

# 基于元启发式算法优化的新乡市洪水风险评价

李红艳<sup>1,2</sup>, 郝景开<sup>1,2</sup>, 刘大为<sup>1,2</sup>, 张峰<sup>1,2</sup>, 毛立波<sup>3</sup>, 王有理<sup>4</sup>

(1. 太原理工大学环境与生态学院, 山西 晋中 030600; 2. 山西省市政工程研究生教育创新中心, 山西 晋中 030600; 3. 山西大地环境投资控股有限公司科创管理部, 山西 太原 030032; 4. 山西山安碧泉海绵城市科技有限公司, 山西 太原 030032)

**摘要:**为提高新乡市洪水风险评价模型的评估能力,采用层次分析法(AHP)、逻辑回归(LR)模型、BP神经网络、随机森林(RF)模型以及结合元启发式算法粒子群优化(PSO)的PSO-BP模型和PSO-RF模型6种方法对新乡市进行洪水风险评估,生成包含200个洪水位置的洪水清单图。选择9个洪水影响因子,采用方差膨胀因子分析了洪水影响因子与洪水发生的相关性。利用混淆矩阵和受试者工作特性曲线对比6种洪水风险评估方法的评估能力,最后获得6种方法的洪水风险分布图。结果表明:PSO-RF、PSO-BP模型的评估效果优于单一算法,受试者工作特性曲线下面积分别为0.953、0.947;根据得到的洪水风险分布图,新乡市至少36.5%的地区被列为高度易受洪水影响的地区,耦合元启发式算法后的洪水风险评价模型具有更高精度。

**关键词:**洪水风险评价模型;元启发式算法;洪水风险分布图;受试者工作特性曲线;新乡市  
**中图分类号:**TV877 **文献标志码:**A **文章编号:**1004-6933(2024)06-0104-09

**Flood risk assessment in Xinxiang City based on meta-heuristic algorithm optimization**//LI Hongyan<sup>1,2</sup>, HAO Jingkai<sup>1,2</sup>, LIU Dawei<sup>1,2</sup>, ZHANG Feng<sup>1,2</sup>, MAO Libo<sup>3</sup>, WANG Youli<sup>4</sup>(1. College of Environment and Ecology, Taiyuan University of Technology, Jinzhong 030600, China; 2. Shanxi Municipal Engineering Graduate Education Innovation Center, Jinzhong 030600, China; 3. Science and Technology Management Department, Shanxi Dadi Environment Investment Holdings Co., Ltd., Taiyuan 030032, China; 4. Shanxi Shan'an Biquan Sponge City Technology Co., Ltd., Taiyuan 030032, China)

**Abstract:** To improve the evaluation capability of the flood risk assessment model in Xinxiang City, six methods including analytic hierarchy process (AHP), logistic regression (LR) model, BP neural network, random forest (RF) model, and PSO-BP model and PSO-RF model combined with meta-heuristic algorithm particle swarm optimization (PSO) were used to conduct flood risk assessment in Xinxiang City, generating a flood inventory map containing 200 flood locations. Nine flood impact factors were selected, and the correlation between flood impact factors and flood occurrence was analyzed using variance inflation factor. The evaluation capabilities of six flood risk assessment methods were compared using confusion matrix and subject working characteristic curve, and finally obtaining the flood risk distribution maps of the six methods. The results show that the evaluation performance of PSO-RF and PSO-BP models is better than that of single algorithms, and the area under curve of the receiver operating characteristic curve is 0.953 and 0.947, respectively. According to the obtained flood risk distribution map, at least 36.5% of the areas in Xinxiang City are classified as highly susceptible to flood impacts, and the flood risk assessment model coupled with meta-heuristic algorithm has higher accuracy.

**Key words:** flood risk assessment model; meta-heuristic algorithm; flood risk distribution map; receiver operating characteristic curve; Xinxiang City

据统计,在所有类型的自然灾害中,洪水是造成最大经济损失和死亡的灾害现象<sup>[1-2]</sup>。长期以来,我国洪水灾害频发,每年都会造成大量财产损失和人员伤亡<sup>[3-4]</sup>。2021年河南省发生历史罕见特大暴雨,作物受灾面积达160万hm<sup>2</sup>,占全国受灾面积的

13.5%<sup>[5]</sup>。分析影响洪水发生的因素,对洪水易发地区进行风险评估,从而减少因洪水事件造成的损失成了从源头预防的重要措施<sup>[6-7]</sup>。洪水是一个复杂的过程,受降水事件、流域特征和自然地理条件的综合影响,洪水过程表现出较强的非线性、非平稳性和

随机性特征<sup>[8]</sup>。洪水风险分布图<sup>[9-10]</sup>能够评估未来发生洪水的概率,对于预测、预防和抵御洪水灾害具有重要意义。借助遥感和 GIS 技术<sup>[11]</sup>进行洪水风险区划分也逐渐成为洪灾风险识别的主要手段<sup>[12]</sup>。许多学者对洪水风险评估开展了研究,如 Khosravi 等<sup>[13]</sup>采用层次分析法(analytic hierarchy process, AHP)对山洪风险进行权重评估;Costache 等<sup>[14]</sup>采用模糊层次分析法来确定山洪风险;Youssef 等<sup>[15]</sup>通过频率比为灾害管理和决策提供准确评估。近年来,机器学习模型已被用于获得高精度的洪水风险分布图,如 Radmehr 等<sup>[16]</sup>使用人工神经网络代替加权过程,进行不同城市地区洪水灾害的脆弱性分析;Khosravi 等<sup>[17]</sup>比较了 4 种基于决策树的机器学习模型的洪水风险评估效果并找到了能力最强的交替决策树模型;Chen 等<sup>[18]</sup>对比分析了朴素贝叶斯树、交替决策树和随机森林(random forest, RF)模型生成的洪水风险分布图。然而,机器学习算法的超参数通常依赖于专家的主观意见,导致难以达到最佳性能。Samantaray 等<sup>[19]</sup>通过比较 BP 神经网络、支持向量机(support vector machine, SVM)模型与结合元启发算法粒子群优化(particle swarm optimization, PSO)的 PSO-SVM 模型,证明经元启发式算法改进的机器学习模型具有更高的准确性。

鉴于此,本文基于河南省新乡市历史数据,选择高程、坡度、坡向、曲率、土地利用类型、归一化植被指数(normalized difference vegetation index, NDVI)、人口密度、夜间灯光、到河流距离 9 个影响因子,结合元启发算法 PSO 对单一算法进行优化,将单一算法(AHP、逻辑回归(logistics regression, LR)模型、BP 神经网络、RF 模型)与集成模型(PSO-BP、PSO-RF)的洪水风险评估结果进行对比,建立研究区的洪水风险分布图,利用受试者工作特性(receiver operating characteristics, ROC)曲线和混淆矩阵对得到的洪水风险分布图进行分析,确定高洪水风险区域,以期新乡市应急管理部门预警监测洪水灾害、风险管理等方面提供参考。

