

引用格式: 贾倩,熊俊楠,尚依炜,肖慧文,王启盛. 一种基于堆叠集成学习的山洪危险性评估方法——以汶川地震扰动区为实证[J]. 山地学报, 2025, 43(2): 321-334.

JIA Qian, XIONG Junnan, SHANG Yiwei, XIAO Huiwen, WANG Qisheng. An innovative approach to risk assessment of flash flood based on Stacked Ensemble Learning: An empirical study of the post-shock area of the Wenchuan earthquake, China [J]. Mountain Research, 2025, 43(2): 321-334.

# 一种基于堆叠集成学习的山洪危险性评估方法 ——以汶川地震扰动区为实证

贾倩<sup>1a</sup>,熊俊楠<sup>1b\*</sup>,尚依炜<sup>1b</sup>,肖慧文<sup>2</sup>,王启盛<sup>3</sup>

(1. 西南石油大学 a. 地球科学与技术学院; b. 土木工程与测绘学院, 成都 610500;

2. 四川水发勘测设计研究院有限公司, 成都 610500; 3. 四川省第四地质大队, 成都 611130)

**摘要:** 山洪危险性评估是灾害链式防控体系的核心环节。利用机器学习算法进行山洪建模及评估山洪风险时, 单一模型普遍存在过拟合风险、泛化能力不足等缺陷。集成模型通过模型融合可提升预测性能和模型稳健性, 但其建模机理与适用性亟待深入探索与验证。本研究提出一种基于堆叠集成学习的山洪危险性评估方法, 以汶川地震重灾区为实证对象, 系统构建了包含随机森林(Random Forest, RF)、逻辑回归(Logistic Regression, LR)、梯度提升决策树(Gradient Boosting Decision tree, GBDT)和多层感知机(Multilayer Perceptron, MLP)的单一模型与堆叠集成模型, 验证集成模型在山洪危险性评估中的对比优势。主要研究发现: (1) 堆叠集成模型的准确率(90.51%)与F1分数(90.37%)显著优于单一模型, 召回率与精确率差异仅1.33%, 展现出更优的均衡性与综合性能。(2) 集成模型的AUC值达96.87%, 较RF、LR、GBDT和MLP分别提升1.55%、2.75%、2.13%和1.00%, 验证了集成策略对预测精度的提升效果。(3) 通过特征重要性分析揭示, 距水系距离(0.047)、归一化植被指数(0.029)、坡度(0.028)构成山洪灾害的关键驱动因子, 水区高差(0.027)、高程(0.017)及不透水面密度(0.016)次之。(4) 在汶川地震扰动区, 山洪极高与高危险区集中分布于东部低洼地带与西部高海拔河流沿岸。本研究可为地震灾区山洪风险识别提供创新方法, 其构建的集成模型框架可为复杂地形区灾害评估研究提供范式参考。

**关键词:** 山洪; 集成学习; 汶川地震扰动区

**中图分类号:** P954; X43

**文献标志码:** A

山洪危险性评估是灾害链式防控体系的核心环节, 其通过解译孕灾环境与致灾因子间的复杂关系, 科学识别与评估潜在风险区域, 进而为防灾规划与应急管理提供决策依据。

当前主流的山洪危险性评估模型主要分为统计学方法、水文学方法、机器学习算法三类。统计学方

法如回归分析、双变量统计和多准则决策分析<sup>[1-3]</sup>, 具可解释性优势, 尤其在结合专家经验时能够提升实用性, 但部分方法因主观赋权或人为干预, 评估结果受限于主观偏差。水文学模型通过水动力学模拟揭示洪水演进规律, 可建立山洪与致灾因子间的响应关系<sup>[4-5]</sup>, 但建模需要获取详尽的流域特征参数,

**收稿日期** (Received date): 2024-08-26; **改回日期** (Accepted date): 2025-01-23

**基金项目** (Foundation item): 国家重点研发计划专项课题(2023YFC3006701); 四川省科技厅重点研发项目(2024YFHZ0134)。[National Key R&D Program of China (2023YFC3006701); Key R&D Project of Sichuan Science and Technology Department (2024YFHZ0134)]

**作者简介** (Biography): 贾倩(1998-), 女, 湖北黄石人, 硕士研究生, 主要研究方向: 机器学习与灾害风险评估。[JIA Qian (1998-), female, born in Huangshi, Hubei Province, M. Sc. candidate, research on machine learning and disaster risk assessment] E-mail: 202221000135@stu.swpu.edu.cn

\* **通讯作者** (Corresponding author): 熊俊楠(1981-), 男, 博士, 教授, 主要研究方向: 地理信息系统与灾害风险分析。[XIONG Junnan (1981-), male, Ph. D., professor, specialized in GIS and disaster risk analysis] E-mail: neu\_xjn@163.com

且受限于计算效率,并不适合大尺度区域。

随着新型计算模型与高性能计算技术的日新月异,机器学习算法已经应用于山洪建模,以随机森林、梯度提升决策树为代表的算法展现出优越的非线性拟合能力,能高效确定山洪与潜在致灾因子间的复杂关系<sup>[6-7]</sup>。然而,这些单一模型普遍存在过拟合风险、泛化能力不足等缺陷。为了克服单一模型常见的弱点,集成模型通过模型融合策略提升预测性能和模型稳健性,常用的方法包括袋装(Bagging)、增强(Boosting)、堆叠(Stacking)和混合(Blending)。现有研究表明,基于 Bagging 和 Boosting 方法(如基于决策树的 Bagging、自适应增强算法 AdaBoost、梯度提升决策树及极限梯度提升等)的集成框架性能优于单一模型<sup>[8-9]</sup>,但堆叠集成模式因其模型架构的复杂性,在山洪风险评估中的应用案例仍较为有限,其建模机理与适用性亟待深入探索与验证。更值得关注的是,不同地理环境下机器学习模型的泛化能力存在显著差异,理论最优模型在特定区域的实际表现可能不尽如人意<sup>[10]</sup>,这凸显了区域适应性研究的必要性。

汶川地震扰动区的特殊孕灾环境是机器学习模型区域适应性研究的典型样本。四川汶川山区是洪灾高发区,1934—2007 年共记录山洪事件 108 起,震后 2008—2015 年新增 68 起,灾变频率提升 63%,震区地表破裂网络导致水文路径重组,显著提高了山洪灾害发生概率<sup>[11]</sup>。现有研究多聚焦于岷江上游单流域或省域尺度分析,缺乏多因子协同作用机制解析<sup>[12-13]</sup>,难以支撑精细化防控需求。

本研究构建多模型融合的评估框架,基于 100 m × 100 m 网格单元,选择随机森林(Random Forest, RF)、逻辑回归(Logistic Regression, LR)、梯度提升决策树(Gradient Boosting Decision Tree, GBDT)和多层感知机(Multilayer Perceptron, MLP)等算法,以汶川地震扰动区为研究样本,采用堆叠集成(Stacking)策略进行模型优化,构建堆叠集成(Stacked Ensemble)模型。通过对比分析独立模型与集成模型的预测性能,系统揭示堆叠集成方法在山洪风险评估中的有效性。研究成果不仅可为震区山洪防治提供科学依据,更重要的是构建了可推广的集成建模范式,所绘制的高精度危险性评估结果可为国土空间规划与灾害风险管理提供量化决策支持。

## 1 研究区概况

汶川地震扰动区地势复杂多样,是川西高山和成都平原交汇的过渡地带。其范围为 102°38′34″E ~ 105°38′20″E, 33°02′32″N ~ 30°44′50″N,包括理县、茂县、汶川县等 12 个市县,总面积约  $3.27 \times 10^4 \text{ km}^2$ ,海拔高度 468 ~ 5810 m。研究区位于巴颜喀拉地块的东部边界,与西南部坚硬的四川盆地相互挤压,形成一系列断层,主要包括汶川—茂县断裂、映秀—北川断裂、灌县—安县断裂和平武—青川断裂,对区内地貌形态有显著的控制作用<sup>[14]</sup>。研究区具有典型的峡谷地形特征,西部中高山较多,地势自西向东呈阶梯式递减。研究区属于亚热带湿润季风气候,年平均降雨约为 1046 mm,兼有高山垂直气候差异,季节干湿明显,夏季常有暴雨。区内河流众多,水系稠密,分布涪江、岷江、安昌河和通口河等。汶川地震扰动区洪水汇流快,成灾时间短,导致区域内山洪事件频发。研究区概况及历史山洪灾害点空间分布位置如图 1 所示。

