

旋转森林模型在滑坡易发性评价中的应用研究

刘渊博¹ 牛瑞卿¹ 于宪煜² 张凯翔³

1 中国地质大学地球物理与空间信息学院,湖北 武汉,430074
2 湖北工业大学土木建筑与环境学院,湖北 武汉,430068
3 中国地质大学信息工程学院,湖北 武汉,430074

摘 要:以三峡库区万州段为研究区,从多源空间数据中提取 29 个致灾因子作为区域滑坡易发性分析的评价指标,在数字高程模型基础上采用集水区重叠法划分斜坡单元,构建旋转森林集成学习模型,定量预测滑坡空间易发性,并生成滑坡易发性分区图。在易发性分区图中,高易发区占 11.6%,主要分布在万州主城区和长江及支流两岸;不易发区占 45.6%,主要分布在人类工程活动低、植被覆盖度高的区域。采用受访者工作特征曲线和曲线下面积对旋转森林模型的滑坡易发性进行评价,结果显示该模型的预测精度为 90.7%,其预测能力优于 C4.5 决策树。研究表明,应用旋转森林进行滑坡易发性评价具有预测能力强、精度高等优点。

关键词:旋转森林;易发性评价;集成学习;滑坡

中图分类号:P208 **文献标志码:**A

滑坡灾害是仅次于地震的第二大地质灾害^[1-2],科学准确的滑坡易发性评价能为减灾防灾提供依据。目前,国内外学者利用多种机器学习模型实现了滑坡易发性评价。Dou 等^[3]和谭龙等^[4]利用人工神经网络、Tehrany 等^[5]和牛瑞卿等^[6]利用支持向量机、Tsangaratos 等^[7]和 Hong 等^[8]利用决策树分别对滑坡易发性进行了定量预测。但这些预测模型均基于单分类器,而在区域滑坡空间预测过程中,涉及到的评价指标包括地形、地质、水文和人类工程活动等多种空间数据,这些数据具有种类多样、结构复杂、数据量大等特点,在构建预测模型时所采集的样本数据集中具有一定的干扰信息,而单分类器模型往往不能达到理想的预测结果,因此,通过组合多个单分类器来提高正确率和泛化能力的集成学习成为滑坡灾害研究的重要方向之一^[9]。

本文介绍的旋转森林算法(rotation forest, ROF)是一种将 C4.5 决策树作为基分类器的集成分类算法,对样本数据集随机分割,采用特征变换有效地增大基分类器间的差异性,从而得到更好的集成效果。它已经被成功应用于遥感影像分类^[10]、医疗诊断^[11]、基因数据分类^[12]、故障诊

断^[13]等多个研究领域,但目前国内外还未将旋转森林应用于滑坡易发性评价中。研究表明,该方法对噪声敏感度低,具有较好的准确率和稳定性^[14],这些特性对于样本数据通常存在干扰信息的滑坡灾害空间预测研究具有很强的针对性。因此,本文以长江三峡库区万州段为研究区,利用 ROF 进行滑坡易发性评价。

1 ROF 模型

ROF 是通过对特征集进行随机分割,然后利用主成分分析(principal components analysis, PCA)对每个特征子集进行旋转变换,并保留所有的主成分进行训练,通过增加分类器之间数据集的差异性来保证基分类器的准确性^[15]。而决策树对特征轴的旋转比较敏感,所以,常选择决策树作为基分类器。

令 $\mathbf{x}=[x_1\ x_2\ \cdots\ x_n]^T$ 为一个具有 n 个属性的样本点, \mathbf{X} 为一个包含 N 个训练样本的训练集,构成 $N\times n$ 的矩阵,同时令 $\mathbf{Y}=[y_1\ y_2\ \cdots\ y_N]$ 是训练集 \mathbf{X} 对应的类别,并且 y_i 属于类别集合 $\{w_1, w_2, \cdots, w_c\}$ 。 D_1, D_2, \cdots, D_L 为选择的 L 个基分