## 1 研究区概况与数据来源

### 1.1 研究区概况

新乡市位于河南省北部,占地面积约 4530 km<sup>2</sup>,南临黄河,北依太行,地势北高南低,属温带大陆性季风气候,年均气温 15.8℃。历史上新乡市多次受到降雨引发的洪水破坏,极端天气和缺乏防洪措施极大地增加了洪水灾害的风险性和破坏性。2021 年新乡市年均降水量达 1377.7 mm,“7·20”

特大暴雨中,新乡市最大降水量达到 907 mm<sup>[5]</sup>,最强时段降水量与郑州市相当,引发的山洪造成了严重的人员伤亡和财产损失。研究区概况如图 1 所示。

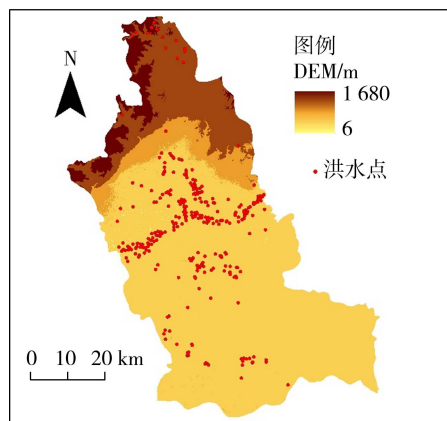


图 1 研究区概况

Fig. 1 Overview of study area

### 1.2 洪水数据清单

洪水清单数据是评估洪水风险的基础,准确细致的历史洪水数据有助于得到更加可靠的洪水风险分布图。本文采用由国家地观测科学数据中心提供的基于 GF-3 号卫星雷达影像的淹没空间分布数据集(<http://www.chinageoss.cn/cddr>),结合从新闻报道和互联网媒体中检索得到的数据,共获得 2021 年“7·20”特大暴雨中 200 个洪水灾害点,在未发生洪水事件的区域内随机选择数量相同的点作为非洪水事件数据集,将包括洪水点和非洪水点的总样本划分为 70% 的训练数据集和 30% 的测试数据集。

### 1.3 洪水影响因子

参考郑德凤等<sup>[20-21]</sup>的研究,选择高程、坡度、坡向、曲率、粗糙度、土地利用类型、降水量、NDVI、人口密度、夜间灯光、到河流的距离用于模拟研究。高程、坡度、坡向、曲率是影响洪水发生的重要地形因素,由于重力作用,洪水更易发生在低洼地区。将根据 ASTER GDEM 数据获取的全球 30 m 分辨率的 DEM 与新乡市行政区划边界裁剪并重新投影,利用 ArcGIS 提取相关地形数据,并根据自然断裂法对高程数据进行分类;坡度是洪水发生的重要因素,随着坡度的增加,渗透时间减少,径流增加,更易引发洪水;粗糙度是指地表单元的曲面面积与其在水平面上的投影面积之比;土地利用类型基于渗透性水平分类,采用由 Li 等<sup>[22]</sup>建立的中国首套 1 m 分辨率的土地覆盖数据,包括 11 个土地利用类型,研究区涉及其中 7 种(交通路线、森林、草地、耕地、人造用地、灌木、水体);降水量数据来源于欧洲中期天气

预报中心发布的 ERA5-Land 数据集;植被密度与洪水发生率呈负相关,NDVI 反映地区的植被覆盖程度,采用 NASA 发布的 MOD13 数据集,负值和 0 分别表示水和荒地,0.2~0.4 表示草地,高于 0.5 表示森林;人口密度对洪水的发生具有显著影响,人口密度大的区域往往也是城镇化进程较快的区域,地表渗透能力降低,更容易导致洪水的爆发;夜间灯光为影像亮区和暗区的比值,可以反映人类活动;到河流距离是洪水风险分析中被广泛使用的影响因素,影响洪水的范围和大小,通过欧几里得距离计算灾害点到主要河流的距离。研究区洪水影响因子数据概况如表 1 所示。

## 2 研究方法

### 2.1 单一模型

AHP 是一种被广泛应用的多准则方法<sup>[23]</sup>,通过一个范围为 1~9 的量对评估标准的相对偏好进行评分,进而构造判断矩阵,将各采样点的洪水影响因子按列进行归一化得到相对权重。求解判断矩阵的最大特征值并计算  $C_R$  值,若  $C_R < 0.1$  则认为判断矩阵满足一致性要求。通过构造判断矩阵、传递矩阵、拟优一致矩阵最终得到特征向量,即对洪水影响因子的赋权。

LR 模型是以 sigmoid 函数为基础的一种机器学习算法,将线性方程的输出转换为概率值,可以用来评估灾害发生的可能性<sup>[24]</sup>。本文中,存在洪水用 1 表示,不存在洪水用 0 表示。洪水发生概率可表示为

$$P = 1 / (1 + e^{-Z}) \quad (1)$$

其中  $Z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$   
 式中: $P$  为洪水发生概率; $Z$  为线性组合; $b_0$  为截距; $b_1, b_2, \dots, b_n$  为斜率系数; $x_1, x_2, \dots, x_n$  为洪水影响因子; $n$  为洪水影响因子数量。

BP 神经网络考虑神经元、权重和激活函数<sup>[25]</sup>,

通过反向传播算法来调整网络中每个神经元的权重和偏置<sup>[26]</sup>,以降低误差、提高网络的精度,可以用于解决分类回归、模式识别等问题。BP 神经网络由输入层、隐藏层、输出层组成。输入层用于接受原始数据特征,每个神经元对应一个特征;隐藏层由神经元组成,通过权重和激活函数传递处理数据的非线性映射关系;输出层用于输出回归问题中的一个连续值或分类问题中的类别概率分布。本文中,输入层的输入神经元为洪水影响因子,输出层的输出神经元由是否为洪水点表示。

RF 模型是一种建立在决策树算法基础上的集成算法,基于 bootstrap 重采样方法在原始数据集中获得样本,生成与原始数据集相等的子集构建决策树,通过融合所有决策树的评估或分类结果来获得最终结果<sup>[27]</sup>。RF 模型在训练过程中不会损失训练数据,对噪声和异常数据具有鲁棒性<sup>[28]</sup>,且可以在模型生成过程中取得真实误差的无偏估计。

### 2.2 集成模型

PSO 算法作为一种元启发算法,通过模拟鸟群的觅食行为来实现算法的优化,粒子通过跟踪两个极值不断迭代更新得到最优解<sup>[29]</sup>。将 PSO 算法与 BP 神经网络和 RF 模型结合,通过 PSO 寻找 BP 神经网络与 RF 模型中的最优参数,构建 PSO-BP 和 PSO-RF 集合模型,结构如图 2 所示。