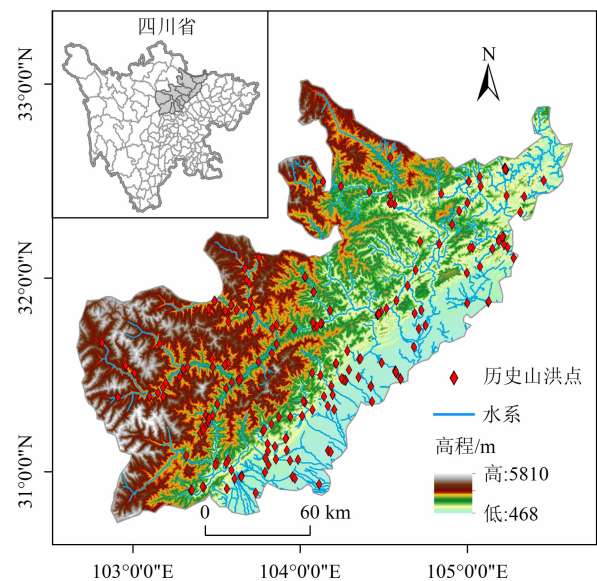


图 1 汶川地震扰动区位置图

Fig. 1 Location of the post-shock area of the Wenchuan earthquake (2008), China

(审图号:GS(2020)4619 号)

## 2 研究方法

通过分析汶川地震扰动区山洪灾害形成机制,

本研究初步选取 15 个山洪危险性评估影响因子建立山洪危险性评估指标体系,通过多重共线性分析对初选因子进行共线性检验,使用方差膨胀系数(Variance Inflation Factor, VIF)和容忍度(Tolerance, TOL)两个统计量开展多重共线性诊断,最后留下 14 个评价指标作为模型的输入数据。本研究选取了 RF、GBDT、LR、MLP 四个模型集成,分别以独立建模和采用 Stacking 方法进行训练。由于常规统计模型在处理平衡数据集时通常表现出更好的预测能力<sup>[15]</sup>,本研究从数据集中选取与正样本数量相等的负样本,构成平衡的训练数据集,对 228 个历史山洪灾害点(正样本)和等量非山洪点(负样本),按 7:3 的比例分层划分为训练样本和验证样本。在精度验证部分,通过 ROC 曲线下面积、精确率、召回率和 F1 分数比较集成模型与独立模型的性能。最后,选择最优模型对研究区的 327 万个格网进行预测,由集成模型得到的山洪危险性评估结果,技术路线如图 2 所示。

## 2.1 数据预处理

山洪灾害发生的影响因素具有多样性和复杂

性,当评价指标内部存在多重共线性问题,可能增加模型复杂度,降低运行速率,或使模型过拟合,影响模型的精度和泛化能力<sup>[16]</sup>。本文采用 VIF 值和 TOL 值作为指标衡量因子的共线程度。VIF 值为存在和不存在多重共线性两种情况下回归系数估计量的方差比值系数。当 TOL 值  $< 0.1$  且 VIF 值  $> 10$  时,表明指标变量之间多重共线性较强;反之,共线性较弱。

皮尔逊相关系数是统计学中用于量化两变量间线性关联程度的指标。该系数取值区间为  $[-1, 1]$ ,当系数趋近于  $\pm 1$  时,表明变量间存在显著线性关联;当系数接近 0 时,则显示线性相关性微弱,系数的正负号分别对应正相关与负相关关系。

## 2.2 模型构建

### 2.2.1 单个机器学习模型

(1)随机森林(RF)是基于多个决策树的非线性集成机器学习模型<sup>[17]</sup>,广泛应用于分类、回归、特征选择和异常检测等各个领域。它包含多个决策树,将输入数据不断分裂成两个独立子数据集,最后生成树结构<sup>[18]</sup>。随机森林使用投票方法处理预测

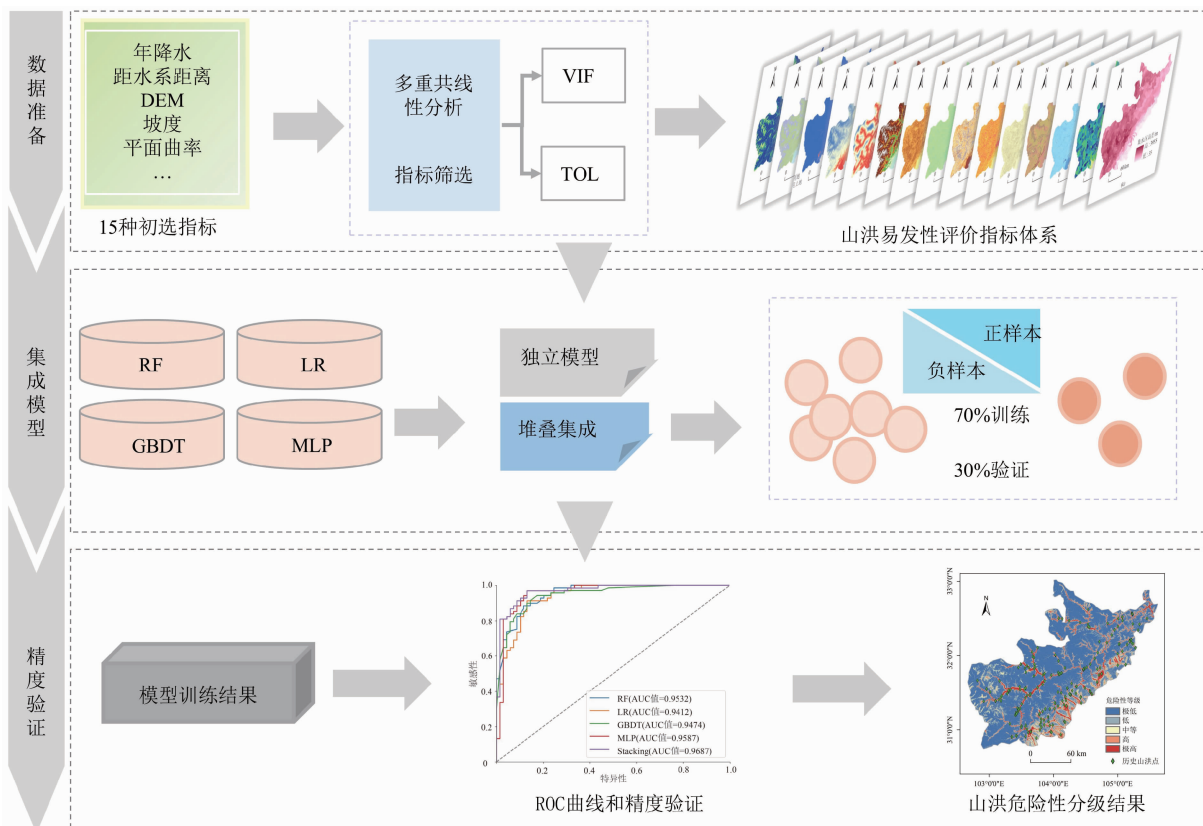


图 2 基于堆叠集成学习的山洪危险性评估方法的技术路线图

Fig. 2 Technical workflow of assessment method for flash flood hazard based on stacked ensemble learning

结果,输出次数最多的分类作为最终输出。它的主要优点是能够有效地减少过拟合,并且对缺失数据、不平衡数据和异常值具有较好的鲁棒性<sup>[19]</sup>,已应用于大量洪水危险性制图研究,并取得了较高的预测精度<sup>[20-22]</sup>。

(2)梯度提升决策树(GBDT)是一种基于集成学习的算法,通过使用多个决策树作为基础学习器来提高预测准确性<sup>[23]</sup>。GBDT 算法的核心思想是通过迭代的方式逐步减小模型的预测误差。在 GBDT 中每次迭代,上一棵树的残差被当作下一棵树的训练数据输入。新的决策树通过对这些残差进行学习,从而减少总体模型的误差。模型的损失在每次迭代中沿负梯度方向减小,从而逐步提高预测精度<sup>[24]</sup>。GBDT 具有较高的准确性,可以在建模过程中识别并纠正误差,但是 GBDT 对异常值比较敏感,在多次迭代中,GBDT 模型会尝试拟合异常值,导致过拟合的风险增加<sup>[25]</sup>。

(3)逻辑回归模型(LR)是一种起源于统计学的广义线性模型,专门用于处理二元分类问题<sup>[26]</sup>。在山洪危险性评估中,LR 是最常用的工具之一,因其计算效率高,能够有效地描述山洪发生与多种条件因素之间的线性关系。LR 通过逻辑函数将山洪事件的发生(1)或不发生(0)与相关条件因素联系起来,逻辑函数如下所示:

$$\text{logit}(P) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$

$$P = \frac{e^{\text{logit}(P)}}{1 + e^{\text{logit}(P)}} \quad (2)$$

式中, $P$  为山洪事件发生的几率; $\text{logit}(P)$  为山洪事件发生的对数几率; $\beta_0$  为截距项; $\beta_1, \beta_2, \dots, \beta_n$  为回归系数; $X_1, X_2, \dots, X_n$  为回归模型的自变量。

(4)多层感知机(MLP)是一种人工神经网络,已广泛应用于图像捕捉等领域的分类和预测<sup>[27]</sup>。MLP 采用分层架构,由输入层、若干隐含层及输出层构成,各层均包含若干计算单元,数据通过前向传播机制在层级间进行非线性变换与特征传递。MLP 通过反向传播算法调整网络权重,具有较强的非线性映射能力,可以实现对输入和输出之间映射关系的学习<sup>[28]</sup>。

