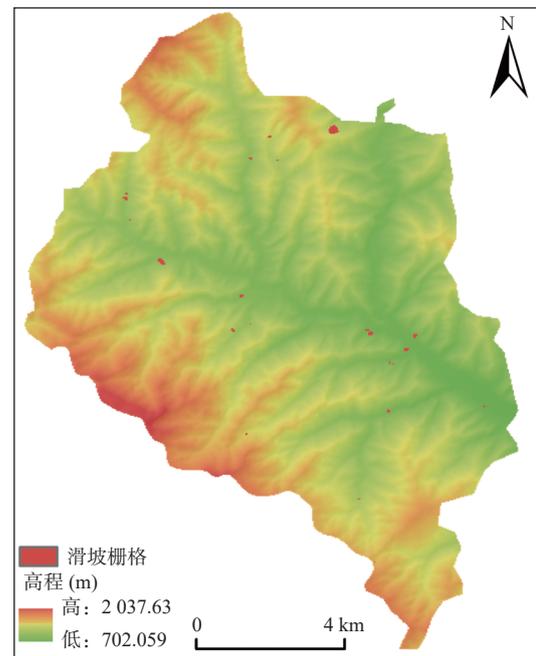


图3 研究区栅格单元划分

Fig. 3 Grid division of Study Area

上, 结合区域孕灾特征, 初步选 9 个孕灾因子作为模型特征因子用于建模, 并依据特征属性, 将特征因子进行离散化后做出属性空间划分(图 5), 具体如下:

①地形因子: 坡度、坡高。

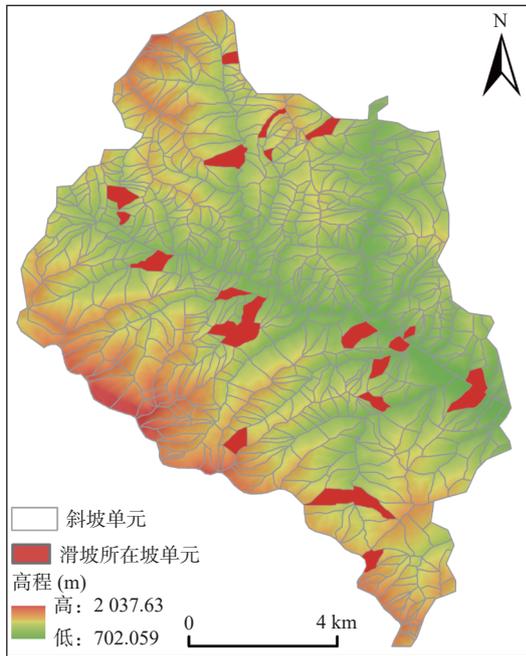


图4 研究区斜坡单元划分

Fig. 4 Slope division of study area

表 3 斜坡单元面积概况

Tab. 3 Overview of slope unit area

| 斜坡单元 面积类型 | 最大面积 (km^2) | 最小面积 (km^2) | 平均面积 (km^2) |
|--------------|---------------------------|---------------------------|---------------------------|
| 面积值 | 0.81 | 0.019 | 0.14 |

滑坡的主要动力来源于自重,坡度则控制着斜坡内部的应力分布,是滑坡的重要影响因素。将研究区斜坡坡度划分为: $0\sim 15^\circ$ 、 $15^\circ\sim 25^\circ$ 、 $25^\circ\sim 35^\circ$ 、 $35^\circ\sim 45^\circ$ 和 $>45^\circ$ 等 5 个属性空间。

在岩土类型大致相似的地质环境下,坡体高度影响着坡体自重大小以及斜坡范围内地形变化,进而影响滑坡的规模与范围(郭芳芳等, 2008)。先将 DEM 重采样为 $5\text{ m}\times 5\text{ m}$ 栅格,再利用分区统计工具得到坡体高度范围,将其分为 $0\sim 20\text{ m}$ 、 $20\sim 50\text{ m}$ 、 $50\sim 100\text{ m}$ 、 $100\sim 300\text{ m}$ 、 $>300\text{ m}$ 等 5 个属性空间。

②斜坡自身特征: 斜坡坡面形态、斜坡结构、斜坡堆积层厚度。

斜坡形态以其剖面曲率划分,剖面曲率定义为曲面在最陡斜坡方向上的曲率,当曲率值 <-0.5 时值判断为凹型坡,曲率值 >0.5 时判断为凸型坡,曲率值在 $-0.5\sim 0.5$ 之间时判断为直线型坡。

据前人研究,斜坡结构在堆积层滑坡的发生中也是一个重要的影响因素(邢林啸等, 2012; 贾琳等, 2021),从研究区已发生的堆积层滑坡来看,39.13%都

发生在顺向斜坡中,而逆向斜坡中无已发生滑坡,故根据野外调查数据,将区内斜坡结构划分为: 顺向斜坡、斜向斜坡、横向斜坡、近水平状斜坡、切向斜坡、逆向斜坡等 6 个属性空间。