### 2.3 混淆矩阵

统计学方法中,混淆矩阵可将分类问题指标可视化<sup>[30]</sup>,用来计算模型的整体准确率,即所有正确预测的样本占总样本的比例。将评估类别与真实类别构造成列联表,利用准确率、精确率、敏感性、特异性、F1 值来度量不同模型进行洪水风险评价的性能。准确率指分类正确的样本数占总样本数的比例;精确率指分类正确的样本个数占分类器判定为正样本个数的比例;敏感性指分类器评估正确的正样本占有所有正样本的比例;特异性指模型在二元分

表 1 洪水影响因子数据概况

Table 1 Overview of flood impact factor data

洪水影响因子	来源	类型	分辨率/m	年份
高程	地理空间数据云 ( <a href="https://www.gscloud.cn">https://www.gscloud.cn</a> )	栅格	30	2022
坡度	根据 DEM 数据计算	栅格	30	2022
坡向	根据 DEM 数据计算	栅格	30	2022
曲率	根据 DEM 数据计算	栅格	30	2022
到河流距离	根据 DEM 数据计算	栅格	30	2022
粗糙度	根据 DEM 数据计算	栅格	30	2022
土地利用类型	SinoLC-1 ( <a href="https://doi.org/10.5281/zenodo.7707461">https://doi.org/10.5281/zenodo.7707461</a> )	栅格	1	2022
降水量	ERA5-Land 数据集	栅格	1000	2021
NDVI	MOD13 数据集 ( <a href="https://search.earthdata.nasa.gov/search">https://search.earthdata.nasa.gov/search</a> )	栅格	250	2021
人口密度	Worldpop 人口栅格数据 ( <a href="https://www.worldpop.org">https://www.worldpop.org</a> )	栅格	1000	2021
夜间灯光	HARVARD 数据集	栅格	1000	2021

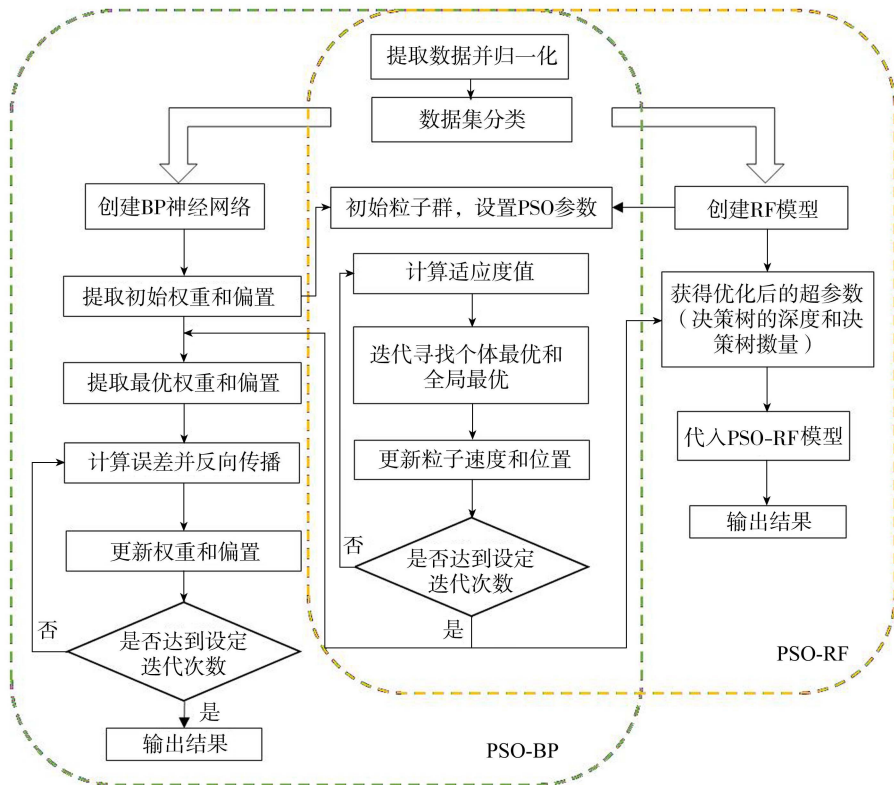


图2 PSO-BP与PSO-RF模型结构

Fig. 2 Structure of PSO-BP and PSO-RF models

类问题中区分两个类别的能力;F1 值为加权调和平均精确率和敏感性后的值。计算公式为

$$A = (T_p + T_n) / (T_p + T_n + F_p + F_n) \quad (2)$$

$$P_R = T_p / (T_p + F_p) \quad (3)$$

$$S_E = T_p / (T_p + F_n) \quad (4)$$

$$S_p = T_n / (T_n + F_p) \quad (5)$$

式中:A 为准确率;P<sub>R</sub> 为精确率;S<sub>E</sub> 为敏感性;S<sub>p</sub> 为特异性;T<sub>p</sub>、T<sub>n</sub>、F<sub>p</sub>、F<sub>n</sub> 分别为真阳性、真阴性、假阳性和假阴性样本数。

## 2.4 ROC 曲线

ROC 曲线是一种用于评价二分类模型准确性的统计工具<sup>[31]</sup>。本文将用于确定不同模型评价洪水风险的能力。曲线下面积 (area under curve, AUC) 可用于衡量不同模型在评价洪水敏感性方面的整体性能, AUC 值的范围为 0.5~1。根据 AUC 值, 可将模型评价性能进行分级: >0.5~0.6 为差、>0.6~0.7 为一般、>0.7~0.8 为好、>0.8~0.9 为非常好、>0.9~1 为极好。

## 3 结果与分析

### 3.1 洪水影响因子选择

多重共线性指变量之间存在强相关性, 会导致回归分析结果不稳定。洪水影响因子具有多样性、复杂性等特征, 为了最大限度地减少模型偏差并优

化评估准确性, 采用公差和方差膨胀因子来测试洪水影响因子之间的相关性<sup>[32]</sup>, 计算公式为

$$T_j = 1 - R_j^2 \quad (6)$$

$$V_j = 1/T_j \quad (7)$$

式中: T<sub>j</sub> 为第 j 个洪水影响因子的公差; V<sub>j</sub> 为第 j 个洪水影响因子的方差膨胀因子; R<sub>j</sub> 为第 j 个洪水影响因子作为因变量时其他影响因子作为自变量的回归拟合系数。

表 2 为 11 个洪水影响因子的公差和方差膨胀因子。由表 2 可见, 大多数洪水影响因子的方差膨胀因子为 1~2, 均低于阈值 10, 但是高程、粗糙度、降水量的方差膨胀因子大于 3。为选择更具代表性

表 2 洪水影响因子公差和方差膨胀因子

Table 2 Tolerance and variance inflation factor of flood impact factors

洪水影响因子	公差	方差膨胀因子
高程	0.202	4.962
坡度	0.471	2.123
坡向	0.947	1.056
曲率	0.977	1.024
粗糙度	0.264	3.783
土地利用类型	0.667	1.499
降水量	0.319	3.138
NDVI	0.682	1.466
人口密度	0.798	1.252
夜间灯光	0.515	1.941
到河流距离	0.512	1.951