### 2.2.2 堆叠集成(Stacking)模型

堆叠(Stacking)模型具有堆栈泛化思想,可以用于构建智能判识模型<sup>[29]</sup>。它将多个个体基学习器的预测结合起来,形成更强、更准确的泛化模

型<sup>[30]</sup>。训练阶段中,每个基础模型被独立训练生成预测结果,随后作为训练数据输入元模型,从而生成最终的预测(图3)。增加基础模型的数量理论上可以通过获取不同的数据特征来提高叠加模型的性能。但是,添加冗余模型可能不会提供额外的帮助,并可能导致过拟合。在本研究中,叠加框架的基础模型为 LR、RF、GBDT,元模型为 MLP。

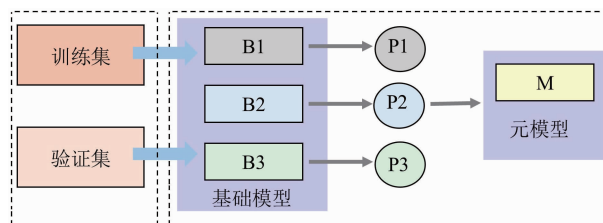


图3 堆叠集成方法原理

Fig. 3 Principle of the stacked ensemble method

## 2.3 精度评价

### 2.3.1 混淆矩阵

混淆矩阵(Confusion Matrix)是以矩阵的形式衡量分类模型的性能的方法。通过混淆矩阵计算出的准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 F1 分数(F1 - Score),非常全面地评估了模型各方面的性能,各精度指标的计算方法如下:

$$A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (3)$$

$$Q = \frac{T_p}{T_p + F_p} \quad (4)$$

$$R = \frac{T_p}{T_p + F_n} \quad (5)$$

$$F_1 = 2 \frac{Q \times R}{Q + R} \quad (6)$$

式中, $A$  为准确率; $Q$  为精确率; $R$  为召回率; $F_1$  为 F1 分数; $T_p$  和  $F_n$  为模型对山洪点的正确预测和错误预测数量; $F_p$  和  $T_n$  分别为模型在非山洪点上的错误预测和正确预测数量。准确率  $A$  是模型正确预测的样本数量占总样本数量的比例;精确率  $Q$  指的是预测为正样本中实际为正的占比;召回率  $R$  是评估模型识别出的正样本在所有正样本中的占比; $F_1$  综合了精确率与召回率的平衡。

### 2.3.2 ROC 曲线

ROC 曲线(Receiver Operating Characteristic Curve)用于图形化地表现二元分类模型的性能<sup>[31]</sup>,它展示了敏感性(False Positive Rate, FPR)与特异

性(True Positive Rate, TPR)之间的平衡。在山洪灾害危险性评估中,它的  $y$  轴和  $x$  轴分别表示敏感性和特异性。ROC 曲线与  $x$  轴和  $y$  轴包围的面积越大,即 AUC(Area Under the Curve)值越接近于 1,表示模型性能越优秀,精度越高。

## 2.4 特征重要性分析

排列重要性(Permutation Importance)是常用于评估机器学习特征重要性的方法,它的运行方式与模型的内部结构无关,而是基于自变量和因变量之间的关系。排列重要性输出的是相对重要性。它的原理是,将某个特征的值打乱,如果模型的性能上升明显,则说明该特征对预测效果非常重要,如果模型的性能提升不大或性能下降,则说明该特征对预测结果不重要。

## 3 数据来源与处理

### 3.1 数据来源

汶川地震扰动区山洪危险性评估基础数据来源如表 1 所示。

### 3.2 历史山洪灾害点

本文使用的历史山洪灾害点分为两个部分,1949—2015 年的山洪清查图来源于中国水利部和财政部全国山洪调查评价项目(National Flash Flood Investigation and Evaluation Project, NFFIEP)。2016—2022 年的灾情数据汇总来自多种渠道,包括《中国水旱灾害公报》<sup>[32]</sup>、洪涝灾情统计数据、网络收集的山洪灾害事件记录、各省市发布的水旱灾害公告和典型灾害复盘报告等<sup>[33]</sup>。

### 3.3 指标体系的构建

从地形地貌、气象水文和土地覆盖三个方面出发,预选汶川地震扰动区山洪危险性评估的 15 个因子(图 4)。气象水文因子包括年均降水量和距水系距离;地形地貌因子包含高程、坡度、平面曲率、高程标准差、地形位置指数、地形湿度指数、水流动力指数、集水区圆度、集水区高差;土地覆盖因子包含土壤类型、归一化植被指数、土地利用、不透水面密度。

气象水文因子中,降雨是诱发山洪的关键因素<sup>[34]</sup>。降雨量的大小直接影响山洪及其次生灾害的发生(图 5a)。距水系距离很大程度影响山洪的

表 1 数据来源  
Tab. 1 Data sources

数据名称	类型	分辨率	时间	数据来源	数据网址
历史山洪灾害点	矢量	—	1949—2015 年	国家山洪灾害调查与评估项目	
降水	栅格	1 km	2022 年	国家地球系统科学数据中心	<a href="https://www.geodata.cn">https://www.geodata.cn</a>
高程	栅格	30 m	2015 年	欧洲航天局	<a href="https://www.esa.int">https://www.esa.int</a>
土地利用	栅格	30 m	2022 年	中国土地利用变化数据库	<a href="http://irsip.whu.edu.cn">http://irsip.whu.edu.cn</a>
土壤	矢量	—	—	国家山洪灾害调查与评估项目	
水系	矢量	—	2022 年	OpenStreetMap	<a href="https://www.openstreetmap.org">https://www.openstreetmap.org</a>
NDVI	栅格	30 m	2022 年	国家生态科学数据中心	<a href="https://nesdc.org.cn">https://nesdc.org.cn</a>
不透水面	栅格	30 m	2022 年	全球人造不透水面数据集	<a href="https://data-starcloud.pcl.ac.cn">https://data-starcloud.pcl.ac.cn</a>