由于区内滑坡大都为堆积层滑坡,堆积层厚度在一定程度上决定了土地利用类型、边坡开挖程度、地下水位埋深,对堆积层滑坡起着重要启发作用,但对于这一指标,目前缺乏经济又精准的测取方法,本研究方法如下: 对每个坡体随机抽点使用插杆进行测量得到平均数,结合依托项目的工程钻探数据对调查结果进行修正并类比到全区,赋值给每个斜坡,将该因子划分为: $0\sim 1\text{ m}$ 、 $1\sim 3\text{ m}$ 、 $3\sim 6\text{ m}$ 、 $6\sim 10\text{ m}$ 、 $>10\text{ m}$ 等 5 个属性空间。

③人类工程相关特征: 距道路距离、距河流距离、距矿区距离。

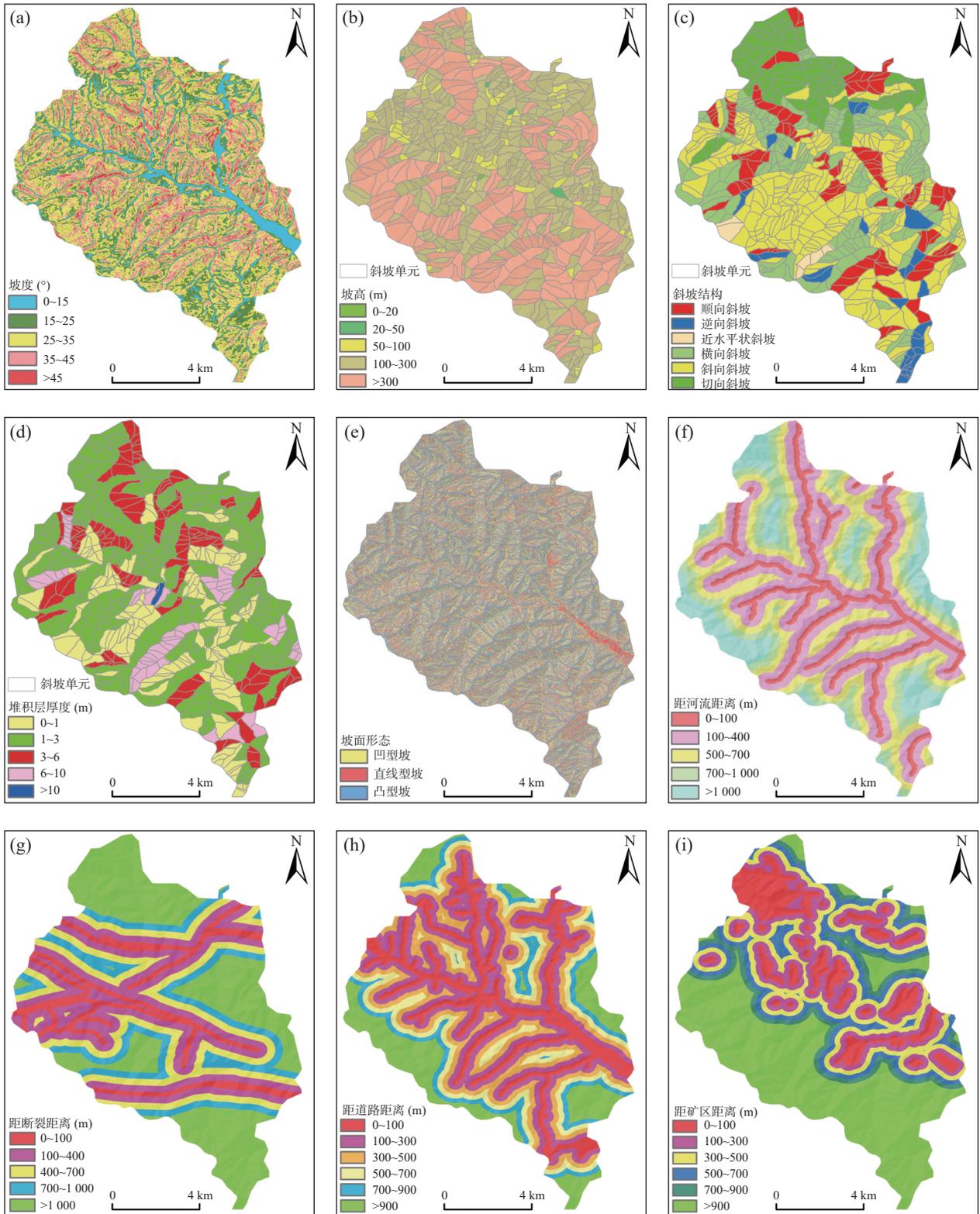
区内平缓地带稀少,道路开通处往往伴随着边坡开挖,形成陡立临空面; 河流对坡脚的侵蚀、依河而居带来的切坡建房; 矿山的开挖、震爆、不合理的矿渣堆积,这些因素都为滑坡的发生提供了有利条件。距河流、道路、矿区的距离均可反应人类工程活动的强弱程度,其各自属性区间划分如图 5 所示。