的因素,消除具有高协方差的候选因素,经分析保留高程这个影响因子。将高程与其余 8 个洪水影响因子再次进行多重共线性分析,结果表明各因素之间不存在强相关性且方差膨胀因子均小于 3。从而确定最终 9 个洪水影响因子:高程、坡度、坡向、曲率、土地利用类型、NDVI、人口密度、夜间灯光、到河流距离,图 3 为这 9 个洪水影响因子的空间分布。

### 3.2 模型特征参数选择

利用洪水影响因子数据进行计算,得到 AHP 法中各洪水影响因子的权重,高程为 0.138、坡度为 0.212、坡向为 0.082、曲率为 0.062、土地利用类型为 0.121、NDVI 为 0.118、人口密度为 0.062、夜间灯光为 0.062、到河流距离为 0.142。

由 SPSS 软件得到 LR 模型的特征参数,其中,对洪水是否发生影响最显著的因素为高程,系数为 -0.965,洪水在平坦的低海拔地区发生较多<sup>[33]</sup>;其次为 NDVI,植被特征会影响一个区域的地表径流和

渗透能力,洪水在植被较少的地方更易发生;第三为到河流的距离,靠近河流的地区更易受到洪水的影响,与 Özay 等<sup>[34]</sup>的研究结果基本一致。去除非显著因素,保证所有影响因素的显著性均小于 0.05,具有统计学意义。最终的 LR 模型为:

$$P = \{1 + \exp[-(6.48 - 0.965x_1 - 0.124x_2 - 0.023x_3 - 0.864x_4 - 0.393x_5)]\}^{-1} \quad (4)$$

式中: $x_1$  为高程; $x_2$  为坡向; $x_3$  为土地利用类型; $x_4$  为 NDVI; $x_5$  为到河流的距离。

BP 神经网络、RF 模型与 PSO-BP 模型、PSO-RF 模型特征参数设置一致。设置种群规模为 100,迭代次数为 50,通过不断更新粒子的位置和速度寻找出最佳的组合。PSO 算法通过 MATLAB 编程语言来实现,适应度值在优化问题中用来评估解决方案的优劣,在最小化问题中,适应度值越低,个体越优。收敛过程中,PSO-BP 模型和 PSO-RF 模型分别在第 32 和第 39 次迭代时适应度值达到了最低值,分别

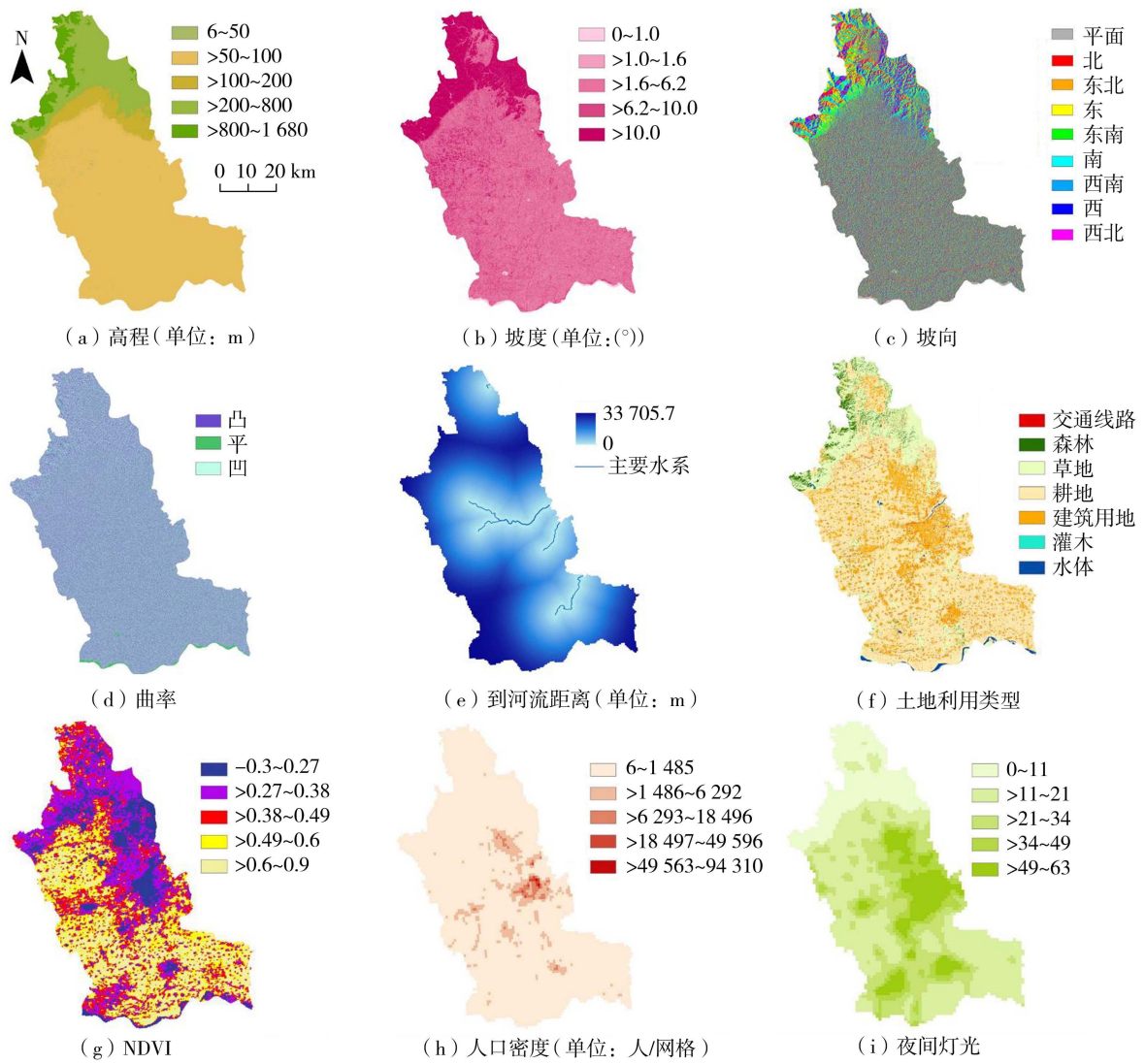


图 3 洪水影响因子的空间分布

Fig. 3 Spatial distribution of flood influencing factors

为 0.236 和 0.295,之后保持稳定,表明模型达到了最佳性能。

### 3.3 模型验证与比较

#### 3.3.1 模型评价性能分析

分别使用训练数据集和测试数据集对 AHP、LR 模型、BP 神经网络、RF 模型、PSO-BP 模型、PSO-RF 模型进行验证,结果如表 3、表 4 所示。由表 3、表 4 可见,单一模型中 BP 神经网络与 RF 模型表现出较好的拟合效果,标准误差均小于 0.03。模拟过程中,RF 模型训练集的准确率较高而测试集的准确率偏低,说明其在数据噪音较大的情况下依然出现了过拟合现象,表明不适宜的超参数会限制机器学习算法的数据学习能力,例如过于简单的网络结构会导致评估结果与实际过程产生较大偏差。与单一算法相比,引入元启发式算法的 PSO-BP 与 PSO-RF 模型能够更准确地评估洪水风险。对于非洪水点,PSO-BP 模型的特异性为 0.883,略大于 PSO-RF 模型,说明在测试集中 PSO-BP 模型将非洪水点错误判断为洪水点的概率较大,在不考虑预防洪成本的前提下对城市的安全性更为有利。