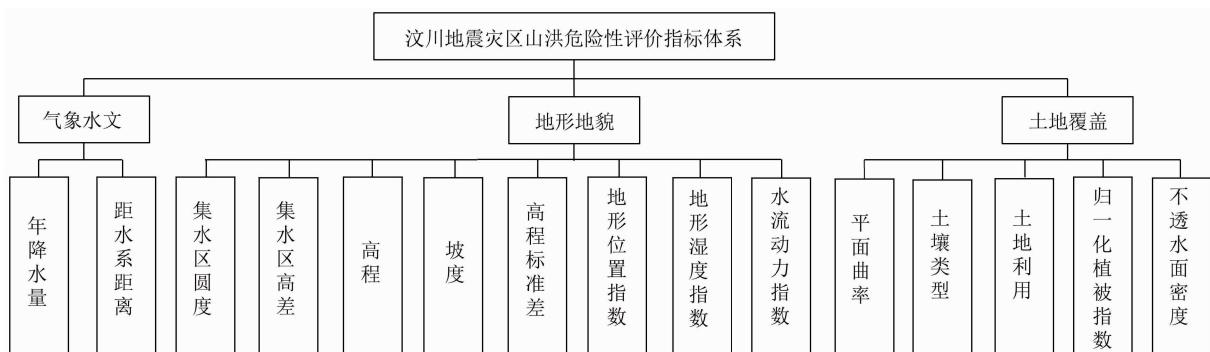


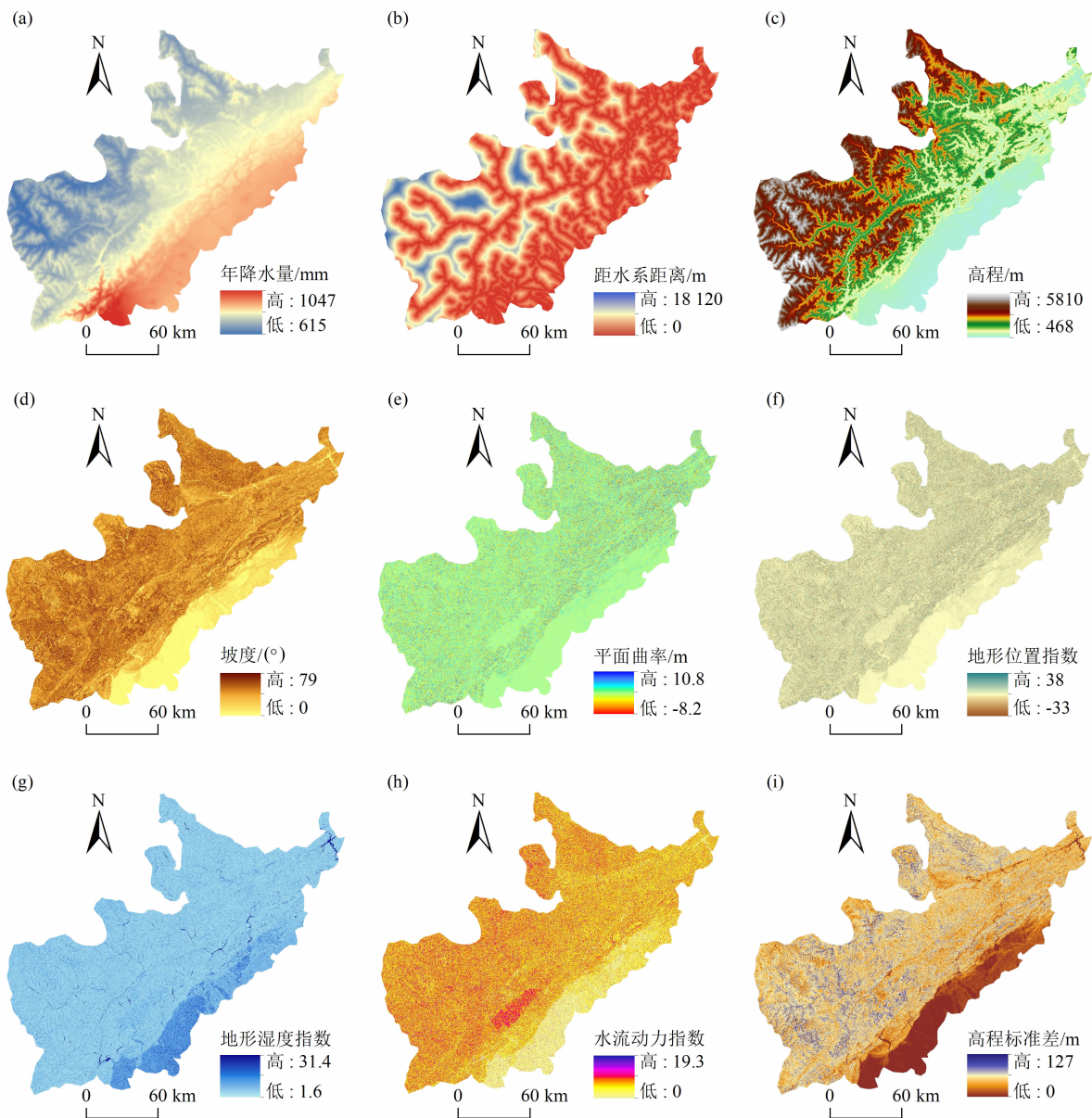
图 4 山洪灾害危险性评估指标体系

Fig. 4 Indicator system of flash flood hazard assessment

强度和发生范围,距离河流越近越容易发生山洪<sup>[35]</sup>(图5b)。

地形地貌因子中,高程是解释山洪暴发的重要因素,山洪的发生和海拔高度呈相反的趋势,海拔低的地区更容易发生洪水<sup>[36]</sup>(图5c)。坡度是影响山洪暴发的重要因素,坡度越高,土壤入渗越少,地表径流量越大<sup>[37]</sup>(图5d)。平面曲率是由数字高程模型的水平面与地表相交得到的,曲率指数有凹(正)、凸(负)、平(零)三种状态,它影响斜坡上的径流产生过程(图5e)。地形位置指数(Topographic Position Index, TPI)可表达为地貌在形状、大小和尺度上的变化<sup>[38]</sup>(图5f)。地形湿度指数(Topographic Wetness Index, TWI)反映岩石土壤的湿度,能够有

效地识别出水流最容易积聚的区域<sup>[39]</sup>(图5g)。水流动力指数(Stream Power Index, SPI)用于衡量流域内特定区域排水侵蚀功率<sup>[40]</sup>,主要用于反映流域中某一位置的水流动力强度<sup>[41]</sup>(图5h)。高程标准差反映地表起伏变化程度的重要指标,具有更高的精度,能够更好地捕捉地形的细微变化(图5i)。集水区圆度是用来衡量流域形状的变量,取值为0~1,靠近1说明集水区接近标准圆形,通常汇聚径流的时间较短,靠近0则说明集水区形状接近长条形,径流汇聚时间可能较长(图5j)。集水区高差是指集水区内最高点和最低点的差值,能够表现流域对径流形成的作用<sup>[42]</sup>(图5k)。其中坡度、平面曲率、TPI、SPI、TWI、高程标准差、集水区圆度、集水区高



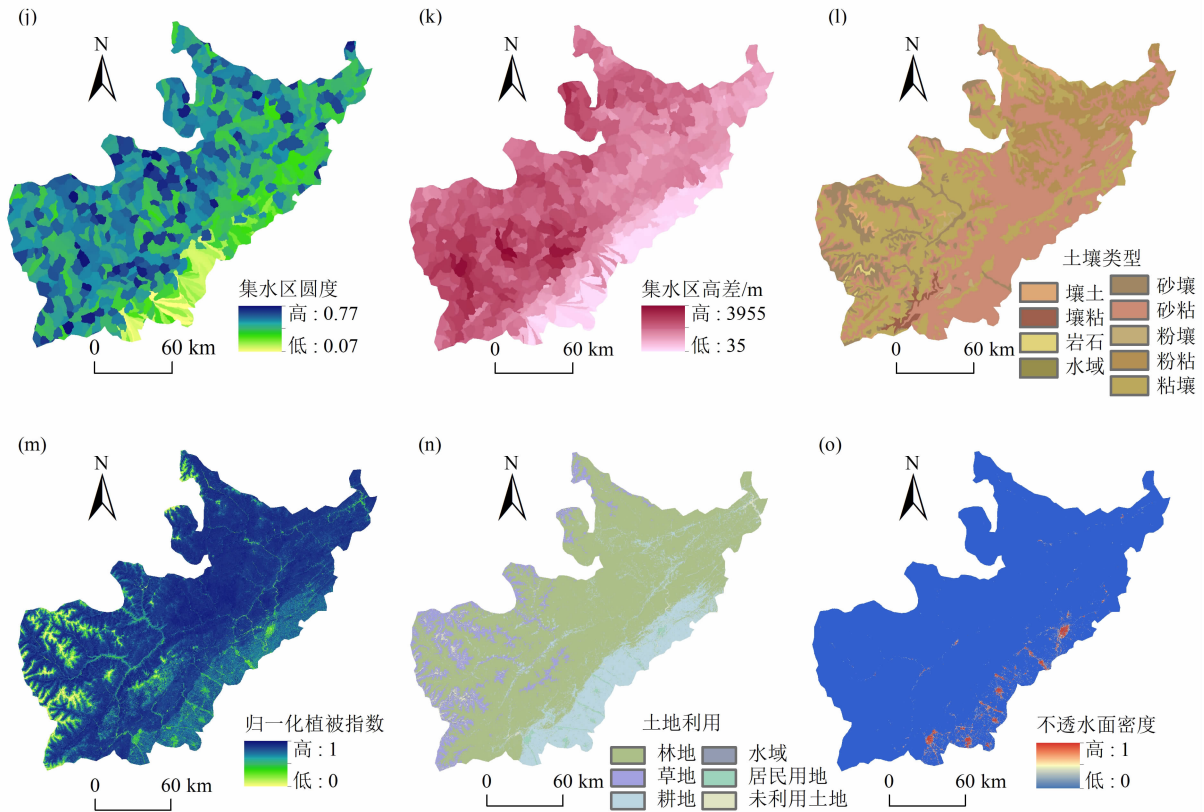


图5 山洪灾害危险性影响因子:(a)年降水量;(b)距水系距离;(c)高程;(d)坡度;(e)平面曲率;(f)地形位置指数;(g)地形湿度指数;(h)水流动力指数;(i)高程标准差;(j)集水区圆度;(k)集水区高差;(l)土壤类型;(m)归一化植被指数;(n)土地利用;(o)不透水面密度  
Fig. 5 Influencing factors of flash flood hazard: (a) annual precipitation; (b) distance from water systems; (c) elevation; (d) slope; (e) plan curvature; (f) topographic position index; (g) topographic wetness index; (h) stream power index; (i) elevation standard deviation; (j) basin circularity; (k) basin relief; (l) soil type; (m) normalized difference vegetation index; (n) land use; (o) impervious surface density

差均由高程计算得到。

土地覆盖因子中,土壤类型对山洪灾害的发生有重要影响。土壤的孔隙度和渗透性直接决定了其排水能力,沙土颗粒粗、透水性强但保水能力弱,而壤土和黏土颗粒细、透水性弱但保水能力强<sup>[43]</sup>(图5l)。植被根茎可以吸收大量水分减少地表径流汇聚,归一化植被指数越高洪水发生概率越低(图5m)。土地利用类型可以直接或间接地影响水文过程,如蒸散发、产流和入渗以及泥沙运输(图5n)。不透水面密度可用来衡量土壤的密封程度,不透水面会阻碍雨水渗入地下,导致雨水沿表面流动,增加地表径流量,并且改变水文路径、流量和流速,引发洪水(图5o)。