④地质条件: 距断裂距离。

区内受山阳—凤镇等深大断裂影响较大,该断裂呈近东西展布,经凤镇横贯研究区,该断裂对区内表层岩体破坏严重,加之研究区内节理裂隙极为发育,极大地破坏了岩土体的完整性,增加了滑坡的发生概率。对研究区地质图进行矢量化,提取研究区内断裂为矢量文件,对其做缓冲分析,划分为 $0\sim 100\text{ m}$ 、 $100\sim 400\text{ m}$ 、 $400\sim 700\text{ m}$ 、 $700\sim 1000\text{ m}$ 、 $>1000\text{ m}$ 等 5 个属性区

将上述特征因子赋值给两种划分单元: ①栅格单元: 在 Arcgis10.7 中,将栅格转化为点数据,用“多值提取至点”工具将各因子值提取至点,与初步选取的 9 个因子构成矩阵 $A_{4,576,590,9}$,作为原始样本 1。②斜坡单元: 除斜坡结构等因子在每个斜坡单元有唯一值外,对于坡面形态等离散型变量,取其众数赋值; 对于坡度等连续型变量,取其平均数赋值,729 个斜坡单元与初步选取的 9 个因子构成矩阵 $B_{729,9}$,作为原始样本 2。

对两种划分单元下的原始样本分别进行频率比分析,得到各因子属性区间的频率比值。频率比值的大小可反应该属性空间滑坡发育的优势程度(郭子正等, 2019),若某属性空间评价单元样本数过少(如斜



a.坡度; b.坡体高度; c.斜坡结构; d.堆积层厚度; e.斜坡形态; f.距河流距离; g.距断裂距离; h.距道路距离; i.距矿区距离

图5 研究区滑坡地质灾害易发性评价指标因子

Fig. 5 Index factors of landslide geological hazard Susceptibility assessment in the study area

坡单元中近水平斜坡数量), 则在选择优势空间时, 剔除该属性空间, 最终统计得各因子的滑坡发育优势空

间(表4)。

随机森林模型对特征因子的相关性有一定要求,

因子之间相关性过高会降低模型的训练精度。对研究区特征因子首先进行正态性检验, 以确定相关性分析的计算方式。若数据满足正态分布, 则选用 Pearson 相关性分析; 若数据不满足正态分布, 则用 Speraman 相关性分析; 针对有序变量数据, 则选用 Kendall's tau-b 等级进行相关性分析。

由表 5 可知, 两种评价单元下各因子显著性 P 值

均为 0.000***, 水平呈现显著性, 拒绝原假设, 因此数据不满足正态分布。由于数据为有序数据, 选用 Kendall's tau-b 等级进行相关性分析(表 6、表 7)。由结果可知, 栅格单元下, 研究区距河流距离与距道路距离特征因子的相关性系数为 0.653 > 0.3; 斜坡单元下, 研究区距河流距离与距道路距离特征因子的相关性系数为 0.606 > 0.3, 说明两种评价单元下道路与河流

表 4 两种评价单元下各因子的滑坡发育优势空间

Tab. 4 Dominant space of landslide development of each factor under two evaluation units

| 特征因子 | 坡度 (°) | 坡高 (m) | 堆积层厚度 (m) | 坡面形态 | 斜坡结构 | 距河流距离 (m) | 距道路距离 (m) | 距矿区距离 (m) | 距断裂距离 (m) |
|------|--------|---------|-----------|------|------|-----------|-----------|-----------|-----------|
| 栅格单元 | 25~35 | 100~300 | 1~3 | 凹型坡 | 顺向斜坡 | 100~400 | 0~100 | 500~700 | >1000 |
| 斜坡单元 | 25~35 | 100~300 | 1~3 | 凹型坡 | 顺向斜坡 | 100~400 | 0~100 | >1000 | >1000 |

表 5 特征因子数据正态性检验结果

Tab. 5 Characteristic factor data Normality test results

| 特征因子 | K-S检验 (栅格单元) | S-W检验 (斜坡单元) |
|-----------|-----------------|-----------------|
| 坡高 (m) | 0.33(0.000***) | 0.762(0.000***) |
| 距河流距离 (m) | 0.343(0.000***) | 0.843(0.000***) |
| 距道路距离 (m) | 0.210(0.000***) | 0.866(0.000***) |
| 距矿区距离 (m) | 0.361(0.000***) | 0.869(0.000***) |
| 堆积层厚度 (m) | 0.268(0.000***) | 0.841(0.000***) |
| 坡度 (°) | 0.293(0.000***) | 0.719(0.000***) |
| 坡面形态 | 0.355(0.000***) | 0.794(0.000***) |
| 斜坡结构 | 0.439(0.000***) | 0.849(0.000***) |
| 距断裂距离 (m) | 0.208(0.000***) | 0.852(0.000***) |