类器, \mathbf{F} 为完整的特征集。

分类过程包括 4 个步骤:①随机将特征集 \mathbf{F} 划分成 K 个不相交的子集, 每个子集将包含 M 个特征;②对于每一个特征子集进行 PCA 后得到

$$\mathbf{R}_i = \begin{bmatrix} a_{i,1}^{(1)}, a_{i,1}^{(2)} \cdots a_{i,1}^{(M_1)} & [0] & \cdots & [0] \\ [0] & a_{i,2}^{(1)}, a_{i,2}^{(2)} \cdots a_{i,2}^{(M_2)} & \cdots & [0] \\ \vdots & \vdots & \ddots & \vdots \\ [0] & [0] & \cdots & a_{i,k}^{(1)}, a_{i,k}^{(2)} \cdots a_{i,k}^{(M_k)} \end{bmatrix} \quad (1)$$

式中, $a_{i,k}$ 为 PCA 变换后的主成分, $i=1,2\cdots L$ 。按照原始特征集的顺序重新排列 \mathbf{R}_i , 形成一个 $N \times n$ 的旋转矩阵 \mathbf{R}_i^a , 则 D_i 分类器所对应的训练集为 $\mathbf{X}\mathbf{R}_i^a$;④在分类时, 对于每一个分类样本 \mathbf{x} , 先经过 $\mathbf{x}\mathbf{R}_i^a$ 变换, 通过分类器 D_i 计算 \mathbf{x} 属于每一个 ω_j 的概率, 然后计算所有分类器的平均值:

$$u_j(\mathbf{x}) = \frac{1}{L} \sum_{i=1}^L d_{i,j}(\mathbf{x}\mathbf{R}_i^a) \quad (2)$$

则 \mathbf{x} 属于概率最大的一类。式(2)中, $u_j(\mathbf{x})$ 为 \mathbf{x} 属于 ω_j 的平均概率; $d_{i,j}(\mathbf{x}\mathbf{R}_i^a)$ 为 D_i 分类器将 \mathbf{x} 归为 ω_j 的概率。

2 研究区概况及数据源

2.1 研究区概况

研究区位于长江三峡库区万州区境内, 地理坐标为 $108^{\circ}11' \sim 108^{\circ}33' \text{E}$ 、 $30^{\circ}24' \sim 30^{\circ}56' \text{N}$, 面积约为 483 km^2 , 三峡库区及各区县边界数据来源为三峡库区地质灾害防治工作指挥部地理信息系统矢量数据, 图 1 为研究区地理位置与滑坡分布示意图。该区位于长江上游地区、重庆东北部, 地处三峡库区腹心, 境内山丘起伏, 丘陵、低山、低中山和山间平地面积约占 $1/2$, 少平坝和台地, 且零星散布, 相对高程达 $106 \sim 1\,762 \text{ m}$ 。气候属亚热带季风湿润带, 具有四季分明, 日照充足, 雨量充沛, 霜雪稀少等特点。根据三峡地质灾害防治工作指挥部 2010 年地质灾害普查资料显示: 万州区内发育的地质灾害约 1 200 处, 预估计威胁居民 8 万多人, 直接经济损失达数百万元。

2.2 工程地质条件

研究区地层划分如图 2 所示, 区内出露的地层总体上从长江两岸往南北由新到老, 多见于中生界侏罗系和三叠系。而岩性多为泥岩或泥岩砂岩相间, 抗剪强度较低, 在水的侵蚀作用下容易产生滑坡。

2.3 数据源

本文采用的主要数据源包括:①Landsat OLI

M' 个主成分, 每个都是 $M \times 1$ 的向量;③对每一个特征子集都进行步骤②的操作, 将得到的所有的主成分系数存入一个系数矩阵 \mathbf{R}_i :

