

基于 MaxEnt 和斜坡单元的崩滑灾害易发性评价

——以云南省孟连县为例

汪木林¹, 郭婷婷^{1*}, 杨彦武², 周旭萌³, 杨宏¹, 张智伟⁴

(1. 昆明理工大学国土资源工程学院, 昆明 650093; 2. 西南有色昆明勘测设计(院)股份有限公司, 昆明 650051;
3. 云南地矿地质工程有限公司, 昆明 650041; 4. 华能澜沧江水电股份有限公司, 昆明 650214)

摘要: 为验证基于 MaxEnt 和斜坡单元崩滑灾害易发性评价模型的可行性与准确性, 以云南省孟连县为例, 选取高程、坡度、坡向、距断层距离、距公路距离、距河流距离、降水量、工程地质岩组、归一化植被指数(NDVI)9个评价因子, 基于 MaxEnt 模型, 对划分的800个斜坡单元进行易发性评价, 将评价结果与基于栅格单元的评价结果进行对比。结果表明:(1)孟连县崩滑灾害极低易发区面积占37.44%; 低易发区占29.08%; 中易发区占17.67%; 高易发区占11.26%; 极高易发区占4.54%。(2)针对因子贡献率, 基于斜坡单元和栅格单元的评价结果具有一定相似性, 均说明人类工程活动和地质构造是孟连县崩滑灾害的主要诱因。(3)基于斜坡单元和栅格单元的评价结果 AUC 值分别为0.858和0.872, z 统计量计算结果表明两者结果精确度相同, 但基于斜坡单元更符合现实调查的实际意义。证明基于 MaxEnt 模型和斜坡单元的崩滑灾害易发性评价模型具有较好的可行性与准确性。

关键词: MaxEnt; 斜坡单元; 崩滑灾害; 易发性评价; 孟连县

中图分类号: P642.21; P642.22; P694; P56 **文献标识码:** A

EVALUATION OF LANDSLIDE SUSCEPTIBILITY BASED ON MAXENT AND SLOPE UNIT

——TAKING MENGLIAN COUNTY OF YUNNAN PROVINCE AS AN EXAMPLE

WANG Mu-lin¹, GUO Ting-ting^{1*}, YANG Yan-wu², ZHOU Xu-meng³, YANG Hong¹, ZHANG Zhi-wei⁴

(1. Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China;
2. Southwest Nonferrous Kunming Survey and Design (Institute) Co. Ltd., Kunming 650051, China;
3. Yunnan Geology and Mineral Engineering Co., Ltd., Kunming 650041, China;
4. Huaneng Lancang River Hydropower Co., Ltd., Kunming 650214, China)

Abstract: In order to verify the feasibility and accuracy of the evaluation model of landslide susceptibility based on MaxEnt and slope unit, taking Menglian County of Yunnan Province as an example, nine evaluation factors of elevation, slope, aspect, distance from fault, distance from highway, distance from river, precipitation, engineering geological rock group and normalized difference vegetation index (NDVI) were selected. Based on the MaxEnt model, the susceptibility of 800 slope units was evaluated, and the evaluation results were compared with the evaluation results based on grid unit. The results show that: (1) The extremely low susceptibility area of landslide disaster in Menglian County accounts for 37.44%

收稿日期: 2024-02-21 **改回日期:** 2024-04-29

基金项目: 国家自然科学基金项目(41d861134008); 云南省 Muhammad Asif Khan 院士工作站(202105AF150076); 云南省重点研发计划项目(202003AC100002); 云南省基础研究计划项目(202001AT070043)

% ; low-prone areas accounted for 29.08 % ; medium-prone areas accounted for 17.67 % ; high-prone areas accounted for 11.26 % ; the extremely high-prone areas accounted for 4.54 % . (2) According to the factor contribution rate, the evaluation results based on slope unit and grid unit have certain similarity, which shows that human engineering activities and geological structure are the main causes of landslide disaster in Menglian County. (3) The AUC values of the evaluation results based on the slope unit and the grid unit are 0.858 and 0.872, respectively. The calculation results of the z statistic show that the accuracy of the two results is the same, but the slope unit is more in line with the practical significance of the actual survey. It is proved that the evaluation model of landslide hazard susceptibility based on MaxEnt model and slope unit has good feasibility and accuracy.