注: ***、**、*分别代表1%、5%、10%的显著性水平。

表 6 特征因子 Kendall's tau-b 等级相关系数矩阵 (栅格单元)

Tab. 6 Characteristic factor Kendall's tau-b rank correlation coefficient matrix (grid units)

| | 距矿区距离 (m) | 距道路距离 (m) | 坡面类型 | 坡度 (°) | 距断裂距离 (m) | 斜坡结构 | 坡高 (m) | 堆积层厚度 (m) | 距河流距离 (m) |
|-----------|-----------|-----------|--------|--------|-----------|--------|--------|-----------|-----------|
| 距矿区距离 (m) | 1 | | | | | | | | |
| 距道路距离 (m) | 0.304 | 1 | | | | | | | |
| 坡面类型 | 0.013 | 0.024 | 1 | | | | | | |
| 坡度 (°) | -0.01 | 0.087 | 0.023 | 1 | | | | | |
| 距断裂距离 (m) | -0.001 | 0.118 | 0.005 | -0.02 | 1 | | | | |
| 斜坡结构 | -0.141 | -0.057 | 0.006 | -0.004 | 0.201 | 1 | | | |
| 坡高 (m) | 0.214 | 0.203 | 0.009 | 0.06 | 0.002 | -0.003 | 1 | | |
| 堆积层厚度 (m) | -0.053 | -0.043 | -0.003 | -0.028 | 0.002 | 0.09 | -0.024 | 1 | |
| 距河流距离 (m) | 0.099 | 0.653 | 0.029 | 0.073 | 0.121 | 0.023 | 0.125 | 0.003 | 1 |

表 7 特征因子 Kendall's tau-b 等级相关系数矩阵(斜坡单元)

Tab. 7 Characteristic factor Kendall's tau-brank correlation coefficient matrix (slope units)

| | 距矿区 距离(m) | 距道路 距离(m) | 坡面 类型 | 坡度 (°) | 距断裂 距离(m) | 斜坡 结构 | 坡高 (m) | 堆积层 厚度(m) | 距河流 距离(m) |
|----------|--------------|--------------|----------|-----------|--------------|----------|-----------|--------------|--------------|
| 距矿区距离(m) | 1 | | | - | | | | | |
| 距道路距离(m) | 0.246 | 1 | | | | | | | |
| 坡面类型 | 0.059 | 0.124 | 1 | | | | | | |
| 坡度(°) | -0.07 | 0.034 | 0.046 | 1 | | | | | |
| 距断裂距离(m) | 0.006 | 0.186 | 0.042 | -0.018 | 1 | | | | |
| 斜坡结构 | -0.112 | -0.06 | 0.032 | -0.083 | 0.199 | 1 | | | |
| 坡高(m) | 0.162 | 0.202 | 0.043 | 0.154 | -0.011 | -0.001 | 1 | | |
| 堆积层厚度(m) | -0.039 | -0.05 | -0.043 | -0.085 | -0.002 | 0.093 | -0.068 | 1 | |
| 距河流距离(m) | 0.02 | 0.606 | 0.122 | 0.004 | 0.138 | 0.011 | 0.125 | -0.002 | 1 |

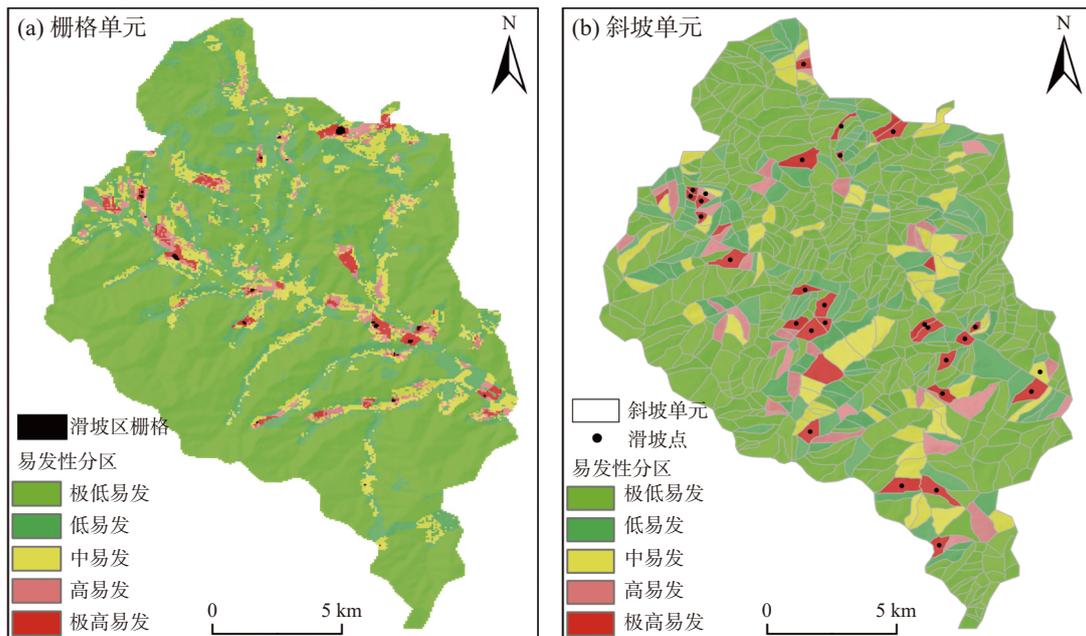
之间的相关性都较高。由于研究区堆积层滑坡主要分布在道路两侧,故剔除河流距离因子,最终采用其他剩余 8 个因子进行建模评价。