分别基于训练集和测试集绘制基于 6 种洪水风险评价模型的 ROC 曲线,如图 4 所示。由图 4 可见,训练集中,PSO-RF 模型 AUC 值最高(1.000),其次为 RF(0.994)、PSO-BP(0.953)、BP(0.946)、LR(0.892)和 AHP(0.793)模型;测试集中,PSO-RF 模型与 PSO-BP 模型仍然具有更好的评估效果,AUC 值分别为 0.953 和 0.947,RF、BP、LR、AHP 的 AUC 值分别为 0.945、0.916、0.858、0.813。综上,基于元启发式算法的 PSO-BP 模型和 PSO-RF 模型表现出更高的 AUC 值以及更好的整体性能,是准确评估洪水易发区、评价洪水风险的首选方法。

#### 3.3.2 评价结果误差分析

采用均方根误差(RMSE)来衡量 BP 神经网络、RF 模型、PSO-BP 模型、PSO-RF 模型在捕捉洪水易发区的空间模式和变化方面的有效性,结果如图 5 所示。RMSE 对最大、最小误差极为敏感,能够有效反映结果的准确性,RMSE 越小表明模型性能越好<sup>[35]</sup>。由图 5 可见,RF 模型在训练集和测试集中的 RMSE 分别为 0.2139 和 0.3145,小于 BP 神经网络(0.3092 和 0.3346),说明 RF 模型在洪水风险评

表 3 各模型训练集统计指标

Table 3 Statistical indicators of each model in training set

模型	$T_P$	$T_N$	$F_P$	$F_N$	敏感性	特异性	精确率	准确率	F1 值	标准误差	95%置信区间
AHP	116	79	61	24	0.829	0.564	0.655	0.696	0.732	0.027	0.739~0.846
LR	107	106	34	33	0.764	0.757	0.759	0.761	0.762	0.019	0.854~0.929
BP	122	126	14	18	0.871	0.900	0.897	0.886	0.884	0.013	0.921~0.970
RF	133	133	7	7	0.950	0.950	0.950	0.950	0.950	0.002	0.990~0.999
PSO-BP	120	130	10	20	0.857	0.929	0.923	0.893	0.889	0.012	0.930~0.977
PSO-RF	140	140	0	0	1	1	1	1	1	0	1

表 4 各模型测试集统计指标

Table 4 Statistical indicators for each model in test set

模型	$T_P$	$T_N$	$F_P$	$F_N$	敏感性	特异性	精确率	准确率	F1 值	标准误差	95%置信区间
AHP	55	31	29	5	0.917	0.517	0.655	0.717	0.764	0.039	0.736~0.888
LR	51	44	16	9	0.850	0.733	0.761	0.792	0.803	0.033	0.793~0.923
BP	52	51	9	8	0.871	0.900	0.897	0.886	0.884	0.027	0.863~0.969
RF	52	54	6	8	0.867	0.900	0.897	0.883	0.881	0.022	0.903~0.987
PSO-BP	52	53	7	8	0.867	0.883	0.881	0.875	0.874	0.020	0.909~0.985
PSO-RF	54	52	8	6	0.900	0.867	0.871	0.883	0.885	0.020	0.912~0.993