## 4 结果与分析

### 4.1 评价因子相关性分析

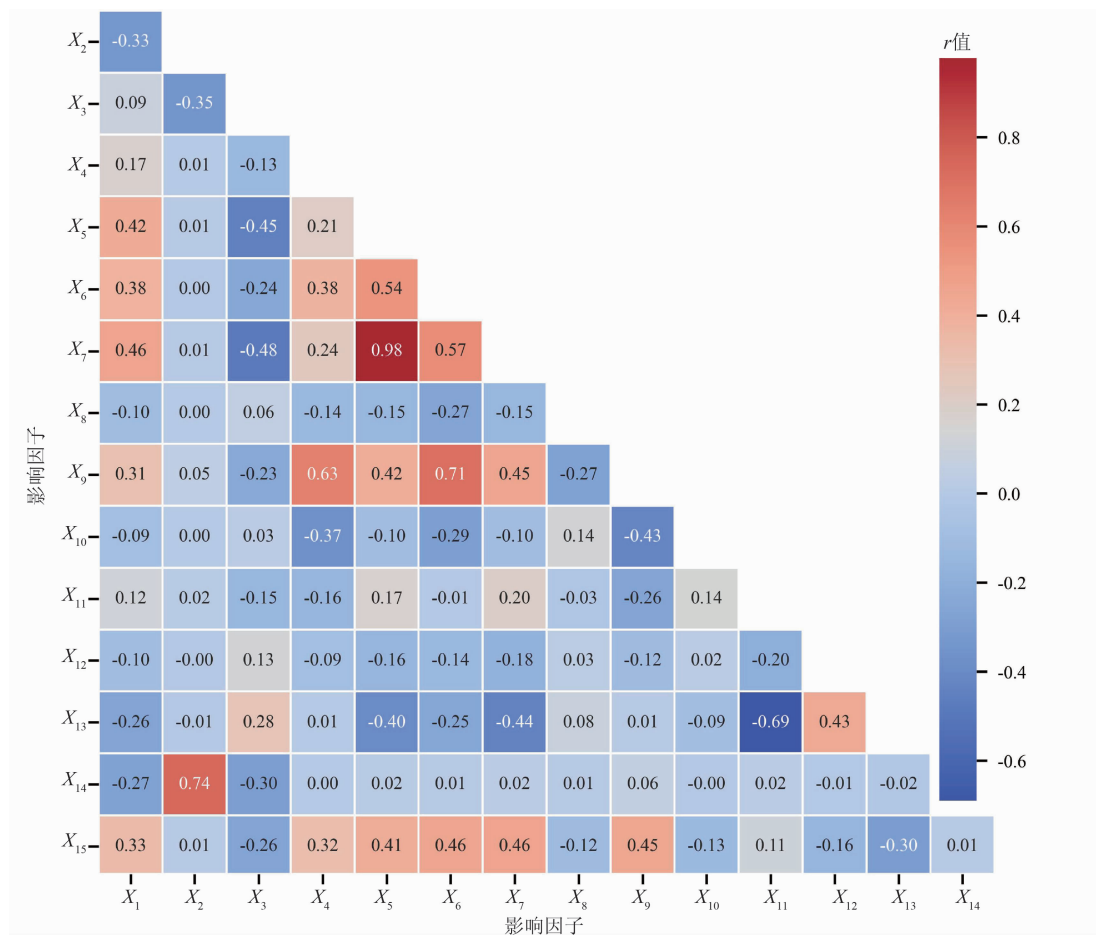
通过对特征变量多重共线性分析,可得到各评

价因子的VIF和TOL值(表2)。当TOL值 $<0.1$ 且VIF值 $>10$ 时因子间存在较大的共线性,大部分特征变量的TOL和VIF值都在可以接受的范围内。坡度的TOL和VIF值分别为0.013和79.199,高程标准差的TOL和VIF值分别为0.014和70.168,说明坡度和高程标准差存在较大的共线性。结合研究区特点,在山洪灾害风险评估中保留具有显著共线性的坡度因子。经重新检验(表2的第4、5列),所有因子的VIF值均小于10且TOL值大于0.1,满足共线性要求。

对15个初始评价指标的原始值进行皮尔逊相关性分析,分析结果如图6所示。当皮尔逊相关系数的绝对值超过0.8时,表明两个变量间存在高度相关性,其中高程标准差和坡度的相关性达到0.98,表示高程标准差与坡度之间存在极强的相关性,与共线性分析结果相同。因此在这两个相关性极强的指标中保留坡度。

表 2 影响因子的膨胀系数和容忍度  
Tab.2 Variance inflation factors and tolerance of influencing factors

变量名称	TOL 值	VIF 值	TOL 值(去除高程标准差后)	VIF 值(去除高程标准差后)
高程	0.198	5.053	0.200	4.996
坡度	0.013	76.693	0.336	2.977
平面曲率	0.641	1.560	0.642	1.558
集水区高差	0.403	2.483	0.403	2.481
集水区圆度	0.630	1.587	0.634	1.578
地形湿度指数	0.399	2.507	0.405	2.469
地形位置指数	0.575	1.738	0.578	1.729
水流动力指数	0.552	1.811	0.552	1.811
年降水	0.426	2.345	0.427	2.344
距水系距离	0.434	2.302	0.436	2.293
归一化植被指数	0.499	2.006	0.499	2.005
土壤类型	0.696	1.436	0.698	1.432
土地利用	0.377	2.650	0.378	2.642
不透水面密度	0.499	2.003	0.504	1.984
高程标准差	0.014	69.852	—	—



X<sub>1</sub> 代表水流动力指数; X<sub>2</sub> 代表地形位置指数; X<sub>3</sub> 代表地形湿度指数; X<sub>4</sub> 代表距水系距离; X<sub>5</sub> 代表高程标准差; X<sub>6</sub> 代表集水区高差; X<sub>7</sub> 代表坡度; X<sub>8</sub> 代表土壤类型; X<sub>9</sub> 代表高程; X<sub>10</sub> 代表年降水; X<sub>11</sub> 代表归一化植被指数; X<sub>12</sub> 代表不透水面密度; X<sub>13</sub> 代表土地利用; X<sub>14</sub> 代表平面曲率; X<sub>15</sub> 代表集水区圆度。

图 6 皮尔逊相关系数

Fig. 6 Pearson correlation coefficient



### 4.2 模型的精度验证与比较

根据建立的危险性评估指标和模型,分别开展汶川地震扰动区山洪危险性评估,各模型精度验证结果如图7与表3显示,五种模型具有很好的拟合精度和预测能力,AUC值均大于0.9。精确率和召回率是评估模型的重要指标,存在一种相互制衡的关系。各模型在分类性能上呈现显著差异:RF精确率和召回率相差不大,仅有1.3%,但精度小于Stacking模型;LR模型的精确率较召回率低5.07%,表明其存在对非山洪样本的误判倾向;而GBDR与MLP模型则表现出相反的偏差模式,其精确率分别较召回率高出6.66%和5.33%,反映了对山洪样本识别能力的不足;在Stacking模型中,精确率和召回率相差1.33%,这种情况得到了很好的解决,表明Stacking模型在山洪灾害危险性评估任务中表现更加均衡,具有更高的稳定性。不同模型的优势不同,Stacking利用这些模型的差异,通过互补的方式减少单个模型的误差,能够更好应对山洪灾害危险性评估任务中数据的复杂性和不确定性。

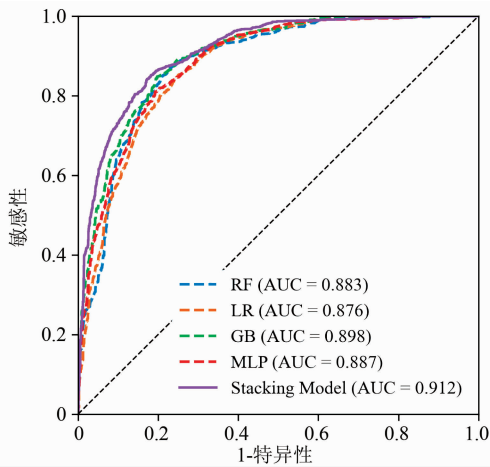


图7 ROC曲线

Fig. 7 Receiver Operating Characteristic (ROC) curve

### 4.3 基于堆叠集成模型的危险性评估

本研究利用Stacking模型得到汶川地震扰动区山洪危险性评估结果,预测的结果是每个网格被预测为1类(受影响)的预测概率,使用自然断点法将其分为五类,即极低危险区[0.034,0.114)、低危险区[0.114,0.276)、中危险区[0.276,0.500)、高危险区[0.500,0.751)、极高危险区[0.751,0.971]。研究区山洪灾害危险性等级如图8所示。

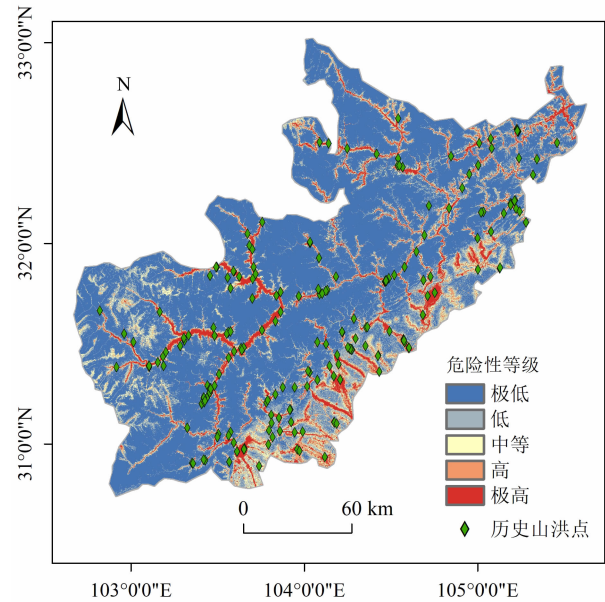


图8 基于Stacking算法的汶川地震扰动区危险性分布图

Fig. 8 Hazard distribution in the post-shock area of the Wenchuan earthquake based on Stacking algorithm