3 评价及结果对比分析

3.1 评价结果

对于栅格单元,将 4 576 590 个栅格单元转化为连接有各因子数据的点后,将 6 192 个滑坡栅格点作为滑坡样本并将其易发性赋值为 1,占总样本的 0.13%,

在滑坡范围外选取 3 倍于滑坡样本的点作为非滑坡样本,以减少滑坡与非滑坡样本之间的数量不平衡与空间相关性,并将其易发性赋值为 0,两类数据合并后以 7:3 的比例划分为训练集与测试集,在 R 语言的 RF 软件包中构建随机森林模型,将合并后数据输入进行训练。训练结束后,将整个研究区点数据输入,得到所有点的易发性值。将点数据转为栅格,利用自然间断法划分出极低、低、中、高、极高等 5 个易发区(图 6)。对于斜坡单元,选取 28 个滑坡所在的斜坡单元作为滑坡样本并将其易发性赋值为 1,占总样本的



a. 栅格单元下研究区滑坡易发性区划图; b. 斜坡单元下研究区滑坡易发性区划图

图 6 研究区滑坡易发性区划

Fig. 6 The division of landslide susceptibility in the study area

3.8%, 在滑坡范围外选取 3 倍于滑坡样本的点作为非滑坡样本, 并将其易发性赋值为 0。剩余操作与栅格单元类似, 结果见图 6。由图 6、表 8 可知: ①两种评价单元下滑坡易发性分区大体相近, 滑坡高、极高易发区主要位于镇域北部矿区附近、西北部道路切坡密集区域、西南部盘山公路切坡密集道路附近, 表明研究区内滑坡的发育与道路、矿区的影响有较强的相关性, 这也与实际情况相符合。②栅格单元下预测的极低、低、中、高和极高易发区占研究区全区比例为: 66.76%、19.50%、8.99%、3.04% 和 1.71%, 已知滑坡样本在高、极高易发样本量中占 95.49%; 斜坡单元下预测的极低、低、中、高、极高易发区占研究区全区比例分别为 61.92%、17.70%、10.09%、5.25% 和 5.04%, 已知滑坡样本在高、极高易发样本量中占 96.42%。③栅格单元下高、极高易发区频率比之和占总频率比值为 0.93, 略低于斜坡单元下比值 0.98, 可知斜坡单元下的滑坡易发性可更好的判断已知滑坡, 且斜坡单元下的评价可很好的表现滑坡的地形分布特点。

表 8 栅格单元与斜坡单元下评价结果频率比

Tab. 8 Frequency ratio of evaluation results under grid unit and slope unit

| 评价单元 | 易发性 | 滑坡单元数 (个) | 滑坡单元比例 (%) | 全区单元数 (个) | 全区单元比例 (%) | 频率比 |
|------|-----|-----------|------------|-----------|------------|-------|
| 栅格单元 | 极低 | 0 | 0.00 | 3 055 331 | 66.76 | 0.00 |
| | 低 | 27 | 0.44 | 892 435 | 19.50 | 0.02 |
| | 中 | 252 | 4.07 | 411 435 | 8.99 | 0.45 |
| | 高 | 603 | 9.73 | 139 129 | 3.04 | 3.20 |
| | 极高 | 5 310 | 85.76 | 78 260 | 1.71 | 50.15 |
| 斜坡单元 | 极低 | 0 | 0.00 | 451 | 35.67 | 0.00 |
| | 低 | 0 | 0.00 | 129 | 28.53 | 0 |
| | 中 | 1 | 3.57 | 79 | 18.38 | 0.19 |
| | 高 | 1 | 3.57 | 38 | 12.62 | 0.28 |
| | 极高 | 26 | 92.86 | 32 | 4.80 | 19.35 |

