

山区公路沿线斜坡物理韧性随机森林评价模型

——以四川省茂县为例

薛蒙蒙¹, 文海家^{1,2,3}, 林渝¹, 孙德亮⁴

(1. 重庆大学 土木工程学院, 重庆 400045; 2. 山地城镇建设新技术教育部重点实验室, 重庆 400045; 3. 库区环境地质灾害防治国家地方联合工程研究中心, 重庆 400045; 4. 重庆师范大学 GIS 应用研究重庆市高校重点实验室, 重庆 401331)

摘要: [目的] 建立山区公路沿线斜坡物理韧性随机森林评价模型, 为山区斜坡抗灾工作提供科学参考。[方法] 以四川省茂县为研究区, 选取高程、坡度、坡向、坡位、微地貌、曲率、顺逆向坡、归一化植被指数、岩性、距水系距离、距断层距离、距道路距离、多年平均降雨 13 个斜坡物理韧性评价因子, 结合公路沿线 498 个历史斜坡破坏点, 构建斜坡物理韧性评价的地理空间信息数据库。将样本数据按照 7:3 的比例分为训练数据和验证数据。采用随机森林方法对训练数据进行训练建模, 将得到的模型分别对训练数据、验证数据和样本整体数据进行预测分析。采用混淆矩阵和 ROC 曲线对模型预测的准确度进行验证。[结果] 评价因子中高程、距道路距离、距水系距离这 3 个评价因子的权重较大。该模型精度较高, 混淆矩阵的精度为 98.9%, 训练数据集、验证数据集和整个研究区的 ROC 曲线下的面积(AUC)值分别为 1.000、0.870 和 0.978。模型仿真到整个研究区中, 将研究区的物理韧性划分为极低、低、中、高、极高 5 个等级。[结论] 基于随机森林方法构建的山区公路沿线斜坡物理韧性评价模型具有较高的稳定性以及可靠性。

关键词: 山区公路沿线斜坡; 物理韧性; 随机森林模型; 地理空间数据

文献标识码: B

文章编号: 1000-288X(2020)04-0168-08

中图分类号: U41, X43, X9

文献参数: 薛蒙蒙, 文海家, 林渝, 等. 山区公路沿线斜坡物理韧性随机森林评价模型[J]. 水土保持通报, 2020, 40(4):168-175. DOI:10.13961/j.cnki.stbctb.2020.04.023; Xue Mengmeng, Wen Haijia, Lin Yu, et al. Random forest evaluation model for physical toughness of slopes along mountain roads [J]. Bulletin of Soil and Water Conservation, 2020, 40(4):168-175.

Random Forest Evaluation Model for Physical Toughness of Slopes Along Mountain Roads

—Taking Maoxian County of Sichuan Province as an example

Xue Mengmeng¹, Wen Haijia^{1,2,3}, Lin Yu¹, Sun Deliang⁴

(1. School of Civil Engineering, Chongqing University, Chongqing 400045, China; 2. Key Laboratory of New Technology for Construction of Cities in Mountain Area of the Ministry of Education, Chongqing 400045, China; 3. National Joint Engineering Research Center of Geohazards Prevention in the Reservoir Area, Chongqing 400045, China; 4. The Key Laboratory of GIS Application Research, Chongqing Normal University, Chongqing 401331, China)

Abstract: [Objective] A random forest evaluation model for physical toughness of slopes along mountain roads was established in order to provide a scientific reference for disaster prevention in mountain areas. [Methods] Taking Maoxian County, Sichuan Province as the research area, this study selected 13 physical toughness assessment factors of slopes, including elevation, aspect, slope direction, slope position, micro landform, curvature, type of slope, normalized vegetation index, lithology, distance from water system, distance from fault, distance from road and annual average rainfall, and combined with 498 historical slope failure points along the highway, to construct a geospatial information database for the evaluation of slope physical toughness. The sample data was divided into training data and validation data according to the

收稿日期: 2020-03-26

修回日期: 2020-05-19

资助项目: 重庆市技术创新与应用示范一般项目(社会民生类)“山地建筑韧性评价大数据分析建模与应用示范”(cstc2018jcsx-msybX0310)

第一作者: 薛蒙蒙(1993—), 女(汉族), 河北省邯郸市人, 硕士研究生, 研究方向为地质资源与地质工程。Email: 1002547483@qq.com。

通讯作者: 文海家(1971—), 男(汉族), 湖南省宁乡县人, 博士, 教授, 主要从事岩土工程防灾减灾方面的研究。Email: jhw@cqu.edu.cn。

proportion of 7 : 3. The random forest method was used to train and model the training data, then the obtained model was used to predict and analyze the training data, the validation data and the overall sample data respectively. And confusion matrix and ROC curve were used to verify the accuracy of the model prediction. [Results] Among the evaluation factors, the weight of elevation, distance from the road, and distance from the water system was higher. The accuracy of the model was high, the accuracy of confusion matrix was 98.9%, and the AUC (area under the ROC curve) values of the training data, the validation data and the overall sample data was 1.000, 0.870 and 0.978, respectively. The model was simulated into the entire study area, and the physical toughness of the study area was divided into 5 levels as: extremely low, low, medium, high and extremely high. [Conclusions] The physical toughness evaluation model of the slope along the mountain highway based on the random forest method has high stability and reliability.