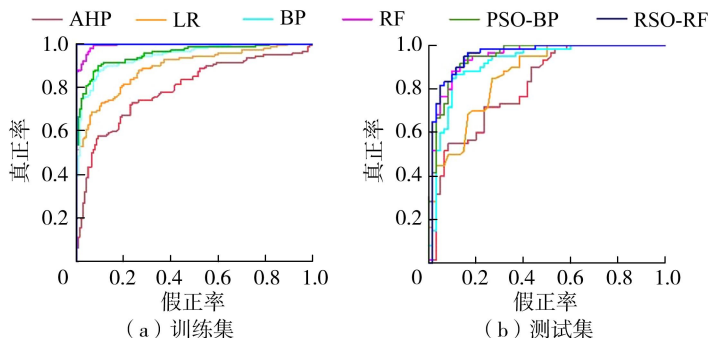


图 4 6 种洪水风险评价模型的 ROC 曲线

Fig. 4 ROC curves of 6 flood risk assessment models

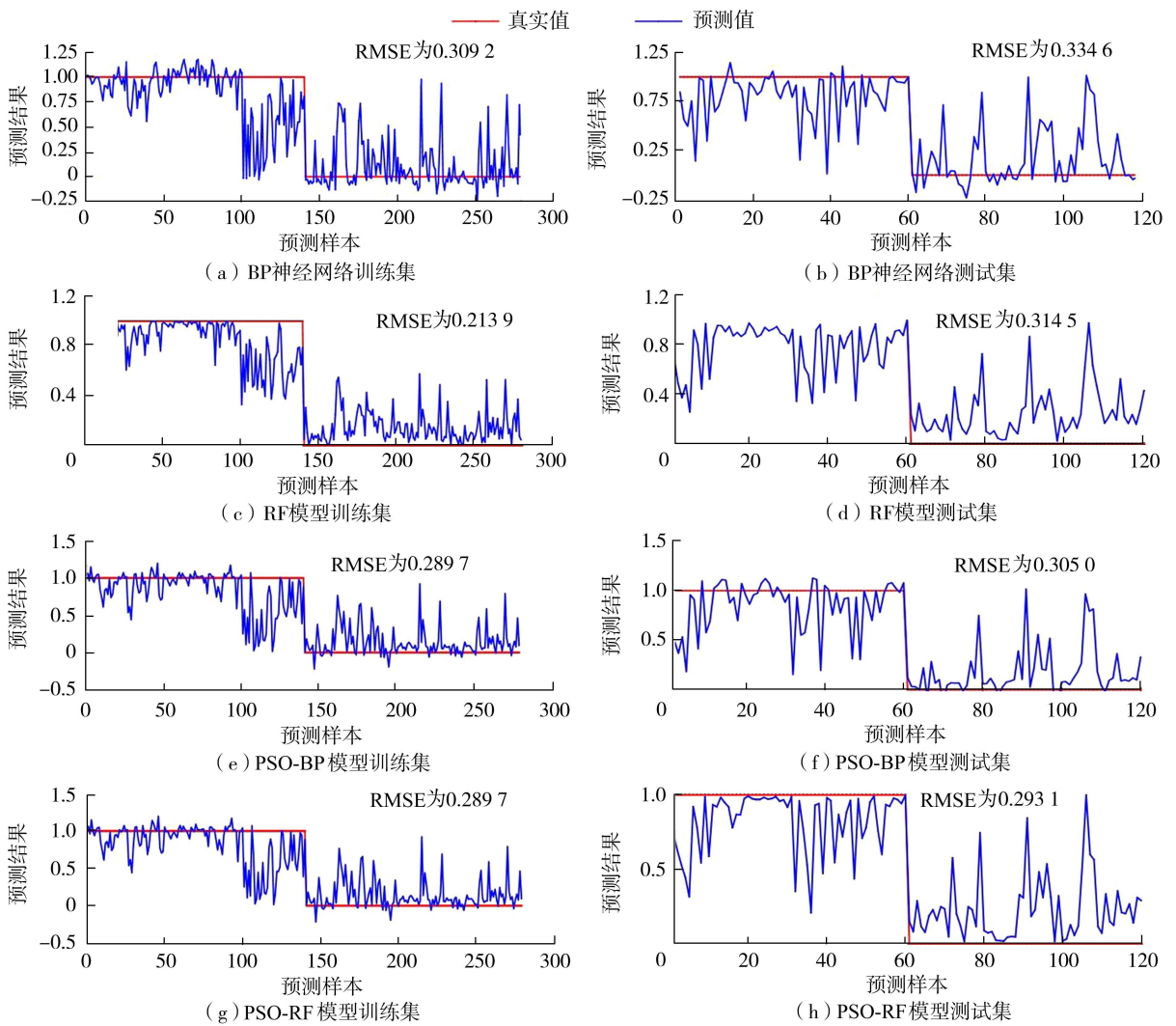


图5 4种模型在训练集和测试集中的预测结果

Fig. 5 Forecast result of 4 models in training and testing sets

估方面误差小于BP神经网络。PSO-BP模型在训练集和测试集中的RMSE分别为0.2897和0.3050, PSO-RF模型在训练集和测试集中的RMSE分别为0.1375和0.2931, 证明经PSO算法优化后的机器学习模型在洪水风险评价方面具有更好的性能, 能够更加准确地分类洪水和非洪水点。

鉴于RF模型的误差分析结果较好, 选择RF模型进行洪水影响因子的重要程度排序工作。通过MATLAB运算得到不同洪水影响因子重要性指数, 结果表明, NDVI的重要性指数为2.066, 为最显著的影响因素, 其次是到河流的距离(1.125)、高程(0.932)、夜间灯光(0.695)、人口密度(0.493)、土地利用类型(0.309)、坡度(0.308)、曲率(0.138)、坡向(-0.07)。

### 3.4 洪水风险分布图

将整个研究区划分为184259个像素, 通过自然断裂法将栅格值代表分为5个等级: 极低、低、中、高

和极高, 以此确定研究区中每个像素的洪水风险等级, 更好地实现洪水风险级别可视化。表5为6种模型评价结果中各洪水风险等级面积占比, 图6为6种模型的洪水风险分布图。由表5可见, PSO-RF模型的评价结果中, 低风险所占面积最大, 为30.68%, 其次是中(21.36%)、高(20.65%)、极高(14.32%)和极低风险(12.99%); PSO-BP模型的评价结果中, 极低、低、中等、高和极高风险的面积占

表5 6种模型评价结果中各洪水风险等级面积占比

Table 5 Proportion of each flood risk level area in evaluation results of 6 models

模型	面积占比/%				
	极低	低	中	高	极高
AHP	10.00	21.66	31.81	25.81	10.72
LR	16.95	22.75	23.80	19.95	16.55
BP	12.74	20.16	25.42	27.16	14.52
RF	21.09	27.65	21.40	16.55	13.31
PSO-BP	31.20	21.89	19.52	17.43	9.96
PSO-RF	12.99	30.68	21.36	20.65	14.32

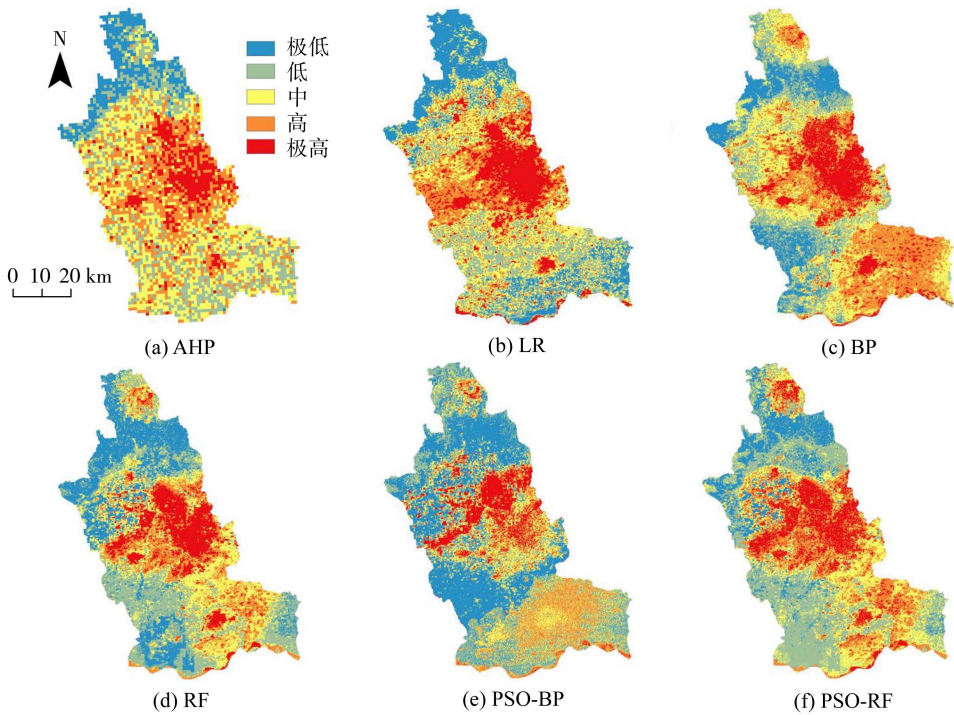


图 6 6 种模型的洪水风险分布图

Fig. 6 Flood risk distribution map of 6 models

比分别为 31.20%、21.89%、19.52%、17.43%、9.96%，具有高和极高的风险区域大多位于城市居民区；RF 模型计的评价结果中，研究区极低、低、中、高和极高风险的面占比分别为 21.09%、27.65%、21.40%、16.55%和 13.31%；BP 神经网络的评价结果中，研究区极低、低、中、高和极高风险的面占比分别为 12.74%、20.16%、25.42%、27.16% 和 14.52%。由图 6 可见，PSO-BP 模型和 PSO-RF 模型在精确定位影响洪水易感性的最关键因素方面效率较高，NDVI、到河流距离和高程是影响洪水风险最重要的因素。研究结果表明，地势平坦、低洼并且靠近河流的地区面临的洪水风险更高。