Key words: MaxEnt; slope unit; landslide disaster; susceptibility evaluation; Menglian County

地质灾害对人类社会和自然环境造成了严重的威胁,其中崩塌与滑坡因其突发性和破坏性备受关注^[1]。为应对地质灾害可能造成的破坏,对潜在地质灾害进行预测,区域性灾害易发性评价是最为有效的手段之一^[2]。因此针对地质灾害易发性评价有众多方法,主要分为定性与定量两种。定性方法因依赖专家的经验水平,主观性较强,缺乏一定科学性。定量方法主要有信息量法^[3]、确定系数法^[4]等基于概率模型的方法,同时随着计算机技术发展,出现了逻辑回归^[5]、神经网络^[6]、支持向量机^[7]、随机森林^[8]等机器学习方法。目前国内外许多学者采用不同方法建立了地质灾害易发性评价模型,均取得了较好的效果,例如孟凡奇等^[9]运用信息量模型和层次分析法进行地质灾害易发性分区,结果良好;李信等^[10]采用 CF、CF-LR、CF-AHP 模型对研究区进行地质灾害易发性评价与精度分析,结果表明 CF-LR 精度更高;Arabameri 等^[11]利用随机森林、交替决策树和费雪线性判别来生成滑坡易发性地图。

MaxEnt 模型基于最大熵原理^[12],主要思想是在满足一组已知约束条件的情况下,选择熵最大的概率分布,应用在地质灾害易发性评价中表现为以利用环境变量推断灾害点的潜在分布。MaxEnt 模型相较于其他模型具有在小样本情况下表现较好、可以处理包含多个变量的复杂数据、更能适应复杂的非线性关系的优点,从而更好地捕捉地质灾害易发性评价中各种环境变量之间的复杂相互作用^[13, 14]。

在滑坡易发性研究中,常用的评价单元有流域单元、栅格单元、地形单元、网格单元和斜坡单元,其中斜坡单元可提高与实际地貌的吻合度,能较为客观准确地体现区域斜坡的整体性和崩滑灾害发育状

况,是较为理想的评价单元^[15]。因此本文以云南省孟连县为例,将 MaxEnt 模型与斜坡单元相结合对崩滑灾害进行易发性评价,计算精确度与因子贡献率,并将最终结果与基于传统栅格单元的评价结果进行对比,验证基于 MaxEnt 和斜坡单元崩滑灾害易发性评价模型的可行性与准确性,为其他相似地区的崩滑灾害易发性评价提供参考。

1 研究区概况

孟连傣族拉祜族佤族自治县(以下称孟连县)位于云南省西南部,隶属普洱市,地理坐标:99°09′~99°46′E, 22°05′~22°32′N,全县总面积 1 893.36 km²(图 1)。孟连县处于横断山系纵谷区的南段,大