Keywords: slopes along mountain roads; physical toughness; stochastic forest model; geospatial data

中国的西南地区大部分是山地地形,交通道路等都是依山而建。道路的修建致使道路沿线斜坡的脆弱性增加,再加上气候和环境的变化,公路沿线斜坡常发生变形和破坏。斜坡的变形和破坏不仅仅取决于其脆弱性,还取决于变形和破坏时的抗灾能力和灾后的恢复能力^[1-2]。抗灾能力和灾后的恢复能力又分为物理层面和社会层面^[3],其中,物理层面的抗灾能力和恢复能力,主要表现为岩土体自身的韧性能力(即物理韧性),受气象条件、地形、植被覆盖等多个方面的影响。了解不同区域中公路沿线斜坡的物理韧性,可为边坡防治措施提供一个合理的依据。国内外关于公路沿线斜坡物理韧性的研究还比较少,且大部分的研究均针对斜坡韧性的定义。2007年,James等人^[4]在加拿大和印度部分地区为研究区,研究恢复力对山区灾害发生的影响,在面临危险时,韧性的社会—生态系统有能力学习和调整,还会与周围的其他社会—生态系统建立有利的联系;2014年,Ayyub^[5]提供了一个可以满足与可靠性和风险相关抽象概念的度量有明确关系的需求的韧性定义,为多个危险性的环境的开发的决策提供了一个依据;Molden等人^[6]综述了提高山区和山区环境的恢复力的现有研究和相关政策,以及恢复力研究的现有问题。虽然现在对斜坡物理韧性的研究还比较少,但应用地理信息技术结合数学模型的研究技术也已经成熟,很多的数学模型被应用到研究中,例如逻辑回归(Logistic)^[7]、熵指数(ED)^[8]、人工神经网络(ANN)^[9]等,这些传统的机器学习模型计算过程都比较简单,计算方法严密等优点,但是也有很多的不足之处,例如逻辑回归模型只处理线性问题,人工神经网络无法对数据量较小的样本正常运算等,而且这些模型的精度也受到限制,会出现过拟合的现象^[10]。为了避免上述问题,提出了一些新的机器学习模型,例如支持向量机(SVM)^[11]、卷积神经网络(CNN)^[12]、随机森林

(RF)^[13]等,这些模型也被陆续运用到斜坡灾害的研究中。2015年,Goetz和Brenning等人^[14]运用交叉验证的方法对支持向量机、随机森林、逻辑回归、证据加权、广义加性模型、bootstrap聚合分类树和BPLDA这7种机器学习方法对滑坡敏感性进行评估,得出结论,随机森林的预测效果相对较好,所以本次研究中选用随机森林建立山区道路沿线斜坡物理韧性评价模型。随机森林算法是一种有多棵决策树组成的集成学习模型^[15],在进行回归应用的时候,其结果取所有决策树的平均值。随机森林对数据挖掘的效果较好,可以处理一些非线性的问题和非高斯问题,且不易出现过拟合的现象,被广泛的应用于各种领域的研究中^[16]。吴润泽和胡旭东^[17]等人运用随机森林和GIS结合,选取高程、坡度、倾坡类型等9个影响因子研究三峡库区湖北段的滑坡易发性,结果表明随机森林模型预测性较好;扈秀宇和秦胜伍^[18]等人运用随机森林模型结合GIS系统,以吉林省洮南市为研究区,选用高程、坡度、坡向等10个因子为影响因子,结合历史泥石流数据,建立泥石流灾害预测模型,结果显示随机森林模型具有较好的预测性。复杂的地形地貌条件、地质构造发育、气候条件以及人类工程活动等的影 响,具有典型山地地貌的茂县是斜坡破坏至灾的多发区,且在茂县县域中,由于道路的修建,道路周围发生变形和破坏的斜坡很多,甚至一个斜坡上会发生多次破坏。因此,本文以四川省茂县公路沿线6 km范围内区域为研究区,以30 m精度栅格建立研究区地理空间数据库,基于随机森林模型,建立茂县公路沿线斜坡物理韧性评价模型。

1 研究区概况与数据来源

1.1 研究区概况

茂县位于四川省西北部(东经102°56′—104°10′,

北纬 $31^{\circ}25' - 32^{\circ}16'$), 青藏高原的东南部边缘。属龙门山断裂带地区^[19], 地跨岷江和涪江上游高山河谷地带^[20]。茂县东部为中山地带, 地势由西北向东南倾斜, 山峰海拔高度在 4 000 m 左右, 相对海拔高度为 1 500~2 500 m。岩土体类型主要由夹沙灰、黑色页岩、砂板岩及火成岩等组成^[21]。气候属于高原性季风气候, 年平均温度为 11°C , 降雨集中在夏季, 常有暴雨发生, 年平均降雨量为 486.3 mm, 植被类型主要是以针叶林为主。茂县的境内有 170 多条河流, 主要的河流有黑水河、土门河和岷江河。