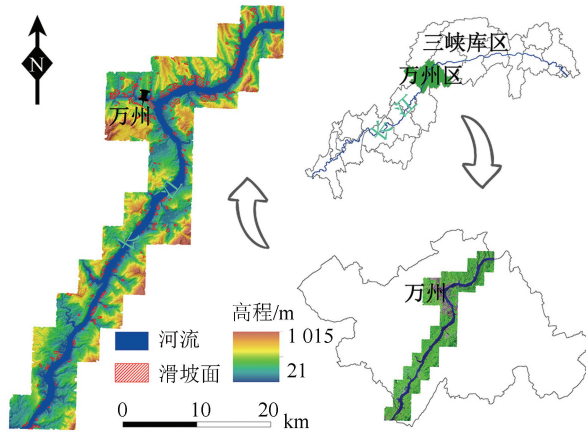


图 1 研究区位置示意图
Fig.1 Location of the Study Area

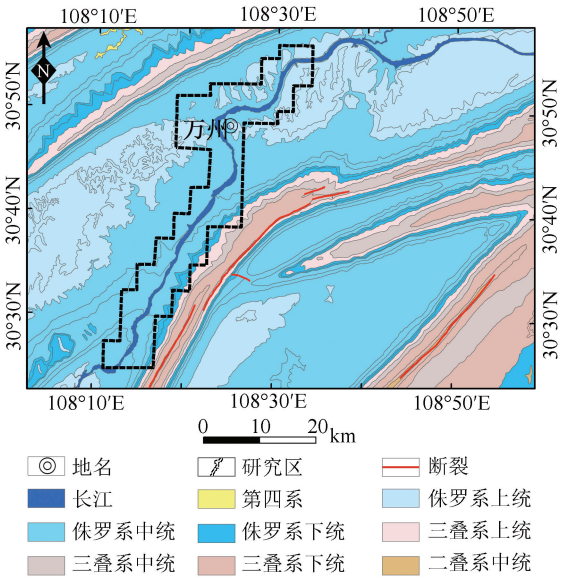


图 2 研究区地层划分
Fig.2 Regional Geological Map of the Study Area

卫星影像数据一景(时间为 2013-02-24, 轨道号为 127/039), 用于提取地表覆盖和人类工程活动等信息;②1:5 万地形图, 用于提取地形地貌、水文条件等相关信息;③1:5 万地质图, 用于提取工程岩组等信息;④该区包括灾害名称、面积、体积等属性的滑坡空间分布数据, 可以从三峡库区地质灾害防治工作指挥部 1:10 000 灾害地质图数

数据库中获取。研究区内共有滑坡 233 处,总面积约 24.06 km²,约占整个研究区总面积的 4.9%。其中,最小滑坡面积约 3 542 m²,最大滑坡面积约 1.1 km²,典型的有驸马滑坡、塘角村滑坡等,其全貌高分一号影像如图 3 所示。

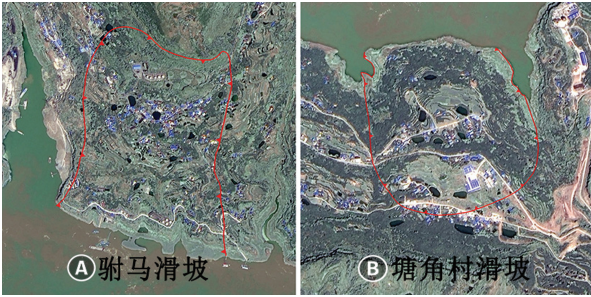


图 3 典型滑坡遥感影像

Fig.3 Remote Sensing Image of Typical Landslide

3 ROF 模型应用的研究方法与步骤

针对滑坡空间预测问题,首先在收集研究区数据资料的基础上选择合理的评价指标;其次针对预测模型划分适当的评价单元;然后建立 ROF 模型,计算滑坡易发性指数,并绘制滑坡易发性分区图;最后对预测模型及预测精度进行检验。

3.1 评价因子

影响滑坡发生的因素包括控制因素和诱发因素。前者指对滑坡发生起控制作用的地质和地形地貌因素等,该类因素在短时期内是基本稳定的,如地层岩性、地质构造、地形等;后者指对滑坡发生起触发作用的因素,其作用会加速滑坡灾害发生的时间,如降雨、人类工程活动等。本文根据研究区的具体特点和前人研究成果^[16-17],选取地形地貌类因子:高程、坡度、坡向、坡高、斜坡形态、地形表面纹理、地形粗糙指数、地形位置指数、地形表面凸率、地形汇聚指数、相对坡位、平面曲率、剖面曲率、地形曲率;基础地质类因子:工程岩组、改进的太阳辐射地形因子计算模型指数^[18];水文条件类因子:水系距离、流域面积、流路长度、地形湿度指数、河谷深度、流域坡度、河流强度指数、距河网垂直距离;地表覆盖类因子:归一化植被指数、归一化湿度指数;诱发因素:降雨、地震、归一化建筑指数。而对于工程岩组和斜坡形态这种名义尺度的变量通常根据对滑坡的危险性的“贡献”程度转换为有序尺度的变量。本文采用专家打分法^[19]并结合统计分析结果进行量化取值(见表 1),其他因子使用实际值。