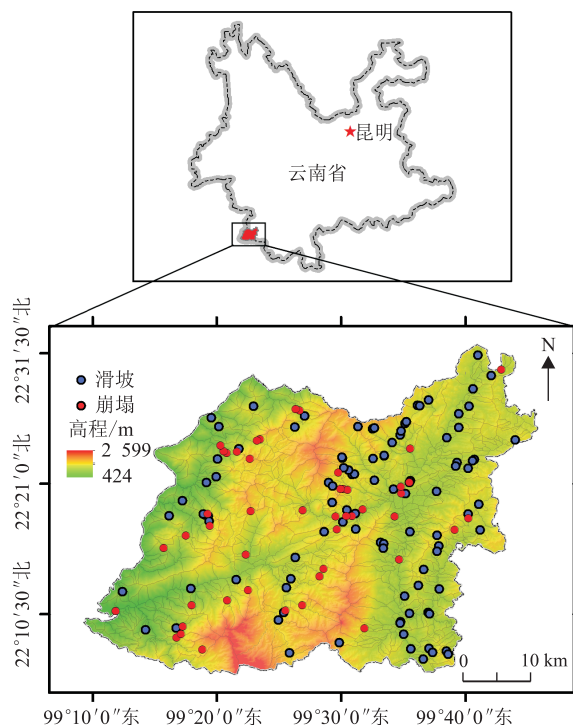


图 1 研究区位置示意图

体为北东-南西走向的山地,夹着宽谷盆地所组成,为贡山余脉所控制形成的宽谷盆地与山地相间的地形地貌。境内最高点位于孟连县南部国界的大黑山山峰,海拔 2 603.5 m,最低处位于南西部的南马河国界处河床,海拔 491.0 m,相对高差 2 112.5 m,总体地势倾斜度较大为 21.9%。根据地貌成因类型及组合形态,将孟连县分为 3 种地貌类型:构造侵蚀地貌、溶蚀地貌和侵蚀堆积地貌。孟连县处于滇西南经向构造活动带内,新构造运动频繁,褶皱构造强烈,断裂构造发育,沉积岩、岩浆岩、变质岩均有出露。受构造破坏,岩层破碎,裂隙发育,岩体抗风化能力较弱。主要人类工程活动有陡坡垦植、公路建设、采矿活动等。根据地质灾害精细化调查显示孟连县境内崩滑灾害共 148 处,其中滑坡 101 处,崩塌 47 处。

2 评价方法

2.1 因子选取及数据来源

在搜集各类孟连县地质灾害野外调查资料及成果的基础上,对孟连县崩滑灾害发育特征及分布规律进行总结,选取高程、坡度、坡向、距断层距离、距公路距离、距河流距离、降水量、工程地质岩组、归一化植被指数(NDVI)等 9 个评价因子进行崩滑灾害易发性分析。如表 1 所示,高程 DEM 数据来源于地理空间数据云,分辨率为 30 m,坡度和坡向数据使用 ArcGIS 从高程数据中提取;断层、公路、工程地质岩组、河流和降雨量数据来源于云南省孟连县地质灾害风险普查数据成果,之后通过 ArcGIS 建立缓冲区,对结果进行裁剪、合并等步骤后得到距断层距离、距公路距离、距河流距离数据;归一化植被指数(NDVI)来源于国家地球系统科学数据中心。

表 1 数据来源

类别	评价因子	数据来源
地形	高程	地理空间数据云(https://www.gscloud.cn/)
	坡度	地理空间数据云(https://www.gscloud.cn/)
	坡向	地理空间数据云(https://www.gscloud.cn/)
地质构造	工程地质岩组	云南省孟连县地质灾害风险普查
	距断层距离	云南省孟连县地质灾害风险普查
水文气象	降雨量	云南省孟连县地质灾害风险普查
	距河流距离	云南省孟连县地质灾害风险普查
植被	NDVI	国家地球系统科学数据中心(http://www.geodata.cn/)
人类工程活动	距公路距离	云南省孟连县地质灾害风险普查

2.2 斜坡单元划分

斜坡单元可以最大限度地反映研究区地质环境条件和地势起伏情况,更贴合现实中地质灾害调查工作要求。基于水文分析法,利用 ArcGIS 软件对 DEM 数据与 DEM 反转数据进行填注、流向提取、流量计算、河网生成和集水流域生成,结合云南省孟连县地质灾害精细化调查(1:50 000)斜坡单元划分结果进行手动调整,剔除噪点,得到斜坡单元栅格图层,具体流程见图 2。最终将研究区划分为 800 个斜坡单元,划分结果见图 3。使用 ArcGIS“分区统计”工具得到基于斜坡单元的评价因子图层,处理结果见图 4。

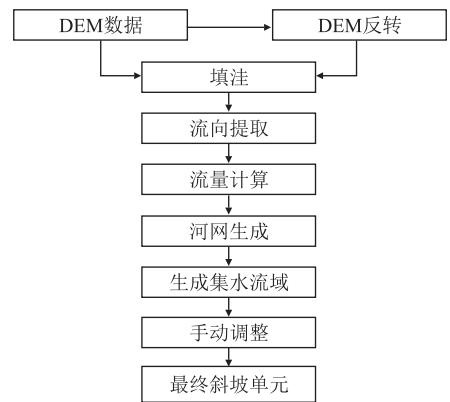


图 2 斜坡单元划分流程

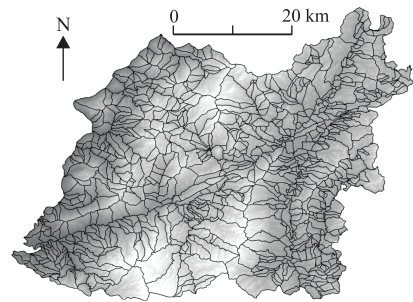


图 3 斜坡单元划分结果

2.3 MaxEnt 建模

MaxEnt 模型早期主要用于生物物种分布预测研究,目前在环境学、金融学等领域也有广泛应用,其核心是最大熵理论,即一组离散随机变量的信息熵在约束条件下达到最大时,即得到其最可能的概率分布^[16]。在崩滑易发性评价中,可将地形地貌、地质构造、人类工程活动等环境条件作为约束,计算崩滑灾害发生概率分布:

$$H(Y | X) = - \sum_{i=1}^n p(x_i) \sum_{j=1}^m p(y_j | x_i) \lg p(y_j | x_i) \quad (1)$$

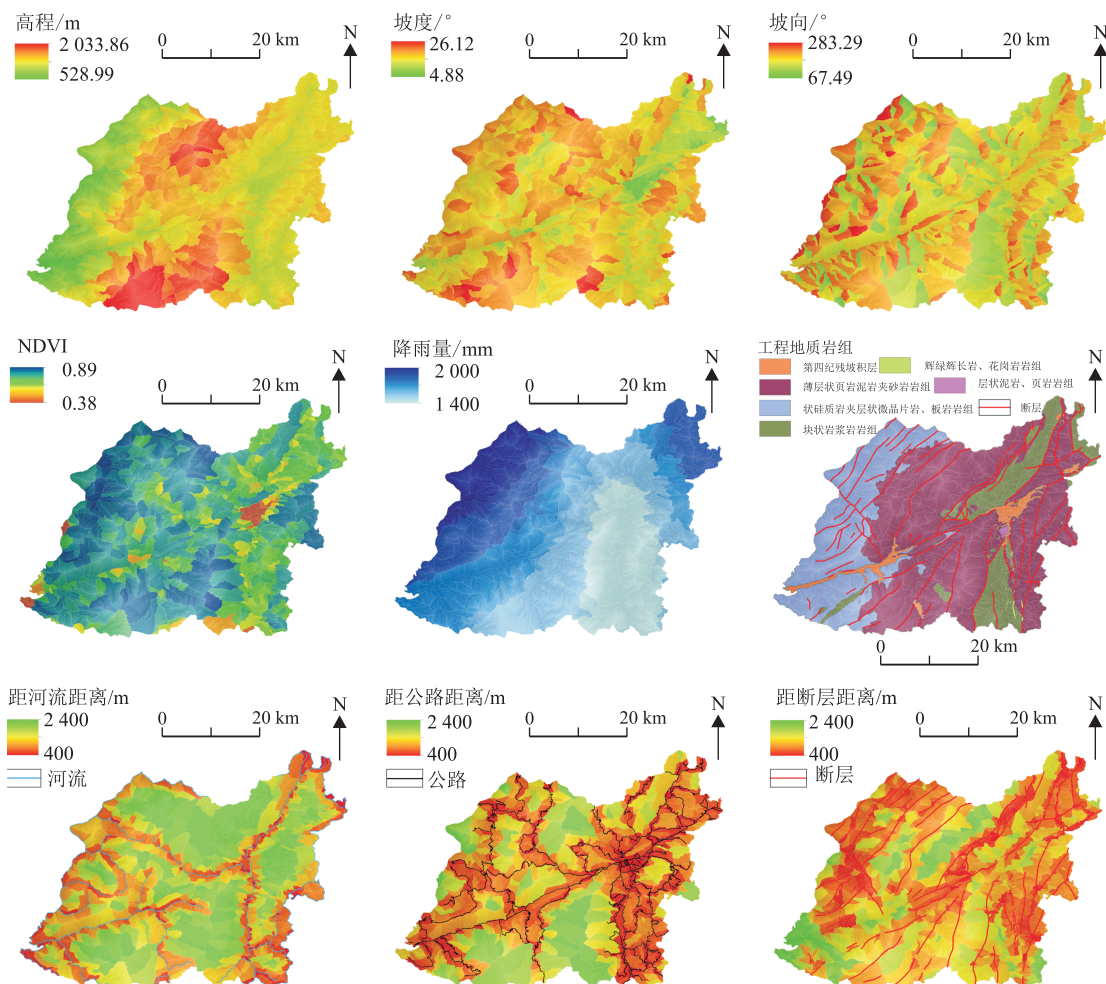


图4 因子提取至斜坡单元结果

其中, Y 为崩滑灾害发生概率分布; X 为环境条件。

将已发生的 148 处崩滑灾害分布坐标数据保存为 .csv 格式, 将高程、坡度、坡向、距断层距离、距公路距离、距河流距离、降水量、工程地质岩组和归一化植被指数 (NDVI) 9 个环境变量栅格数据的分辨率及范围调整统一, 转换为 .acs 数据, 导入 MaxEnt 模型软件。设置 70% 数据点为训练集, 30% 为验证集, Bootstrap 重复计算为 10 次, 选择 Random seed, 绘制响应曲线 (response curves), 用 jackknife 功能来分析影响崩滑灾害分布的环境因子, 输出预测数据形式为 Logistic 概率栅格值, 以表示崩滑灾害在各栅格发生概率。模型生成 ROC 曲线, 以 AUC 值检验模型精度: $AUC < 0.6$, 预测精度差; $0.6 \leq AUC < 0.7$, 预测精度一般; $0.7 \leq AUC < 0.8$, 预测精度较好; $0.8 \leq AUC < 0.9$, 预测精度很好; $AUC \geq 0.9$, 预测精度极高, 效果最佳。

3 评价结果

将 MaxEnt 模型计算得到的 ASC 数据导入

ArcGIS 中并转为栅格数据, 根据自然断点法^[17], 将崩滑灾害易发性分为 5 级: 极低易发区 (< 0.18)、低易发区 ($0.18 \sim 0.31$)、中易发区 ($0.31 \sim 0.45$)、高易发区 ($0.45 \sim 0.62$) 和极高易发区 (> 0.62), 生成易发性分区图 (图 5), 对各等级面积和崩滑灾害数量进行统计得到易发性分区统计图 (图 6)。基于斜坡单元的崩滑灾害易发性评价中, 极低易发区面积占 37.44%, 灾害占比 3.4%; 低易发区占 29.08%, 灾害占比 18.37%; 中易发区占 17.67%, 灾害占比 22.45%; 高易发区占 11.26%, 灾害占比 24.49%; 极高易发区占 4.54%, 灾害占比 31.29%。