## 4 结论

a. 经元启发算法 PSO 算法优化后的模型在洪水风险评价方面具有更好的性能，能够更准确地地区分洪水点和非洪水点，并对风险等级进行更精确地评估。PSO-BP 模型与 PSO-RF 模型在测试集中的准确率、精确率、敏感性、特异性、F1 值均在 0.85 以上，标准误差均为 0.020，且具有更低的 RMSE。

b. 机器学习模型的洪水风险评价效果明显优于传统方法，且使用元启发算法优化后的超参数能够提高模型的分类精度，集成模型在训练集和测试集上的分类准确率明显高于单一模型。PSO-RF 模型、PSO-BP 模型、RF 模型、BP 神经网络、LR 模型、AHP 的 AUC 值分别为 0.953、0.947、0.945、0.916、

0.858 和 0.812，可见 PSO-BP 模型和 PSO-RF 模型具有更高的评价能力。

c. 洪水影响因子的重要程度由大到小依次为：NDVI、到河流距离、高程、夜间灯光、人口密度、土地利用类型、坡度、曲率、坡向。新乡市有至少 36.5% 的区域被列为洪水高风险区，大多位于坡度较缓、地势较为平坦的地区。

## 参考文献：

- [ 1 ] COSTACHE R, ARABAMERI A, COSTACHE I, et al. Flood susceptibility evaluation through deep learning optimizer ensembles and GIS techniques [ J ]. Journal of Environmental Management, 2022, 316: 115316.
- [ 2 ] 卢兴超, 徐宗学, 李永坤, 等. 基于多智能体模型的城市洪涝灾害动态风险评估 [ J ]. 水资源保护, 2024, 40 ( 4 ): 36-47. ( LU Xingchao, XU Zongxue, LI Yongkun, et al. Assessment on dynamic risk of urban flooding and waterlogging disaster based on multi-agent models [ J ]. Water Resources Protection, 2024, 40 ( 4 ): 36-47. ( in Chinese ) )
- [ 3 ] 王晟, 张珂, 晁丽君, 等. 基于集合模拟的汉江上游洪水与滑坡灾害风险评估 [ J ]. 水资源保护, 2023, 39 ( 6 ): 70-76. ( WANG Sheng, ZHANG Ke, CHAO Lijun, et al. Assessment of flood and landslide hazard risk based on ensemble simulation over upstream of the Han River [ J ]. Water Resources Protection, 2023, 39 ( 6 ): 70-76. ( in Chinese ) )
- [ 4 ] 任宏昌, 张恒德. 郑州“7·20”暴雨的精细化特征及主



- 要成因分析[J]. 河海大学学报(自然科学版), 2022, 50(5): 1-9. (REN Hongchang, ZHANG Hengde. Refined features and main causes of "7·20" rainstorm in Zhengzhou [J]. Journal of Hohai University (Natural Sciences), 2022, 50(5): 1-9. (in Chinese))
- [5] 国家统计局. 2022 中国统计年鉴[M]. 北京: 中国统计出版社, 2022.
- [6] 刘建芬, 王慧敏, 张行南. 城市化背景下城区洪涝灾害频发的原因及对策[J]. 河海大学学报(哲学社会科学版), 2012, 14(1): 73-75. (LIU Jianfen, WANG Huimin, ZHANG Xingnan. Causes of flood-stricken urban areas in the context of urbanization and its countermeasures [J]. Journal of Hohai University (Philosophy and Social Sciences), 2012, 14(1): 73-75. (in Chinese))
- [7] 高玉琴, 王慧, 刘钺, 等. 基于空间信息格网的南京市洪水风险评估[J]. 水利水电科技进展, 2024, 44(6): 6-12. (GAO Yuqin, WANG Hui, LIU Yue, et al. Flood disaster risk assessment in Nanjing City based on spatial information grids [J]. Advances in Science and Technology of Water Resources, 2024, 44(6): 6-12. (in Chinese))
- [8] XU Yuanhao, HU Caihong, WU Qiang, et al. Research on particle swarm optimization in LSTM neural networks for rainfall-runoff simulation[J]. Journal of Hydrology, 2022, 608: 127553.
- [9] CHOWDHURI I, PAL S C, CHAKRABORTTY R. Flood susceptibility mapping by ensemble evidential belief function and binomial logistic regression model on river basin of eastern India[J]. Advances in Space Research, 2020, 65(5): 1466-1489.
- [10] 王如锴, 练继建, 苑希民, 等. 基于 PSO-SVM 的鹤盛流域山洪风险评价[J]. 水资源保护, 2024, 40(2): 46-54. (WANG Rukai, LIAN Jijian, YUAN Ximin, et al. Risk assessment of torrential floods in the Heshengxi Watershed based on PSO-SVM [J]. Water Resources Protection, 2024, 40(2): 46-54. (in Chinese))
- [11] 王军, 梁忠民, 施晔. 基于 GIS 的水库洪水风险图编制[J]. 河海大学学报(自然科学版), 2010, 38(1): 20-25. (WANG Jun, LIANG Zhongmin, SHI Ye. Mapping of flood risk of reservoirs using GIS technology [J]. Journal of Hohai University (Natural Sciences), 2010, 38(1): 20-25. (in Chinese))
- [12] 曾子悦, 许继军, 王永强. 基于遥感空间信息的洪水风险识别与动态模拟研究进展[J]. 水科学进展, 2020, 31(3): 463-472. (ZENG Ziyue, XU Jijun, WANG Yongqiang. Advances in flood risk identification and dynamic modelling based on remote sensing spatial information [J]. Advances in Water Science, 2020, 31(3): 463-472. (in Chinese))
- [13] KHOSRAVI K, NOHANI E, MAROUFINIA E, et al. A GIS-based flood susceptibility assessment and its mapping in Iran; a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique[J]. Natural Hazards, 2016, 83(2): 947-987.
- [14] COSTACHE R, BARBULESCU A, PHAM Q B. Integrated framework for detecting the areas prone to flooding generated by flash-floods in small river catchments[J]. Water, 2021, 13(6): 758.
- [15] YOUSSEF A M, PRADHAN B, JEBUR M N, et al. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia[J]. Environmental Earth Sciences, 2015, 73(7): 3745-3761.
- [16] RADMEHR A, ARAGHINEJAD S. Developing strategies for urban flood management of Tehran city using SMCDM and ANN[J]. Journal of Computing in Civil Engineering, 2014, 28(6): 05014006.
- [17] KHOSRAVI K, PHAM B T, CHAPI K, et al. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran[J]. Science of the Total Environment, 2018, 627: 744-755.
- [18] CHEN Wei, LI Yang, XUE Weifeng, et al. Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods [J]. Science of the Total Environment, 2020, 701: 134979.
- [19] SAMANTARAY S, SAHOO A, AGNIHOTRI A. Prediction of flood discharge using hybrid PSO-SVM algorithm in Barak River Basin[J]. MethodsX, 2023, 10: 102060.
- [20] 郑德凤, 高敏, 李钰, 等. 基于 GIS 的大连市暴雨洪涝灾害综合风险评估[J]. 河海大学学报(自然科学版), 2022, 50(3): 1-8. (ZHENG Defeng, GAO Min, LI Yu, et al. Comprehensive risk assessment of rainstorm-flood disaster in Dalian City based on GIS[J]. Journal of Hohai University (Natural Sciences), 2022, 50(3): 1-8. (in Chinese))
- [21] 刘高峰, 龚艳冰, 黄晶. 基于流域系统视角的城市洪水风险综合管理弹性策略研究[J]. 河海大学学报(哲学社会科学版), 2020, 22(3): 66-73. (LIU Gaofeng, GONG Yanbing, HUANG Jing. Research on resilient strategies of urban flood risk comprehensive management from the perspective of river basin system [J]. Journal of Hohai University (Philosophy and Social Sciences), 2020, 22(3): 66-73. (in Chinese))
- [22] LI Zhuohong, HE Wei, CHENG Mofan, et al. SinoLC-1: the first 1 m resolution national-scale land-cover map of China created with a deep learning framework and open-access data [J]. Earth System Science Data, 2023, 15(11): 4749-4780.