表 1 评价因子量化

Tab.1 Evaluation Factor Quantization

因子	评价标准	危险性	评分值
工程岩组	硬岩	不危险	1
	软硬相间	较重危险	2
	软岩	重度危险	3
斜坡形态	GE/GR	不危险	1
	GE/V	轻度危险	2
	GE/X	轻度危险	3
	V/GR	轻度危险	4
	X/GR	轻度危险	5
	X/V	重度危险	6
	V/X	重度危险	7
	X/X	极度危险	8
	V/V	极度危险	9

3.2 ROF 模型滑坡易发性评价方法

基于 ROF 模型滑坡易发性评价的具体步骤如下。

1)划分评价单元。斜坡单元是滑坡、崩塌等地质灾害发育的基本单元,并且在致灾因素中,河流和沟谷的发育阶段对滑坡的形成具有明显的控制作用,因此采用基于幼年期沟谷划分的斜坡单元作为评价单元,可以与地质环境条件紧密联系,综合体现致灾因素的作用,使评价结果更贴近于实际^[20]。本文采用斜坡单元作为评价单元,其划分方法为在数字高程模型基础上采用“集水区重叠法”。首先对正、反数字高程数据进行洼地填充;然后分别对河流流向、累积流量、河流网络进行计算、并绘制集水区;最后通过正反集水区重叠切割以及人工编修生成斜坡单元。通过此方法将研究区划分为 1 909 个评价单元。

2)构建特征集。由 29 个评价因子作为条件属性和是否滑坡作为决策属性(1 代表滑坡或易发,0 代表非滑坡或不易发)形成一张二维表,每一行描述一个评价单元,每一列描述对应评价单元的某中属性,因此,该二维表包含了 1 909 行和 30 列。

3)选择模型训练样本。为消除无量纲影响,将所有评价因子归一化到 0~1 之间。用全部的滑坡数据和相同数量的非滑坡数据随机选取 80%形成训练样本集,20%作为验证样本集。

4)建立 ROF 模型。对同一数据集, M 的取值与分类器的精度并不存在单调对应关系,其精度也会随着 M 取值的不同而有所变化。但是通过实验发现,当 $M=1$ 或者 n ,所得到的效果最差,其他取值时变化不大,因为 $M=1$ 时,经过旋转变换和重组后,所有基分类器数据集都是相同的,不属于集成分类器;当 $M=n$ 时没有对属性集进行分割。而对于分类器的个数,当 $L=10$ 时

即可达到很好的分类效果^[21]。本文选择 $M=3$, $L=10$ 进行旋转森林模型训练。

5)滑坡易发性评价。利用 ROF 模型对研究区内所有评价单元进行计算,得到每个评价单元属于各分类的概率值,利用属于滑坡类的概率值生成滑坡易发性指数图,根据易发性指数生成易发性分区图。

4 结果分析

通过 ROF 模型计算得到的滑坡易发性指数值为 0~1,对应滑坡易发性从低到高。为了区分滑坡的易发程度,根据分布直方图的分布特点,采用自然间断点法将 ROF 预测的滑坡易发性指数划分为 $[0, 0.22)$ 、 $[0.22, 0.49)$ 、 $[0.49, 0.76)$ 、 $[0.76, 1.0)$,分别对应为不易发区、低易发区、中易发区和高易发区,结果见图 4(a)。图中显示高易发区主要分布在苎溪河两岸、大周镇到万州城区长江左岸、太龙镇长江右岸和万州主城区;不易发区主要分布在武陵镇长江左岸以及远离长江两岸、人类工程活动低、植被覆盖度高的区域。