茂县的地形地貌条件复杂, 各种节理和裂隙发育较多, 而且岩土体中页岩和板岩的抗风化能力较弱, 因此其斜坡的稳定性较低。而且由于夏季经常有暴雨发生, 雨水对斜坡岩土体的静水压力作用、冲刷作用、泥化作用和动水压力作用等使得斜坡的稳定性进一步的降低, 易于发生变形甚至滑动。茂县对外交通以公路为主, 国道 213 线、省道 302 线纵贯全境, 大部分道路依山而建, 道路两旁的斜坡经人工改造较多, 岩土体较为松散, 且过多的人工开挖改变了地应力的分布, 致使道路沿线斜坡破坏至灾多发。如 2014 年 7 月 17 日, 国道 213 线的 K774+600 m 处路段发生大体积斜坡破坏, 100 m 的道路被掩埋, 造成 21 人受伤, 10 人遇难。而在对该斜坡破坏点采集的过程中, 发现这些破坏点集中分布在道路沿线一定范围内。这可能是由于道路修建过程中, 沿线斜坡一定范围内的地应力发生改变。因此, 经过实际勘察和研究^[20], 选取道路沿线 6 km 范围内作为研究区。

1.2 数据来源

本研究的斜坡破坏点数据来源于四川省国土资源厅的 2019 年斜坡隐患点排查数据, 包括了斜坡破坏点的编号、类型和位置坐标; 高程图、路网、河网来源于谷歌遥感影像, 地质图由全国地质资料馆的 1:20 万地质图矢量获得, 植被数据来源于 NASA 的 MODIS 遥感成品数据, 年平均降雨数据来源于国家气象局网站。

2 地理空间数据库

2.1 影响因子

山区公路斜坡物理韧性受到多个因素的影响, 影响因子的选择是基于物理韧性机理进行选择。通过对茂县地质条件和环境条件的研究分析, 可以了解到茂县的斜坡物理韧性受到地形地貌、地质构造、环境条件以及人类工程活动的影响。经过对现有研究中的评价因子的统计, 发现地形地貌中高程、坡度、坡向这 3 个因子是经常被选用的, 因为这 3 个因子是描述

地形地貌的基本指标。除此之外, 曲率^[22-23]、微地貌和坡位^[20]也是经常考虑的地形地貌因素, 曲率展现了地表的起伏度, 微地貌和坡位表现了一个斜坡的微观地形变化。岩性也是影响斜坡物理韧性的重要因素之一, 岩土体的抗风化能力会影响到斜坡的稳定性。由于茂县位于龙门山断裂带地区, 地层活动性较强, 断层节理发育较多, 这些构造周围的斜坡的岩土体破碎程度较大, 所以斜坡的物理韧性较低。岩质边坡中顺逆向坡也是影响斜坡物理韧性的一个重要因素, 一般来说, 顺向坡的物理韧性较差, 易于发生斜坡失稳。茂县的气候条件为高原性季风气候, 夏季经常有强降雨天气, 降雨会降低的斜坡的物理韧性。而且茂县的县域中水系较多, 河流对坡岸具有冲刷作用, 降低河流沿线的斜坡稳定性。在现有的研究中, 植被条件也是一个在研究斜坡稳定时经常被考虑的因素, 一方面, 植被可以降低雨水等对斜坡的冲刷作用, 另一方面植物的根对岩土体具有锚固作用, 增强斜坡的物理韧性。归一化植被指数可以描述植被的覆盖率和生长状况, 所以一般选用它作为植被指标。综上所述, 选取了高程、坡度、坡位、坡向、曲率、微地貌、岩性、顺逆向坡、归一化植被指数、距断层距离、距水系距离、距道路距离以及多年平均降雨量这 13 个因素作为本次研究的评价因子。

2.2 数据处理

高程图是在谷歌 0.5 m 影像通过裁剪获得; 坡度、坡位、坡向、曲率、微地貌是将高程图通过 GIS 软件处理获得; 顺逆向坡是地层倾向与斜坡坡向叠加分析后按照相应的规律重分类获得; 岩性、断层均对 1:20 万地质图矢量化获得。由于高程图的栅格图层是 $30\text{ m} \times 30\text{ m}$, 所以本次研究中以 $30\text{ m} \times 30\text{ m}$ 栅格单元作为最小研究单元, 将整个研究区按照栅格单元法进行划分, 即在 ArcMap 中建立 $30\text{ m} \times 30\text{ m}$ 的渔网, 得到了 208 828 个栅格单元。为了方便后期的数据处理, 对坡向、坡位、微地貌、顺逆向坡、岩性这 5 个非量化的评价因子进行量化处理并重分类。同时对断层、道路、水系进行多级缓冲, 并采用重分类数据, 重分类标准详见表 1。其他 5 个定量评价指标直接采用原始数据。