4 结果分析

为方便对比研究, 将基于斜坡单元的崩滑灾害易发性评价结果与基于传统栅格单元的易发性评价结果进行比较, 基于传统栅格单元的易发性评价结果见图 7。

4.1 因子贡献率

根据 MaxEnt 模型输出结果, 在基于斜坡单元

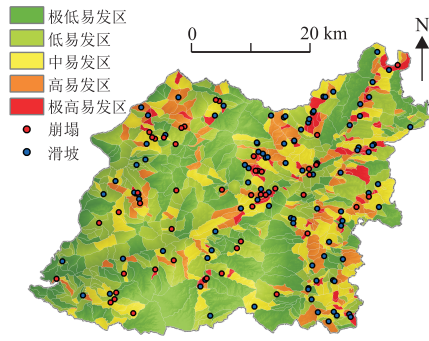


图 5 崩滑灾害易发性分区图

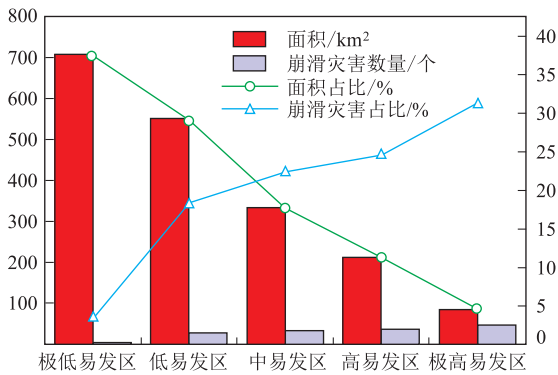


图 6 崩滑灾害易发性分区统计

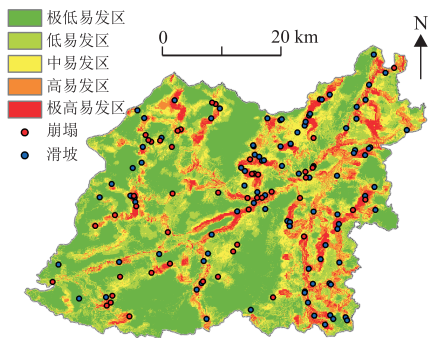


图 7 基于栅格单元的易发性评价结果

易发性评价中,距公路距离贡献率最高,为 40.4%,距断层距离为 14.7%,坡向为 11.1%,距河流距离为 8.9%,工程地质岩组为 6.8%,坡度为 6.3%,降雨量、高程、NDVI 均未超过 5%;基于栅格单元易发性评价中距公路距离贡献率为 45.9%,NDVI 为 16.3%,距断层距离为 7.6%,坡度为 6.6%,坡向为 6.1%,距河流距离为 5.9%,工程地质岩组为 5.5%,降雨量、高程贡献率均未超过 5%,见图 8。NDVI 数据因子贡献率具有较大差异,这是由于 NDVI 数据源分辨率为 1 : 500,远高于其他评价因子矢量数据,在进行分区统计时,可能造成数据波动。总的来说,两项评价结果具有一定相似性,均说明人类工程活动和地质构造是孟连县崩滑灾害的主

要诱因,表明了基于 MaxEnt 和斜坡单元的崩滑灾害易发性评价模型具有可行性。

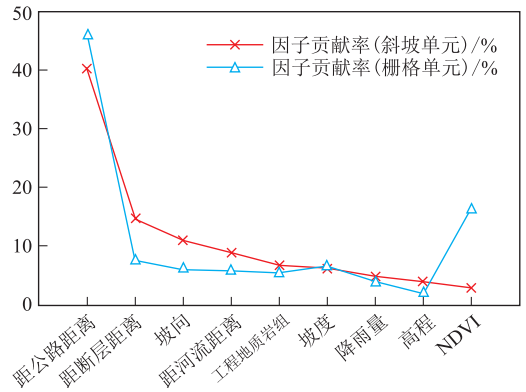