而为了分析 ROF 的预测结果,利用基分类器 C4.5 决策树同样进行了预测,其计算得到的滑坡易发性指数采用相同的方法划分出易发性分区图,结果见图 4(b)。

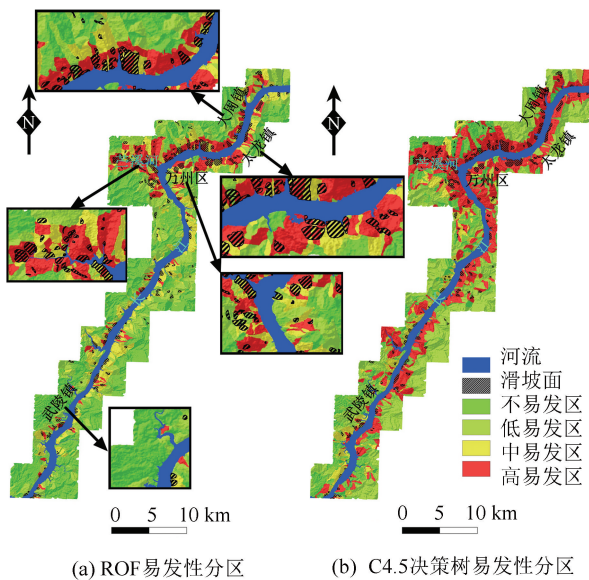


图 4 滑坡易发性分区图

Fig.4 Landslide Susceptibility Zoning Map

为了比较 ROF 和 C4.5 决策树的预测结果,通过计算两种模型易发性分区图的每个类别中存在滑坡的个数与每个类别个数的比值,即特定类别精度^[22],公式如下:

$$P_j = \frac{A_j}{B_j} \times 100\% \tag{3}$$

式中, $j=1,2\cdots M$ (M 为滑坡易发性分人类别的总数); A_j 和 B_j 分别为第 j 类滑坡易发性分区中存在滑坡的斜坡单元个数和斜坡单元总数。两种模型的特定类别精度如图 5 所示。由图 5 可见,ROF 模型在高易发区的特定类别精度远远高于 C4.5 模型,由此说明,ROF 模型预测的滑坡易发性分区中高易发区包含了更多先前调查的滑坡。

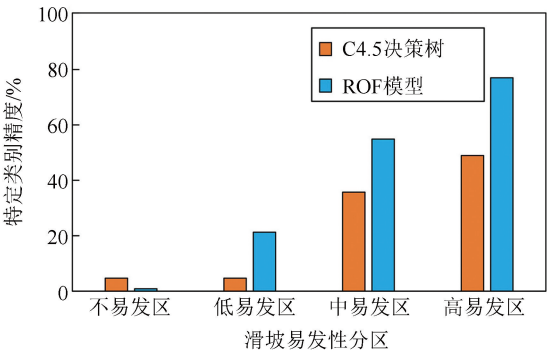


图 5 不同预测模型的特定类别精度

Fig.5 The Class-Specific Accuracies by Different Prediction Models

为进一步比较 ROF 模型和 C4.5 决策树的预测能力,本文采用了受访者工作特征(receiver operator characteristic, ROC)曲线和曲线下的面积(area under curve, AUC)对其进行精度评价^[23]。ROC 曲线以只有正类(滑坡)和负类(非滑坡)的二分类方式为例,以敏感度(滑坡而被预测为滑坡)为纵坐标,以 1-特异性(不是滑坡而被预测为滑坡)为横坐标绘制的曲线。ROC 曲线越靠近左上角即 AUC 越大,说明预测精度越高。两种模型预测结果的 ROC 曲线和 AUC 值如图 6 所示,由图 6 可见,ROF 模型对滑坡预测结果的 ROC 曲线更靠近左上角,其 AUC 值为 0.907,即预测精度为 90.7%,高于 C4.5 决策树模型的 79.8%。通过检验 20% 的样本集得到了 ROF 模型、C4.5 决策树模型的准确率分别为 84.3% 和 72.3%。因此,ROF 模型的预测能力优于 C4.5 决策树,并且其预测结果可靠。

5 结 语

本文在 C4.5 决策树的基础上构建了 ROF 模型,定量预测了滑坡易发性,生成滑坡易发性分区图,并采用 ROC 曲线和 AUC 评价了预测结果。研究表明,ROF 模型将多个 C4.5 决策树分类器集成起来,并将每个分类器的分类结果加权平均

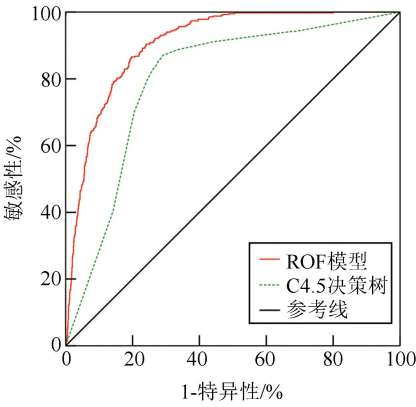


图 6 受访者工作特征曲线

Fig.6 Receiver Operator Characteristic Curve

决定最终的分类,相比 C4.5 决策树单分类器模型,ROF 模型具有更好的精度和可靠性,同时也具有更好的分类能力,在一定程度上解决了滑坡易发性分析中评价因子具有干扰信息造成精度不高的问题。因此,将 ROF 应用在滑坡易发性评价中具有广泛的应用前景。