为了降低数据离散性, 对全部 13 个因子进行归一化, 将每个因子的值进行线性变换, 使数值归一到 $[0, 1]$ 之间, 归一化的方法为:

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

式中: X^* 为归一化后的数据; X 为原始数据; X_{\min} 为数据的最小值; X_{\max} 为数据的最大值。

并在 ArcGIS 中构建研究区 13 个斜坡物理韧性评价因子专题图层(如图 1 所示)。最终得到含 13 个

斜坡物理韧性评价因子所有栅格单元的地理空间数据库。

表 1 山区公路沿线斜坡物理韧性影响因子分类

影响因子	分级数	分级标准
坡向	9	1. 平面;2. 北;3. 东北;4. 东;5. 东南;6. 南;7. 西南;8. 西;9. 西北
坡位	6	1. 山谷;2. 下坡;3. 平坡;4. 中坡;5. 上坡;6. 山脊
微地貌	10	1. 深流、峡谷;2. 中坡处浅谷、水系;3. 高地水源、水系;4. U 型山谷;5. 平原;6. 空旷斜坡;7. 台地、上斜坡;8. 局部山谷中的山脊;9. 在平原中坡处的山脊、小山;10. 山脊、山顶高处
顺逆向坡	6	1. 顺向/外倾坡;2. 斜交外倾坡;3. 斜交坡;4. 横交/切向坡;5. 逆向/反向/内倾坡;6. 平叠坡
岩性	10	1. Q_p ; 2. P_2 ; 3. S_{mx2}, T_{1b} ; 4. $S_{mx1}, S_{mx3}, S_{mx4}, D_{wg1}, D_{wg2}$; 5. $P_1, S_{mx5}, O, D_{y12}, T_{3x}$; 6. Z_{bdn}, P_{thn3}, D_3 ; 7. T_{2z}, T_{3zh} ; 8. $C+P$; 9. ϵ ; 10. $\delta_{23}, \gamma_{24}, \gamma_{51-2}, \eta_{751-2}$
距断层距离/m	11	1. <1 000;2. 1 000~2 000;3. 2 000~3 000;4. 3 000~4 000;5. 4 000~5 000;6. 5 000~6 000;7. 6 000~7 000;8. 7 000~8 000;9. 8 000~9 000;10. 9 000~10 000;11. >10 000
距道路距离/m	16	1. <200;2. 200~400;3. 400~600;4. 600~800;5. 800~1 000;6. 1 000~1 200;7. 1 200~1 400;8. 1 400~1 600;9. 1 600~1 800;10. 1 800~2 000;11. 2 000~2 200;12. 2 200~2 400;13. 2 400~2 600;14. 2 600~2 800;15. 2 800~3 000;16. >3 000
距水系距离/m	16	1. <200;2. 200~400;3. 400~600;4. 600~800;5. 800~1 000;6. 1 000~1 200;7. 1 200~1 400;8. 1 400~1 600;9. 1 600~1 800;10. 1 800~2 000;11. 2 000~2 200;12. 2 200~2 400;13. 2 400~2 600;14. 2 600~2 800;15. 2 800~3 000;16. >3 000

3 斜坡物理韧性评价

3.1 随机森林模型

随机森林方法是美国科学家 Leo Breiman^[24-26]于 2001 年提出的一种机器学习方法,它是由并行式集成学习理论的 Bagging 方法以及随机子空间方法相结合而形成。随机森林包含了多个由 bagging 集成学习技术训练得到的决策树,是一种基于决策树基本分类器的一个集成学习模型^[27]。随机森林模型各个回归树之间的相关性较低,对噪声数据和异常性的容忍度较高,具有较好的准确性^[28]。其主要思想是通过自助法从原始数据集中随机且有放回地抽取 K 个训练集,形成 K 棵决策树;每棵决策树内部的节点从 N 个韧性评价因子中随机选取 n 个因子作为分裂标准($n \leq N$);每棵树都不做剪枝处理,任其自由生长;最后,使 K 棵决策树组成随机森林,对 K 个结果采用投票的方式得到最终分类结果^[29],计算公式为:

$$p(c|v) = \sum_{k=1}^K p_k(c|v) \quad (2)$$

式中: K 为 RF 中决策树的数目; c 为斜坡破坏与否($c=0,1,0$ 表示斜坡发生破坏,1 表示斜坡没有发生破坏); P 为概率函数。

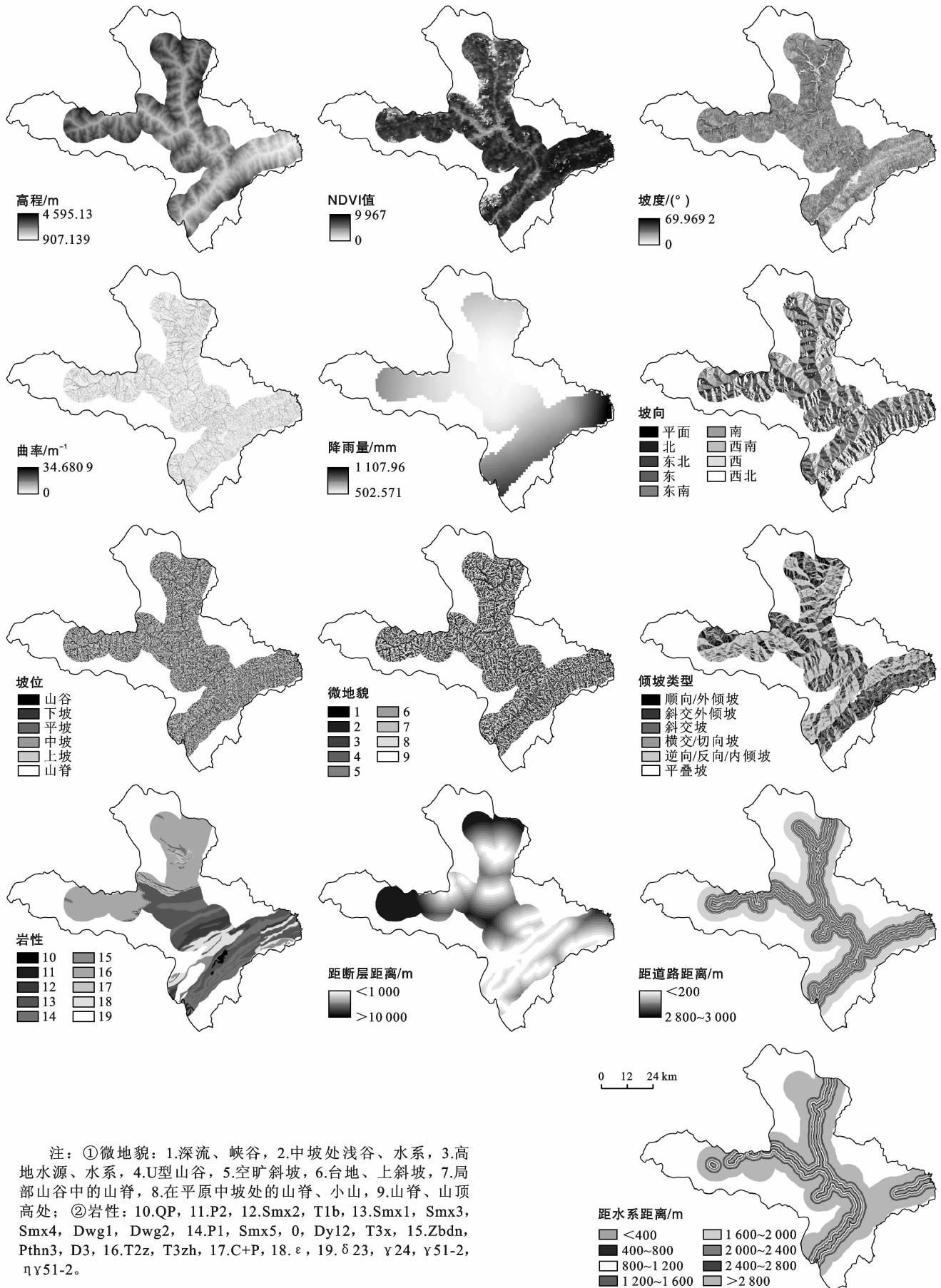
3.2 模型构建

通过对斜坡破坏灾害的数据统计,共有 498 栅格单元发生脆性破坏,将其记为 0。在实际的情况下,

斜坡破坏灾害并不是以点的形式存在,而是发生在一定范围内。在 ArcGIS 中对历史破坏点进行缓冲,缓冲的范围选为 600 m,缓冲范围以外的区域未受到地灾的影响发生变形破坏,记为 1。随机选取 2 000 个未受影响点和 200 个破坏点作为一组样本,总共选取 5 组训练样本。将训练样本通过 R 语言进行随机森林训练,在建立随机森林模型之前,将样本数据按照 7:3 的比例分为训练样本数据和验证样本数据。在运用训练样本数据建立随机森林模型,获得斜坡物理韧性预测的训练模型。再对验证样本数据进行预测,选择预测效果最好的一组作为最终训练模型。

3.3 评价因子权重

在随机森林模型中对 13 个评价因子的重要性进行排序,得到的结果如图 2 所示。%IncMSE 是指在对一个指标进行随机赋值,如果这个指标重要的话,那么预测时误差也会增大。IncNodePurity 中的 Node Purity 是 RSS(残差平方和)的减少,Node Purity 的增减也就是基尼指数的减少。%IncMSE 指标中高程、距道路距离、距水系距离、NDVI 以及降雨这 5 个评价因子的作用比重较大,而 IncNodePurity 中距道路距离、距水系距离、高程、距断层距离和坡向这 5 个评价因子的作用比重较大。综合上述两个指标的排序可知高程、距道路距离、距水系距离这 3 个评价因子对斜坡物理韧性的影响最大。