由模型得到两种评价单元下各特征因子的重要性比例(图 7), 可知道路开挖、斜坡结构对滑坡的发生起主要作用, 事实上, 现有滑坡多发生在道路两旁及顺向结构斜坡中。

3.2 模型精度及评价结果对比

结合前人研究(刘坚等, 2018; 郑迎凯等, 2020), 笔者选取 ROC 曲线来进行模型精度验证。将全部样本滑坡易发性概率递减排序, 阈值从 1 至 0 变更, 计

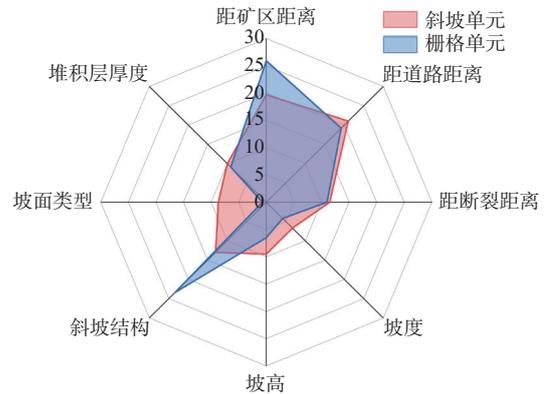


图 7 各特征因子贡献值

Fig. 7 Contribution value of each characteristic factor

算各阈值下对应的(FPR、TPR)数值对, 将数值对绘制于直角坐标系中。横坐标 FPR 代表非滑坡预测为滑坡的数量/真实非滑坡数量; 纵坐标 TPR 代表滑坡预测为滑坡的数量/真实滑坡数量。曲线下面积(AUC)越接近 1, 则说明模型预测性能越高。根据两种评价单元下预测结果, 绘制各自 ROC 曲线(图 8), 可以看出, 基于栅格单元、斜坡单元下随机森林模型的 AUC 值分别为 0.864、0.921。表明 5 m 高精度栅格单元下随机森林模型的预测精度小于基于斜坡单元下模型的预测精度。

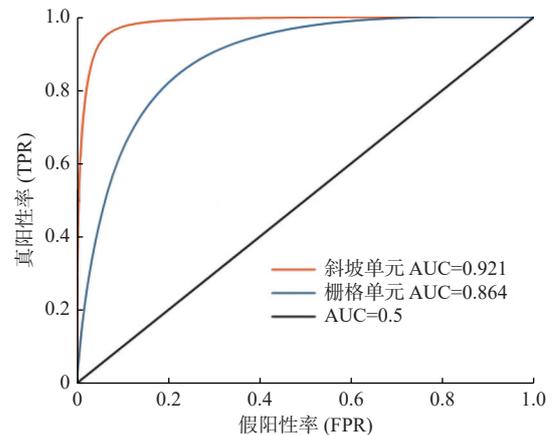


图 8 ROC 曲线

Fig. 8 ROC curve

滑坡的发生相对于整个区域是一个小概率事件, 整个区域内滑坡易发性概率在预测模型精度较高的情况下, 均值越低, 表明预测的不确定性越低; 标准差越低, 表明预测区分度更优, 由表 9 可知, 栅格单元下预测滑坡易发性概率均值为 0.10 小于斜坡单元下滑坡易发性概率均值 0.13; 栅格单元下预测滑坡易发性概率标准差为 0.13, 小于坡单元下滑坡易发性概率标

准差 0.18。说明栅格单元作为评价单元, 预测结果具有较差的区分度, 斜坡单元反之。

表 9 不同评价单元下易发性概率均值与标准差
Tab. 9 Mean and standard deviation of probability of Susceptibility under different evaluation units

| 评价单元 | 均值 | 标准差 |
|------|------|------|
| 栅格单元 | 0.10 | 0.13 |
| 斜坡单元 | 0.13 | 0.18 |

4 结论

(1) 结合研究区孕灾特征, 选取了坡度、坡高、坡面形态、斜坡结构类型、堆积层厚度、道路、矿区、区域地质断裂等 8 个孕灾因子作为模型特征因子, 栅格单元、斜坡单元两种评价单元的预测结果, 均发现模型主要根据道路、斜坡结构等因素区分滑坡、非滑坡区域。据各因子频率比值可知, 研究区内, 在距道路 100 m 内的顺向浅层堆积层斜坡中, 最易发生滑坡。

(2) 在栅格单元、斜坡单元两种评价单元下, 研究区 95.49%、96.42% 的滑坡落在随机森林模型预测分区出的高、极高易发区内; 通过 AUC 值、易发性概率均值、易发性概率标准差对比表明: 随机森林是一种可靠、全面的滑坡易发性预测模型, 评价单元为斜坡单元时预测精度更高、准确性更强。