图 8 评价因子贡献率

4.2 精确度检验

利用 AUC 值评估模型预测精度,生成 ROC 曲线,红色代表训练数据平均值,蓝色代表训练数据标准差,黑色代表随机预测见图 9。基于斜坡单元的崩滑灾害易发性评价训练结果 AUC 均值为 0.858,标准差为 0.035 8;基于栅格单元的崩滑灾害易发性评价训练结果 AUC 均值为 0.872,标准差为 0.028 5。根据 AUC 值检验模型精度标准,两种训练结果为预测精度较好,均可用于崩滑灾害易发性评价,其中基于斜坡单元的训练结果 AUC 值略低于基于栅格单元的训练结果,差值为 0.014。

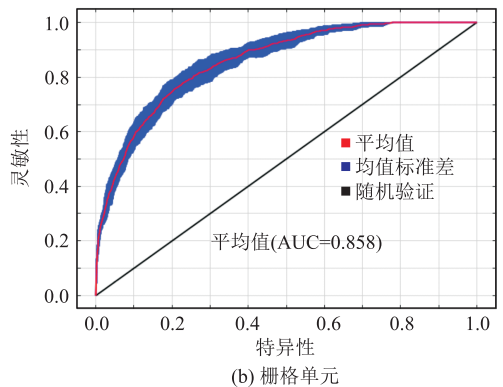
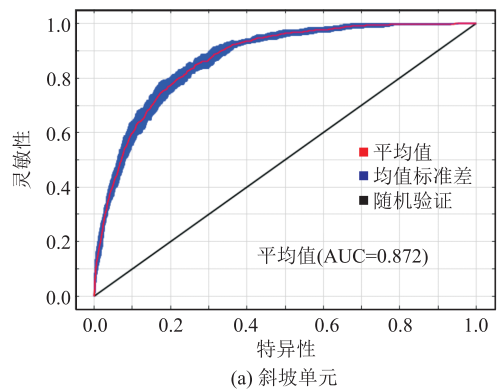


图 9 ROC 曲线

4.3 讨论

为更准确比较两者差异性,使用标准差计算 AUC 差异的 z 统计量^[18]:

$$z = \frac{AUC_1 - AUC_2}{SE(SUC_1 - SE(AUC_2))} \quad (2)$$

式中, $SE(AUC_1)$ 为 AUC_1 的标准误差。

提出零假设 $H_0: AUC_1 = AUC_2$, 即两个模型的性能相等, 标准正态分布的分位数取 1.96, 对应于显著性水平 $\alpha = 0.05$ 。计算结果为 $|z| = 0.308$, 小于 1.96, 无法拒绝零假设, 因此认定基于斜坡单元的训练结果和基于栅格单元的训练结果 AUC 值在统计学上无显著差异, 即精确度相同。