注：①微地貌：1.深流、峡谷，2.中坡处浅谷、水系，3.高地水源、水系，4.U型山谷，5.空旷斜坡，6.台地、上斜坡，7.局部山谷中的山脊，8.在平原中坡处的山脊、小山，9.山脊、山顶高处；②岩性：10.QP，11.P2，12.Smx2，T1b，13.Smx1，Smx3，Smx4，Dwg1，Dwg2，14.P1，Smx5，0，Dy12，T3x，15.Zbdn，Pthn3，D3，16.T2z，T3zh，17.C+P，18.ε，19.δ23，γ24，γ51-2，ηγ51-2。

图 1 茂县公路沿线斜坡物理韧性评价因子图层

3.4 模型验证

为检测模型的有效性,常采用混淆矩阵来分析模型的预测精度。即将无破坏看作正类(positive)、有破坏视为负类(negative),实例是正类且被预测为正类,称为真正类(true positive, TP);实例是负类但被预测为正类,称之为假正类(false positive, FP),如果实例是负类且被预测为负类,称为真负类(true negative, TN);实例是正类但被预测成负类,则称之为假负类(false negative, FN)。如表 2 所示,1 代表正类,无破坏发生;0 代表负类,有破坏发生。将本文的训练样本数据代入随机森林模型中,计算得到样本数据的抵抗灾变破坏的能力 P 。以 0.5 为阈值, P 大于 0.5,表示抵抗灾变破坏能力大,预测结果为无灾变; P 小于 0.5,表示抵抗灾变破坏能力小,预测结果有灾变。从表 2 中可以看出,训练样本的总体预测精度达到 98.9%,召回率和精密度分别为 1.000 和 0.988,效果较好。

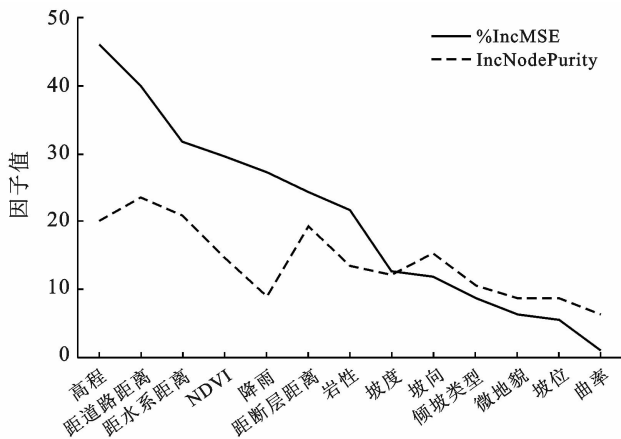


图 2 各评价因子的重要性排序

表 2 训练样本观测值与预测值的对比

项目	预测		召回率	精密度	准确度
	1(无破坏)	0(破坏)			
实际 1(无破坏)	1 400(TP)	0(FN)	1	0.988	0.989
实际 0(破坏)	17(FP)	123(TN)			

除此之外,也可用 ROC 曲线对模型的准确度进行研究。ROC 曲线下的面积(AUC),可以定量的表示模型预测的准确率^[30]。将训练数据、验证数据和整个研究区的数据进行 ROC 曲线分析,分析结果显示,训练数据集、验证数据集和研究区全体数据集仿真结果的 AUC 值分别为 1.000,0.870 和 0.978,验证数据集和其他两个数据集差异相差不大,且 3 个 AUC 值都较高,为 1 或接近于 1,说明该模型具有良好的可靠性。

3.5 模型应用

将训练模型对验证数据进行训练,并仿真到整个研究区中,得到整个研究区的斜坡物理韧性概率图。为了方便研究将得到的概率图分为 5 个等级,经过反复的试算,划分的标准为极低($p < 0.85$)、低($0.85 < p < 0.90$)、中($0.90 < p < 0.98$)、高($0.98 < p < 0.99$)、极高($p > 0.99$)。得到研究区的斜坡物理韧性等级图(图 3)。通过图 3 可以看出大部分的破坏点分布在斜坡物理韧性较低和低的地方,而韧性高的地方发生的破坏点较少,这个结果是和理论分析一致的。