为了验证危险性评估结果的合理性,通常从两个方面进行考虑:低危险区应占据研究区域总面积的最大比例;其次,位于极高危险区的山洪点数应在所有历史山洪点中占据最高比例<sup>[44]</sup>。结果如表4所示,极高危险区面积占整个研究区面积的7.08%、高危险区占5.17%、中危险区占7.54%、

表3 各模型精度评估结果

Tab. 3 Accuracy assessment results of selected models

模型	准确率/%	精确率/%	召回率/%	F1分数/%	AUC/%
RF	87.59	88.06	86.76	87.41	95.32
LR	88.32	86.11	91.18	88.57	94.12
GBDT	87.59	90.48	83.82	87.02	94.74
MLP	88.32	90.62	85.29	87.88	95.87
Stacking	90.51	91.04	89.71	90.37	96.87

低危险区占 15.82%、极低危险区占 64.40%。在极高危险区的历史山洪灾害数为 192 个,占历史山洪灾害总数的 83.8%,山洪点密度最大,高达 0.083 个/km<sup>2</sup>,落在极低危险区的山洪点个数为 2 个,山洪点密度最小,低至 0.0001 个/km<sup>2</sup>,山洪危险性等级与山洪密度成正比。结果表明集成模型评价得到的汶川地震扰动区山洪灾害危险性结果较为合理。

表 4 危险性分区合理性检验结果

Tab. 4 Results of rationality assessment of hazard zoning

危险区	面积/km <sup>2</sup>	山洪点数/个	山洪点密度/ (个/km <sup>2</sup> )
极低	21050.41	2	0.0001
低	5171.69	10	0.0019
中	2464.60	9	0.0037
高	1688.24	15	0.0089
极高	2314.17	192	0.0830

极高和高危险区主要集中在研究区东部平原和西部的河流附近,东部的山洪危险性最高,未来发生山洪的危险性最大,这些地区人口密度较高,需要更多的投入和管理,做好抗汛减灾工作。西部的极高危险区和高危险区主要集中在河流附近的区域,这里人口密度较低,需要加强这些区域的山洪灾害预警工作。

#### 4.4 基于堆叠集成模型的特征影响评估

基于 Stacking 模型对汶川地震扰动区山洪灾害危险性影响因子的重要性进行分析,研究表明,危险性评估因子中,距水系距离(0.047)、归一化植被指数(0.029)、坡度(0.029)、水区高差(0.027)、高程(0.017)、不透水面密度(0.016)是影响山洪危险性的重要因素,水流动力指数和土壤类型对该地区山洪发生的影响较小(图 9)。距水系距离作为影响山洪发生的关键水文因子,得分最高。这表明在距河流较近的区域,由于水流的汇聚和集中,山洪发生的可能性显著增加。这与历史山洪灾害发生情况较为一致,汶川地震扰动区的山洪高发区和极高危险区大多集中在河流附近。归一化植被指数作为植被覆盖的代表指标,表明在汶川地震扰动区植被稀少的区域更容易发生山洪灾害,因为植被覆盖较少时,降水更容易形成地表径流,从而增加该地区山洪

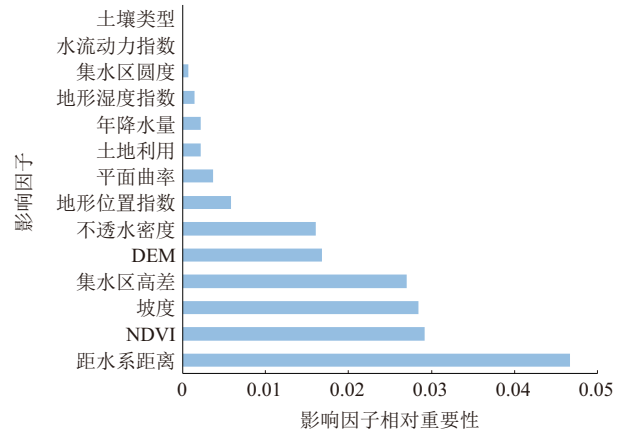


图 9 山洪危险性影响因子相对重要性

Fig. 9 Relative importance of factors contributing flash flood risk

的风险。地形因子如坡度、集水区高差和高程决定了水流的路径和流速以及汇聚情况,从而影响了山洪的发生。不透水面密度反映了城市建设发展和人类活动,该因子的重要性得分显示,人类活动对山洪灾害的发生有着显著的促进作用,尤其是在城市化进程较快的地区,不透水地面增多导致水流难以渗透,地表径流迅速增加,进一步加剧了山洪的风险。

## 5 结论与讨论

### 5.1 结论

本文考虑了三种机器学习算法、一种深度学习算法和基于这四种模型的集成算法,利用 ROC 曲线、混淆矩阵对五种算法的性能进行比较,并对汶川地震扰动区危险性分布结果进行验证,得到的结果如下。

(1) Stacking 模型在汶川地震扰动区山洪危险性评估中表现更优秀,AUC 值达到 96.87%,较单一模型提升了 1.55%、2.75%、2.13%、1.00%。

(2) 四个独立模型各有长短,出现精确率偏低或召回率偏低以及两个指标都不高情况。Stacking 模型通过互补的方式减少单个模型的误差,准确率(90.51%)和 F1 分数(90.37%)相较于独立模型有小幅提升,精确率和召回率为 91.04% 和 89.71%,相差仅 1.33%。这表明 Stacking 模型融合了四个模型各自的优点,相比于独立模型在整体上表现更加均衡,具有更好的整体性能。

(3) 汶川地震扰动区山洪灾害相对重要性较高的影响因子为距水系距离(0.047)、归一化植被指

数(0.029)、坡度(0.028)、集水区高差(0.027)、高程(0.017)和不透水面密度(0.016)。这些因子反映了水文特征、地形地貌、植被覆盖和人类活动对汶川地震扰动区山洪发生的主导作用。

(4)基于 Stacking 模型的山洪风险区划图显示,汶川地震扰动区的极高和高危险区主要集中在东部平原区域以及西部山区河流附近,这些地区需要加强山洪防御和管理工作。

## 5.2 讨论

Stacking 模型的评价精度在很大程度上取决于所选基础模型的性能和多样性。多样化的基础模型能够为 Stacking 模型提供更为丰富和互补的信息,从而提升集成模型的预测能力。然而,单个基础模型的高性能同样至关重要,只有当基础模型本身具备较强的预测能力时,Stacking 模型才能在此基础上进一步优化,表现出更优异的效果。在本研究中,所使用的四个独立模型在汶川地震扰动区的山洪危险性评估任务中均展现了较好的性能。正是由于这些基础模型的优秀表现,Stacking 模型得以进一步整合各模型的优势,达到了更高的精度。这不仅验证了 Stacking 模型在该特定区域中的有效性,也表明通过选择性能优越且互补的基础模型,可以显著提升集成模型的表现。因此,本文的研究结果进一步说明 Stacking 模型作为一种可靠的工具,在山洪危险性评估中具有较高的应用价值。

山洪是由多方面的因素导致的,山洪危险性评估工作不仅依赖模型的性能,水文气候、地形地貌、人为因素等等都是影响山洪建模的重要部分,一些相关数据的获取存在一定难度,高质量数据的可用性仍然是山洪建模的一大难点。由于本研究的重点在叠加集成模型在山洪危险性评估工作的表现,因此没有深入分析各因子与山洪之间的相关性。在未来的研究中,需要更加关注因子与山洪的相关性,以避免无关因子对结果的干扰。本文采用了四个基本模型,因此堆叠模型的学习能力受到这些基本模型的限制。未来可以探索与其他基础模型的多种组合,寻求模型性能的提升。

## 参考文献 (References)

[1] 肖扬帆. 空间多准则决策与异构空间数据集成的洪水危险性评估[D]. 武汉: 华中科技大学, 2019: 2-10. [XIAO Yangfan. Flood hazard assessment based on spatial multi-criteria decision

making and heterogeneous spatial data integration [D]. Wuhan: Huazhong University of Science and Technology, 2019: 2-10]