由于栅格单元是将地面数据抽象为规则的方格, 无法直观体现地形变化, 同时地形信息的离散化可能导致信息表达不完整甚至丢失。斜坡单元通常是由地形数据直接提供的, 可以更准确地反映地表的实际特征与地形变化, 且在崩滑灾害野外调查的实际工作中, 能够帮助调查人员更好地理解与辨识灾点, 提高对潜在危险的认识, 同时以斜坡单元为单位开展工作, 提高了调查任务的可行性与实用性。因此基于 MaxEnt 模型与斜坡单元的崩滑灾害易发性评价在保证了精确度与基于栅格单元易发性评价精确度相同的同时, 拥有更好的实用性, 更符合地质灾害调查与防治工作的实际意义。

5 结论

(1) 以云南省孟连县崩滑灾害作为研究对象, 结合灾害发育特征及分布规律, 选取了高程、坡度、坡向、距断层距离、距公路距离、距河流距离、降水量、工程地质岩组、归一化植被指数 (NDVI) 9 个评价因子, 将研究区划分为 800 个斜坡单元。基于 MaxEnt 模型, 对 148 个崩滑灾害点和评价因子栅格数据进行建模分析后, 得到孟连县崩滑灾害易发性分区结果并导入 ArcGIS 中。

(2) 根据自然断点法, 将研究区划分为极低易发区、低易发区、中易发区、高易发区和极高易发区, 并统计得到极低易发区面积占 37.44%, 灾害占比 3.4%; 低易发区占 29.08%, 灾害占比 18.37%; 中易发区占 17.67%, 灾害占比 22.45%; 高易发区占 11.26%, 灾害占比 24.49%; 极高易发区占 4.54%, 灾害占比 31.29%。

(3) 将基于斜坡单元的崩滑灾害易发性评价结果与基于传统栅格单元的易发性评价结果进行比较, 距公路距离、距断层距离均属于贡献率较高因子, 两项评价结果具有一定相似性, 均说明人类工程

活动和地质构造是孟连县崩滑灾害的主要诱因。

(4) 对基于斜坡单元与栅格单元的评价结果利用 AUC 值评估模型预测精度, 分别为 0.858 和 0.872, 均为预测精度较好, 可用于崩滑灾害易发性评价。使用标准差计算 AUC 差异的 z 统计量, 认定两者结果精确度相同。由于斜坡单元可以更准确地反映地表的实际特征与地形变化, 因此基于 MaxEnt 模型与斜坡单元的崩滑灾害易发性评价在保证了精确度与基于栅格单元易发性评价精确度相同的同时, 拥有更好的实用性, 更符合地质灾害调查与防治工作的实际意义。