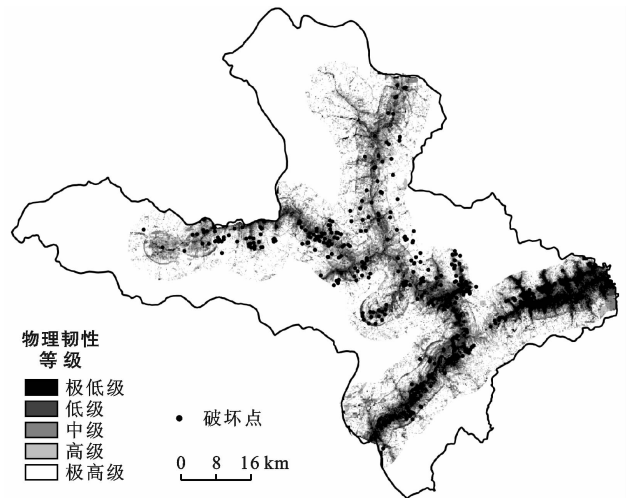


图 3 茂县公路沿线斜坡物理韧性等级

对得到的茂县公路沿线斜坡物理韧性等级图进行进一步的定量的研究。统计每个物理韧性等级的面积占比以及其中的斜坡破坏点和破坏点占比,再根据破坏点占比和面积占比计算滑坡密度(表 3)。通过表 3 可以看出,随着物理韧性等级的增加,其面积占比也在随之增加,而其区域的破坏点占比则随之减少,滑坡密度也是随之减少的。物理韧性极低和低等级区域的占比比较少,只有 19.57%,但是其中的斜坡破坏点却占到了占到了整个研究区的 49.8%,斜坡破坏点的密度也较高。这表示训练模型的仿真程度较好。

表 3 茂县公路沿线斜坡物理韧性分级统计

韧性等级	栅格单元/个	比例/%	破坏点/个	破坏点比例/%	滑坡密度
极低	17 633	8.44	138	27.71	3.28
低	23 236	11.13	110	22.09	1.99
中	35 278	16.89	94	18.88	1.12
高	39 089	18.72	79	15.86	0.85
极高	93 592	44.82	77	15.46	0.34

4 结论

(1) 本研究对山区公路沿线斜坡物理韧性进行了研究,确认地形地貌、地质环境、降雨等因素会影响斜坡物理韧性。因此选取高程、坡度、坡位、坡向、曲率、微地貌、岩性、顺逆向坡、归一化植被指数、距断层距离、距水系距离、距道路距离以及多年平均降雨量这 13 个因子作为斜坡物理韧性的评价因子,构建地理空间数据库。

(2) 山区公路沿线斜坡物理韧性受多种因素耦合影响,简单的线性模型的预测准确度较低,所以选择随机森林模型。将随机森林模型结合地理空间数据库和历史斜坡破坏点进行建模,建立山区道路沿线斜坡物理韧性评价模型,并对评价因子进行排序,高程、距道路距离、距水系距离这 3 个评价因子的权重较大。运用混淆矩阵和 ROC 曲线验证建立的物理韧性评价模型的准确性。其混淆矩阵的召回率和精密度分别为 1.000 和 0.988,精度达到 98.9%;训练数据集、验证数据集以及区域仿真的 ROC 曲线 AUC 值分别为 1.000,0.870 和 0.978 模型具有较高的预测性。

(3) 将模型仿真到真个研究区域中,得到山区道路斜坡物理韧性评价图。运用统计分析验证仿真的精确度,极低和低等级区域的占比比较少,只有 19.57%,但是其中的历史灾变点却占到了整个研究区的 49.8%,历史斜坡破坏点主要分布在物理韧性低和较低的区域中。综合表明该模型具有较高的可靠性与稳定性。

[参 考 文 献]

- [1] Manuel Herrera1, Edo Abraham, Ivan Stoianov. A graph-theoretic framework for assessing the resilience of sectorised water distribution networks [J]. *Water Resources Management*, 2016,30(5):1685-1699.
- [2] 孙德亮,吴健平,文海家,等.山地斜坡灾变韧性大数据分析模型与应用研究:以城口县为例[J].*重庆师范大学学报(自然科学版)*,2019,36(3):64-71.
- [3] Zhang Xiwen, Song Jing, Peng Jian, et al. Landslides-oriented urban disaster resilience assessment: A case study in Shenzhen, China [J]. *The Science of the Total Environment*, 2019,661(APR. 15):95-106.
- [4] James S, Gardner Julie Dekens. Mountain hazards and the resilience of social-ecological systems: Lessons learned in India and Canada [J]. *Natural Hazards*, 2007,41(2):317-336.
- [5] Ayyub M Bilal. Systems resilience for multihazard environments: Definition, metrics, and valuation for decision making [J]. *Risk Analysis*, 2014,34(2):340-355.
- [6] Molden D, Hurni H, Von Dach S W, et al. Focus issue: Vulnerability and resilience to natural hazards and disasters in mountains [J]. *Mountain Research and Development*, 2015,35(2):103-103.
- [7] Yang Jintao, Song Chao, Yang Yang, et al. New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector: A case study of Duwen Highway Basin, Sichuan Province, China [J]. *Geomorphology*, 2019,324:62-71.
- [8] Subrata Mondal, Sujit Mandal. Landslide susceptibility mapping of Darjeeling Himalaya, India using index of entropy (IOE)model [J]. *Applied Geomatics*, 2019,11(2):129-146.
- [9] 朱崇浩,张建经,马东华,等.基于 DInSAR-BP 神经网络的震后区域滑坡危险性综合评价研究[J].*工程地质学报*,2020,28(3):530-540.
- [10] 刘坚,李树林,陈涛.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(7):1085-1091.
- [11] Romulus C, Mihnea C P, Dieu T B, et al. Spatial predicting of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning [J]. *Journal of Hydrology*, 2020,585(6):69-79.
- [12] Fang Zhice, Wang Yi, Peng Ling, et al. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping [J]. *Computers and Geosciences*, 2020, 139(6): 104470.
- [13] Khaled T, Tao Cheng, Yang Zhang. Mapping landslide susceptibility and types using random forest [J]. Taylor & Francis, 2018,2(2):159-178.
- [14] Goetz J N, Brenning A, Petschko H, et al. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling [J]. *Computers and Geosciences*, 2015,81(8):1-11.
- [15] 李亭,田原,邬伦,等.基于随机森林方法的滑坡灾害危险性区划[J].*地理与地理信息科学*,2014,30(6):25-30,2.
- [16] 赖成光,陈晓宏,赵仕威,等.基于随机森林的洪灾风险评价模型及其应用[J].*水利学报*,2015,46(1):58-66.
- [17] 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/OL].*地球科学*:1-16[2020-05-05]. <http://kns.cnki.net/kcms/detail/42.1874.P.20200313.1618.002.html>.
- [18] 扈秀宇,秦胜伍,窦强,等.基于 GIS 和随机森林模型的