参 考 文 献

[1] Huang Runqiu, Xu Qiang. China Typical Catastrophic Landslide [M]. Beijing: Science Press, 2008(黄润秋,许强. 中国典型灾难性滑坡[M]. 北京: 科学出版社, 2008)

[2] Yin Kunlong, Zhang Guirong, Chen Lixia, et al. Landslide Risk Assessment [M]. Beijing: Science Press, 2010(殷坤龙,张桂荣,陈丽霞,等. 滑坡灾害风险分析[M]. 北京: 科学出版社, 2010)

[3] Dou J, Yamagishi H, Pourghasemi H, et al. An Integrated Artificial Neural Network Model for the Landslide Susceptibility Assessment of Osado Island, Japan [J]. Natural Hazards, 2015, 78(3): 1 749-1 776

[4] Tan Long, Chen Guan, Zeng Runqiang, et al. Application of Artificial Neural Network in Landslide Susceptibility Assessment [J]. Journal of Lanzhou University (Natural Sciences), 2014, 50(1): 15-20 (谭龙,陈冠,曾润强,等. 人工神经网络在滑坡敏感性评价中的应用[J]. 兰州大学学报(自然科学版), 2014, 50(1): 15-20)

[5] Tehrany M, Pradhan B, Mansor S, et al. Flood Susceptibility Assessment Using GIS-based Support Vector Machine Model with Different Kernel Types [J]. Catena, 2015, 125(125): 91-101

[6] Niu Ruiqing, Peng Ling, Ye Runqing, et al. Landslide Susceptibility Assessment Based on Rough Sets and Support Vector Machine [J]. Journal of

Jilin University (Earth Science Edition), 2012, 42 (2): 430-439(牛瑞卿,彭令,叶润青,等. 基于粗糙集的支持向量机滑坡易发性评价[J]. 吉林大学学报(地球科学版), 2012, 42(2): 430-439)

[7] Tsangaratos P, Ilia I. Landslide Susceptibility Mapping Using a Modified Decision Tree Classifier in the Xanthi Perfection, Greece [J]. Landslides, 2016, 13(2): 305-320

[8] Hong H, Pradhan B, Xu C, et al. Spatial Prediction of Landslide Hazard at the Yihuang Area (China) Using Two-class Kernel Logistic Regression, Alternating Decision Tree and Support Vector Machines [J]. Catena, 2015, 133(133): 266-281

[9] Li Ting, Tian Yuan, Wu Lun, et al. Landslide Susceptibility Mapping Using Random Forest [J]. Geography and Geo-Information Science, 2014, 30 (6): 25-30(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)

[10] Liu Ben. The Research on Ensembles Method and Its Application on Remote Sensing Image Classification [D]. Dalian: Dalian University of Technology, 2014(刘贲. 集成方法研究及其在遥感分类中的应用[D]. 大连: 大连理工大学, 2014)

[11] Ozcift A, Gulten A. Classifier Ensemble Construction with Rotation Forest to Improve Medical Diagnosis Performance of Machine Learning Algorithms [J]. Computer Methods and Programs in Biomedicine, 2011, 104(3): 443-451

[12] Liu Yaqing, Lu Huijuan, Du Bangjun, et al. Study on Classifier Algorithm of Genetic Data Based on Rotation Forest [J]. Journal of China University of Metrology, 2015, 26(2): 227-231(刘亚卿,陆慧娟,杜帮俊,等. 面向基因数据分类的旋转森林算法研究[J]. 中国计量学院学报, 2015, 26(2): 227-231)

[13] Ao Pei, Li He, Zhao Sifang, et al. Analog Circuit Fault Diagnosis Based on Improved Integrated ELM [J]. Electronics World, 2014, 20(20): 89-90(敖培,李贺,赵四方,等. 基于集成改进 ELM 的模拟电路故障诊断[J]. 电子世界, 2014, 20(20): 89-90)