(3) 滑坡易发性评价中, 栅格单元与斜坡单元在评价结果上, 两种评价单元的预测结果都有良好的表现。栅格单元的样本数远高于斜坡单元, 意味着样本数据中滑坡样本比例减小, 样本的不均衡性增大, 使得随机森林模型的预测结果性能变弱, 但以栅格单元为评价单元可减少特征因子的数据损失, 在灾害防治具体空间部署上有着更精细的参考; 斜坡单元在特征因子赋值时, 会造成部分数据损失, 但其能更好体现滑坡灾害的地形空间属性, 相较于栅格单元其样本的不均衡性更小, 使得其基于随机森林模型的预测性能高于栅格单元。

参考文献(References):

常志璐, 黄发明, 蒋水华, 等. 基于多尺度分割方法的斜坡单元划分及滑坡易发性预测[J]. 工程科学与技术, 2023, 55(1): 184-195.
CHANG Zhilu, HUANG Faming, JIANG Shuihua, et al. Slope Unit

Extraction and Landslide Susceptibility Prediction Using Multi-scale Segmentation Method[J]. Advanced Engineering Sciences, 2023, 55(1): 184-195.

方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. 统计与信息论坛, 2011, 26(03): 32-38.

FANG Kuangnan, WU Jianbin, ZHU Jianping, et al. Review of random forest methods[J]. Statistics and Information Forum, 2011, 26(03): 32-38.

范立民, 何进军, 李存购. 秦巴山区滑坡发育规律研究[J]. 中国地质灾害与防治学报, 2004, 15(01): 47-51.

FAN Limin, HE Jinjun, LI Cungou. Study on the development law of landslides in Qinba Mountains[J]. Chinese Journal of Geological Hazards and Prevention, 2004, 15(01): 47-51.

郭芳芳, 杨农, 孟晖, 等. 地形起伏度和坡度分析在区域滑坡灾害评价中的应用[J]. 中国地质, 2008, 35(1): 131-143.

郭子正, 殷坤龙, 黄发明, 等. 基于滑坡分类和加权频率比模型的滑坡易发性评价[J]. 岩石力学与工程学报, 2019, 38(02): 287-300.

GUO Zizheng, YIN Kunlong, HUANG Faming, et al. Landslide susceptibility evaluation based on landslide classification and weighted frequency ratio model[J]. Chinese Journal of Rock Mechanics and Engineering, 2019, 38(02): 287-300.

黄发明, 胡松雁, 闫学涯, 等. 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, 2022, 41(02): 79-90.

HUANG Fangming, HU Songyan, YAN Xueya, et al. Landslide susceptibility prediction modeling and main control factor identification based on machine learning[J]. Bulletin of Geological Science and Technology, 2022, 41(02): 79-90.

黄润秋. 20 世纪以来中国的大型滑坡及其发生机制[J]. 岩石力学与工程学报, 2007, 26(3): 433-454.

HUANG Runqiu. Large-scale landslides and their mechanism in China since the 20th century[J]. Journal of Rock Mechanics and Engineering, 2007, 26(3): 433-454.

贾琳, 蔡静森, 晏鄂川, 等. 基于地质环境分区的南漳县城区滑坡易发性评价[J]. 人民长江, 2021, 52(05): 86-94.

JIA Lin, CAI Jingsen, YAN Echuan, et al. Evaluation of landslide susceptibility in Nanzhang County based on geological environment zoning[J]. People's Yangtze River, 2021, 52(05): 86-94.

谷天峰, 王家鼎, 付新平, 等. 基于斜坡单元的区域斜坡稳定性评价方法[J]. 地理科学, 2013, 33(11): 1400-1405.

GU Tianfeng, WANG Jiading, FU Xinping, et al. Regional slope stability evaluation method based on slope unit[J]. Scientia Geographica Sinica, 2013, 33(11): 1400-1405.

刘彬. 基于 GIS 与随机森林算法的斜坡单元类型划分方法[J]. 经纬天地, 2021, (04): 82-86.

LIU Bin. Classification method of slope unit types based on GIS and random forest algorithm[J]. Jingwei Heaven and Earth, 2021,