- Sciences), 2023, 51(1): 1-9. (in Chinese))
- [25] ZHAO Renjun. The Xinanjiang model applied in China [J]. Journal of Hydrology, 1992, 135(1/2/3/4): 371-381.
- [26] 陆旻皎. 新安江模型研究的回顾和展望[J]. 水利学报, 2021, 52(4): 432-441. (LU Minjiao. Recent and future studies of the Xinanjiang model[J]. Journal of Hydraulic Engineering, 2021, 52(4): 432-441. (in Chinese))
- [27] 吴勇拓, 李致家, 戚振亚, 等. 基于水文模型的缺资料流域设计洪水计算[J]. 河海大学学报(自然科学版), 2023, 51(6): 1-8. (WU Yongtuo, LI Zhijia, QI Zhenya, et al. Design flood calculation of watershed with lack of data based on hydrological model [J]. Journal of Hohai University (Natural Sciences), 2023, 51(6): 1-8. (in Chinese))
- [28] ZHAO Pengcheng, WANG Qianjun, WU Wenyan, et al. Spatial-mode-based calibration (SMoC) of forecast precipitation fields with spatially correlated structures: an extended evaluation and comparison with gridcell-by-gridcell postprocessing [J]. Journal of Hydrometeorology, 2023, 24(9): 1509-1525.
- [29] 董磊华, 熊立华, 万民. 基于贝叶斯模型加权平均方法的水文模型不确定性分析[J]. 水利学报, 2011, 42(9): 1065-1074. (DONG Leihua, XIONG Lihua, WAN Min. Uncertainty analysis of hydrological modeling using the Bayesian model averaging method [J]. Journal of Hydraulic Engineering, 2011, 42(9): 1065-1074. (in Chinese))

(收稿日期: 2024-01-02 编辑: 施业)

(上接第 112 页)

- [23] ALMANZI K Y, SHARIFF A R M, KALANTAR B, et al. Performance evaluation of hospital site suitability using multilayer perceptron (MLP) and analytical hierarchy process (AHP) models in malacca, malaysia [J]. Sustainability, 2022, 14(7): 3731.
- [24] ZENG Z P, LAN J Y, HAMIDI A R, et al. Integrating Internet media into urban flooding susceptibility assessment: a case study in China [J]. Cities, 2020, 101: 102697.
- [25] NANEHKARAN Y A, ZHU L C, CHEN J D, et al. Application of artificial neural networks and geographic information system to provide hazard susceptibility maps for rockfall failures [J]. Environmental Earth Sciences, 2022, 81(19): 475.
- [26] HU Caihong, ZHANG Xueli, LI Changqing, et al. Real-time flood classification forecasting based on k-means++ clustering and neural network [J]. Water Resources Management, 2022, 36(1): 103-117.
- [27] RAZAVI-TERMEH S V, SADEGHI-NIARAKI A, SEO M, et al. Application of genetic algorithm in optimization parallel ensemble-based machine learning algorithms to flood susceptibility mapping using radar satellite imagery [J]. Science of the Total Environment, 2023, 873: 162285.
- [28] RAZAVI-TERMEH S V, SADEGHI-NIARAKI A, CHOI S M. A new approach based on biology-inspired metaheuristic algorithms in combination with random forest to enhance the flood susceptibility mapping [J]. Journal of Environmental Management, 2023, 345: 118790.
- [29] PARYANI S, BORDBAR M, JUN C, et al. Hybrid-based approaches for the flood susceptibility prediction of Kermanshah Province, Iran [J]. Natural Hazards, 2023, 116(1): 837-868.
- [30] 黄钦, 谭翠, 杨波. 基于 XGBoost 算法的亚热带地区生态旅游适宜性评价方法研究 [J]. 地球信息科学学报, 2024, 26(2): 303-317. (HUANG Qin, TAN Cui, YANG Bo. Research on the evaluation method of ecotourism suitability in subtropical regions based on XGBoost algorithm [J]. Journal of Geo-Information Science, 2024, 26(2): 303-317. (in Chinese))
- [31] LIU Jun, XIONG Junnan, CHEN Yangbo, et al. A new avenue to improve the performance of integrated modeling for flash flood susceptibility assessment: applying cluster algorithms [J]. Ecological Indicators, 2023, 146: 109785.
- [32] 徐奎, 潘昊, 宾零陵, 等. 基于强化学习的海南岛山洪灾害易发性评估 [J]. 水资源保护, 2023, 39(2): 95-100. (XU Kui, PAN Hao, BIN Lingling, et al. Evaluation of mountain torrent disaster vulnerability in Hainan Island based on reinforcement learning [J]. Water Resources Protection, 2023, 39(2): 95-100. (in Chinese))
- [33] TIEN BUI D, HOANG N D, MARTÍNEZ-ÁLVAREZ F, et al. A novel deep learning neural network approach for predicting flash flood susceptibility: a case study at a high frequency tropical storm area [J]. Science of the Total Environment, 2020, 701: 134413.
- [34] ÖZAY B, ORHAN O. Flood susceptibility mapping by best-worst and logistic regression methods in Mersin, Turkey [J]. Environmental Science and Pollution Research, 2023, 30(15): 45151-45170.
- [35] 赵二峰, 李章寅, 袁冬阳. 基于双阶段注意力机制的大坝变形深度学习预测模型 [J]. 河海大学学报(自然科学版), 2023, 51(6): 44-52. (ZHAO Erfeng, LI Zhangyin, YUAN Dongyang. Deep learning model for deformation prediction of dam based on dual-stage attention mechanism [J]. Journal of Hohai University (Natural Sciences), 2023, 51(6): 44-52. (in Chinese)).

(收稿日期: 2024-04-11 编辑: 王芳)