[14] Shao Liangshan, Ma Han. Study on Classifier Ensemble Algorithm Based on Rotation Forest [J]. Computer Engineering and Applications, 2015, 51 (23): 149-154(邵良杉,马寒. 基于旋转森林的分类器集成算法研究[J]. 计算机工程与应用, 2015, 51 (23): 149-154)

[15] Rodríguez J, Kuncheva L, Alonso C J. Rotation Forest: A New Classifier Ensemble Method [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2006, 28(10): 1 619-1 630

[16] Hu Qiufen, Wang Gang, Wang Shimei, et al. Evaluation and Regionalization of Landslide Geological Disaster in Wanzhou City, Three Gorges Reservoir Area[J]. *Journal of China Three Gorges Univ (Natural Sciences)*, 2006, 38(1): 20-25(胡秋芬, 王刚, 王世梅, 等. 三峡库区万州滑坡地质灾害危险性评价与区划[J]. 三峡大学学报(自然科学版), 2016, 38(1): 20-25)

[17] Zhang Jun, Yin Kunlong, Wang Jiajia, et al. Evaluation of Landslide Susceptibility for Wanzhou District of Three Gorges Reservoir[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2016, 35(2): 284-296(张俊, 殷坤龙, 王佳佳, 等. 三峡库区万州区滑坡灾害易发性评价研究[J]. 岩石力学与工程学报, 2016, 35(2): 284-296)

[18] Zeng Zhongping, Fu Xiaolin, Liu Xuemei, et al. Mapping and Quantitative Analysis on the Correlation Between Landslide and Geometric Alignment of Strikes of Slope and Stratum Based on GIS[J]. *Geography and Geo-Information Science*, 2006, 22(1): 22-25(曾忠平, 付小林, 刘雪梅, 等. GIS支持下滑坡斜坡类型量化及制图研究[J]. 地理与地理信息科学, 2006, 22(1): 22-25)

[19] Huang Dengqi. Regional Landslide Stability Evaluation Based on FUZZY of GIS[J]. *Prospect of Science and Technology*, 2016, 6(6): 128-129(黄登琪. 基于GIS的FUZZY在区域滑坡稳定性评价中的应用[J]. 科技展望, 2016, 6(6): 128-129)

[20] Wu Xueling, Ren Fu, Niu Ruiqing, et al. Landslide Spatial Prediction Based on Slope Units and Support Vector Machines[J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(12): 1499-1503(武雪玲, 任福, 牛瑞卿, 等. 斜坡单元支持下的滑坡易发性评价支持向量机模型[J]. 武汉大学学报·信息科学版, 2013, 38(12): 1499-1503)

[21] Kuncheva L, Rodriguez J. An Experimental Study on Rotation Forest Ensembles[M]. Heidelberg, Berlin: Springer, 2007

[22] Yu X, Wang Y, Niu R, et al. A Combination of Geographically Weighted Regression, Particle Swarm Optimization and Support Vector Machine for Landslide Susceptibility Mapping: A Case Study at Wanzhou in the Three Gorges Area, China[J]. *International Journal of Environmental Research and Public Health*, 2016, 13(5): 487-521

[23] Hanley J, Mcneil B. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve[J]. *Radiology*, 1982, 143(1): 29-36

Application of the Rotation Forest Model in Landslide Susceptibility Assessment

LIU Yuanbo¹ NIU Ruiqing¹ YU Xianyu² ZHANG Kaixiang³

- 1 Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China
2 School of Civil Engineering and Architecture, Hubei University of Technology, Wuhan 430068, China
3 Faculty of Information Engineering, China University of Geosciences, Wuhan 430074, China

Abstract: Focusing on Wanzhou region of the Three Gorges Reservoir, 29 hazard factors were extracted from the multi-source spatial data used in evaluation factors in landslide susceptibility analysis. The study area was partitioned into slope units from digital elevation model to resample the conditioning factors. A rotation forest model was trained and used to map landslide susceptibility with the best accuracy being 90.7%, according to the receiver operator characteristic (ROC) curve and area under the curve (AUC). The higher susceptibility zones were about 11.6% of the total area, and primarily distributed in the main Wanzhou city zone, and along both sides of the Yangtze River and its tributaries. The stability zones are accounted for about 45.6%, mainly distributed in the areas of low human engineering activities and high surface cover degree. The results show that application of rotation forest in the landslide susceptibility assessment exhibits both excellent prediction ability and high accuracy.

Key words: rotation forest; susceptibility assessment; ensemble learning; landslide