- [2] YOUSSEF A M, PRADHAN B, SEFRY S A. Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models [J]. *Environmental Earth Sciences*, 2016, **75**(1): 12. DOI: 10.1007/s12665-015-4830-8
- [3] CAO Yifan, JIA Hongliang, XIONG Junnan, et al. Flash flood susceptibility assessment based on geodetector, certainty factor, and logistic regression analyses in Fujian province, China [J]. *ISPRS International Journal of Geo-Information*, 2020, **9**(12): 748. DOI: 10.3390/ijgi9120748
- [4] 王雪梅, 郭良, 翟晓燕. 降雨时空不确定性对小流域洪峰模拟的影响评估[J]. *人民黄河*, 2024, **46**(4): 49-54. [WANG Xuemei, GUO Liang, ZHAI Xiaoyan. Assessing the impact of rainfall spatial-temporal uncertainty on flood peak simulation in small mountainous catchment [J]. *Yellow River*, 2024, **46**(4): 49-54] DOI: 10.3969/j.issn.1000-1379.2024.04.008
- [5] 王雨潇, 刘波, 王文鹏, 等. 基于 HEC-HMS 模型的三峡区间洪水模拟[J]. *长江科学院院报*, 2024, **41**(6): 76-83. [WANG Yuxiao, LIU Bo, WANG Wenpeng, et al. Simulation of flood in three gorges region based on HEC-HMS model [J]. *Journal of Changjiang River Scientific Research Institute*, 2024, **41**(6): 76-83] DOI: 10.11988/ckyyb.20230047
- [6] MOSAVI A, OZTURK P, CHAU K. Flood prediction using machine learning models; Literature review [J]. *Water*, 2018, **10**(11): 1536. DOI: 10.3390/w10111536
- [7] YAO Jing, ZHANG Xiaoxiang, LUO Weicong, et al. Applications of stacking/blending ensemble learning approaches for evaluating flash flood susceptibility [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, **112**: 102932. DOI: 10.1016/j.jag.2022.102932
- [8] BUI D T, TSANGARATOS P, NGO P T T, et al. Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods [J]. *Science of the Total Environment*, 2019, **668**: 1038-1054. DOI: 10.1016/j.scitotenv.2019.02.422
- [9] ARABAMERI A, SAHA S, CHEN Wei, et al. Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques [J]. *Journal of Hydrology*, 2020, **587**: 125007. DOI: 10.1016/j.jhydrol.2020.125007
- [10] CHEN Jialei, HUANG Guoru, CHEN Wenjie. Towards better flood risk management: Assessing flood risk and investigating the potential mechanism based on machine learning models [J]. *Journal of Environmental Management*, 2021, **293**: 112810. DOI: 10.1016/j.jenvman.2021.112810
- [11] 彭立, 刘邵权, 刘淑珍, 等. 汶川地震重灾区 10 县资源环境承载力研究[J]. *四川大学学报(工程科学版)*, 2009, **41**(3): 294-300. [PENG Li, LIU Shaoquan, LIU Shuzhen, et al.

- Studies on bearing capacity of resources and environment of 10 counties in area hit by Wenchuan Earthquake [J]. *Journal of Sichuan University (Engineering Science Edition)*, 2009, **41**(3): 294 - 300] DOI: 10.15961/j.jsuese.2009.03.043
- [12] 叶帮苹, 冯汉中, 刘志红, 等. 基于 Logistic 模型的四川山洪流域危险性评价[J]. *成都信息工程大学学报*, 2020, **35**(5): 573 - 578. [YE Bangping, FENG Hanzhong, LIU Zhihong, et al. Risk assessment of Sichuan mountain flooding based on logistic model [J]. *Journal of Chengdu University of Information Technology*, 2020, **35**(5): 573 - 578] DOI: 10.16836/j.cnki.jcuit.2020.05.016
- [13] 王钧, 宇岩, 欧国强, 等. 岷江上游汶川地震重灾区山洪灾害危险分区研究[J]. *长江科学院院报*, 2017, **34**(1): 54 - 60. [WANG Jun, YU Yan, OU Guoqiang, et al. Flash flood risk zoning of areas hit by Wenchuan Earthquake in the upper reach of Minjiang River [J]. *Journal of Yangtze River Scientific Research Institute*, 2017, **34**(1): 54 - 60] DOI: 10.11988/kyyb.20150956
- [14] 张圆圆, 毛爽, 张淑伟. 2000—2017 年龙门山断裂带生态脆弱性演变研究[J]. *天津农业科学*, 2020, **26**(2): 22 - 28. [ZHANG Yuanyuan, MAO Shuang, ZHANG Shuwei. Evolution of ecological vulnerability in Longmenshan fault zone from 2000 to 2017 [J]. *Tianjin Agricultural Sciences*, 2020, **26**(2): 22 - 28] DOI: 10.3969/j.issn.1006-6500.2020.02.006
- [15] 马旭廷. 基于大数据和机器学习的山洪灾害高风险区识别研究[D]. 山东: 山东建筑大学, 2024: 39 - 42. [MA Xuting. Research on high-risk area identification of flash flood disasters based on big data and machine learning [D]. Shandong: Shandong Jianzhu University, 2024: 39 - 42] DOI: 10.27273/d.cnki.gsajc.2023.000807
- [16] 石辉, 邓念东, 周阳. 随机森林赋权层次分析法的崩塌易发性评价[J]. *科学技术与工程*, 2021, **21**(25): 10613 - 10619. [SHI Hui, DENG Niandong, ZHOU Yang. Evaluation of collapse susceptibility based on random forest weighted analytic hierarchy process [J]. *Science Technology and Engineering*, 2021, **21**(25): 10613 - 10619]
- [17] 杜鹏, 陈宁生, 伍康林, 等. 基于随机森林模型的藏东南地区滑坡易发性评价及主控因素分析[J]. *成都理工大学学报(自然科学版)*, 2024, **51**(2): 328 - 344. [DU Peng, CHEN Ningsheng, WU Kanglin, et al. Evaluation of landslide susceptibility in southeast Tibet based on a random forest model [J]. *Journal of Chengdu University of Technology (Science & Technology Edition)*, 2024, **51**(2): 328 - 344] DOI: 10.3969/j.issn.1671-9727.2024.02.12
- [18] 高泽民, 丁明涛, 杨国辉, 等. 川藏铁路孜热—波密段泥石流灾害危险性评价[J]. *工程地质学报*, 2021, **29**(2): 478 - 485. [GAO Zemin, DING Mingtao, YANG Guohui, et al. Hazard assessment of debris flow along Zire-Bomi section of Sichuan-Tibet Railway [J]. *Journal of Engineering Geology*, 2021, **29**(2): 478 - 485] DOI: 10.13544/j.cnki.jeg.2021-0160
- [19] 王启盛, 熊俊楠, 程维明, 等. 耦合统计方法、机器学习模型和聚类算法的滑坡易发性评价方法[J]. *地球信息科学学报*, 2024, **26**(3): 620 - 637. [WANG Qisheng, XIONG Junnan, CHENG Weiming, et al. Landslide susceptibility mapping methods coupling with statistical methods, machine learning models and clustering algorithms [J]. *Journal of Geo-information Science*, 2024, **26**(3): 620 - 637] DOI: 10.12082/dqxxkx.2024.230427
- [20] 吴小君, 方秀琴, 任立良, 等. 基于随机森林的山洪灾害风险评估——以江西省为例[J]. *水土保持研究*, 2018, **25**(3): 142 - 149. [WU Xiaojun, FANG Xiuqin, REN Liliang, et al. Risk assessment of mountain torrents disaster based on random forest: A case study in Jiangxi province [J]. *Research of Soil and Water Conservation*, 2018, **25**(3): 142 - 149] DOI: 10.13869/j.cnki.rswc.2018.03.021
- [21] COSTACHE R, HONG Haoyuan, PHAM Q B. Comparative assessment of the flash-flood potential within small mountain catchments using bivariate statistics and their novel hybrid integration with machine learning models [J]. *Science of the Total Environment*, 2020, **711**: 134514. DOI: 10.1016/j.scitotenv.2019.134514
- [22] HOSSEINI F S, CHOUBIN B, MOSAVI A, et al. Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: Application of the simulated annealing feature selection method [J]. *Science of the Total Environment*, 2020, **711**: 135161. DOI: 10.1016/j.scitotenv.2019.135161
- [23] FRIEDMAN J H. Greedy function approximation: A gradient boosting machine [J]. *The Annals of Statistics*, 2001, **29**(5): 1189 - 1232. DOI: 10.1214/aos/1013203451
- [24] LIANG Weizhang, LUO Suizhi, ZHAO Guoyan, et al. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms [J]. *Mathematics*, 2020, **8**(5): 765. DOI: 10.3390/math8050765
- [25] 祝元丽, 冯向阳, 闫庆武, 等. 基于 GBDT 的望奎县农田土壤有机碳主控因子研究[J]. *中国环境科学*, 2024, **44**(3): 1407 - 1417. [ZHU Yuanli, FENG Xiangyang, YAN Qingwu, et al. Spatial distribution and main controlling factors of soil organic carbon under cultivated land based on GBDT model in black soil region of Northeast China [J]. *China Environmental Science*, 2024, **44**(3): 1407 - 1417] DOI: 10.19674/j.cnki.issn1000-6923.2024.0054
- [26] 范天程, 贾云飞, 李云飞, 等. 基于遥感影像与逻辑回归模型的延河流域沟壑分布概率预测[J]. *水土保持研究*, 2022, **29**(4): 316 - 321. [FAN Tiancheng, JIA Yunfei, LI Yunfei, et al. Prediction of gully distribution probability in Yanhe basin based on remote sensing image and logistic regression model [J]. *Research of Soil and Water Conservation*, 2022, **29**(4): 316 - 321] DOI: 10.13869/j.cnki.rswc.20220414.007