参考文献

- [1] 黄润秋. 灾害性崩滑地质过程的全过程模拟[J]. 中国地质灾害与防治学报, 1994, 5(S1): 11-17.
- [2] NHU V, HOANG N, NGUYEN H, et al. Effectiveness assessment of Keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area[J]. CATENA, 2020, 188(7): 104458.
- [3] 金朝, 费雯丽, 丁卫, 等. 基于信息量模型和 Logistic 回归模型的地质灾害易发性评价——以十堰市郧阳区为例[J]. 资源环境与工程, 2021, 35(6): 845-850.
- [4] 付树林, 梁丽萍, 刘延国. 基于 CF-Logistic 模型的雅砻江新龙段地质灾害易发性评价[J]. 水土保持研究, 2021, 28(4): 404-410.
- [5] 许冲, 徐锡伟. 逻辑回归模型在玉树地震滑坡危险性评价中的应用与检验[J]. 工程地质学报, 2012, 20(03): 326-333.
- [6] 常晁瑜, 薄景山, 李孝波, 等. 地震黄土滑坡滑距预测的 BP 神经网络模型[J]. 地震工程学报, 2020, 42(6): 1609-1614.
- [7] 赵铮, 陈建华, 甘先霞, 等. 基于信息量法和支持向量机的芦山县滑坡危险性评价[J]. 物探化探计算技术, 2022, 44(01): 96-101.
- [8] 杨硕, 李德营, 严亮轩, 等. 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, 2021, 28(4): 131-138.
- [9] 孟凡奇, 高峰, 林波, 等. 基于 AHP 和信息量模型的地质灾害易发性评价——以鲁东片区为例[J]. 灾害学, 2023, 38(3): 111-117.
- [10] 李信, 薛桂澄, 夏南, 等. 基于 CF、CF-LR 和 CF-AHP 模型的国家热带雨林公园地质灾害易发性研究: 以海南保亭为例[J]. 现代地质, 2023, 37(4): 1033-1043.
- [11] ARABAMERI A, SAHA S, ROY J, et al. Landslide Susceptibility Evaluation and Management Using Different Machine Learning Methods in The Gallicash River Watershed, Iran[J]. Remote Sensing, 2020, 12(12): 475-504.
- [12] E. T. JAYNES. Information theory and statistical mechanics [J]. Reprinted from THE PHYSICAL PEVIEW, 1981, 106(4): 620-630.
- [13] RECOPUERTO-MEDINA L M, AGUADO A B M, BALDONADO B M M, et al. Predicting the potential

- nationwide distribution of the snail vector, *Oncomelania hupensis quadrasi*, in the Philippines using the MaxEnt algorithm. [J]. *Arasitology Research*, 2023, 123(1): 41-45.
- [14] FELICÍSIMO Á M C A. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study[J]. *Landslides*, 2013, 10(2): 175-189.
- [15] 蒋文学,李益敏,杨雪,等. 基于斜坡单元的怒江州滑坡易发性研究[J]. *水土保持学报*, 2023, 37(5): 160-167.
- [16] 张玺国,周雄冬,徐梦珍,等. 西藏地质灾害易发性及对水能开发适宜度影响[J]. *地理学报*, 2022, 77(7): 1603-1614.
- [17] WOZNIAKOWSKA P, EATON D W. Machine Learning - Based Analysis of Geological Susceptibility to Induced Seismicity in the Montney Formation, Canada[J]. *Geophysical Research Letters*, 2020, 47(22): e2020GL089651.
- [18] 周步祥,黄河,刘治凡,等. 基于假设检验的快速事件检测算法[J]. *工程科学与技术*, 2020, 52(4): 42-48.

作者简介: 汪木林(1998—),男,安徽亳州人,硕士研究生,主要研究方向为工程地质与灾害地质。E-mail:1653815234@qq.com

通讯作者: 郭婷婷(1984—),女,内蒙古呼和浩特人,博士,副教授,主要研究方向为灾害地质与水文地质。E-mail:181604785@qq.com

《地质灾害与环境保护》征稿启事

本刊主管单位:四川省教育厅;主办单位:成都理工大学、地质灾害防治与地质环境保护国家重点实验室。学术性期刊,季刊。1990年9月创刊,国内外公开发行人。本刊已加入《中国学术期刊(光盘版)》和“中国期刊网”,“万方数据网络系统(ChinaInfo)数字化期刊群”,“中文科技期刊数据库”。欢迎投稿,欢迎订阅。

征稿内容: 自然灾害(特别是地质灾害)与人为灾害防治,生态环境、资源环境和地质环境保护,水文地质与工程地质,岩土工程与工程勘察,与本学科有关的边缘和交叉学科等方面的优秀论文,最新科技成果,新理论、新方法、新技术研讨,及工程技术经验总结,新成就、新动向方面的综述、述评,新书刊、新设备评介等。

来稿注意事项:

- (1) 每篇稿件一般以中文 6 000 字为宜(包括图表等所占版面字数)。
- (2) 每篇稿件必须有:中英文篇名、摘要(摘要以第三人称书写,200 字左右)、图表名,所在单位中英文名称;第一作者简介,主要包括:姓名、性别、出生年、职务、职称,最高学历、专业、研究方向,主要科技成果及其获奖情况,以及出版专著、发表论文等简况。
- (3) 来稿请交电子文档,另可附打印件 1 份,可电子邮件投稿,地址:dzzh@cdu.edu.cn。

《地质灾害与环境保护》编辑部

2024 年 9 月 25 日