- (04): 82–86.
- 刘坚, 李树林, 陈涛, 等. 基于优化随机森林模型的滑坡易发性评价[J]. 武汉大学学报(信息科学版), 2018, 43(07): 1085–1091.
- LIU Jian, LI Shulin, CHEN Tao, et al. Evaluation of landslide susceptibility based on optimized random forest model[J]. Geomatics and Information Science of Wuhan University, 2018, 43(07): 1085–1091.
- 刘伟, 袁湘秦, 连海波, 等. 陕南地区堆积层滑坡发育规律研究[J]. 安徽农业科学, 2021, 49(10): 194–197+224.
- LIU Wei, YUAN Xiangqin, LIAN Haibo, et al. Research on the development law of accumulation landslide in southern Shaanxi[J]. Journal of Anhui Agricultural Sciences, 2021, 49(10): 194–197+224.
- 强菲, 赵法锁, 段钊, 等. 陕南秦巴山区地质灾害发育及空间分布规律[J]. 灾害学, 2015, 30(2): 193–198.
- QIANG Fei, ZHAO FaSuo, DUAN Zhao, et al. Development and spatial distribution of geological disasters in Qinling-Data Mountains of south Shaanxi[J]. Journal of Catastrophology, 2015, 30(2): 193–198.
- 陕西省自然资源厅. 陕西省城镇(乡镇)地质灾害风险调查评价技术要求(1: 10000)(试行)[R]. 西安: 陕西省自然资源厅, 2021.
- 唐睿旋. 堆积层滑坡单体稳定性评估及区域易发性评价研究[D]. 北京: 中国地质大学(北京), 2017.
- TANG Ruixuan. Study on stability assessment and regional vulnerability assessment of accumulative landslide[D]. Beijing: China University of Geosciences (Beijing), 2017.
- 田述军, 张珊珊, 唐青松, 等. 基于不同评价单元的滑坡易发性评价对比研究[J]. 自然灾害学报, 2019, 28(06): 137–145.
- TIAN Shujun, ZHANG Shanshan, TANG Qingsong, et al. Comparative study of landslide susceptibility assessment based on different evaluation units[J]. Journal Of Natural Disasters, 2019, 28(06): 137–145.
- 吴润泽, 胡旭东, 梅红波, 等. 基于随机森林的滑坡空间易发性评价: 以三峡库区湖北段为例[J]. 地球科学, 2021, 46(01): 321–330.
- WU Runze, HU Xudong, MEI Hongbo, et al. Spatial susceptibility assessment of landslides based on random forest: A case study in the Hubei section of the Three Gorges Reservoir Area[J]. Earth Science, 2021, 46(01): 321–330.
- 邢林啸. 三峡库区典型堆积层滑坡成因机制与预测预报研究[D]. 武汉: 中国地质大学, 2012.
- XING Linxiao. Research on the genetic mechanism and prediction of typical accumulation landslide in the Three Gorges Reservoir Area[D]. Wuhan: China University of Geosciences, 2012.
- 赵力行, 范文, 柴小庆, 等. 秦巴山区地质灾害发育规律研究——以镇巴县幅为例[J]. 地质与资源, 2020, 29(02): 187–195.
- ZHAO lixing, FAN Wen, CHAI Xiaoqing, et al. A study on the development law of geological hazards in qinling-bashan mountains: taking the Zhenba County area as an example[J]. Geology and Resources, 2020, 29(02): 187–195.
- 张林梵, 王佳运, 张茂省, 等. 基于BP神经网络的区域滑坡易发性评价[J]. 西北地质, 2022, 55(02): 260–270.
- ZHANG Linfan, WANG Jiayun, ZHANG Maosheng, et al. Evaluation of regional landslide susceptibility based on BP neural network[J]. Northwestern Geology, 2022, 55(02): 260–270.
- 段钊, 彭建兵, 陈伟, 等. 泾河下游黄土台塬区滑崩灾害空间分异研究[J]. 西北地质, 2018, 51(03): 214–222.
- DUAN Zhao, PENG Jianbing, CHEN Wei, et al. Spatial differentiation of landslide disasters in the loess plateau area of the lower reaches of the Jinghe River[J]. Northwestern Geology, 2018, 51(03): 214–222.
- 张世林. 秦巴山区斜坡结构类型及变形破坏模式研究[D]. 西安: 长安大学, 2020.
- ZHANG Shilin. Study on slope structure type and deformation failure mode in qinling-bashan mountain area [D]. Xi'an: Chang'an University, 2020.
- 郑迎凯, 陈建国, 王成彬, 等. 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, 2020, 39(06): 131–144.
- ZHENG Yingkai, CHEN Jianguo, WANG Chengbin, et al. Application of certainty coefficient and Random forest model in landslide susceptibility assessment in Mangshi, Yunnan[J]. Bulletin of Geological Science and Technology, 2020, 39(06): 131–144.
- Ahmed Mohamed Youssef, Hamid Reza Pourghasemi, Zohre Sadat Pourtaghi, et al. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at WadiTayyah Basin, Asir Region, Saudi Arabia[J]. Landslides, 2016, 12(02): 839–856.
- Merghadi Abdelaziz, Abderrahmane Boumezeur, Tien Bui Dieu, et al. Landslide Susceptibility Assessment at Mila Basin (Algeria): A Comparative Assessment of Prediction Capability of Advanced Machine Learning Methods[J]. ISPRS International Journal of Geo-Information, 2018, 7(7): 268.
- Saro Lee, Joong-Sun Won, Seong Woo Jeon, et al. Spatial Landslide Hazard Prediction Using Rainfall Probability and a Logistic Regression Model[J]. Mathematical Geosciences, 2015, 47(5): 565–589.