- 泥石流敏感性分析:以吉林省洮南市北部山区为例[J].水土保持通报,2019,39(5):204-210,217,2.
- [19] 廖炳勇,何晓飞,曾强,等.龙门山构造带茂汶断裂在茂县、汶川一带构造特征[J].四川地质学报,2019,39(3):374-378.
- [20] 文海家,李洋,薛靖元,等.基于大数据挖掘的山区公路沿线滑坡易发性小区划[J].自然灾害学报,2018,27(4):159-165.
- [21] 袁宏.茂县地质灾害的主要成因和预防措施[J].能源与环境,2019(6):108-109.
- [22] Norbazlan M Y, Biswajeet P, Helmi Z M S, et al. Spatial landslide hazard assessment along the Jelapang corridor of the north-south expressway in Malaysia using high resolution airborne LiDAR data [J]. Arabian Journal of Geosciences, 2015,8(11): 9789-9800.
- [23] 杨城,林广发,张明锋,等.基于DEM的福建省土质滑坡敏感性评价[J].地球信息科学学报,2016,18(12):1624-1633.
- [24] Breiman L. Bagging predictors [J]. Machine Learning, 1996,24(2):123-140.
- [25] Ho T. The random subspace method for constructing decision forests [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998,20(8):832-844.
- [26] Breiman L. Random forests [J]. Machine Learning, 2001,45(1):5-32.
- [27] 董师师,黄哲学.随机森林理论浅析[J].集成技术,2013,2(1):1-7.
- [28] 王仁超,朱品光.基于随机森林回归方法的爆破块度预测模型研究[J].水力发电学报,2020,39(1):89-101.
- [29] 赖成光,陈晓宏,赵仕威,等.基于随机森林的洪灾风险评估模型及其应用[J].水利学报,2015,46(1):58-66.
- [30] YiLmaz I. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine [J]. Environmental Earth Sciences, 2010,61(4):821-836.

(上接第167页)

- [19] 鲍士旦.土壤农业化学分析方法[M].北京:中国农业科技出版社,2000.
- [20] 王文锋,李春花,黄绍文,等.不同施肥模式对设施菜田土壤微生物量碳、氮的影响[J].植物营养与肥料学报,2016,22(5):1286-1297.
- [21] 符鲜,杨树青,刘德平,等.施氮水平对河套灌区套作小麦—玉米土壤微生物量碳、氮的影响研究[J].生态环境学报,2018,27(9):1652-1657.
- [22] Mandal A, Patra A K, Singh D, et al. Effect of long-term application of manure and fertilizer on biological and biochemical activities in soil during crop development stages [J]. Bioresource Technology, 2007, 98(18):3585-3592.
- [23] Lovell R D, Jarvis S C, Bardgett R D. Soil microbial biomass and activity in long term grassland: Effects of management changes [J]. Soil Biology and Biochemistry, 1995,27(7):969-975.
- [24] 陶朋闯,陈效民,靳泽文,等.生物质炭与氮肥配施对旱地红壤微生物量碳、氮和碳氮比的影响[J].水土保持学报,2016,30(1):231-235.
- [25] Tian Yongqiang, Zhang Xueyan, Liu Jun, et al. Effects of summer cover crop and residue management on cucumber growth in intensive Chinese production systems: Soil nutrients, microbial properties and nematodes [J]. Plant and Soil, 2011,339(1/2):299-315.