- [27] QI C R, SU Hao, NIESBNER M, et al. Volumetric and multi-view CNNs for object classification on 3D data [A]. 2016: 5648–5656. DOI: 10.1109/CVPR.2016.609
- [28] 曾营, 张迎宾, 张钟远, 等. 基于X-多层感知器耦合模型的滑坡易发性评价——以贵州省松桃自治县为例[J]. 山地学报, 2023, **41**(2): 280–294. [ZENG Ying, ZHANG Yingbin, ZHANG Zhongyuan, et al. Landslide susceptibility evaluation based on coupled X-multilayer perceptron model—a case study of Songtao Autonomous county of Guizhou province, China [J]. Mountain Research, 2023, **41**(2): 280–294] DOI: 10.16089/j.cnki.1008–2786.000748
- [29] 蒲未来, 刘敦龙, 桑学佳, 等. 融合多源异构数据的滑坡变形阶段智能判别方法[J]. 灾害学, 2023, **38**(4): 179–186. [PU Weilai, LIU Dunlong, SANG Xuejia, et al. An intelligent identification method of landslide deformation stage based on multi-source heterogeneous data [J]. Journal of Catastrophology, 2023, **38**(4): 179–186] DOI: 10.3969/j.issn.1000–811X.2023.04.029
- [30] SMYTH P, WOLPERT D. Linearly combining density estimators via stacking [J]. Machine Learning, 1999, **36**(1): 59–83. DOI: 10.1023/A:1007511322260
- [31] SWETS J A. Measuring the accuracy of diagnostic systems [J]. Science, 1988, **240**(4857): 1285–1293. DOI: 10.1126/science.3287615
- [32] 中华人民共和国水利部. 中国水旱灾害公报: 2016–2022 [EB/OL]. 北京: 中华人民共和国水利部, (2024-05-31) [2024-08-01]. <http://www.mwr.gov.cn/sj/tjgb/zgshzhgb/>. [Ministry of Water Resources of the People's Republic of China. China flood and drought disaster bulletin: 2016–2022 [EB/OL]. Beijing: Ministry of Water Resources of the People's Republic of China, (2024-05-31) [2024-08-01]. <http://www.mwr.gov.cn/sj/tjgb/zgshzhgb/>]
- [33] 涂勇, 吴泽斌, 何秉顺. 2011—2019年全国山洪灾害事件特征分析[J]. 中国防汛抗旱, 2020, **30**(9/10): 22–25. [TU Yong, WU Zebin, HE Bingshun. Analysis on the characteristics of flash flood disasters in China from 2011 to 2019 [J]. China Flood & Drought Management, 2020, **30**(9/10): 22–25] DOI: 10.16867/j.issn.1673–9264.2020239
- [34] 时开鑫, 陈跃红, 张晓祥, 等. 基于图聚类神经网络的江西省山洪灾害危险性区划研究[J]. 地理与地理信息科学, 2023, **39**(3): 7–15. [SHI Kaixin, CHEN Yuehong, ZHANG Xiaoxiang, et al. Flash flood hazard regionalization based on graph clustering neural network in Jiangxi province, China [J]. Geography and Geo-Information Science, 2023, **39**(3): 7–15] DOI: 10.3969/j.issn.1672–0504.2023.03.002
- [35] FERNANDEZ D S, LUTZ M A. Urban flood hazard zoning in Tucumán province, Argentina, using GIS and multicriteria decision analysis [J]. Engineering Geology, 2010, **111**(1): 90–98. DOI: 10.1016/j.enggeo.2009.12.006
- [36] CHAPI K, SINGH V P, Shirzadi A, et al. A novel hybrid artificial intelligence approach for flood susceptibility assessment [J]. Environmental Modelling & Software, 2017, **95**: 229–245. DOI: 10.1016/j.envsoft.2017.06.012
- [37] TEHRANY M S, PRADHAN B, JEBUR M N. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS [J]. Journal of Hydrology, 2013, **504**: 69–79. DOI: 10.1016/j.jhydrol.2013.09.034
- [38] MACMILLAN R A, SHARY P A. Chapter 9 Landforms and landform elements in geomorphometry [M]// HENGL T, REUTER H I. Developments in soil science. Amsterdam: Elsevier, 2009, **33**: 227–254. DOI: 10.1016/S0166–2481(08)00009–3
- [39] WANG Yi, HONG Haoyuan, CHEN Wei, et al. Flood susceptibility mapping in Dingnan county (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm [J]. Journal of Environmental Management, 2019, **247**: 712–729. DOI: 10.1016/j.jenvman.2019.06.102
- [40] POUDYAL C P, CHANG Chandong, OH H J, et al. Landslide susceptibility maps comparing frequency ratio and artificial neural networks: A case study from the Nepal Himalaya [J]. Environmental Earth Sciences, 2010, **61**(5): 1049–1064. DOI: 10.1007/s12665–009–0426–5
- [41] CAO Chen, XU Peihua, WANG Yihong, et al. Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas [J]. Sustainability, 2016, **8**(9): 948. DOI: 10.3390/su8090948
- [42] KAISER M, GUNNEMANN S, DISSE M. Regional-scale prediction of pluvial and flash flood susceptible areas using tree-based classifiers [J]. Journal of Hydrology, 2022, **612**: 128088. DOI: 10.1016/j.jhydrol.2022.128088
- [43] 祁靓雯. 基于GIS技术的紫阳县山洪灾害风险评估[D]. 西安: 长安大学, 2018: 40–41. [QI Liangwen. Mountain flood risk assessment of Ziyang county based on GIS [D]. Xi'an: Chang'an University, 2018: 40–41]
- [44] 熊俊楠, 曹依帆, 程维明, 等. 福建省山洪灾害危险性评价[J]. 山地学报, 2019, **37**(4): 538–550. [XIONG Junnan, CAO Yifan, CHENG Weiming, et al. Risk assessment of mountain torrent disasters in Fujian province, China [J]. Mountain Research, 2019, **37**(4): 538–550] DOI: 10.16089/j.cnki.1008–2786.000446

# An Innovative Approach to Risk Assessment of Flash Flood Based on Stacked Ensemble Learning: An Empirical Study of the Post-Shock Area of the Wenchuan Earthquake, China

JIA Qian<sup>1a</sup>, XIONG Junnan<sup>1b\*</sup>, SHANG Yiwei<sup>1b</sup>, XIAO Huiwen<sup>2</sup>, WANG Qisheng<sup>3</sup>

(1. *a. College of Geosciences and Technology; b. School of Civil Engineering and Geomatics, Southwest Petroleum University, Chengdu 610500, China*; 2. *Sichuan Water Development Investigation, Design & Research Co., Ltd., Chengdu 610500, China*; 3. *The 4th Geological Brigade of Sichuan, Chengdu 611130, China*)

**Abstract:** Risk assessment of flash flood is a prerequisite of flood-related geo-disaster chain prevention and control. Although machine learning algorithms were introduced for flash flood modeling to assess their susceptibility with high efficiency, but a single model generally suffer from the defects of overfitting or insufficient generalization capacity. Some integrated model improve the prediction performance and model robustness through model fusion, but its modeling mechanism and applicability need to be deeply explored and verified.

In this study, it proposed an innovative approach to flash flood risk assessment based on Stacking algorithm, and took an empirical study of the post-shock area of the Wenchuan earthquake in 2008, china. It constructed four single machine learning models (Random Forest, RF; Logistic Regression, LR; Gradient Boosting Decision Tree, GBDT; Multilayer Perceptron, MLP) and a Stacked Ensemble (SE) model which evolved from a fusion of the four machine learning algorithms. Then it systematically validated the comparative advantages of the SE model in the process of flash flood susceptibility assessment.

(1) The stacked ensemble model significantly outperformed single models in terms of accuracy (90.51%) and F1 score (90.37%), with only a 1.33% difference between recall and precision rates, demonstrating superior balance and overall performance.

(2) The AUC (Area Under the Curve) value obtained in the SE model reached 96.87%, an improvement of 1.55%, 2.75%, 2.13%, and 1.00% compared to RF, LR, GBDT, and MLP, respectively, verifying the effectiveness of the ensemble approach in enhancing prediction accuracy.

(3) Feature importance analysis revealed that distance to water systems (0.047), Normalized Difference Vegetation Index (NDVI) (0.029), and slope (0.028) were the key driving factors for flash flood occurrences, followed by water area elevation difference (0.027), elevation (0.017), and impervious surface density (0.016).

(4) In the post-shock area of Wenchuan earthquake, the extremely high and high-risk of flash flood zones were concentrated in the eastern low-lying areas and along high-altitude rivers in the west.

This study provides an innovative approach to flash flood risk identification in earthquake-affected areas, and the ensemble model framework constructed can serve as a paradigm reference for disaster assessment research in complex terrain areas.

**Key words:** flash flood; ensemble learning; the post-shock area of the Wenchuan earthquake

(责任编辑 朱颖彦 钟雨倩)